IRSA 2025

TAGUNGSBAND
PROCEEDINGS



5th International Railway Symposium Aachen

Aachen, Germany

19 - 20 November 2025

VERANSTALTER / ORGANISER

Eurail press

PARTNER

Research Center Railways

RWTH AACHEN UNIVERSITY

Herausgeber / *Editors*

Univ.-Prof. Dr.-Ing. Nießen, Nils
Verkehrswissenschaftliches Institut (VIA) /
Institute of Transport Science

Mies-van-der-Rohe-Straße 1
52074 Aachen, Germany
WWW: http://www.via.rwth-aachen.de
E-Mail: via@via.rwth-aachen.de

Univ.-Prof. Dr.-Ing. Schindler, Christian
Univ.-Prof. Dr. Pfaff, Raphael
Institut für Schienenfahrzeuge (IFS) /
Institute of Rail Vehicles

Seffenter Weg 8
52074 Aachen, Germany
WWW: http://www.ifs.rwth-aachen.de
E-Mail: sekretariat@ifs.rwth-aachen.de

DVV Media Group GmbH
Heidenkampsweg 75
20097 Hamburg, Germany
WWW: http://www.dvvmedia.com
E-Mail: info@dvvmedia.com

Verfügbar über das institutionelle Repositorium
der RWTH Aachen University /
Available via the institutional repository
of RWTH Aachen University
DOI: 10.18154/RWTH-2026-00357

Publikationsserver der Universitätsbibliothek /
Publication Server of the University Library
RWTH Aachen University

Templergraben 61
52062 Aachen Germany
WWW: http://www.ub.rwth-aachen.de

Aachen, Germany, January 2026

Inhalt / Contents

Session 1

KAPAZITÄT I
RAILWAY CAPACITY I

11 **Digitalisierung im Kapazitätsmanagement der Eisenbahn: Herausforderungen und Erwartungen**
Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

31 **Servicebasierte Infrastrukturanpassung zur Optimierung von EBW- und (LST-) Planungsprozessen**
Alexander Kuckelberg

44 **Approximation von Bediensystemen zur Leistungsfähigkeitsbestimmung von Eisenbahninfrastruktur in Bahnhöfen**
Tamme Emunds, Nils Nießen

65 **Capacity and Quality Assessment of Track Groups in Railway Nodes**
Sirikorn Satitrattanashewin, Jan Eisold, Beate Lux

Session 2

KAPAZITÄT II
RAILWAY CAPACITY II

84 **Targeted infrastructure investments for more efficient use of railway capacity by different market segments**
Harald Buschbacher

100 **Evaluating network capacity effects of ETCS L2 in a timetable-based capacity assessment model**
Cédric Kekes, Franziska Theurich, Karl Nachtigall

116 **Railway Infrastructure Capacity Evaluation of Modular Rail Pods under ETCS Level 2 Signalling System**
Zheng Ning, Egidio Quaglietta, Mahnam Saeednia

Session 3

BAHNSTEIGE
RAILWAY PLATFORMS

133 **Rahmenbedingungen für eine Implementierung von Bahnsteigbarrieren und Zugangskontrollen im deutschen Vollbahnsystem**
Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

149 **Der Höhenunterschied zwischen Fahrzeug und Bahnsteigkante. Erfassung der Einstiegssituationen im deutschen SPNV.**
Sebastian Herwartz-Polster, Aaron Paz Martinez, Johannes Pagenkopf, Marc Sparsbrod, Felix Heizler

Session 4

GÜTERVERKEHR
FREIGHT TRAFFIC

166 **Baukasten für einen zukunftsfähigen Schienengüterverkehr in der Fläche**
Manfred Enning, Bernd Schmidt

181 **DAK: Quantensprung für den Schienengüterverkehr oder technologisch Sackgasse**
Kurt Fuchs

Session 5

ZUSTANDSÜBERWACHUNG I
CONDITION MONITORING I

192 **Entwicklung eines Zustandsüberwachungssystems für Losradachsen**
Mirco Janßen, Sven Jenne, Martin Möhring, Pascal Gil, Henning Seeger, Christian Schindler

205 **Effizientes EN 50318-konformes Simulationsmodell der Interaktion zwischen Pantographen und Oberleitung**
Matthias Zelinka, Martin Joch, Alexander Schirrer, Stefan Jakubek

221 **Entwicklung eines integrierten digitalen Zwillings auf Basis von Open Data für die Gleisinstandhaltung mittels Smartphone Sensorik**
Philipp Leibner, Jannik Goersch, Benedikt Neubauer, Thomas Hempel, Christian Schindler, Raphael Pfaff

237 **Wheel/rail adhesion-based support of automated train operation (ATO)**
Marcus Fischer, Felix Kröger, Thomas Rasel, Norman Kreisel, Sebastian Stumm

Session 6

ASSISTIERTES UND AUTOMATISIERTES FAHREN
ASSISTED AND AUTOMATED DRIVING

256 **Vorentwicklung kamerabasierter onsight-ATO-Funktionen – Schienenkopfund Weichenlagendetektion sowie infrastrukturseitige Unterstützung der Lichtraumprofilüberwachung bei geschobenen Einheiten im skalierten Modell**
Tobias Hofmeier, Kilian Otto, Yan Cao, Martin Cichon

272 **Hybride 5G Echtzeitübertragung in der Anschlussbahn im Kontext von SAMIRA 2.0**
Klaus Mößner, Sam Münchow, Charbel Lahoud

Session 7

RAD-SCHIENE INTERAKTION
WHEEL-RAIL INTERACTION

288 **Vorstellung einer neuen Radform zur Verminderung thermisch verursachter Zugeigenspannungen und deren möglichen Folgen im Radkranz klotzgebremster Räder**
Sören Barteldes, Andreas Bittner, Franz Murawa

302 **Kontaktlose Messung von Rad-Schiene Kräften auf Schienenfahrzeugen**
Jakob Moder, Erik Gerstl

Session 8

FAHRPLANANALYSE & -VERBESSERUNG
TIMETABLE ANALYSIS AND IMPROVEMENT

313 **Automated Delay Detection in Railway Networks**
Thomas Graffagnino

328 **Estimation of Train Dwell Times using a Data-Driven Approach**
Takuya Kaneko, Toshihiro Hiura, Kazushige Yonemoto, Kiyoshi Narita, Hideyuki Yabuki, Shuho Watanabe, Norio Tomii

345 **Improving operational quality through global optimization of reserve and buffer times**
Frédéric Weymann, David Janecek, Katrien Pauwels, Anneleen Van de Mergel, Quentin Van Leeckwyck

Session 9

FAHRKOMFORT
RIDE COMFORT

361 **The sensitivity of simulated ride comfort of the tail coach in a high-speed EMU to unsteady aerodynamic loading and variation in suspension parameters**
Guozhen Jing, Hauke Schmidt, Xin Ding, Chutan Xue, Gang Chen

378 **Active vertical secondary suspension using low power actuator and variable damper for railway vehicles**
Yoshiki Sugahara

Session 10

ZUSTANDSÜBERWACHUNG II
CONDITION MONITORING II

395 **Landing data science and AI in condition monitoring for bogie components and track**
Yan Niu, Dietmar Kraft, Kai Heimann

Session 11

FAHRZEUG UND FAHRWEG
VEHICLE AND CHASSIS

412 **Optimierung eines selbststeuernden Einzelachsfahrwerks mittels eines genetischen Algorithmus**
Stefan Lipinski, Gustav Hellings, Christian Schindler

429 **Herausforderungen des Engineerings für das Redesign/Refurbishment der S-Bahn Rheinland (BR422 und BR423)**
Oliver Hagemann, Martin Kuhlins, Leon Beckmann

444 **Digitales Instandhaltungsregelwerk bei der Deutschen Bahn, Personenverkehr**
Margit Olbrich, Toni Schnell

Session 12

SICHERHEIT + BAHNÜBERGANG
SAFETY + LEVEL CROSSING

459 **Konzepte zur netzunabhängigen Energieversorgung für eine alternative Sicherung von Bahnübergängen**
Carolin Lang, Lukas Leicht, Patrick Ziesel, Aaron Zdarsky, Martin Cichon

476 **Security-Management: Risiko-orientierte Bewertung von physischen und Cyber-Angriffen auf das Eisenbahnsystem**
Matthias Jelinski, Walter Schmitz, Stefan Pickl

496 **Development and field experiment of the safety assisting system for passive level crossings**
Shota Miyoshi, Makoto Yagi, Daisuke Yamaguchi, Tomonori Hasegawa

Session 13

INFRASTRUKTURZUSTAND I
INFRASTRUCTURE CONDITION I

510 **KI-Anwendungen zur regelmäßigen Überwachung der Eisenbahninfrastruktur**
Kira Zschiesche, Johannes Pelz, Alexander Reiterer

537 **Quantifizierung der Wechselwirkungen von Instandhaltungs- und Erneuerungsmaßnahmen auf den Zustand des Schienennetzes**
Felix Lampe, Felix Doose, Christian Schaper, Björn Dickenbrok

553 **Bedarfsermittlung bei der DB InfraGO AG – Chancen und Herausforderungen der Integration des Anlagenzustands**
Sebastian Ude, Alexander Rhein, Tobias Jacke, Sascha Willgardt-Levin

Session 14

INFRASTRUKTURZUSTAND II
INFRASTRUCTURE CONDITION II

568 **Causal Discovery for Railway Health Condition Monitoring – A Case Study**
Thorsten Neumann, Martin Rabel, Oana-Iuliana Popescu, Andreas Gerhardus

580 **Monitoring Data Mapping for Infrastructure-oriented Visualization of Railway Track Condition**
Laura T. Rodríguez-Bayona, Héctor A. Fernández-Bobadilla, Ullrich Martin

596 **Decision support for tamping activities**
Ute Gläser, Joseph Feng, Mirjana Jovanovic

Session 15

DATEN UND PLANUNG
DATA AND PLANNING

609 **LeiDa-D – Versionierte, zentralisierte Datenhaltung für Basis- und Infrastrukturdaten**
Mario Trappein, Dennis Götz

626 **Traffic Management System (TMS) in der Schweiz – Neue Herausforderungen bez. Sicherheit und Zuverlässigkeit**
Ernst Zollinger, Gabrio Caimi, Vanessa Hänni

638 **Optimierung der betrieblichen Planungsprozesse bei der DB Fernverkehr AG zur Umsetzung der Anforderungen des AnnexVII**
Tobias Körner, Danny Rohrberg, Tobias Mohn, Heiko Güthenke

Session 16

FAHRZEUGTECHNIK
VEHICLE TECHNOLOGY

650 **Load-dependent influences on rail damage mechanisms**
Dieter Knabl, Alexander Plotho, Stefan Marschnig

666 **Proposal of Reliable Onboard Train Localization Method by Integration of RTK-GNSS, IMU, and Multiple One-dimensional LiDAR Sensors**
Kensuke Nagai, Haw-Shyang Chang, Wataru Ohnishi, Takafumi Koseki, Shota Miyoshi, Tomonori Hasegawa

682 **Small, Lightweight, Driverless Rail Vehicle for Rural Areas in Europe**
Jens Koenig, Thomas Mueller, Rickard Persson, Libor Lochman, Pauline Guicheney

Session 17

BAHNBETRIEB
RAILWAY OPERATIONS

703 **A SUMO-Based Study of Urban Rail Operations on Frankfurt's Corridor A**
Paula von der Heide, Lars Schnieder

725 **Modelling Rail Carrier Assignment and Relocation in Multimodal Pod System**
Nina D. Versluis, Mahnam Saeednia

740 **Impact of incorrect and incomplete transport statistical data: A Case Study of the Rhineland – ZARA Ports Corridor**
Wolf-Dietrich Geitz, Upendra Atre

IRSA 2025

SESSION 1

KAPAZITÄT I

RAILWAY CAPACITY I

Session 1
Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

Digitalisierung im Kapazitätsmanagement der Eisenbahn: Herausforderungen und Erwartungen

von Heel, Emanuel[1]; Scherer, Philipp[1]; Büker, Thorsten[1]; Benz, Alexandra[2]; Pfeifer, Andreas[2]; Nießen, Nils[2]; Beck, Michael J.[3]; Kramer, Urs[4]; Hack, Linus[4]; Ziegler, Niklas[4]

[1]quattron GmbH
[2]Verkehrswissenschaftliches Institut der RWTH Aachen
[3]neXt Capacity GmbH
[4]Institut für Rechtsdidaktik der Universität Passau

Zusammenfassung

Das Digitale Kapazitätsmanagement (DCM) wird als eine Grundvoraussetzung für eine optimierte Infrastrukturnutzung im Eisenbahnwesen angesehen. Durch DCM soll es möglich sein, bereits vorhandene, nicht genutzte Kapazitäten einfach zugänglich und nutzbar zu machen sowie die Kapazitätsausschöpfung im Betrieb zuverlässiger zu gestalten. Hierzu wird im Projekt „Digitalisierung im Kapazitätsmanagement des Gesamtsystems Schiene“ des Deutschen Zentrums für Schienenverkehrsforschung beim Eisenbahn-Bundesamt (DZSF) ein Stand des heutigen Kapazitätsmanagements gemeinsam mit Branchenakteuren aufgenommen und ein Idealbild für eine zukünftige, digitale Lösung entwickelt. Zur Ableitung eines Umsetzungskonzepts für die ausgearbeitete Lösung werden mögliche Herausforderungen und Erwartungen der Branche aufgezeigt. Übergeordnetes Ziel ist es, Empfehlungen zur Realisierung eines DCM bereitzustellen.

Keywords: Digitalisierung, Kapazitätsmanagement, Kapazitätssteigerung, Herausforderungen, Erwartungen

Session 1
Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

1 Zielsetzung

Verkehrsverlagerungen auf die Schiene sind ein zentrales Element deutscher und europäischer Verkehrs- und Klimapolitik. Um die ambitionierten Ziele – wie die Verdopplung der Fahrgastzahlen im Schienenpersonenverkehr (SPV) und die Erhöhung des Marktanteils des Schienengüterverkehrs (SGV) auf 25 % bis 2030 – zu erreichen, ist eine signifikante Steigerung der Netzkapazität erforderlich. Um Kapazitäten auf dem deutschen Schienennetz im erforderlichen Maße bereitzustellen, ist ein Dreiklang aus infrastrukturellen Maßnahmen, betrieblich/technischen Ansätzen im Systemverbund und einer Digitalisierung der Prozesse mitsamt Effizienzsteigerungen notwendig. Ein Kernelement stellt DCM dar, das vielfältige und spezifische Funktionalitäten voraussetzt. National werden die Vorteile von DCM bereits durch die Beschleunigungskommission Schiene attestiert [1, 2]. Auch auf europäischer Ebene wird eine Modernisierung und Digitalisierung des Kapazitätsmanagements durch RailNetEurope (RNE) angestrebt. Dies umfasst insbesondere auch eine Neuorganisation aller Prozesse der Kapazitätsplanung, die im Timetable Redesign (TTR) aufgehen [3].

2 Definition von DCM und Methodik

Der Begriff DCM wird derzeit vielfältig interpretiert, sodass eine eindeutige Definition erforderlich ist. Dazu werden Branchenentwicklungen analysiert und Stakeholder befragt, um ein einheitliches Verständnis zu erzeugen. Eine Zusammenfassung der gesammelten Erkenntnisse liefert die grundlegende Definition, dass DCM als ein durchgängiger, digital unterstützter Prozess angesehen wird, der alle Planungshorizonte – von der langfristigen Kapazitätsstrategie bis zur Betriebsdurchführung – umfasst. Im Detail bedeutet dies:

- DCM umfasst den gesamten Planungshorizont vom langfristigen Kapazitätsmanagement bis zur Planung und Betriebsdurchführung jedes einzelnen Betriebstags. Sowohl die Fahrplangestaltung als auch die Betriebsdurchführung sind integrale Bestandteile des Gesamtsystems.
- DCM wird durch geeignete IT-Infrastruktur unterstützt. Dabei wird der Zugriff auf die Kapazität in Echtzeit geplant, optimiert und gesteuert. Hierbei kann Automatisierung unterstützend wirken.
- DCM bedient sich eines digitalen, topologischen Abbilds der europäischen Infrastruktur mit zielgerichteter Abstraktion passend zum jeweiligen Anwendungshorizont.

- DCM ist ein nicht-proprietärer Prozess mit standardisierten Schnittstellen.
- Alle Schritte von DCM – von der Kapazitätsplanung bis Disposition – bilden einen widerspruchsfreien und europaweit einheitlichen Prozess mit gleichen Fristen und Regularien. DCM ermöglicht eine effiziente Bewirtschaftung und optimierte Ausnutzung der Eisenbahninfrastruktur inklusive der Etablierung geeigneter Regelkreise, um für nachfolgend auftretende, sich wiederholende Betriebssituationen zu lernen.
- Wesentliche Beteiligte an DCM sind Infrastrukturbetreiber (von Schienenwegen und Serviceeinrichtungen) sowie Eisenbahnverkehrsunternehmen bzw. Netzzugangsberechtigte.
- Der Prozess DCM ist transparent für alle Beteiligten und diskriminierungsfrei.

Das Projekt untersucht darüber hinaus den aktuellen Stand einzelner Umsetzungsprojekte im Rahmen der Digitalisierung des Kapazitätsmanagements sowie Herausforderungen und Erwartungen der Branche bei der Implementierung von DCM. Daraus wird ein Idealbild von DCM abgeleitet, welches nur durch spezifische Fähigkeiten umgesetzt werden kann. Abschließend erfolgt im Rahmen des Projektes die Formulierung von Empfehlungen zur Umsetzung eines DCM.

3 Stand der Technik

Die Fahrplankonstruktion bei der DB InfraGO AG erfolgt derzeit rechnergestützt mit dem System RUT-K (Rechnerunterstütztes Trassenmanagement – Konstruktion), wobei viele Arbeitsschritte manuell durchgeführt werden. Eine auf Optimierungsalgorithmen basierende automatisierte Fahrplanerstellung erfolgt bislang nur in Einzelfällen. Die Kommunikation zwischen Infrastrukturbetreibern (EIU) und Eisenbahnverkehrs-unternehmen (EVU) ist in Deutschland, aber auch im Ausland, durch eine Vielzahl digitaler und analoger Schnittstellen geprägt und damit heterogen. Der Blick auf die rechtlichen Grundlagen des Kapazitätsmanagements in Deutschland zeigt, dass die Trassenzuweisung auf der EU-Richtlinie 2012/34/EU sowie dem Eisenbahnregulierungsgesetz (ERegG) basiert. Trassenanträge können sowohl im Rahmen des Netzfahrplans als auch ad-hoc gestellt werden. Bei konkurrierenden Anträgen greift ein Koordinierungsverfahren. Für Serviceeinrichtungen gelten ergänzend die Vorgaben der EU-Durchführungsverordnung 2017/2177, die teilweise nationales Recht verdrängt.

Ein zentrales Digitalisierungsprojekt in diesem Kontext ist „#Einfachbahn“, das zahlreiche Tools zur Trassenanmeldung, Bauinformation und Betriebssteuerung

bereitstellt. Allerdings führt die Vielzahl an Anwendungen zu einer fragmentierten Systemlandschaft, die einer besseren Integration bedarf. Ein weiterer wichtiger Baustein in der nationalen Kapazitätsplanung ist der neue Kapazitätsplanungs- und Zuweisungsprozess „KaZu Novum“. Dieser verfolgt das Ziel, durch Teilprodukte wie das Kapazitätsnutzungskonzept (KNK), den Kapazitätsnutzungsplan (KNP) und Kapazitätsrahmenverträge (KRV) eine durchgehende Planung von der langfristigen Strategie bis zur Trassenvergabe zu ermöglichen. Die Umsetzung erfolgt schrittweise und orientiert sich an der Zielsetzung des Deutschlandtakts. Parallel dazu werden unter dem Label „neXt“ digitale Werkzeuge zur automatisierten Fahrplanerstellung entwickelt, wie etwa das System „Click&Ride“ für den Gelegenheitsverkehr. Langfristig soll ein integriertes Capacity & Traffic Management System (CTMS) entstehen, das Planung, Betrieb und Instandhaltung verknüpft. [4] [5]

Auf europäischer Ebene zeigen sich ähnliche Entwicklungen. In Österreich nutzt die ÖBB Infrastruktur AG das System M-AMA (Modular Network Access Order Management System) zur Trassen- und Serviceeinrichtungsbuchung sowie das Planungstool TRAKSYS. In der Schweiz übernimmt die unabhängige Trassenvergabestelle (TVS) die zentrale Trassenvergabe. Dort kommen Systeme wie NeTS-AVIS (Netzweites Trassensystem) und das Rail Control System (RCS) zum Einsatz. Frankreich setzt mit dem System SIPH (Système Industriel de Production Horaire) auf eine industrialisierte Fahrplankonstruktion, die auf dem Tool TPS basiert. [6] [7] [8] [9] [10] [11]

Ein zentrales europäisches Projekt ist das „Timetable Redesign for Smart Capacity Management“, das von RailNetEurope koordiniert wird. TTR verfolgt eine grundlegende Neugestaltung des Kapazitätsmanagements in vier Phasen: Capacity Strategy, Capacity Model, Capacity Supply und Path Requests. Ziel ist ein durchgehender, transparenter Prozess, der unterschiedliche Verkehrsarten und Planungshorizonte berücksichtigt. Die Umsetzung erfordert eine neue EU-Verordnung (2023/0271/COD), die derzeit in Vorbereitung ist. Zur digitalen Unterstützung des TTR-Prozesses werden verschiedene Tools entwickelt. Das Path Coordination System (PCS) dient der internationalen Trassenanmeldung, während das European Capacity Management Tool (ECMT) die Visualisierung von Kapazitäten ermöglicht. Ergänzt wird dies durch das Temporary Capacity Restrictions (TCR) Tool zur Koordination temporärer Kapazitäts-einschränkungen. Hinsichtlich der Baustellen- bzw. TCR-Koordination gilt künftig der Annex VII der Richtlinie 2012/34/EU [12]. Die Umsetzung der genannten Systeme erfolgt schrittweise, wobei der Jahresfahrplan 2028 als Ziel für die vollständige TTR-Implementierung gilt. [13] [14]

Auch die technischen Spezifikationen für Telematikanwendungen (TAF/TAP TSI) durch die EU spielen für die Schnittstellen eine zentrale Rolle. Sie standardisieren die Kommunikation zwischen EIU und EVU, etwa bei Trassenanmeldungen, Zuglaufmeldungen und Verspätungsinformationen. Die DB InfraGO AG plant, ihr bestehendes Trassenportal Netz (TPN) bis 2027 durch TSI-konforme Schnittstellen zu ersetzen. Dies wird in den kommenden Jahren zur Ablösung unterschiedlicher bestehender Systeme wie dem Trassenportal Netz (TPN) sowie zu einer deutlichen Veränderung der IT-Infrastruktur führen. [15]

Insgesamt zeigt sich, dass der aktuelle Stand der Technik im Wesentlichen durch eine rechnerunterstützte, aber noch stark manuelle Fahrplankonstruktion geprägt ist. Die Digitalisierung ist in vielen (Teil-)Bereichen fortgeschritten, jedoch oft fragmentiert und proprietär. Europäische Initiativen wie TTR, PCS und TCR bieten Potenzial für ein einheitliches, digitales Kapazitätsmanagement. Die größte Herausforderung liegt in der Integration bestehender Systeme, der Vereinheitlichung von Schnittstellen und der Einbindung aller relevanten Akteure in einen durchgängigen, transparenten und diskriminierungsfreien Prozess.

4 Herausforderungen und Erwartungen an DCM

Zur Gewährleistung der Praxistauglichkeit von DCM wird die Branche mittels Interviews, Fragebögen und auch in Form von Workshops eingebunden – es gibt insgesamt 39 Teilnehmende bzw. Rückmeldungen. Dabei werden auf Grundlage auftretender Probleme sowie den geäußerten Erwartungen, die im Zielbild aufgehen, die aktuellen Herausforderungen für eine potenzielle Umsetzung von DCM abgeleitet. Zur Strukturierung werden diese verschiedenen Dimensionen (prozessual, technologisch, wirtschaftlich und juristisch) zugeordnet. Eine Herausforderung kann dabei einerseits ein Widerspruch zwischen den Vorstellungen der Stakeholder oder von Prozessschritten oder anderseits ein Problem sein, welches bei Stakeholdern in der betrieblichen Praxis auftritt. Insgesamt lassen sich im Rahmen der Bearbeitung 33 Herausforderungen ableiten: Der Großteil der Herausforderungen betrifft dabei die prozessuale Dimension (20 Herausforderungen), gefolgt von der technologischen Dimension (fünf Herausforderungen), wirtschaftlichen Dimension (vier Herausforderungen) und der juristischen Dimension (vier Herausforderungen). Dies ist allerdings zunächst ein Trend, der sich vermutlich aus der Gruppenzusammensetzung der Stakeholder ergibt. Die Herausforderungen erhalten eine Bezeichnung gemäß ihrer Zugehörigkeit zu einer Dimension mittels Buchstaben (P = prozessual, W = wirtschaftlich, T = technologisch)

und eine fortlaufende Nummerierung innerhalb der Dimension. Die Nummerierung ist nicht mit einer Rangfolge innerhalb einer Dimension gleichzusetzen, sondern erfolgt rein zufällig. Darüber hinaus findet sich in dieser Spalte eine Bezeichnung der durch die Herausforderung betroffenen Stakeholder (Eisenbahninfrastrukturunternehmen = EIU, Eisenbahnverkehrsunternehmen = EVU, Aufgabenträger = AT, Verbände = VB). In der vierten Spalte findet sich eine Aussage dazu, ob im Rahmen der Erhebungen ein Hinweis auf einen Konflikt zwischen den Stakeholdern in dieser Herausforderung identifiziert werden konnte. Nachfolgend werden die vier Dimensionen einzeln aufgezeigt und deren wesentliche Herausforderungen beschrieben.

Prozessuale Dimension: Aktuell liegt in der Branche kein gemeinsames Verständnis des Begriffs „Kapazitätsmanagement“ vor – die Ausgestaltung von DCM ist daher nicht definiert. Zum anderen muss die Rolle eines Systemintegrators verortet werden. Schnittstellen und Kommunikation bei der Trassenplanung sind lückenhaft, Informationen über verfügbare Trassen sind nicht einsehbar und kurzfristige Änderungen müssen aufwendig eingearbeitet werden. Eine Trasse kann darüber hinaus nicht im Ganzen bestellt werden, da Strecke, Verkehrsstation und Abstellung getrennt behandelt und abgerechnet werden. Zudem fehlen einheitliche internationale Regeln und Fristen, die Prozesse standardisieren können. Überdies erschweren eine Vielzahl auf spezielle und abgegrenzte Use Cases zugeschnittene Tools den Überblick und es fehlen geregelte Fahrplan-Rückfallebenen zur kurzfristigen, situationsgerechten Reaktion und damit verbunden eine ausreichende Resilienz des Gesamtsystems. Eine integrierte Berücksichtigung von Bauzuständen und konkreten technischen Eigenschaften von Fahrzeugen findet nicht in aufwandsreduzierendem Maße statt.

Tabelle 1: Auflistung der identifizierten Herausforderungen in der prozessualen Dimension

Bezeichnung und Betroffenheit	Problemstellung/ Herausforderung im Status Quo	Idealvorstellung der Stakeholder	Konflikte zwischen betroffenen Stakeholdern?
P.1 EIU, EVU	Es existiert kein gemeinsames Verständnis des Begriffs "Kapazitätsmanagement".	Schärfung des Begriffs des Kapazitätsmanagements im Sinne des Gesamtsystems Bahn	Ja

Bezeichnung und Betroffenheit	**Problemstellung/ Herausforderung im Status Quo**	**Idealvorstellung der Stakeholder**	**Konflikte zwischen betroffenen Stakeholdern?**
P.2 EIU, EVU	Es existiert kein einheitlicher Prozess zur Berücksichtigung und Bekanntmachung von Bauzuständen.	Beherrschen von Infrastrukturvarianten und Bauzuständen	Nein
P.3 EIU, EVU, AT	Keine Berücksichtigung der Abstellanlagen und der Serviceeinrichtungen sowie von Betriebsfahrten bei der Kapazitätsplanung.	Ganzheitliche Betrachtung von Zugfahrten, auch über den kommerziellen Teil hinaus	Ja
P.4 EIU, EVU	Technische Eigenschaften der Fahrzeuge werden in Kapazitätsplanung nicht ausreichend berücksichtigt.	Möglichst exakte Abbildung des Kapazitätsverbrauchs einer Zugfahrt	Ja
P.5 EIU, EVU	Manueller Anpassungsaufwand je Trasse während der Fahrplanerstellung ist wegen Interessenskonflikten der EVU sehr hoch.	Automatisierung der Fahrplanerstellung und -optimierung	Ja
P.6 EIU, EVU	Wiederkehrende, manuelle Arbeitsschritte je Trasse während der Disposition erzeugen hohen Aufwand.	Eine Automatisierung der Betriebsführung, der Kommunikation und des Reschedulings soll zur Reduktion des Arbeitsaufwandes und zu schnellerer Reaktion beitragen	Ja
P.7 EIU, EVU, AT, VB	Über den gesamten Planungsprozess existiert eine Vielzahl an Tools, die lückenhaft miteinander verbunden sind.	DCM soll die Planungsphasen von der Langfristplanung bis zur Disposition am Betriebstag widerspruchsfrei und konsistent einbinden und die damit einhergehenden Beteiligten im erforderlichen Maße berücksichtigen	Ja
P.8 EIU, EVU	Planungsparameter der Zugtrassen werden im Laufe der Planungsphasen immer weiter fixiert.	Variationsmöglichkeit von Planungsparametern in allen Planungshorizonten	Ja

Session 1
Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

Bezeichnung und Betroffenheit	Problemstellung/ Herausforderung im Status Quo	Idealvorstellung der Stakeholder	Konflikte zwischen betroffenen Stakeholdern?
P.9 EIU, EVU	Prozesse zur Aufnahme geeigneter Regelkreise, Planungsparametern und Live-Daten im gewünschten Format zur kontinuierlichen Verbesserung/Optimierung des Fahrplans existieren nicht.	Kontinuierliche Optimierung von Fahrplanvarianten und Auslastungssteuerung über alle Planungshorizonte unter Beachtung verschiedener Zielfunktionen und unter Einbindung von Echtzeitdaten, um Idealzustand aus genutzter Kapazität und Betriebsqualität zu erzielen	Nein
P.10 EIU, EVU	Geregelte Fahrplan-Rückfallebenen im Betrieb existieren im System nicht. Dies führt zu Arbeitsaufwand bei EIU und EVU durch unflexible Neutrassierung bei Störungen.	Einspielen und Beherrschen von Rückfallebenen des Systems	Nein
P.11 EIU, EVU, AT, VB	Es existieren keine einheitlichen, widerspruchsfreien, internationalen Regeln und Systeme zur Trassenvergabe. Zudem stößt die Abgabe von nationalen Kompetenzen an eine höhere Ebene auf Widerstand.	Einheitliche, integrierte und europäische Regeln, Prozesse und Systeme zur Trassenvergabe und Optimierung von Fahrplanvarianten	Ja
P.12 EIU, EVU	Zeitverzögerte Weitergabe von Informationen - teilweise auf Papier - während der Planungsphasen und fehlende Schnittstellen erhöhen Zeit- und Personalaufwand.	Digitalisierung der Arbeitsschritte in und zwischen den einzelnen Planungsphasen zur schnelleren Trassen-bestellung und einfacheren Suche nach Alternativen	Ja
P.13 EIU, EVU, AT	Lückenhafte bis hin zu fehlende grenzüberschreitende Kommunikation bei Trassenplanung/-vergabe (sowohl international als auch national) bzw. fehlende Vorgaben; starke Abhängigkeit von Bearbeitenden.	Grenzüberschreitende (nationale /internationale) Kommunikationsvorgaben für die Schnittstelle zwischen den Beteiligten bei der Trassenplanung und -vergabe zur Verbesserung der Einbindung aller Beteiligten	Nein
P.14 EIU, EVU	Gestaltungsspielraum der EIU in der Kapazitätsplanung im Rahmen des Trassenbestellprozess ist an die Wünsche der EVU gekoppelt, Kapazitäten können nicht vorgeplant werden.	Kapazitätsplanung erfolgt durch das EIU unter dezidierten Gestaltungsspielräumen	Ja

Bezeichnung und Betroffenheit	**Problemstellung/ Herausforderung im Status Quo**	**Idealvorstellung der Stakeholder**	**Konflikte zwischen betroffenen Stakeholdern?**
P.15 EIU, EVU, AT	Getrennte Bestellung und Abrechnung von Trasse, Station und Abstellung.	Etablierung eines gemeinsamen Bestell- und Abrechnungsprozesses für alle Infrastrukturen (Fahrweg und Serviceeinrichtungen)	Ja
P.16 EIU, EVU, AT, VB	Es sind keine Informationen über verfügbare Trassen einsehbar; Unterschiedliche Regeln, Fristen und Prozesse zur Informationsbereitstellung existieren (bspw. zu Fahrplan-aktualisierungen).	Transparente Informationsbereitstellung zwischen allen Beteiligten an definierten Zeitpunkten und in ausreichender Informationstiefe (bspw. über verfügbare Trassen, Fahrplanaktualisierungen)	Nein
P.17 EIU, EVU	Interne Bedürfnisse in der EVU-seitigen Produktionsplanung führen zu zusätzlichen Iterationsschleifen.	Weiterführende Berücksichtigung der Bedürfnisse aus der Produktionsplanung der EVU im Rahmen von DCM	Ja
P.18 EIU, EVU	Prozesse zur Bekanntmachung von Baustellen sind nicht eindeutig; Baustellenankündigungen sind teils widersprüchlich.	Rechtzeitige, widerspruchsfreie Baustellenankündigungen und Etablierung eines Abweichungsmanagements	Nein
P.19 EVU	Kurzfristige Änderungen an Trassen sind nur mit viel Aufwand durchführbar.	Möglichkeit zur ad hoc Anmeldung und Änderung von Trassen zur kurzfristigen Adaption an aktuelle Marktbedürfnisse und Produktionsbedingungen	Nein
P.20 EIU, EVU	Vielzahl an Tools macht es für Bearbeitende schwer, den Überblick zu behalten. Ausgestaltung eines DCM noch nicht final geklärt: Verbessert und verbindet ein DCM bestehende Prozesse oder wird ein komplett eigenes Tool geschaffen?	Verbindung von Tools oder Neuerschaffung eines Tools zu effizienterem Kapazitätsmanagement des Systems Eisenbahn	Ja

Wirtschaftliche Dimension: Der aktuelle Prozess enthält Wertschöpfungsverluste, da vorhandenes Personal nicht wertschöpfende Tätigkeiten wie bspw. manuelle Datenübertragung durchführen muss. Zudem wird aufgrund von Widersprüchen zwischen optimaler Kapazitätsausnutzung und Kundenbedürfnissen nicht die gesamte verfügbare Kapazität des Schienennetzes wirtschaftlich genutzt. Hierzu zählt auch die nicht ausreichend auf Kapazitätsschonung ausgerichteten Trassenpreissysteme. Darüber hinaus ist während der Umsetzung von DCM u. a. bei der Integration und Zusammenführung der Systeme sowie die Schulung des Personals mit zusätzlichen Kosten zu rechnen. Hierbei sollte eine asymmetrische Kostenverteilung auf die Stakeholder vermieden werden, um einzelne Stakeholder nicht übermäßig zu belasten.

Tabelle 2: Auflistung der identifizierten Herausforderungen in der wirtschaftlichen Dimension

Bezeichnung und Betroffenheit	**Problemstellung/ Herausforderung im Status Quo**	**Idealvorstellung der Stakeholder**	**Konflikte zwischen betroffenen Stakeholdern?**
W.1 EIU, EVU	Vorherrschender Personalmangel wird dadurch verschärft, dass sich vorhandenes Personal mit nicht wertschöpfender Arbeit beschäftigen muss (bspw. Bestellungen über Papier und manueller Datenübertragung).	Reduzierung des Personalaufwands durch DCM	Nein
W.2 EIU, EVU, AT	Die Finanzierung von DCM kann sich asymmetrisch auf EIU verteilen und ist nicht für alle EIU wirtschaftlich; EVU müssen Systemanpassungen selbst finanzieren.	Die Finanzierung eines DCM muss für alle Akteure gesichert sein, ohne einzelne Beteiligte übermäßig zu belasten	Ja
W.3 EIU, EVU, AT	Aufgrund von Widersprüchen zwischen optimaler Kapazitätsausnutzung und Kundenbedürfnissen wird nicht die gesamte verfügbare Kapazität des Schienennetzes genutzt. Zudem sind Trassenpreissysteme nicht ausreichend auf Kapazitätsschonung ausgelegt.	Berücksichtigung eines Yield Managements/Dynamic Pricing im DCM zwecks Kapazitätssteuerung und -schonung	Ja
W.4 EIU, EVU	Heutiges Personal ist im Umgang mit einem zukünftigen DCM nicht erfahren.	Ausbildung/ Umschulung von Personal	Nein

Technologische Dimension: Es existiert kein europaweit einheitliches, je nach Planungsstufe abstrahiertes Datenmodell der Infrastruktur sowie Vorgaben dazu. Stattdessen gibt es für die verschiedenen Planungsstufen eine Vielzahl an Tools, die nicht miteinander verknüpft oder aufeinander abgestimmt sind und zudem nicht jeden Prozessschritt abbilden können. Darüber hinaus existieren keine IT-Standards für ein einheitliches, integriertes System.

Tabelle 3: Auflistung der identifizierten Herausforderungen in der technologischen Dimension

Bezeichnung und Betroffenheit	**Problemstellung/ Herausforderung im Status Quo**	**Idealvorstellung der Stakeholder**	**Konflikte zwischen betroffenen Stakeholdern?**
T.1 EIU, EVU, AT	Es existiert kein europaweit einheitliches, je nach Planungsstufe abstrahiertes Datenmodell der Infrastruktur sowie Vorgaben dazu.	Einheitliche, gemeinsame und konsistente Datenbasis zur Verbesserung des Informationsaustausches und der Datenhaltung	Nein
T.2 EIU, EVU, AT	Für verschiedene Planungsstufen existiert eine Vielzahl an Tools, die nicht miteinander verknüpft oder aufeinander abgestimmt sind. Bestehende Systeme können nicht jeden Prozessschritt abbilden.	Lückenlose Verbindungen der Tools und aller Phasen durch geeignete Schnittstellen	Ja
T.3 EIU, EVU, AT	Umfang und Ausgestaltung der DCM-Funktionalitäten ist (zwischen Stakeholdern) nicht final geklärt.	Optimale Funktionalitäten für effiziente Arbeitsweisen mit DCM	Nein
T.4 EIU	DCM unterscheidet sich hinsichtlich IT-Verfahren von den aktuell genutzten Systemen.	Erarbeitung und Einführung neuer IT-Verfahren	Nein

Bezeichnung und Betroffenheit	Problemstellung/ Herausforderung im Status Quo	Idealvorstellung der Stakeholder	Konflikte zwischen betroffenen Stakeholdern?
T.5 EIU, EVU	IT-Standards für ein einheitliches, integriertes System existieren nicht.	Einheitliche, europäische Standards für IT, insbesondere zur Unterstützung des Lieferantenmarkts	Ja

Juristische Dimension: Der Zeitpunkt der Verabschiedung sowie der genaue Inhalt der Kapazitätsverordnung auf EU-Ebene (Kap-VO-E, 2023/0271/COD [16]) sind noch unklar. Hinsichtlich der Diskriminierungsfreiheit und der Transparenz, insbesondere bei der Trassenvergabe sowie der zur Verfügung gestellten Informationen, äußern die Stakeholder – insbesondere die EVU – erhöhte Ansprüche. Zudem muss die Umlage der Implementierungskosten von DCM im Vorhinein geklärt sein. Die Vereinbarkeit mit einschlägigen Rechtsakten nationalen und europäischen Rechts muss bei der Prüfung der Umsetzbarkeit im Fokus stehen.

Tabelle 4: Auflistung der identifizierten Herausforderungen in der juristischen Dimension

Bezeichnung	Problemstellung/ Herausforderung im Status Quo	Idealvorstellung der Stakeholder	Konflikte zwischen betroffenen Stakeholdern?
J.1	Zeitpunkt der Verabschiedung sowie genauer Inhalt der Kapazitätsverordnung ist noch unklar.	Eindeutiger Inhalt mit entsprechend eindeutigen Verpflichtungen zur rechtssicheren Umsetzung im geforderten Umfang gepaart mit Vorschriften zur effektiven Durchsetzung	Ja
J.2	Die Diskriminierungsfreiheit – insbesondere bei der Trassenvergabe – muss gewährleistet sein.	Faire Bedingungen im Wettbewerb; insbesondere die Vergabe erfolgt angemessen, fair und transparent, ohne dass Unternehmen (vor allem als Teil eines Monopolisten) bevorteilt werden	Ja
J.3	Die Transparenz – insbesondere der Trassenvergabe sowie der zur Verfügung gestellten Informationen – muss gewährleistet sein.	Die Informationen gewährleisten die Nachvollziehbarkeit und Nachprüfbarkeit (auch beim Einsatz Künstlicher Intelligenz) und damit die Sicherstellung fairer Bedingungen im Wettbewerb	Ja

Session 1
Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

Bezeichnung	Problemstellung/ Herausforderung im Status Quo	Idealvorstellung der Stakeholder	Konflikte zwischen betroffenen Stakeholdern?
J.4	Die Einzelheiten der Umlage der Kosten der Implementierung von DCM nach dem Vollkostenprinzip sind noch nicht abschließend geklärt.	Klare Vorgaben für die Kostenverteilung und möglicher Subventionierung ohne (schwerwiegende) Belastungen	Ja

Die Herausforderungen werden anschließend zu Erwartungen abstrahiert und stellen damit erste Anforderungen an das Idealbild dar (siehe Abbildung 1).

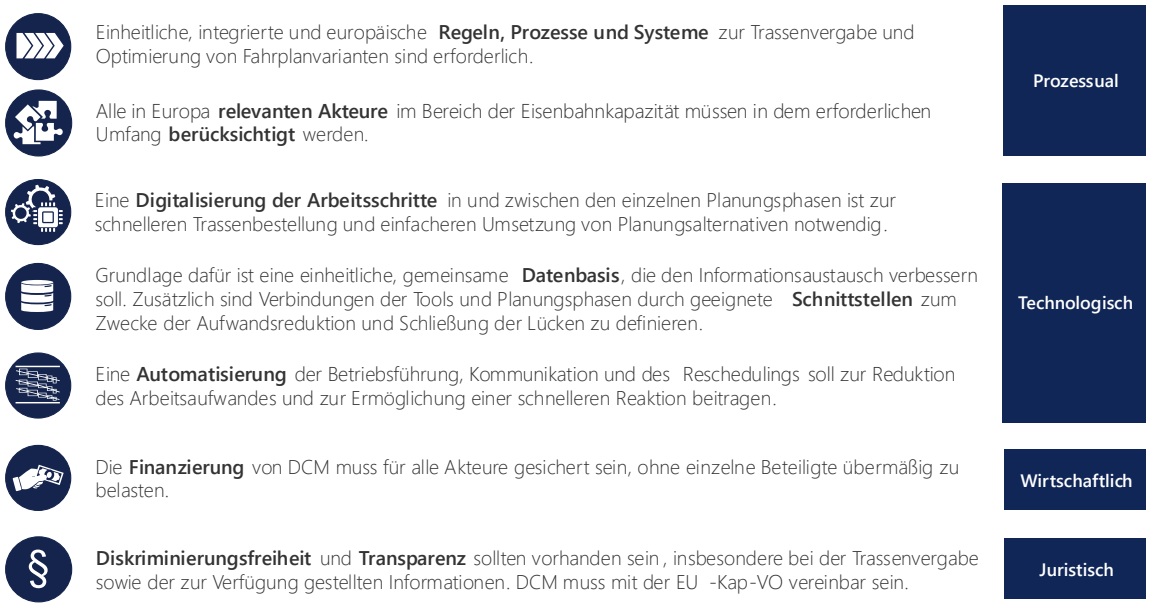


Abbildung 1: Aus Herausforderungen abstrahierte Erwartungen bei der Umsetzung des DCM für den Verkehrsträger Schiene

Generell ist es für DCM von Bedeutung, einen allgemeingültigen, verlässlichen und ständig aktuellen Bestand aller relevanten Daten (z. B. Infrastruktur, Kapazitätszugriffe) für alle Prozesse der Planungsphasen vorzuhalten. Darüber hinaus wird die Kommunikation der verschiedenen Akteure untereinander vereinfacht und beschleunigt. Dadurch sowie durch eine Echtzeitsteuerung können Prozesse in Bezug auf die Herstellung der Durchgängigkeit zwischen Kapazitätsplanung, Trassenzuweisung und Betriebsmanagement optimiert werden. Somit lassen sich Kommunikationslücken über den gesamten Prozess schließen, Fehler bei der Datenübertragung vermeiden und vorhandenes Personal bei allen Akteuren besser einsetzen. Diese Erwartungen werden im

weiteren Projektverlauf als Anforderungen in einem Idealbild zur Umsetzung von DCM festgehalten.

5 Juristische Rahmenbedingungen

Viele Herausforderungen sind nur in geringem Umfang oder gar nicht juristischer Natur. Häufig finden sich schlicht keine Regelungen zu DCM. Unzureichende Information und Transparenz beispielsweise sind nicht notwendigerweise eine juristische Herausforderung; sie ergäbe sich allenfalls daraus, dass die Anforderungen an die Infrastrukturbetreiber nicht hoch genug sind. Vielmehr sind die Herausforderungen eher im technischen Bereich angesiedelt.

Unsicherheiten aus juristischer Sicht ergeben sich aber aus dem Verordnungsentwurf, einschließlich des unbekannten Zeitpunkts seines Inkrafttretens. Der europäische Rechtsrahmen sollte mangels Abweichungsmöglichkeit und aufgrund langwieriger Änderungsverfahren als gegeben angesehen werden. Zentrale Fragen dieses umfassenden Rechtsaktes werden zur Klärung auf das European Network of Infrastructure Managers (ENIM) delegiert, wobei diese voraussichtlich erst nach Inkrafttreten geklärt werden. Die Verordnung benötigt zudem außerhalb ihres Anwendungsbereichs Anpassungsvorschriften.

Eine Vereinfachung der Rahmenbedingungen für DCM folgt aus der Verordnung nicht zwangsläufig. Aspekte wie das fortlaufende Planungsverfahren, Art. 33 VO-E, oder die vorgeplante Kapazität, Art. 20 VO-E, dürften nicht ohne umfangreiche Änderungen in den Systemen der Infrastrukturbetreiber möglich sein. Konkrete Anhaltspunkte für die Umsetzung fehlen. Die tiefgreifenden Änderungen der Methoden der Kapazitätszuweisung werden zudem wohl nicht ohne erhebliche Reibungspunkte zeitnah in den Buchungstools umsetzbar sein. Inwieweit der Infrastrukturbetreiber auf Änderungen in den unterschiedlichen Planungsstufen schnell genug reagieren kann, um eine reibungslose Implementierung von DCM zu gewährleisten, ist unklar. Fragen hinsichtlich der Kostenverteilung stellen sich, wobei gerade bei einem Konzernverbund die Kosten für die Implementierung des DCM strikt nach Unternehmen getrennt werden müssen.

Der mögliche Einsatz von KI sorgt für Konflikte in Bezug auf Transparenz und Diskriminierungsfreiheit, wobei es maßgeblich auf die Art ihres Einsatzes ankommt. Das betrifft etwa die Erstellung von Prognosen, die Auslegung von Rechtsbegriffen sowie die Vorbereitung bzw. das direkte Treffen einer Entscheidung. Insbesondere können sich Herausforderungen aus der KI-Verordnung ergeben.

Inwieweit die geplanten Änderungen durch den Verordnungsentwurf zu einer Erleichterung oder zu einer Erschwerung für DCM führen, kann noch nicht abschließend beurteilt werden. Als wenig aufschlussreich hat sich bislang eine rechtsvergleichende Betrachtung erwiesen. Juristisch liegt der Fokus auf den zu entwickelnden Maßnahmen, die zur Implementierung von DCM erforderlich sind. Abstrakt können dabei nur wenige Maßstäbe aufgestellt werden. Da das DCM auf einen Gewinn an Effizienz, auf Diskriminierungsfreiheit sowie (erhöhte) Transparenz bei der Kapazitätszuweisung abzielt, bestehen keine abstrakten Probleme mit diesem Zielbild. Entscheidend ist die Einhaltung der (unionalen) Rechtsvorschriften bei dessen Umsetzung.

6 Weiteres Vorgehen

Im weiteren Projektvorgehen erfolgt die Erstellung eines Idealbildes eines digitalisierten und konsistenten Kapazitätsplanungs- und -zuweisungsprozesses. Europäische Rahmenbedingungen, die durch die Richtlinie 2012/34/EU und den Entwurf der neuen europäischen Kapazitätsverordnung gegeben sind, gelten bei der Konzeptionierung als fest einzuhaltende Vorgaben. Auf Grundlage dieses Idealbildes werden zentrale Geschäftsfähigkeiten abgeleitet, die zur Umsetzung eines digitalisierten Gesamtprozesses und zur Auflösung bzw. Abmilderung der identifizierten Herausforderungen erforderlich sind. Diese umfassen dabei sowohl grundlegende Elemente wie das Vorhandensein einer zentralen Datenbasis, standardisierter Schnittstellen oder einer erforderlichen Datentransparenz als auch prozessspezifischere Fähigkeiten wie z. B. die Fahrplanung in Echtzeit. Der potenzielle Nutzen der Geschäftsfähigkeiten wird anschließend anhand verschiedener Kriterien und einer qualitativen Nutzwertanalyse bewertet. Die Kriterien bilden verschiedene Dimensionen ab und berücksichtigen sowohl zeitliche Umsetzungsaspekte als auch juristische Restriktionen oder Wirkungsbereiche. Zudem wird analysiert, inwieweit bereits existierende Digitalisierungsansätze im Bereich des Kapazitätsmanagements die Umsetzung dieser Geschäftsfähigkeiten unterstützen können. Insgesamt sollen die Ergebnisse eine fundierte Einschätzung darüber liefern, welche Geschäftsfähigkeiten als essenziell für ein DCM einzustufen sind und welche eher einen ergänzenden optionalen Mehrwert besitzen.

Abschließend werden in Abstimmung mit der Branche etappierte und priorisierte Empfehlungen zur Umsetzung eines DCM abgeleitet, welche dabei helfen können, das skizzierte Idealbild umzusetzen.

Literatur

[1] Bundesministerium für Digitales und Verkehr (BMDV), „Beschleunigungskommission Schiene - Abschlussbericht," Bonn, 2022.

[2] Dr.-Ing. Michael J. Beck, „Infrastrukturkapazität und Trassen müssen so einfach zu buchen sein wie ein Hotelzimmer," *Rail Business,* pp. 4-5, August 2023.

[3] RailNetEurope (RNE), „TTR Fact Sheets - Complementary information document to Descripton of the Timetabling and Capacity Redesign Process," RailNetEurope, Wien, 2022.

[4] D. Pöhle, A.-L. Frank, S. Kühn und J. Schlaich, „Automatisierte Fahrplanerstellung bei der DB Netz," in *Tagungsdokumentation zur „HEUREKA – Optimierung in Verkehr und Transport"*, Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV) e.V., 2021.

[5] DB InfraGO AG, „KaZu Novum – Kapazitätsplanung und -zuweisung der Zukunft," [Online]. Available: https://www.dbinfrago.com/web/schienennetz/kazu-novum-11909200. [Zugriff am 07 10 2024].

[6] ÖBB Infrastruktur AG, „M-AMA - Handbuch OBS Version 1.4," [Online]. Available: https://infrastruktur.oebb.at/de/geschaeftspartner/it-tools/m-ama/m-ama-handbuch.pdf. [Zugriff am 01 12 2023].

[7] ÖBB Infrastruktur AG, „TRAKSYS," [Online]. Available: https://trackfwd.oebb.at/de/artikel/traksys. [Zugriff am 07 12 2023].

[8] Schweizerische Trassenvergabestelle (TVS), „Die Schweizerische Trassenvergabestelle kurz erklärt," 2023. [Online]. Available: https://www.tvs.ch/tvs-kurz-erklaert. [Zugriff am 30 11 2023].

[9] Schweizerische Bundesbahnen (SBB), „NeTS Plan: Trassenplanung für Europas dichtestes Bahnnetz," 2023. [Online]. Available: https://bahninfrastruktur.sbb.ch/de/produkte-

dienstleistungen/bahninformatiksysteme/fahrplanplanung/nets-plan.html. [Zugriff am 30 11 2023].

[10] Schweizerische Bundesbahnen (SBB), „RCS – Verkehrsmanagenent für Europas dichtestes Bahnnetz," [Online]. Available: https://bahninfrastruktur.sbb.ch/content/dam/internet/bahninfrastruktur/downloads/de/produkte-dienstleistungen/bahninformatiksysteme/bahnbetrieb/Broschuere-RCS.pdf.sbbdownload.pdf. [Zugriff am 08 12 2023].

[11] S. Reseau, „SIPH : Ouverture du service," 10 2019. [Online]. Available: https://www.actualites-si-clients-sncfreseau.fr/N11/article.html. [Zugriff am 04 12 2023].

[12] DB InfraGO AG, „Annex VII – Der neue Prozess zur Baukommunikation," [Online]. Available: https://www.dbinfrago.com/web/schienennetz/fahren_und_bauen/annex-vii-richtlinie-2012-34-eu-11857512. [Zugriff am 28 10 2024].

[13] RailNetEurope (RNE), TTR Fact Sheets - Complementary information document to Description of the Timetabling and Capacity Redesign Process, Version 2.0, 2022.

[14] RailNetEurope (RNE), DCM Fact Sheet, Version 1.0, 2022.

[15] DB InfraGO AG, „Kurzpräsentation TAF/TAP TSI bei der DB InfraGO AG," [Online]. Available: https://www.dbinfrago.com/web/schienennetz/netzzugang-und-regulierung/taf-tap-tsi/allgemein-11089206#. [Zugriff am 07 10 2024].

[16] Europäische Kommission, „Vorschlag für eine Verordnung des Europ. Parlaments und des Rates über die Nutzung von Fahrwegkapazität der Eisenbahn im einheitl. europ. Eisenbahnraum, zur Änderung der Richtl. 2012/34/EU und zur Aufhebung der VO (EU) 913/2010; Verfahren 2023/0271/COD," Amtsblatt der Europäischen Union, Brüssel, 2023.

Session 1

Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler

Autoren



Von Heel, Emanuel

Emanuel von Heel studierte bis 2018 Mobilität und Verkehr an der RWTH Aachen mit der Vertiefung im Schienen- und Straßenverkehrsbereich. Er ist seit 2022 als Berater bei der quattron GmbH in den Bereichen Kapazitätsplanung, Finanzierung und Resilienz tätig und war zuvor vier Jahre als Projektleiter in verschiedenen Verkehrsprojekten bei der DTV-Verkehrsconsult GmbH aktiv.



Scherer, Philipp

Philipp Scherer studierte bis 2019 Physik an der TU Dortmund und ist seit Januar 2020 bei der VIA Consulting & Development GmbH beschäftigt, die 2024 in der quattron GmbH aufging. Hier beschäftigt er sich insbesondere mit eisenbahnbetriebswissenschaftlichen Fragestellungen, der Auswertung großer Datenmengen und der Pflege verschiedener Softwaretools.



Büker, Thorsten

Thorsten Büker schloss 2006 das Studium des Bauingenieurwesen an der RWTH Aachen ab und promovierte dort 2010. Er war 2008 Mitgründer der VIA Consulting & Development GmbH, die 2024 in der quattron GmbH aufging. Hier ist er als Geschäftsführer tätig. Projekterfahrungen aus vierzehn Ländern erlauben ihm eine ganzheitliche Betrachtung des Systems Bahn mit Fokus im Dreieck Infrastrukturplanung – Betrieb – Technologie.



Benz, Alexandra

Alexandra Benz studierte bis 2024 Bauingenieurwesen mit Schwerpunkt Verkehrswesen an der RWTH Aachen und arbeitet seitdem als Wissenschaftliche Mitarbeiterin am Verkehrswissenschaftlichen Institut der RWTH Aachen im Bereich der Eisenbahnbetriebswissenschaft.

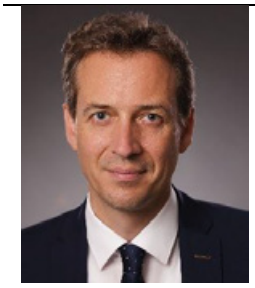


Pfeifer, Andreas

Dr.-Ing. Andreas Pfeifer ist seit 2021 Oberingenieur am Verkehrswissenschaftlichen Institut der RWTH Aachen. Nach dem Studium des Wirtschaftsingenieurwesens mit der Fachrichtung Bauingenieurwesen an der RWTH Aachen promovierte er dort 2019. Von 2014 bis 2020 war er als wissenschaftlicher Mitarbeiter am Verkehrswissenschaftlichen Institut tätig.

Session 1

Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler



Nießen, Nils

Univ.-Prof. Dr.-Ing. Nils Nießen leitet seit 2013 das Verkehrswissenschaftliche Institut der RWTH Aachen. Nach dem Studium des Bauingenieurwesens und der Wirtschaftsgeographie an der RWTH Aachen promovierte er dort 2008. Anschließend war er Projektingenieur bei HaCon in Hannover und Geschäftsführer der VIA Consulting & Development GmbH.



Beck, Michael

Michael Beck ist CEO der neXt Capacity GmbH. Ihn treibt an, Digital Capacity Management zum Game Changer für die Eisenbahn in Europa zu machen.

Nach dem Studium des Wirtschaftsingenieurwesens an der TU Darmstadt beriet er bei McKinsey Eisenbahnunternehmen in Europa, Afrika, Russland und Arabien. In dieser Zeit wurde er an der Ruhr-Universität Bochum zur Hierarchischen Planung von Taktfahrplänen promoviert.

Zwischen 2010 und 2019 war er bei der Deutschen Bahn als Leiter Fahrplan und Kapazitätsmanagement verantwortlich für den strategischen Fahrplan, Netzfahrplan und Baufahrplan. Er war Initiator und erfolgreicher Umsetzer einer umfassenden Transformation der Geschäfts- und IT-Architektur.



Kramer, Urs

Prof. Dr. Urs Kramer ist seit 2009 Inhaber der Lehrprofessur für Öffentliches Recht an der Universität Passau. Sein Forschungsschwerpunkt liegt unter anderem im Eisenbahnrecht. Nach seinem Studium der Rechtswissenschaften an der Universität Marburg promovierte er dort 2002 zum Recht der Eisenbahninfrastruktur.



Hack, Linus

Linus Hack, Jurist (Univ.) ist Rechtsreferendar und seit 2023 wissenschaftlicher Mitarbeiter an der Lehrprofessur für Öffentliches Recht von Prof. Dr. Urs Kramer im Bereich Eisenbahnrecht. Er studierte Rechtswissenschaften mit Schwerpunkt „Rechtsdurchsetzung im Wirtschaftsrecht" an der Universität Passau.

Session 1

Emanuel von Heel, Philipp Scherer, Thorsten Büker, Alexandra Benz, Andreas Pfeifer, Nils Nießen, Michael J. Beck, Urs Kramer, Linus Hack, Niklas Ziegler



Ziegler, Niklas

Niklas Ziegler, Jurist (Univ.), LL.B. Legal Tech studierte bis 2024 Rechtswissenschaften an der Universität Passau mit Schwerpunkt im Bereich Digitalisierung. Seit Ende 2024 ist er wissenschaftlicher Mitarbeiter im Eisenbahnrecht und Doktorand an der Lehrprofessur für Öffentliches Recht von Prof. Dr. Urs Kramer an der Universität Passau.

Session 1
Alexander Kuckelberg

Servicebasierte Infrastrukturanpassung zur Optimierung von EBW- und (LST-) Planungsprozessen

Kuckelberg, Alexander

quattron GmbH

Zusammenfassung

Über die letzten Jahre haben sich stabile und etablierte Planungsprozess entwickelt, in denen sich unter anderem Aspekte Eisenbahnbetrieblicher Untersuchungen (EWBU) und PT1-Plaungen als „zwei Welten“ ergänzen. Ihre Koppelung ist jedoch nicht ohne Aufwand, Erfahrung und entsprechende Tools einfach herstellbar, was ggf. zu zeitaufwändigen Iterationszyklen führt, bspw. zur betrieblichen Bestätigung einer PT1-Planung.

Mit LaaS (LUKS as a Service) steht eine Sammlung verschiedener, aus eisenbahnbetrieblichen Kontexten stammender Funktionalitäten und Methoden als modulare Dienste zur Verfügung, um bewährte EBW-Funktionen dediziert und unabhängig nutzen zu können, um sie in neuen Kontexten über Schnittstellen einsetzen zu können. Eine dieser Services ist die entstehende skriptbasierte Infrastrukturanpassung, mit der es möglich wird, Modifikationen an Infrastrukturdaten bei großer Menge mit überschaubarem Aufwand durchzuführen, um die sich ergebenden Varianten in EBW-Untersuchungen nutzen zu können.

Mit Blick auf diese Services ist in einem Pilotprojekt zunächst die Machbarkeit der Kopplung von PT1-Planung und EBWU untersucht und bestätigt worden, was es grundsätzlich ermöglicht, EBW-Kennwerte in der PT1-Planung ohne bisher notwendige Iterationsschritte unmittelbar berechnen und berücksichtigen zu können. Eine automatisierte Kapselung und Anbindung an bestehende Planungstools steht aus, doch zeigen die ersten Ergebnisse eine einfache und schnelle Möglichkeit, betriebliche Kennwerte in der Planung berücksichtigen zu können.

Keywords: Infrastrukturplanung; Variantenbildung; EBWU; Prozessoptimierung

1 Einführung

Planungs- und Umsetzungsprozesse im Eisenbahnbereich betreffen in der Regel viele verschiedene Tätigkeiten, umspannen zum Teil längere Zeiträume und sind oft gewachsene und weiterentwickelte, damit aber auch bewährte und etablierte bzw. belastbare Prozesse. Einen erheblichen Teil nimmt dabei die Erneuerung und Neuerrichtung von Leit- und Sicherungstechnischen Komponenten und Infrastrukturbereichen ein.

Im Zusammenspiel von lang- und mittelfristiger betrieblicher Planung sowie der Umsetzungsplanung (konkret PT1-Planung zur LST-Planung) kommt es zu verschiedenen Berührungspunkten Eisenbahnbetriebswissenschaftlicher (EBW) Ansätze, Verfahren und Untersuchungen sowie der PT1-Planung: Frühzeitig legen EBW-Untersuchungen (EBWU) den Grundstein und konkretisieren die Idee, wie Eisenbahninfrastruktur durch neue oder angepasste Leit- und Sicherungstechnik optimierter, effizienter oder kostengünstiger zu nutzen ist und weisen bspw. den betrieblichen Nutzen von Baumaßnahmen nach. Die konkrete PT1-Planung ist dann ein eigener Schritt, diese betrieblichen Visionen in konkreten Baumaßnahmen abzubilden zu können. Dabei ergeben sich oftmals begründete und offenkundige Abweichungen von den EBWU-Ansätzen, z.B. aufgrund von Regelwerken, physischen Restriktionen und weiteren, in EBW-Untersuchungen unberücksichtigten Rahmenbedingungen, weshalb die resultierende Umsetzungsplanung zuweilen erneut mit einer angepassten EBWU bewertet und bestätigt oder aber zu modifizieren klassifiziert werden kann. Gerade dieses letzte Zusammenspiel ist zeitaufwändig und mit absehbar zunehmenden Bau- und Modernisierungstätigkeiten werden diese Prozesse zunehmend zeitraubend und in Summe Performance hemmend.

Die Motivation des in diesem Paper vorgestellten Ansatzes, EBWU-Ansätze und PT1-Planung enger miteinander zu verzahnen, resultiert genau aus diesen Überlegungen, dass EBWU-PT1-Iterationsschleifen reduziert werden oder im optimalen Zielzustand ganz entfallen können, weil die sonst in EBWU gewonnenen Erkenntnisse und Indikationen bereits „in Echtzeit“ während der PT1-Planung zumindest in Teilen ersichtlich werden und damit die PT1-Planung näher an betriebliche Bewertungen rücken und diese berücksichtigen kann. So wird in [1] beispielsweise explizit während der Erarbeitung einer neuen Blockteilung von EBW-Ansätzen Abstand genommen, „da diese […]

deutlich aufwendiger wären, keine schnellen Iterationen ermöglichen und auch keine zusätzlichen Erkenntnisse erwarten lassen.“, aber trotzdem final eine EBWU als Nachweis der korrekten Planung vorgesehen ist (und hoffentlich keine erneute Bearbeitungsschleife bedingt). Die nachfolgend vorgestellte Koppelung von EBW-Ansätzen mit der PT1-Welt ist entwickelt worden, um (auch) solche Überlegungen und Entscheidungen obsolet machen zu können und bei der Menge absehbarer Planungsleistungen im Rahmen einer breiten Digitalisierung Prozesse optimieren und beschleunigen, mithin auch Ressourcen im EBW-Bereich gewinnbringender einsetzen zu können.

Der aktuelle Stand der Technik sieht hier derzeit viele manuelle Schritte und Interaktionen zwischen Tools der jeweils betrachteten Welten vor, ein Austausch von Daten wird derzeit rudimentär über mehr, eher weniger zueinander passende Datenformate ermöglicht und auch manuelle Nacharbeiten und Ausführungen von Prozessschritten sind in einem Umfang üblich, der eine Automatisierung gerade im Angesicht der bevorstehenden Herausforderungen neuer Planungen nahelegt.

2 Ausführungsplanung, Planteil 1 (PT1)

Die Ausführungsplanung bildet sowohl fachlich als auch umfänglich den zentralen Teil der Planungsleistungen. Auf Basis der Entwurfs- und Genehmigungsplanung stellt sie die Realisierung des Bauvorhabens sicher. Ziel dieser Planungsphase ist die Erstellung der Ausführungspläne als Grundlage für die Mengenermittlung und damit als Vorbereitung der Vergabe.

Für den betrachteten Kontext des schienengebundenen Verkehrs wird nach Planteil 1 (PT1) und Planteil 2 (PT2) unterschieden, wobei PT1 allgemeingültige Planunterlagen der Sicherungsanlage umfasst, der PT2 die herstellerbezogene Planunterlagen gemäß Planteil 1 und den Anwendungsrichtlinien des Herstellers. Zusätzlich sind verschiedene Richtlinien wie die Ril 819 beim PT1 relevant.

Bei der Ausführungsplanung haben sich verschiedene Tools und Datenstandards etabliert. Daneben entstand über die letzten Jahre auch mit PlanPro ein inzwischen etablierter und sehr hilfreicher Ansatz, eine durchgängige elektronische Datenhaltung ohne Medienbrüche in der Leit- und Sicherungstechnik zu ermöglichen und damit Planungsprozesse auf dieser Ebene stark zu beschleunigen. Das Datenmodell enthält alle Planungsparameter für elektronische/digitale Stellwerke (ESTW, DSTW), ETCS,

streckenseitige Ausrüstungen für den automatisierten Bahnbetrieb (ATO TS), Bahnübergangssicherungsanlagen, Zugnummernmeldeanlagen und Zuglenkung.

PlanPro repräsentiert die geplante Infrastruktur als Knoten-Kantenmodell, welches für Kanten beispielsweise eine Kilometrierung und Streckenzugehörigkeit kennt. Die Nutzung von PlanPro-Daten während der Planungsphase ist daher auch der gewählte Ansatz, die Kopplung mit der EBWU-Welt zu realisieren.

Weitere Ausführungen zu PlanPro und dessen bisheriger Beitrag zur durchgängigen Datenhaltung in LST-Planungsprozessen sind bspw. [2] zu entnehmen.

3 Eisenbahnbetriebliche Untersuchungen (EBWU)

Ziel eisenbahnbetriebswissenschaftlicher Untersuchungen (EBWU) ist zumeist, Fahrwegkapazität und Betriebsqualität unter Variation der genannten Einflussgrößen und Beachtung des Systemverbunds Eisenbahn zu bewerten, womit je nach Kontext unterschiedliche Fragestellungen beantwortet werden: Machbarkeitsstudien, Analyse von Infrastruktur- und Betriebsdaten, Anforderungsmanagement oder Regelwerkserstellung und -begleitung etc.

Wenn im Kontext von Planungsprozessen und deren verschiedenen Stufen von Eisenbahnbetrieblichen Untersuchungen gesprochen wird, handelt es sich hierbei in der Regel um ein Themenfeld, für das große Infrastrukturbetreiber wie die DB InfraGO AG eigene Abteilungen und Kompetenzen aufgebaut haben und die auch bei PT1-Planungen wie bereits angeklungen einbezogen werden.

Typisch sind hier mikroskopische Spurplanmodelle, die als Knoten-Kantenmodell die Topologie modellieren, dabei Elemente der Kapazitätsplanung berücksichtigen und auch eigene Layer zur Modellierung betrieblich relevanter Aspekte der Sicherungstechnik und logisch-operativer Aspekte berücksichtigen, z.B. die bei einer Fahrplanung eingesetzten Betriebsstellenfahrwegen. Typisch ist hier der Begriff der Betriebsstelle und der Zugehörigkeit der meisten Daten zu einer Betriebsstelle. Zur Modellierung dieser Daten ist bei der DB InfraGO AG auch seit vielen Jahren eine eigene XML-Spezifikation im Einsatz, die diese Welt in ihrer Komplexität gut abdeckt und seit langem bereits einen Datenaustausch zwischen unterschiedlichen Tools unterstützt.

Für konstruktive Daten (d.h. geplante Zugfahrten bzw. Fahrplan), ist bei der DB InfraGO AG ebenfalls eine XML-Spezifikation im Einsatz, die seit Jahren (weiter-) entwickelt wird und auf dem Infrastrukturmodell aufsetzt und Anleihen der RailML-

Spezifikation aufweist. Speziell die konsequente Orientierung auf Betriebsstellen findet sich hier wieder, da bspw. die Sequenz der befahrenen Betriebsstellen einen essenziellen Anteil der Fahrplanung ausmacht und wegen eines Zwangs zu Wiederholungsfreiheit von Betriebsstellen innerhalb eines Zuglaufweges unmittelbar die Modellbildung von z.B. Konstruktionszügen mit bedingt.

4 Transformation und Kopplung der Datenwelten

Die beiden vorangegangenen Kapitel haben in kurzem Abriss die Grundlagen und Grundstrukturen der Datenmodelle der beiden zu verknüpfenden Welten aufgezeigt und angerissen.

Beide Datenwelten – also die Spurplanwelt der EBW und die PlanPro-Welt der anzubindenden PT1-Planung – sind für ihre Kontexte und Zwecke über Jahre gewachsen und weiterentwickelt worden, weisen Ähnlichkeiten aber auch deutliche Differenzen auf und eine Abbildung bzw. Wiedererkennung von Elementen zwischen den beiden Datenwelten ist der Kern einer Koppelung aber mitnichten unmittelbar augenscheinlich.

Dabei kommt in der Gesamtbetrachtung der Abbildung zwischen den Welten zugute, dass sie in der Regel gleiche Wurzeln haben und bei den lang- und mittelfristigen EBWU-Planungen bereits die topologischen Knoten (also Weichen, Kreuzungen und Kreuzungsweichen) vorhanden sind, die in der nachfolgenden PT1-Planung so übernommen und nicht mehr grundsätzlich verändert werden. Damit reduziert sich die primäre Verknüpfung auf Erkennung von gleichen Elementen auf deren Positionierung entlang von Strecken und Gleisen. Mit der Annahme, dass die „initialen“ EBWU-Daten nach wie vor auch die Basis der PT1-Planung sind, kann auch hier davon ausgegangen werden, dass beide Welten mindestens ähnliche Daten verwenden, speziell Strecken und Streckengleise der PlanPro-Seite noch der EBW-Welt entsprechen.



Abbildung 1: Navigation in PlanPro zwischen Objekten, Kanten, Abschnitten, Knoten.

Die Verknüpfung der Datenwelten erfolgt in verschiedenen Stufen (exemplarisch in Auszügen in vorheriger Abbildung dargestellt):

- Zuordnung zugehöriger Kanten der beiden Datenwelten über Streckenabschnitte und deren Streckennummern;
- Ableitung und Identifikation des Streckengleises durch Analyse der Kanten zwischen topologischen Knoten, ggf. mehrstufig über mehrere Knoten hinweg entlang des Stammgleises;
- Zuordnung einzelner Knoten aufgrund von Typ und Kilometrierung;
- Ggf. Name (dieses Kriterium ist nach ersten Erfahrungen aber kaum verlässlich).

Abgesehen von vielen Knotentypen, die die EBW-Welt kennt, sind die in diesem Kontext betrachteten Knotenelemente der LST in der Regel gut lokalisierbar und zwischen den Datenwelten assoziierbar, was die Transformation von Abschnittsknoten zwischen den Datenwelten mit ausreichender Qualität ermöglicht.

5 Konstruktionsdaten

Anders als Infrastrukturdaten sind Fahrplan- bzw. Konstruktionsdaten nur in der EBW-Welt vorgesehen, die PT1-Planung hat konzeptuell mit Zugfahrten oder konkreten Zugläufen zunächst keine Berührungspunkte.

Daher ist es für eine betriebliche Bewertung von Ist- und Plan-Zuständen auf EBW-Seite wichtig, einen für die betrachteten Bereiche relevanten Fahrplan mit langfristig repräsentativen Verkehren aufzustellen, abhängig vom verwendeten KPI möglicherweise

auch nur als Mengengerüste statt auskonstruiertem Fahrplan. Die Konstruktionsdaten sollten auf den Zielzustand abbilden, für den die Planung erfolgt und bei Angabe nutzbarer LST-Techniken diesen stets auch einbeziehen, also z.B. „ETCS nutzbar“ aktivieren oder weitere „Obermengen“ (z.B. „kann LZB und ETCS – ja nach Infrastruktur nutzbar) an Einstellungen für beide Szenarien bilden.

Als Ausgangspunkt für die prototypische Evaluierung dieses Ansatzes wurde auf den geltenden Jahresfahrplan zurückgegriffen, um damit die gewählten EBW-KPI zu berechnen.

In zukünftigen Anwendungsfällen kann ggf. auch auf konstruktive Daten „von extern“ verzichtet werden und mit dem nachfolgend skizzierten Serviceansatz auch direkt konkrete Zugfahrten mit übergeben werden, für die z.B. eine Mindestzugfolgeberechnung auf der (angepassten) Infrastruktur berechnet werden kann. Beispiel dieses Ansatzes sind repräsentative Zugkompositionen und denen Mengengerüste, mit denen beispielsweise der Deutschland-Takt betrachtet wurde.

6 Prototyp zur Kopplung

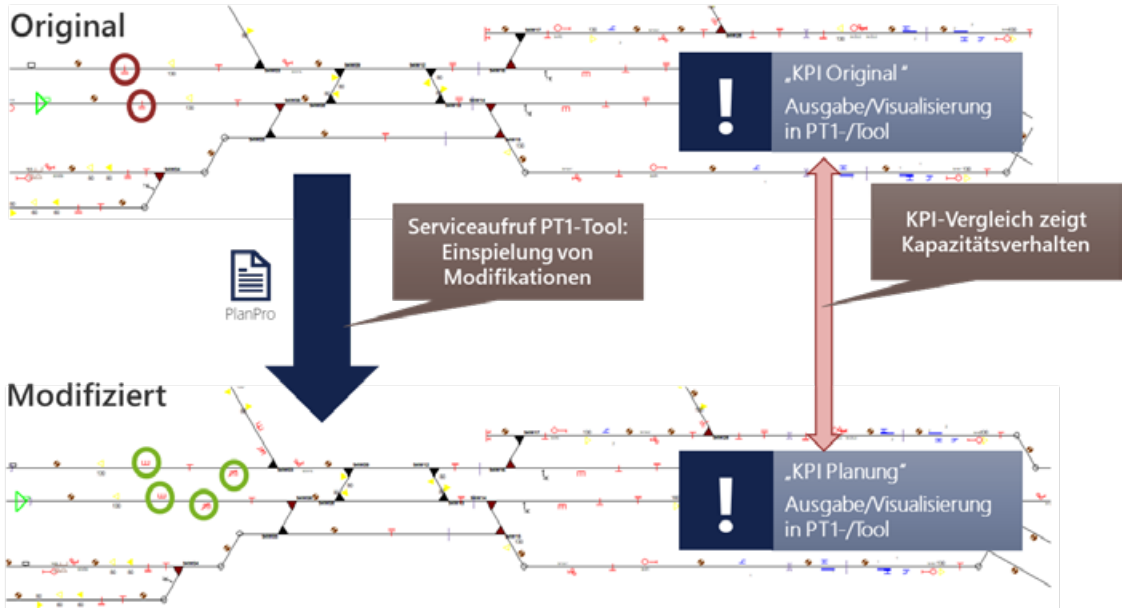


Abbildung 2: „Grundidee“ zur Koppelung und Ermittlung von EBW-Indikatoren

Als Vorstufe zur serviceorientierten Architektur und Anbindung wurde im Rahmen einer konkreten LST-PT1-Planung der Ansatz der Kopplung und die Nutzbarkeit von EBW-Bewertungen und KPI-Werten innerhalb der Planungsphase zunächst prototypisch und manuell nachgestellt, um den grundsätzlichen Nutzen abschätzen zu können bevor weitere Automatisierungs- und Anbindungsschritte unternommen werden.

Dazu wurde die EBW-Suite LUKS [5] verwendet und die Kopplung für das Planungsprojekt grob wie nachfolgend beschrieben nachgestellt.

Aufbau einer Datenbasis:

- Import der Ist-Infrastrukturdaten für den betrachteten Korridor plus Zulaufstrecken aus existieren XML-Datenquellen.
- Import der im Korridor verkehrenden Zugfahrten aus dem gültigen Jahresfahrplan.

Mit diesem Grundsetting wurde ein erster Indikatorwert zum Vergleich der Ist- und Plan-Situation gewählt. Neben verschiedenen simulativen Möglichkeiten aus LUKS wurde die Nennleistung als Leistungskennwert [5] einer analytischen Streckenbewertung gewählt,

da er zum einen fachlich zur Planung einer Streckenneuausrüstung passt, zum anderen aber auch eine sehr einfache Interpretation zulässt, die gut in Tendenzen interpretierbar und in ihrer Berechnung sehr schnell ist, was für interaktive ad-hoc-Auswertungen und Anfragen auch ein zentrales Kriterium sein wird. Notwendige Streckendefinitionen wurden vorgenommen und als Referenz für Ist- und alle Plan-Fälle hinterlegt.

Daneben wurde die in LaaS erprobte „Modify"-Funktionalität nach LUKS übernommen sowie das Kernstück einer Konvertierung, die Suche nach gleichen Bereichen und Elementen in Spurplan- und PlanPro-Daten und die Verknüpfung derselben zu späteren Auskunft und Verwendung, in ersten Versionen nutzbar gemacht und ebenfalls in LUKS eingebunden.

Zur prototypischen Kopplung wurden damit dann die folgenden Schritte durchgeführt:

- Ermittlung des gewählten Kennwertes für den aktuellen Infrastrukturstand, die vorgesehenen Konstruktionsdaten und Streckendefinitionen. Dieser Wert ist die Referenz für alle folgenden Plan-Zustände und basiert auf „der alten" Infrastruktur und deren LST-Ausrüstung.
- Einspielung einer PlanPro-Datei in den Ist-Spurplan: Dazu werden vorher bestimmte Elemente der PlanPro-Datei und die übertragene Lokation im Spurplan ermittelt, ein neues gemäß PlanPro-Element erstellt und in den Spurplan eingefügt sowie zuvor aufgrund der Streckendefinitionen „Platz geschaffen", indem in begrenzt-bestimmten Raum ausgewählte Elemente entfernt werden, um sie durch die in den PlanPro-Dateien enthaltenen Informationen zu ersetzen. Diese Übertragung beachtet räumliche und inhaltliche Filter und kann auf einzelne Streckenabschnitte, aber auch gezielt auf einzelnen LST-Elementen angewendet werden.
- Automatische Anpassung der Konstruktionsdaten, falls Laufwege neu gebildet werden müssen, jedoch keine Anpassung weiterer Angaben und Eigenschaften, um die Vergleichbarkeit von Ist und Plan zu gewährleisten.
- Erneute Kennwertberechnung mit modifizierter Infrastruktur unter ansonsten identischen Rahmenbedingungen. Dieser Wert ist dann der jeweilige Plan-Wert zum Stand der PT1-Planung und kann gegen den initial bestimmten Ist-Wert verglichen werden.

Der im letzten Schritt ermittelte Kennwert ist folglich der Indikator für die PT1-Planung, welche Änderungen bei Kapazitäten beim aktuellen Planungsstand zu erwarten sind. Mit diesem Vergleich, so die Grundidee dieses Ansatzes, werden spätere „vollwertige" EBW-

Untersuchungen weniger „Überraschungen“ ergeben und Iterationsschleifen im Prozess selbst reduziert, da EBW-Erkenntnisse auf diese Weise sehr einfach bereits innerhalb der PT1-Planungsprozesse Berücksichtigung finden können.

Der prototypische Ansatz hat bereits in erster „Version“ gezeigt, dass hier die ursprünglichen Erwartungen erfüllt wurden und nachvollziehbare und erklärbare KPI-Veränderungen sichtbar wurden, was sich später auch in einer „vollwertigen“ EBWU gezeigt hat. Eine exemplarische Gegenüberstellung beispielhafter Indikatoren des betrachteten Planungsprojektes zeigt Abbildung 3.

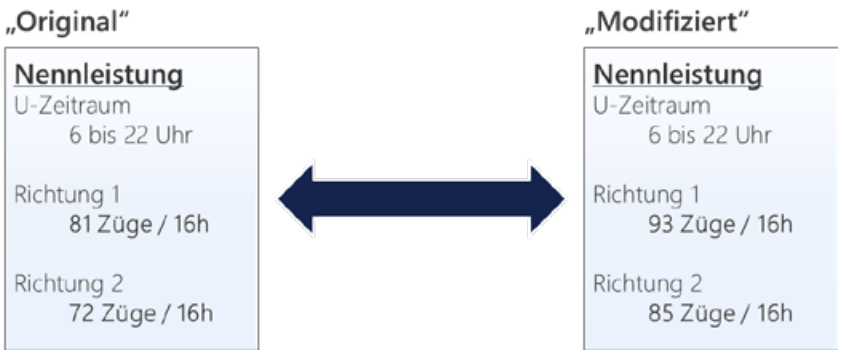


Abbildung 3: „Entwicklungstendenz“ der EBW-Indikatoren (Beispielprojekt) aufgrund vorgegebener Konfigurationen und beibehaltener Konstruktionsdaten

7 Servicebasierte Kopplung

Auch wenn der beschriebene prototypische Ansatz bereits gezeigt hat, dass die Kopplung grundsätzlich möglich ist und hilfreiche Resultate und Erkenntnisse ermittelt werden können, ist der manuelle Ansatz des Prototyps auf Basis von LUKS nicht praktikabel und muss im Sinn einer „Massenoperation“ zwingend automatisiert werden.

In der PT1-Planung werden aktuell wenige, etablierte Tools eingesetzt, die auch die PlanPro-Datenschnittstelle bedienen können, z.B. [3], [4]. Zur Kopplung und Automatisierung bieten sich nun zwei Ansätze an:

- Vollständige Integration der EBW-Ansätze in die genutzten Tools sowie
- Anbindung bestehender EBW-Tools an die PT1-Planungstools über entsprechende Schnittstellen.

Der erste Ansatz scheidet aufgrund des hohen Aufwandes und der erwartet schlechten Nachhaltigkeit solch einer „monolithischen Lösung“ de facto aus, er entspricht aus vielerlei Gründen auch nicht mehr Ansätzen moderner(er) Systemarchitekturen. Zudem sind Konstruktionsdaten konzeptuell auch nicht Teil der PT1-Planungstools.

Der zweite Ansatz reduziert die Anpassung der PT1-Tools auf wenige Schnittstellenaufrufe und die (überschaubar aufwändige) Integration derselben in die Oberflächen des Tools, womit aber die Trennung in einzelne Komponenten auch bei Verwendung anderer KPIs oder Anpassung von Konstruktionsdaten gelingt, so dass nachhaltigere Pflege- und Bearbeitungsprozesse an wenigen, jeweils aber zuständigen Stellen möglich werden.

Mit LaaS (LUKS as a Service) steht seit einigen Jahren zudem eine serviceorientierte Suite zur Verfügung, die grundsätzlich viele der angesprochenen Funktionalitäten bereits jetzt bietet und die im angerissenen Sinn einfach weiterentwickelt werden kann und damit als „EBW-Backend“ für die vorgesehenen Aufgaben eingesetzt werden kann.

Die Weiterentwicklung von LaaS zur automatisierten KPI-Auswertung ist aktuell in einer frühen Phase, eine Koppelung mit einem der PT1-Tools sinnvoll und im Sinn der vorgesehenen Nutzung und Optimierung notwendig, wenngleich der Service aufgrund der offenen Aufrufmöglichkeiten sicherlich auch in vielen weiteren Anwendungen, bei denen Infrastrukturanpassungen als Massenoperation oder per Skript erfolgen soll, bereits jetzt viel zur Effizienz- und Performancesteigerung beitragen kann.

8 Bewertung und weitere Arbeiten

Mit dem vorgestellten Ansatz wurde gezeigt, dass die Grenzen der PT1-Planungswelt und der EBW-Untersuchungen zumindest verschwimmen können und es Wege gibt, eisenbahnbetriebliches Wissen in Planungsprozessen frühzeitig zu berücksichtigen und mindestens Tendenzen der betrieblichen Auswirkungen einer Planung aufzeigen zu können.

Dennoch ist dieser Ansatz vor allem bei „echten“ EBW-Experten insofern fraglich, weil EBW-Untersuchungen in der Regel immer noch viel Erfahrung, Fachwissen und Kenntnis der Materie und angewandter Verfahren erfordert, um nachhaltig valide und belastbare Resultate erzielen und Berechnungen richtig interpretieren zu können. Diese Hürde bleibt, aber sie ist im gezeigten Verfahren zumindest in allen EBW-Berechnungen gleich hoch, was mindestens eine Abschätzung betrieblicher Tendenzen dann doch valide erscheinend macht.

Dies mag auch Stimmen, die eine Gefahr eben in der Interpretation und Deutung von EBW-Ergebnissen befürchten, eine erste Einschätzung geben: Der Ansatz macht EBW-Untersuchungen mitnichten obsolet, aber er kann dazu beitragen, Iterationsschleifen deutlich zu reduzieren und die Qualität von PT1-Planungen mit Blick auf betriebliche

Bewertungen zu erhöhen und damit die Menge anstehender Planungsaufgaben letztlich effizienter bewältigen zu können.

Zudem realisieren LUKS, LaaS und andere EBW-Tools bereits jetzt verschiedene KPI (Simulationen, UIC406, Warteschlangenbasierte Ansätze) für verschiedene Anwendungen und Spezialdomänen und auch die Standardisierung der Fahrplanleistung als spezifiziertem KPI ist bei der DB InfraGO AG im Fluss, so dass bei Art und Aussage des ermittelten KPI weitere Entwicklungen zu erwarten sind und ein gezielter Einsatz nach Situation und Szenario aufgrund der Serviceorientierung einfach möglich wird.

Mit der aktuellen Umsetzung als manuell zu bedienender Prototyp ist die Handhabung des Prozesses sperrig, aber es sind an dieser Stelle „nur noch“ Prozessschritte, die automatisiert und in Services gebracht werden müssen. Die Arbeiten schreiten in LaaS voran, doch stehen bisher Beispielprojekte und Pilotanwendungen zur direkten Kopplung und wirklichen Erprobung bzw. Demonstration der Nutzbarkeit im Planungsalltag aus.

Konkret stehen in der nächsten Zeit die Weiterentwicklung und Evaluation der PlanPro-Spurplan-Verknüpfung an, die Spezifikation der Serviceaufrufe und deren weitere Implementierung und ggf. erste Schritte zur Einbindung in bestehende PlanPro-Tools.

Als eine weitere Erkenntnis der prototypischen Umsetzung erscheinen zudem weitere Entwicklungsschritte sinnvoll und hilfreich: Ein Konverter von Spurplan in PlanPro, um aus dem ersten Schritt einer lang- bis mittelfristigen EBWU-basierten Planung eine in Topologie und Benennung konsistente initiale Grundlage der PT1-Planung zu schaffen, was die Verknüpfungsergebnisse einfacher und verlässlicher macht und damit die durchgängige Datenhaltung auch zwischen unterschiedlichen Datendomänen fördert.

Ein weiterer Aspekt dient der Weiterentwicklung und Bereitstellung konstruktiver Daten, die sicherlich eine wichtige Rolle bei der letztlichen KPI-Berechnung spielt und kaum aus der PT1-Planung bedient werden kann. Denkbar sind bspw. Ansätze, Zugzahlen oder Fahrpläne als Eingaben direkt zu übergeben bzw. vorzuhalten oder auch Mindestzugfolgezahlen (statt fortgeschrittener EBW-KPI-Zahlen) speziell auf per Aufruf definierbare Zugfahrten auszuwerten.

Referenzen

1. J. Denißen, M. Flieger, M. Kümmling, M. Küpper, S. Wanstrath, *Optimierung der Blockteilung mit ETCS Level 2 im Digitalen Knoten Stuttgart*, SIGNAL + DRAHT (113) 7+8 / 2021

2. Christoph Klaus, „Die digitale LST-Planung (PlanPro) als Wegbereiter der Digitalen Schiene“, SIGNAL + DRAHT (112) 1+2 / 2020
3. ProVi LST, https://www.provi-cad.de/produkte/lst/
4. ProSig/Stellwerkslogik http://www.prosig.de/
5. F. Weymann, S. Kurby, D. Janecek, *Deutschlandweite Kapazitätsaussagen mittels Automatisierung von eisenbahnbetriebswissenschaftlichen Untersuchungen*, Proceedings of the 2nd International Railway Symposium Aachen 2019 (IRSA 2019)

Session 1
Tamme Emunds, Nils Nießen

Approximation von Bediensystemen zur Leistungsfähigkeitsbestimmung von Eisenbahninfrastruktur in Bahnhöfen

Emunds, Tamme[1] and Nießen, Nils[1]

[1]RWTH Aachen University, Verkehrswissenschaftliches Institut

Zusammenfassung

Die existierenden bedientheoretischen Methoden zur Leistungsfähigkeitsbetrachtung von Eisenbahninfrastruktur in Bahnhöfen beschreiben bisher insbesondere die isolierte Leistungsfähigkeit von Gleisgruppen oder Gesamtfahrstraßenknoten. In diesem Beitrag wird eine Methode zur Analyse der gesamten Infrastruktur im Bahnhofsbereich vorgestellt. Dazu werden die Ankunfts- und Abfahrtsprozesse in den Gesamtfahrstraßenknoten approximiert und zusammen mit den Verkehrshalten in der Gleisgruppe in einem kombinierten Bediensystem beschrieben. Dieses kann als zeitkontinuierliche Markov-Kette dargestellt und mit Hilfe von probabilistischem Model-Checking analysiert werden. Mit der hier beschriebenen Approximationsmethode können erstmals Kennzahlen zur Leistungsfähigkeit gesamter Bahnhöfe mit einer hohen Anzahl an Fahrstraßen berechnet werden. Die errechneten Kennzahlen bieten eine wichtige Entscheidungsgrundlage für Infrastrukturbetreiber zur Langfrist-Analyse von Infrastrukturmaßnahmen.

Keywords: Leistungsfähigkeit, Infrastrukturdimensionierung, Bediensystem, zeitkontinuierliche Markov-Kette

1. Einleitung

Zur strategischen Planung der Weiterentwicklung von Eisenbahnnetzen ist die frühzeitige Quantifizierung der Leistungsfähigkeit von Eisenbahninfrastruktur von großer Bedeu-

tung. In vielen Fällen wird die Infrastruktur für mehrere Dekaden geplant, es ist dementsprechend essentiell fahrplanunabhängige Methoden zur Bewertung von Eisenbahninfrastruktur einzusetzen. Strecken und Knoten können zur fahrplanunabhängigen Planung mit Bediensystemen modelliert werden. Während Strecken als einkanalige Systeme analysiert werden können, erfordert die Bewertung von Bahnhöfen die Modellierung von komplexeren Bediensystemen.

Zur Berechnung der planmäßigen Wartezeiten in Bahnhöfen werden diese in Eisenbahnbetriebswissenschaftlichen Untersuchungen (vgl. [4]) in *Gleisgruppen* (GG) und *Gesamtfahrstraßenknoten* (GFK) aufgeteilt. Eine Gleisgruppe beschreibt dabei den Bereich eines Knotens, in welchem sich die Gleise für Verkehrshalte und Wartevorgänge befinden. Die Gleisgruppe schließt an einem (Kopfbahnhof) oder mehreren Enden an Gesamtfahrstraßenknoten an, in welchen die eingehenden Streckengleise auf die unterschiedlichen Gleise der Gleisgruppe und *durchgehenden Hauptgleise* (HG) aufgeteilt werden. Im Gegensatz zu den Gleisen in der Gleisgruppe sind in großen Knoten auf den durchgehenden Hauptgleisen keine planmäßigen Halte zulässig.

Die Leistungsfähigkeit von Gleisgruppen kann mit mehrkanaligen Bediensystemen [15, 11] untersucht werden, in welchen jedes Gleis einen parallelen Bedienkanal darstellt. Für die Berechnung der planmäßigen Wartezeiten auf Gesamtfahrstraßenknoten müssen nicht nur parallele Bedienungen, sondern insbesondere auch Konflikte von unterschiedlichen Fahrstraßen modelliert werden. Die Belegungskonflikte werden entweder über die Verwendung eines einkanaliges Ersatzsystems [18, 19] aufgelöst oder detailliert auf Ebene von Teilfahrstraßenknoten [13, 16] oder Fahrwegen [6] betrachtet.

Um eine Berechnung der planmäßigen Leistungsfähigkeit im Gesamtsystem aus Gleisgruppe und Gesamtfahrstraßenknoten zu ermöglichen, wird in [7] der Ansatz aus [6] zur Leistungsfähigkeitsanalyse von Abzweigstellen weiterentwickelt. Dabei wird für jeden Fahrweg ein eigenes Bediensystem (mit den drei Bedienprozessen Einfahrt, Halt und Ausfahrt und eigenem Ankunftsprozess) eingeführt, um eine genaue Engpassanalyse zu ermöglichen. Andere bedientheoretische Methoden zur Analyse der Leistungsfähigkeit von gesamten Eisenbahnknoten nutzen entweder Simulationen zur Kalkulation eines Warteschlangennetzwerks in Güterverkehrsknoten [1], oder approximieren die außerplanmäßigen Wartezeiten durch die Einführung zusätzlicher Wartezeiten zur Modellierung von Bedienkonflikten in Gesamtfahrstraßenknoten [8].

Das existierende Modell aus [7] ermöglicht zwar die Berechnung fahrwegscharfer Leistungskennzahlen, der Berechnungsaufwand steigt aber exponentiell mit der Anzahl der betrachteten Fahrwege. Im Folgenden soll dieses Modell deshalb so angepasst werden, dass eine Berechnung unabhängig von der Größe des betrachteten Knotens möglich ist.

In diesem Beitrag stellen wir ein Verfahren zur Analyse der planmäßigen Leistungsfähigkeit von gesamten Bahnhöfen vor. Dabei werden die Gesamtfahrstraßenknoten mit Hilfe der Verkettungsmethode nach Schwanhäußer [17] abstrahiert und die entstehenden Bedienprozesse in einem gemeinsamen Modell mit den Bedienprozessen auf der Gleisgruppe kombiniert. Das Gesamtsystem wird in einem Bediensystem modelliert und als zeitkontinuierliche Markov-Kette dargestellt. Darüber können unterschiedliche Kennzahlen zur planmäßigen Leistungsfähigkeit der analysierten Bahnhofsinfrastruktur errechnet werden.

Abschnitt 2 führt dazu zunächst die Modellierung von Eisenbahnknoten als Bediensystem ein. Ein Fokus liegt darin auf der Approximation des Gesamtsystems (2.2) und der Formulierung als Markov-Kette (2.3.1). Anschließend wird die Methode am Beispiel des Knotens *Aachen Hbf* getestet und mit den etablierten Verfahren zur isolierten Betrachtung der GFKs und GG verglichen (Abschnitt 3). In Abschnitt 4 folgt eine Diskussion der erzielten Ergebnisse, sowie die Entwicklung der sich aus den Ergebnissen ergebenden Forschungsfragen.

2. Bediensysteme zur Analyse von Eisenbahnknoten

In diesem Abschnitt wird das bedienungstheoretische Modell zur Leistungsanalyse von Bahnhöfen eingeführt.

2.1. Bediensysteme zur Kapazitätsanalyse

Ein Bediensystem ist ein formales Modell zur Beschreibung von Ankunfts- und Bedienprozessen von den verschiedenen Objekten in einem System. Im Rahmen der Anwendung zur Bestimmung der planmäßigen Leistungsfähigkeit werden dabei Anfragen der Eisenbahnverkehrsunternehmen zur Belegung der zur Verfügung stehenden Infrastruktur als Objekte betrachtet. Eine Belegungsanfrage stellt somit die Forderung nach einer Trasse für eine Zugfahrt dar, im Folgenden bezeichnen wir Anfragen auch als Züge. Die Anfra-

gen beschreiben sowohl einen Ankunftszeitpunkt, zu dem die Belegung starten soll, als auch die geplante Dauer der Belegung, die *Bedienzeit*. Die Zeit zwischen zwei Ankunftszeitpunkten wird im Folgenden als *Zwischenankunftszeit* bezeichnet. Da sowohl der Ankunftszeitpunkt als auch die Bedienzeit der Anfragen als zufällig verteilt angenommen werden, werden Ankunfts- und Bedienprozess eines Bediensystems mit Wahrscheinlichkeitsverteilungen modelliert. Außerdem können für das Bediensystem eine Anzahl k an parallelen Bedienkanälen und eine Anzahl m an Wartepositionen, an denen eingetroffene aber noch nicht bediente Anfragen warten können, vorgegeben werden. Eine grundlegende Einführung in Bediensysteme und Anwendungen in der Verkehrswissenschaft ist beispielsweise in [9] zu finden.

In analytischen Kapazitätsuntersuchungen wird der Ankunftsprozess durch die mittlere Zwischenankunftszeit t_A sowie den Variationskoeffizienten v_A, und der Bedienprozess durch die mittlere Bedienzeit t_B sowie den Variationskoeffizienten v_B beschrieben. Das Modell unterstellt dabei Ankunftsraten $\lambda = \frac{n_{\text{ges}}}{U}$, welche die Anzahl aller Anfragen n_{ges} im Zeithorizont U darstellen, und Bedienraten $\mu = \frac{1}{t_B}$.

Eine wichtige Kennzahl, der *Belegungsgrad* $\rho = \frac{\lambda}{k \cdot \mu}$, beschreibt die Auslastung des Bediensystems. Aufgrund der Zufälligkeit der Ankunfts- und Bedienprozesse kann sich eine Warteschlange vor der Bedienung im System bilden, die erwartete *Länge der Warteschlange* L ist eine weitere Kennzahl für das Bediensystem. Für ein Bediensystem mit einer begrenzten Anzahl an Wartepositionen $m < \infty$ können nur begrenzt viele Forderungen in der Warteschlange warten; Forderungen, die bei voller Warteschlange ankommen, verlassen das System ohne Bedienung. Die Wahrscheinlichkeit, dass eine ankommende Forderung nicht direkt bedient werden kann, wird als *Wartewahrscheinlichkeit* p_W bezeichnet.

Für einkanalige ($k = 1$) Bediensysteme mit exponentialverteilten Zwischenankunfts- und Bedienzeiten (M/M) lassen sich die geschlossene Formel $L_{M/M/1} = \frac{\rho^2}{1-\rho}$ zur Berechnung der Warteschlangenlänge L und $p_W = \rho$ als Ausdruck der Wartewahrscheinlichkeit angeben.

Zur Analyse von Bediensystemen mit beliebigen, stoch. unabhängigen Wahrscheinlichkeitsverteilungen für den Ankunfts- und Bedienprozess (GI/GI) existieren Approximationsformeln. Beispielsweise kann mit der Formel von Pollaczek-Chintchin [14, 9] die

erwartete Länge der Warteschlange

$$L_{M/GI/1} = \frac{\rho^2}{1-\rho} \cdot \frac{1+v_B^2}{2} = L_{M/M/1} \cdot \frac{1+v_B^2}{2} \tag{1}$$

in einem einkanaligen M/GI System berechnet werden. Für mehrkanalige $k \in \mathbb{N}$ GI/GI Systeme kann die Formel von Hertel verwendet werden. Sie beschreibt die erwartete Länge der Warteschlange

$$L_{GI/GI/k} \approx L_{M/M/k} \cdot \frac{1}{\gamma}, \tag{2}$$

in Abhängigkeit von

$$\gamma = \frac{2}{c \cdot v_B^2 + v_A^2} \tag{3}$$

und

$$c = \left(\frac{\rho}{k}\right)^{1-v_A^2} \cdot (1+v_A^2) - v_A^2. \tag{4}$$

Um die Kennzahlen von komplexeren Bediensystemen zu berechnen, können Bediensysteme mit exponentialverteilten Ankunfts- und Bedienprozessen (M/M) als zeitkontinuierliche Markov-Ketten dargestellt werden. Die Kennzahlen können dann als Kombination von Zustandswahrscheinlichkeiten im stationären Zustand ausgedrückt werden.

Eine entsprechende Markov-Kette zur Analyse der Kennzahlen des hier betrachteten Bahnhof-Bediensystems wird in Abschnitt 2.3.1 beschrieben. Zunächst wird die Modellierung der Prozesse innerhalb eines Eisenbahnknotens eingeführt.

2.2. Modellierung von Eisenbahnknoten

Betrachten wir die Infrastruktur eine Beispielbahnhofs, der aus zwei Gesamtfahrstraßenknoten (GFK) West (W) und Ost (O), einer Gleisgruppe (GG) und zwei durchgehenden Hauptgleisen (HG) besteht.

In Abbildung 1 wird die Modellierung des Beispiel-Bahnhofs aus Abschnitt 3 als kombiniertes Bediensystem skizziert. Die verschiedenen Fahrstraßen eines Bahnhofs können anhand des in ihrer Einfahrt verwendeten GFK, sowie der Angabe ob sie in der Gleisgruppe halten (bzw. wenden) oder auf den Hauptgleisen ohne Halt durchfahren, zu Strömen zusammengefasst werden. Dabei kann entweder ein GFK zur Einfahrt und der andere zur Ausfahrt genutzt werden, oder, durch die Durchführung einer Wende in der Gleisgruppe,

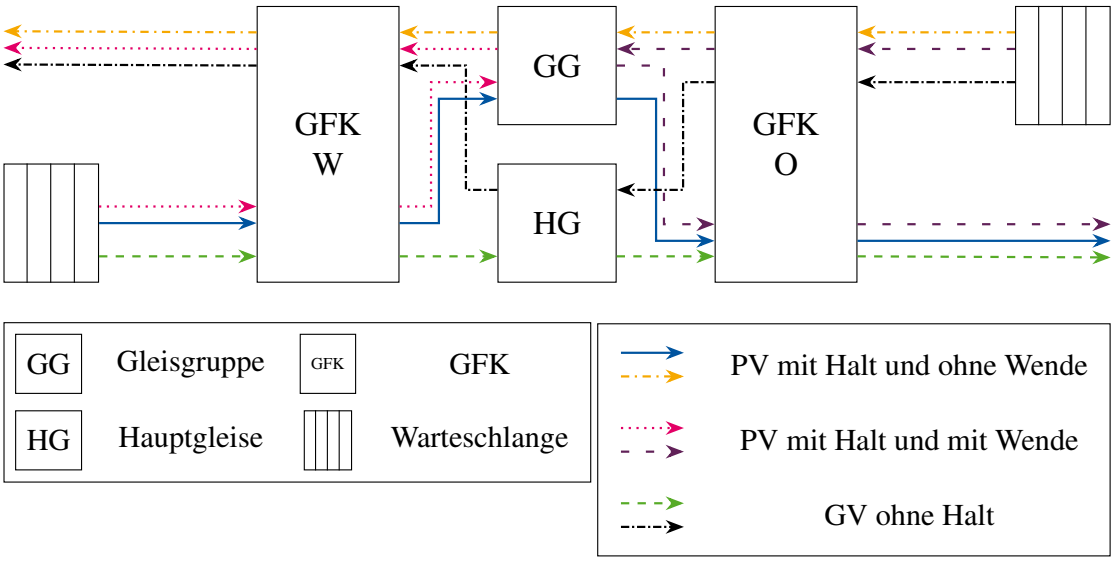


Abbildung 1: Skizze der möglichen Ströme für den Bahnhof aus Abschnitt 3 mit einer Gleisgruppe, zwei durchgehenden Hauptgleisen und zwei Gesamtfahrstraßenknoten (GFKs).

derselbe GFK für Ein- und Ausfahrt benutzt werden. Die mehrkanalige Gleisgruppe steht dabei für alle haltenden Anfragen zur Verfügung; Züge ohne planmäßigen Halt werden auf den durchgehenden Hauptgleisen geplant.

Der Bedienprozess einer Anfrage lässt sich dementsprechend in die drei Phasen Einfahrt, Halt und Ausfahrt aufteilen. Für Anfragen ohne planmäßigen Halt werden nur die beiden Phasen zur Ein- und Ausfahrt betrachtet. Wir bezeichnen die Bedienzeiten in den einzelnen Phasen als Einfahrsperrzeit t_B^{Ein}, Haltezeit t_B^{Halt} und Ausfahrsperrzeit t_B^{Aus}. Die Sperrzeiten sollten typischerweise aus einem mikroskopischen Modell errechnet werden, siehe bspw. [10]. Falls keine mikroskopischen Daten verfügbar sind, müssen die Sperrzeiten geschätzt werden. Die Haltezeiten können beispielsweise aus historischen Daten oder Fahrplänen in ähnlichen Knotenpunkten für die einzelnen Verkehrsarten und unter Beachtung der geforderten Wenden geschätzt werden.

Insgesamt können Forderungen aus zwei verschiedenen Ankunftsströmen in das System eintreten, wir bezeichnen die Anzahl Anfragen mit n_{W}^{An} (GFK W) und n_{O}^{An} (GFK O). Anhand der *Durchfahrtswahrscheinlichkeit* $p_{\text{G}}^{\text{HG}} = n_{\text{G}}^{\text{HG}}/n_{\text{G}}^{\text{An}}$, welche den Anteil der durchfahrenden Züge $n_{\text{G, HG}}$ beschreibt, können die ankommenden Anfragen an einem GFK G auf die Gleisgruppe und die Hauptgleise aufgeteilt werden. Analog beschreibt die *Wendewahrscheinlichkeit* p_G^{Wende} den Anteil der wendenden Zugfahrten in der Gleisgruppe an

der Gesamtzahl der über GFK G ankommenden Anfragen.

In der hier diskutierten Modellierung des Gesamtsystems werden die Bedienprozesse der Ein- bzw. Ausfahrt auf den Gesamtfahrstraßenknoten A und B durch jeweils ein einkanaliges Bediensystem approximiert.

2.3. Kombinierte Modellierung des Gesamtsystems

Das in Abbildung 1 dargestellte Gesamtsystem eines Bahnhofs wird im Folgenden als Markov-Kette beschrieben. Im Anschluss wird die Methode zur Berechnung der Leistungsfähigkeitskennzahlen kurz erläutert.

Im Modell werden die Ankunftsraten $\lambda_{\mathrm{W}}^{\mathrm{An}} = n_{\mathrm{W}}^{\mathrm{An}}/U$ und $\lambda_{\mathrm{O}}^{\mathrm{An}} = n_{\mathrm{O}}^{\mathrm{An}}/U$ an den beiden GFKs, sowie der dreigeteilte Bedienprozess mit Bedienraten μ_B^{Ein}, μ_B^{Halt} und μ_B^{Aus} betrachtet. Dabei ist zu beachten, dass aufgrund der Approximation aus Abschnitt A für einen GFK G dieselbe Bedienrate $\mu_{B,\mathrm{G}}$ sowohl für die Ein- als auch für die Ausfahrt unterstellt wird. Weiterhin bezeichnen wir die Anzahl der Gleise in der Gleisgruppe mit k_{GG}.

2.3.1. Formulierung als zeitkontinuierliche Markov-Kette

Die zeitkontinuierliche Markov-Kette zur Modellierung des zuvor beschriebene Bediensystems kann über ihre Zustandsmenge S und allen Übergängen U zwischen den Zuständen beschrieben werden. Dabei kann ein Zustand $s = (s_{\mathrm{GFK\,O}}, s_{\mathrm{GFK\,W}}, s_{\mathrm{GG}}, s_{\mathrm{HG}}) \in S$ als Tupel der Einzelsystemzustände $s_{\mathrm{GFK\,O}}$, $s_{\mathrm{GFK\,W}}$, s_{GG}, s_{HG} beschrieben werden.

Die Eigenschaften des Teilzustands $s_{\mathrm{GFK\,O}} = \left(q_O, q_O^n, e_O, d_O\right)$ von GFK O (und analog von GFK W) beschreiben die Anzahl der aktuell auf Einfahrt wartenden Forderungen $q_O \in [m]_0 = \{0, \ldots, m\}$, sowie die Tatsache, ob die nächste Forderung in der Warteschlange durchfährt $q_O^n = 1$, oder in der Gleisgruppe hält $q_O^n = 0$. Außerdem wird für haltende Züge beschrieben, ob der GFK zum aktuellen Zeitpunkt über eine einfahrende $e_O = 1$ oder ausfahrende $d_O = 1$ Zugfahrt belegt ist oder gerade keine entsprechende Belegung ($e_O = d_O = 0$) erfolgt.

Für den Gleisgruppenzustand $s_{\mathrm{GG}} = \left(tr_O, tr_W, tr_O^w, tr_W^w\right)$ wird die Anzahl der aktuell haltenden Züge mit Ausfahrt über GFK O $tr_O \in [k_{GG}]$ oder GFK W $tr_W \in [k_{GG}]$ nachgehal-

ten. Forderungen, die bereits ihre Mindesthaltezeit eingehalten haben und auf Ausfahrt warten, werden ebenfalls nach Ausfahrrichtung getrennt über $tr_O^w, tr_W^w \in [k_{GG}]$ modelliert.

Über den Zustand der Hauptgleise $s_{\mathrm{HG}} = (hg_O, hg_W)$ lässt sich nachhalten, in welcher Bedienphase sich die Hauptgleise aus Richtung GFK O (oder analog GFK W) befinden. Sie können entweder aktuell nicht in Bedienung sein $hg_O = 0$, oder einen durchfahrenden Zug $hg_O \in \{1,2\}$ in einer der beiden Bedienungsphasen beschreiben. Dabei ist die Bedienung in zwei Phasen aufgeteilt, um die Bedienzeiten der einkanaligen Ersatzsysteme für die beiden GFKs getrennt zu betrachten. Zur Durchfahrt eines Zuges sind aber immer beide GFKs für anderweitige Nutzung gesperrt.

Die in den Zuständen beschriebenen Eigenschaften der mehrkanaligen Gleisgruppe und der als einkanalig approximierten GFKs werden über die Übergänge im Markov-Ketten-Modell miteinander verknüpft. Eine genaue Auflistung aller Übergänge und der entsprechenden Übergangsraten ist in Anhang B angegeben.

2.3.2. Berechnung von Leistungsfähigkeitskennzahlen

Mit Hilfe der im vorherigen Abschnitt beschriebenen zeitkontinuierlichen Markov-Kette, können Kennzahlen zur Beschreibung der Leistungsfähigkeit des Gesamtsystems berechnet werden. Dafür werden die Kennzahlen in Abhängigkeit von den Zustandswahrscheinlichkeiten der Markov-Kette im stationären Zustand ausgedrückt. In dieser Arbeit verwenden wir die Methodik aus [6, 7] und benutzen das Python-Modul `stormpy` [12] um die Kennzahlen mit Hilfe von Algorithmen zum probabilistischem Model-Checking zu erhalten.

Das über die Markov-Kette formalisierte System beschreibt dabei lediglich exponentialverteilte Übergangszeiten. Zur Approximation der unterstellten *GI* Ankunfts- und Bedienprozesse benutzen wir die Formeln von Hertel aus (2)-(4). Für die Schätzung der erwarteten Warteschlangenlänge verwenden wir einen Parameter von $k = 1$ und für die Schätzung der Wartewahrscheinlichkeit unterstellen wir ein $k = k_{GG}$-kanaliges Bediensystem. Hierbei ist anzumerken, dass die Approximationsformeln (2)-(4) von Hertel primär zur Abschätzung der Warteschlangenlänge geeignet ist. Da sich die Wartewahrscheinlichkeit proportional zur Warteschlangenlänge verhält, benutzen wir die Formel aus (2) aber auch zur Approximation der Wartewahrscheinlichkeit.

Um eine Einschätzung über die Leistungsfähigkeit zu ermitteln, vergleichen wir die errechneten Warteschlangenlängen und Wartewahrscheinlichkeiten mit den Grenzwerten aus [4]. Dabei wird für die Warteschlangenlänge ein Grenzwert (vgl. [18, 4]) von

$$L_{\max} = 0.479 \cdot \exp(-1.3 p_{\mathrm{Rz}}) \cdot 1/\varphi \,, \tag{5}$$

in Abhängigkeit von dem Anteil der Personenzüge p_{Rz} am Betriebsprogramm und der Verkettungszahl φ des Gesamtfahrstraßenknotens (siehe Anhang A), unterstellt. Für die Wartewahrscheinlichkeit in Personenbahnhöfen wird ein Grenzwert von $p_{W,\max} = 0.05$ festgesetzt. Der Quotient aus errechneter Kennzahl und Grenzwert wird im Folgenden als *Qualitätsfaktor* bezeichnet. Die optimale Qualität ist bei einem Qualitätsfaktor von 1.0 erreicht, Werte darüber deuten auf eine zu geringe Leistungsfähigkeit für das unterstellte Betriebsprogramm, während ein Qualitätsfaktor deutlich unter 1.0 für eine unwirtschaftliche Planung spricht.

Im nächsten Abschnitt werden die für ein Anwendungsbeispiel so errechneten Kennzahlen des Gesamtsystems mit den Kennzahlen zu isoliert betrachteter Gleisgruppe und Gesamtfahrstraßenknoten verglichen.

3. Anwendungsbeispiel

Zur Anwendung der skizzierten Methode betrachten wir im Folgenden die Infrastruktur von *Aachen Hbf*. Im Anhang C sind der Gleisplan und ein Beispiel-Fahrplan für eine Stunde im Juli 2025 zu finden. Damit wurde ein Betriebsprogramm mit zugehörigen Fahrstraßen und entsprechenden Konflikten konstruiert. Außerdem wurden die Sperrzeiten für Aus- und Einfahrten geschätzt. Die hier ermittelten Aussagen zur Leistungsfähigkeit sind also nicht unbedingt als repräsentativ für die tatsächliche Infrastruktur zu bewerten.

Insgesamt sind für die analysierte Spitzenstunde 17 Zugfahrten angefragt. Davon kommen 7 über den GFK O aus Richtung Rothe Erde (ARE) und 10 über den GFK W aus Richtung Belgien (BEL) oder Aachen-Schanz (ASZ) in das Gesamtsystem an.

Die entsprechenden Ankunftsraten können zusammen mit anderen Eingangsparametern der Tabelle 1 entnommen werden. Alle Parameter wurden auf Grundlage der Daten aus Anhang C konstruiert. Für den Variationskoeffizient der Bedienung in den GFKs wurden die Variationskoeffizienten der drei Teilprozesse Einfahrt, Halt und Ausfahrt mit der Me-

Tabelle 1: Parameter für GFK Ost, GFK West und GG

Parameter	GFK Ost	GFK West	GG
Ankunftsrate [λ]	0.117	0.167	0.250
Bedienrate [μ]	0.690	0.632	0.078
Wendewahrscheinlichkeit [p_G^{Wende}]	0.333	0.556	–
Durchfahrtswahrscheinlichkeit [p_G^{HG}]	0.143	0.100	–
Variationskoeff. Ankunft	0.494	0.733	–
Variationskoeff. Bedienung	0.949	0.958	0.876
Variationskoeff. Haltezeit	–	–	0.876
Anzahl Gleise	–	–	7
Verkettungszahl [φ]	0.6327	0.52	–

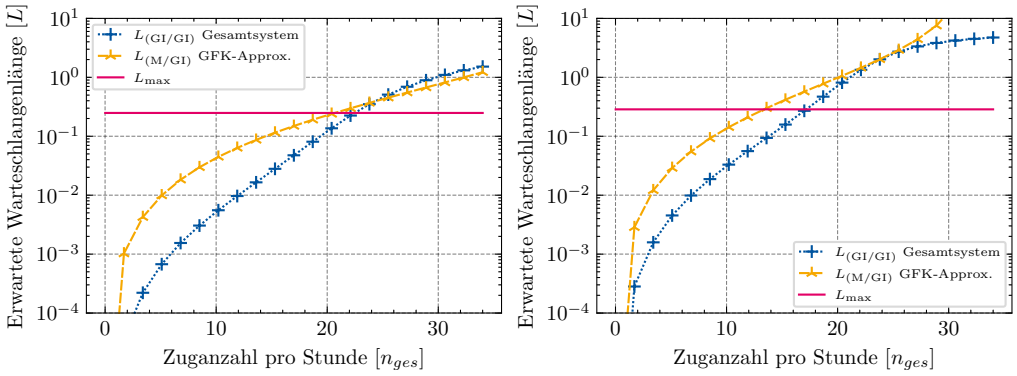
thode aus [7] kombiniert. Im Fall der isolierten Betrachtung mit der Verkettungs-Methode nach Abschnitt A wurde für jeden GFK der entsprechende Variationskoeffizient der verketteten Belegungen im GFK benutzt. Dieser beträgt 0.973 für GFK West und 0.855 für GKF Ost.

Das Gesamtsystem wurde mit Hilfe der in Abschnitt 2.3.1 vorgestellten zeitkontinuierlichen Markov-Kette mit einer Anzahl von $m = 10$ Wartepositionen pro GFK modelliert, die entsprechende Markov-Kette besteht aus $1,449,096$ Zuständen und $7,798,208$ Übergängen. Das probabilistische Model-Checking benötigte auf einem Computer (i7-1260P mit 8x 2.1 Ghz und 16GB RAM) insgesamt ca. 40s zur Berechnung aller verwendeten Kennzahlen pro Szenario. Zur Visualisierung der Staueffekte wurden die Kennzahlen für Szenarien mit verschiedenen Zugzahlen zwischen 1.7 und 34 berechnet, wobei lediglich die Ankunftsraten proportional verändert wurden. Die relative Aufteilung der Verkehrsmenge auf die verschiedenen Ströme durch den Knoten wurde dabei nicht verändert.

In den folgenden Abschnitten werden die berechneten Kennzahlen beschrieben und mit den Kennzahlen der isolierten Betrachtungen verglichen.

3.1. Vergleich der erwarteten Warteschlangenlänge

Zunächst betrachten wir die über das Gesamtsystem errechnete Länge der Warteschlange $L_{GI/GI}$ vor den beiden Gesamtfahrstraßenknoten. In Abbildung 2 sind die Ergebnisse für GFK Ost (Abbildung 2a) und West (Abbildung 2b) jeweils im Vergleich mit der Warteschlangenlänge $L_{M/GI}$ des isolierten einkanaligen Ersatzsystems (vgl. Anhang A)



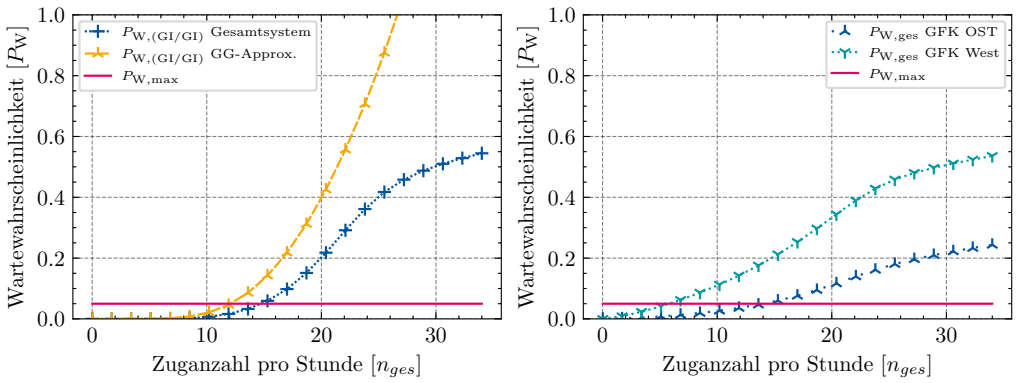
(a) GFK Ost *(b) GFK West*

Abbildung 2: Vergleich der erwarteten Warteschlangenlänge für die GFKs Ost und West in der kombinierten Berechnung mit der Ersatzsystem-Approximation für den jeweils einzelnen GFK.

dargestellt.

Die erwarteten Warteschlangenlängen sind für beide Methoden am GFK West deutlich höher als am GFK Ost. Das ist in erster Linie durch die Unterschiede in der Zugzahl des unterstellten Betriebsprogramms zu erklären, im GFK West müssen ca 40% mehr Anfragen bearbeitet werden. Zusätzlich ist die verkettete Bedienrate von GFK West noch etwas geringer als die Bedienrate in GFK Ost.

Im Vergleich der beiden Methoden lässt sich für beide GFKs außerdem noch beobachten, dass mit der Betrachtung des Gesamtsystems eine geringere Warteschlangenlänge als mit der isolierten Betrachtung des einzelnen GFKs errechnet wurde. Die Erklärung dafür liegt in der Modellierung des Gesamtsystems: Anfragen, welche den betrachteten GFK zur Ausfahrt aus der Gleisgruppe belegen möchten, können auf ihrem Haltegleis auf die Ausfahrt warten. Die Gleisgruppe dient also neben dem Verkehrshalt auch als zusätzlicher Wartebereich, der die Staueffekte vor stark ausgelasteten GFKs abfedern kann. In der isolierten Betrachtung werden die auf Ausfahrt wartenden Forderungen hingegen zusammen mit den einfahrenden Forderungen in der gemeinsamen Warteschlange vor der Bedienstelle gesammelt. Diese Abhängigkeitseffekte zwischen GFK und GG sind somit lediglich durch eine integrierte Betrachtung von GFK und GG analysierbar.



(a) Gleisgruppe *(b) Vor GFK*

Abbildung 3: Vergleich der Wartewahrscheinlichkeit in der Gleisgruppe zwischen dem Gesamtsystem und der isolierten Betrachtung (a), sowie Vergleich der Wartewahrscheinlichkeit bei Ankunft an GFK Ost/West in der kombinierten Berechnung (b).

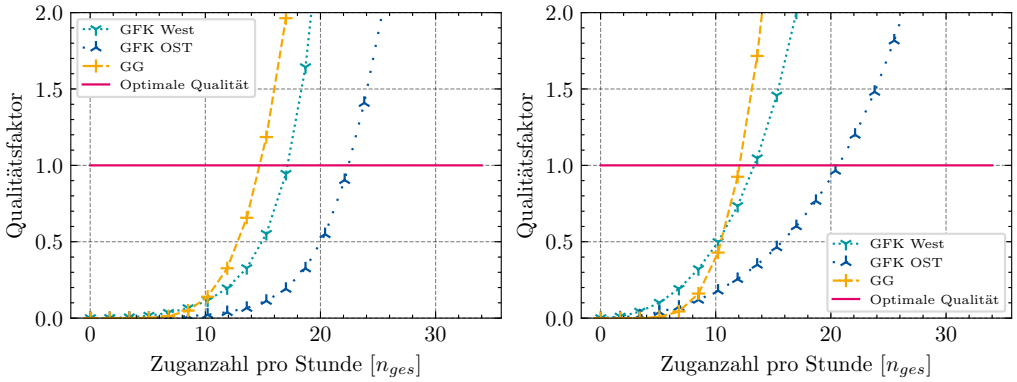
3.2. Vergleich der Wartewahrscheinlichkeit

Die Leistungsfähigkeit der Gleisgruppe wird in [4] über die Kennzahl der Wartewahrscheinlichkeit beschrieben. In Abbildung 3a ist diese für die beiden untersuchten Methoden in Abhängigkeit von der Gesamtzugzahl aufgeführt. Darin ist im Vergleich mit der Gesamtsystembetrachtung für die isolierte Berechnung ein stärkerer Anstieg der Wartewahrscheinlichkeit zu beobachten.

Beide Methoden geben die Wahrscheinlichkeit wieder, dass alle Gleise der Gleisgruppe durch haltende, auf Ausfahrt wartende oder bereits einfahrende Züge belegt sind. In der isolierten Betrachtung besteht die Bedienzeit in der mehrkanaligen Gleisgruppe dabei nicht nur aus der Mindesthaltezeit, wie im Gesamtsystem, sondern zusätzlich aus der Ein- und Ausfahrsperrzeit. Im Gesamtsystem sind die Gleise der GG zwar auch für die Ein- und Ausfahrt der haltenden Züge belegt, allerdings ist lediglich nur eine Ein- oder Ausfahrt pro GFK gleichzeitig möglich.

Dementsprechend entstehen Staueffekte schon vor den GFKs, eine Auswertung der entsprechenden Wartewahrscheinlichkeiten ist in Abbildung 3b gegeben. Ähnlich zur Wartewahrscheinlichkeit (Abschnitt 3.1) ist auch hier die höhere Auslastung des westlichen GFKs deutlich zu beobachten.

3.3. Bestimmung der planmäßigen Leistungsfähigkeit



(a) Kombiniert *(b) Elementscharf*

Abbildung 4: Der erwartete Qualitätsfaktor nach Zugzahl im Gesamtsystem für die kombinierte Berechnung und im Vergleich für die Approximation des einzelnen Elements.

Zur Bestimmung der planmäßigen Leistungsfähigkeit können die errechneten Kennzahlen der Komponenten mit den entsprechenden Grenzwerten kombiniert werden (Abschnitt 2.3.2), um den Qualitätsfaktor auszurechnen. Die planmäßige Leistungsfähigkeit bezeichnet dann die Zugzahl für einen Qualitätsfaktor von 1.0.

In Abbildung 4 ist der Qualitätsfaktor nach Teilelementen in der kombinierten (4a) und isolierten (4b) Betrachtung dargestellt. Beide Methoden weisen die höchsten Qualitätsfaktoren für die Gleisgruppe, gefolgt von GFK West und schließlich GFK Ost auf. Allerdings liegt die Auslastung von GFK West in der isolierten Betrachtung sehr nah an der Auslastung der Gleisgruppe und wird bei kleineren Verkehrsmengen sogar zum rechnerischen Flaschenhals.

Im Vergleich der beiden Methoden ist zu erkennen, dass die Modellierung als Gesamtsystem deutlich höhere Zugzahlen für das gleiche Qualitätsniveau erlaubt. Während für die kombinierten Betrachtung in der maßgebenden GG eine Leistungsfähigkeit von 14 Zugfahrten pro Stunde errechnet wird, kommt die isolierte Berechnung der Knotenelemente lediglich auf 12 Zugfahrten in der Gleisgruppe. Im Gesamtsystem profitieren die Warteschlangen vor den beiden GFKs dabei von dem zusätzlichen Warteraum der GG. Aber auch für die Gleisgruppe werden durch die gemeinsame Modellierung geringere Qualitätsfaktoren berechnet, da ankommende Forderungen bereits zur Einfahrt vor den GFKs synchronisiert werden.

4. Zusammenfassung und Diskussion

In diesem Beitrag wurde eine neue Methode zur Berechnung der Leistungsfähigkeit von Eisenbahnknoten im Gesamtsystem vorgestellt. Dazu werden die Gesamtfahrstraßenknoten mit der Verkettungsmethode als einkanaliges Bediensystem abstrahiert und gemeinsam mit der Gleisgruppe in einem kombinierten Bediensytem modelliert. Über die Formalisierung als zeit-kontinuierliche Markov-Kette ist eine Berechnung von Leistungsfähigkeits-Kennzahlen durch Methoden des probabilistischen Model-Checking möglich. Die Ergebnisse eines Beispielknoten zeigen eine höhere Leistungsfähigkeit des Gesamtsystems bei Verwendung derselben Grenzwerte wie sie in der Praxis für die isolierte Betrachtung der Knotenelement eingesetzt werden.

Dabei können insbesondere die Abhängigkeitseffekte zwischen Ein- und Ausfahrt in GFKs und Verkehrshalten in der GG im Gesamtsystem dargestellt werden. So kann die Gleisgruppe beispielsweise als Warteraum für ausfahrende Forderungen genutzt werden, um den Qualitätsfaktor der GFKs zu verkleinern. Umgekehrt wirken die einkanalig modellierten GFKs bereits als eine Art Filter, um die an der Gleisgruppe ankommenden Forderungen zu synchronisieren.

Welche der beiden Modellierungen dabei näher an der Realität der auftretenden Warteeffekte in der Fahrplanerstellung liegen, sollte in weiterer Forschung gesondert untersucht werde. Ein besonderer Fokus könnte dabei auf der Entwicklung von Grenzwerten zu Kennzahlen des Gesamtknotens liegen. Die hier vorgestellte Modellierung als zeitkontinuierliche Markov-Kette erlaubt dazu die Formulierung weiterer Kennzahlen, die über die klassische Warteschlangenlänge an den GFKs und Wartewahrscheinlichkeit zur Nutzung der GG hinausgehen.

Mit der eingeführten Betrachtung lassen sich Kennzahlen für das Gesamtsystem berechnen, allerdings ist eine detaillierte Identifizierung von Engpässen im GFK-Bereich durch die einkanalige Approximation nicht ohne weiteres möglich. Neben der Weiterentwicklung der Grenzwerte und Kennzahlen könnte deshalb auch die Entwicklung einer Zwischenstufe der Abstraktionsebenen zwischen Fahrstraßen und ganzen GFKs für weitere Forschung interessant sein.

Insgesamt bietet die vorgestellte Methode Infrastrukturbetreibern erstmals die Möglichkeit zur Berechnung von Kennzahlen eines gesamten Eisenbahnknotens, unabhängig von der Anzahl der Fahrstraßen. Damit werden sie in die Lage versetzt, bereits in frühen Pla-

nungsstufen die Leistungsfähigkeit der gesamten Knoten-Infrastruktur einzuschätzen, ohne zeitaufwendige Simulationen für einzelne Fahrplan-Szenarien durchzuführen.

Förderung

Dieser Beitrag wird von der Deutschen Forschungsgemeinschaft (DFG) – 2236/2 gefördert.

Literatur

[1] Bychkov, Igor ; Kazakov, Alexander ; Lempert, Anna ; Zharkov, Maxim: Modeling of Railway Stations Based on Queuing Networks. In: *Applied Sciences* 11 (2021), Nr. 5, 2425. http://dx.doi.org/10.3390/app11052425. – ISSN 2076–3417

[2] DB InfraGO AG: *Trassenfinder – Railway Path Request and Conflict Search Tool*. https://trassenfinder.de/, 2025. – Zugriff am 02.07.2025

[3] DB InfraGO AG ; SMA und Partner AG: *Entwurfsstand mKoK 2026ff Schienengüterverkehr Bereich Nord-West*. https://www.dbinfrago.com/web/schienennetz/kazu-novum-11909200, 2024. – Stand: 06.05.2024. Zugriff am 04.07.2025

[4] DB Netz AG: *Richtlinie Fahrwegkapazität*. 2019

[5] Deutsche Bahn / HAFAS: *HAFAS Arrival Board for Station 8000001 at 07:00 – Archived Version*. https://tinyurl.com/fppnjeuu, 2025. – Zugriff am 02.07.2025

[6] Emunds, Tamme ; Niessen, Nils: Evaluating Railway Junction Infrastructure: A Queueing-Based, Timetable-Independent Analysis. In: *Transportation Research Part C: Emerging Technologies* 165 (2024). http://dx.doi.org/10.1016/j.trc.2024.104704

[7] Emunds, Tamme ; Niessen, Nils: A Queueing-Based Approach for Timetable-Independent Railway Station Performance Analysis. In: *Journal of Rail Transport Planning & Management* 36 (2025). http://dx.doi.org/10.1016/j.jrtpm.2025.100539

[8] Fink, Alexander ; Hantsch, Fabian ; Martin, Ullrich ; Schmidhäuser, Stefan ; Fietze, Mario: Ermittlung der Leistungsfähigkeit für Zugmeldestellen in Eisen-

bahnnetzen. In: *Proceedings of the 4th International Railway Symposium Aachen 2023*. Aachen, 2023, S. 408–427

[9] FISCHER, Klaus ; HERTEL, Günter: *Bedienungsprozesse im Transportwesen: Grundlagen und Anwendungen der Bedienungstheorie*. Transpress-Verlag, 1990

[10] HANSEN, Ingo ; PACHL, Jörn: *Railway Timetabling and Operations: Analysis, Modelling, Optimisation, Simulation, Performance, Evaluation*. Eurail press, 2014. – ISBN 3–7771–0462–0

[11] HERTEL, Günter: Exakte Lösung Zur Berechnung Der Wartegleiszahl Vor Im Einrichtungsbetrieb Befahrenen Streckengleisen Bei Nicht-Poisson-Ankünften (G/M/1-Wartesystem). In: *Wissenschaftliche Zeitschrift der Hochschule für Verkehrswesen* 31 (1984), S. 195–205

[12] JUNGES, Sebastian ; VOLK, Matthias: *Stormpy - Python Bindings for Storm: Version 1.7.0*. `https://moves-rwth.github.io/stormpy/index.html`. Version: 2023

[13] NIESSEN, Nils: *Leistungskenngrößen Für Gesamtfahrstraßenknoten*, Verkehrswissenschaftliches Institut der RWTH Aachen, Diss., 2016

[14] POLLACZEK, Felix: Über eine Aufgabe der Wahrscheinlichkeitstheorie. In: *Mathematische Zeitschrift* 32 (1930), Nr. 1, 64–100. `http://dx.doi.org/10.1007/bf01194620`. – ISSN 0025–5874, 1432–1823

[15] POTTHOFF, Gerhart: *Verkehrsströmungslehre - Die Zugfolge auf Strecken und in Bahnhöfen*. Transpress, 1970

[16] SCHMITZ, Christoph ; WEIK, Norman ; ZIEGER, Stephan ; NIESSEN, Nils ; SCHMEINK, Anke: Markov Models for the Performance Analysis of Railway Networks. In: *International Conference on Railway Operations Modelling and Analysis (RailLille)*. Lille, France, 2017, S. 23

[17] SCHWANHÄUSSER, Wulf: Die Ermittlung der Leistungsfähigkeit von großen Fahrstraßenknoten und von Teilen Des Eisenbahnetzes. In: *Archiv für Eisenbahntechnik* 1978 (1978), Nr. 38, S. 7–18

[18] SCHWANHÄUSSER, Wulf ; SCHULTZE, Kurt: *Ermittlung von Qualitätsmaßstäben für die Berechnung der Leistungsfähigkeit eines Streckenabschnittes und Entwicklung eines Rechenverfahrens zur Ermittlung von Endverspätungen: Forschungsarbeit für die Deutsche Bundesbahn*. 1982

[19] WEIK, Norman ; BOHLIN, Markus ; NIESSEN, Nils: *Long-Term Capacity Planning of Railway Infrastructure: A Stochastic Approach Capturing Infrastructure Unavailability*, Verkehrswissenschaftliches Institut der RWTH Aachen, Diss., 2020

Autoren



Emunds, Tamme
Tamme Emunds studierte bis 2021 Mathematik an der RWTH Aachen und arbeitet seitdem als Wissenschaftlicher Mitarbeiter am Verkehrswissenschaftlichen Institut der RWTH Aachen im Bereich Eisenbahnbetriebswissenschaft. In 2024 übernahm er zusätzlich die Rolle des Gruppenleiters für Modellierung und Analyse.



Nießen, Nils
Univ.-Prof. Dr.-Ing. Nils Nießen leitet seit 2013 das Verkehrswissenschaftliche Institut der RWTH Aachen. Nach dem Studium des Bauingenieurwesens und der Wirtschaftsgeographie an der RWTH Aachen promovierte er dort 2008. Anschließend war er Projektingenieur bei HaCon in Hannover und Geschäftsführer der VIA Consulting & Development GmbH.

A. Verkettungsmethode zur Analyse von Gesamtfahrstraßenknoten

Zur Abstraktion der Bedienprozesse auf den Gesamtfahrstraßenknoten orientieren wir uns in dieser Arbeit an Schwanhäußer [17]. Dazu werden alle Zugfahrten i in einem GFK betrachtet und für je zwei Zugfahrten i, j die Wahrscheinlichkeit $p_{\varphi,ij}$ bestimmt, dass diese sich ausschließen.

Unter Berücksichtigung des Anteils $p_i = n_i / n_{\mathrm{G}}^{\mathrm{ges}}$ einer Zugfahrt i mit Zugzahl n_i an der Anzahl $n_{\mathrm{G}}^{\mathrm{ges}}$ aller Zugfahrten im GFK G, kann dann

$$p_{\varphi,ij} = \begin{cases} p_i \cdot p_j, & \text{Zugfahrt } i \text{ und } j \text{ schließen einander aus} \\ 0, & \text{sonst} \end{cases} \tag{6}$$

bestimmt werden. Dabei setzt sich die Anzahl $n_{\mathrm{G}}^{\mathrm{ges}}$ aller Zugfahren im GFK G aus den Anzahlen der ankommenden Züge $n_{\mathrm{G}}^{\mathrm{An}}$ und der ausfahrenden Züge $n_{\mathrm{G}}^{\mathrm{Aus}}$ zusammen. Die *Verkettungszahl* φ des Gesamtfahrstraßenknoten berechnet sich dann aus der Summe $\varphi = \sum_i \sum_j p_{\varphi,ij}$. Der Kehrwert der Verkettungszahl kann als Anzahl gleichzeitig tätiger Bedienkanäle interpretiert werden [17], und damit genutzt werden, um den Grenzwert

einkanaliger Bediensysteme auch auf die einkanaligen Ersatzsysteme für GFKs anzuwenden, siehe Gleichung (5).

Die mittlere Bedienzeit des einkanaligen Ersatzsystems im GFK G kann über

$$t_{B,\mathrm{G}} \approx \sum_i \sum_j p_{\varphi,ij} \cdot t_{B,i} \tag{7}$$

durch Verrechnen der Ausschlusswahrscheinlichkeit $p_{\varphi,ij}$ mit der Sperrzeit für Zugfahrt i berechnet werden. Für die Bestimmung des Variationskoeffizienten $v_{B,\mathrm{G}}$ ist außerdem das zweite Moment der Bedienzeit $\hat{t}_{B,\mathrm{G}} \approx \sum_i \sum_j p_{\varphi,ij} \cdot t_{B,i}^2$ zu berechnen.

B. Übergänge in der zeitkontinuierlichen Markov-Kette

Im Folgenden werden die Übergänge innerhalb der zeitkontinuierlichen Markov-Kette aus Abschnitt 2.3.1 dargelegt.

Ein Übergang $u = (s, s') \in U$ beschreibt dabei die Rate mit der sich das System von einem Zustand $s \in S$ zu einem andern Zustand $s' \in S$ verändert. Dabei werden zur Beschreibung der Übergänge die folgenden Abbildungen benutzt.

- Mit $k_{\mathrm{Belegt}} : s \mapsto tr_O + tr_W + tr_O^w + tr_W^w$ wird die Anzahl an belegten Gleisen beschrieben.
- $\mathrm{GG}_{\mathrm{frei}} : s \mapsto k_{\mathrm{Belegt}} < (k_{GG} - e_O - e_W - d_O - d_W)$ beschreibt ob in einem Zustand s ein Gleis der Gleisgruppe belegt werden kann.
- $\mathrm{HG}_{\mathrm{frei}} : s \mapsto (hg_O = 0) \wedge (hg_W = 0)$ gibt an, ob die durchgehenden Hauptgleise gerade frei sind.
- $\mathrm{O}_{\mathrm{frei}}^{\mathrm{Ein}} : s \mapsto \mathrm{GG}_{\mathrm{frei}}(s) \wedge \mathrm{HG}_{\mathrm{frei}}(s) \wedge (e_O = 0) \wedge (d_O = 0)$ beschreibt ob eine Einfahrt von GFK O zur Gleisgruppe möglich ist.
- $\mathrm{W}_{\mathrm{frei}}^{\mathrm{Ein}} : s \mapsto \mathrm{GG}_{\mathrm{frei}}(s) \wedge \mathrm{HG}_{\mathrm{frei}}(s) \wedge (e_W = 0) \wedge (d_W = 0)$ beschreibt ob eine Einfahrt von GFK W zur Gleisgruppe möglich ist.
- $\mathrm{O}_{\mathrm{frei}}^{\mathrm{Aus}} : s \mapsto \mathrm{HG}_{\mathrm{frei}}(s) \wedge (e_O = 0) \wedge (d_O = 0)$ beschreibt ob eine Ausfahrt von der Gleisgruppe über GFK O möglich ist.
- $\mathrm{W}_{\mathrm{frei}}^{\mathrm{Aus}} : s \mapsto \mathrm{HG}_{\mathrm{frei}}(s) \wedge (e_W = 0) \wedge (d_W = 0)$ beschreibt ob eine Ausfahrt von der Gleisgruppe über GFK W möglich ist.
- $\mathrm{HG}_{\mathrm{frei}}^{\mathrm{Durch}} : s \mapsto \mathrm{O}_{\mathrm{frei}}^{\mathrm{Aus}} \wedge \mathrm{W}_{\mathrm{frei}}^{\mathrm{Aus}}$ gibt an, ob eine Durchfahrt gerade möglich ist.

Wir formulieren die Übergänge $u = (s, s')$ in der Form $A \rightarrow_\phi A'$, wobei in A eine logische

Formel angegeben wird, welche die Ausgangszustände s definiert, und in A' die Änderung von s zu s' beschrieben werden. Mit ϕ wird die Übergangsrate beschrieben. Wie in den Modellen aus [6, 7] nutzen wir eine künstliche Übergangsrate $D >> 1$ zur Modellierung von quasi direkten Übergängen. Für die Berechnungen in Abschnitt 3 wird eine Übergangsrate von $D = 600$ unterstellt, was eine durchschnittliche Verzögerung von 0.1s bedeutet.

Die Modellierung des Ankunftsprozess an einem GFK $G \in \{O, W\}$ erfolgt über die Übergänge

$$(q_G < m) \wedge (q_G > 0) \rightarrow_{\lambda_G} (q_G + 1) \tag{8}$$

$$(q_G = 0) \rightarrow_{\lambda_G \cdot p_{\mathrm{G}}^{\mathrm{HG}}} (q_G + 1) \wedge (q_G^n = 0) \tag{9}$$

$$(q_G = 0) \rightarrow_{\lambda_G \cdot (1 - p_{\mathrm{G}}^{\mathrm{HG}})} (q_G + 1) \wedge (q_G^n = 1). \tag{10}$$

Dabei ist zu beachten, dass ankommenden Anfragen zu einer leeren Warteschlange über (9) und (10) auf die GG und die HG aufgeteilt werden, indem die Ankunftsrate λ_G mit der Durchfahrtswahrscheinlichkeit $p_{\mathrm{G}}^{\mathrm{HG}}$ multipliziert wird.

Der Start eines Einfahrprozesses wird dann über Übergänge der Form

$$(q_G > 0) \wedge (\mathrm{G}_{\mathrm{frei}}^{\mathrm{Ein}}(s)) \wedge (q_G^n = 0) \rightarrow_{D \cdot p_{\mathrm{G}}^{\mathrm{HG}}} (q_G - 1) \wedge (e_G = 1) \wedge (q_G^n = 0) \tag{11}$$

$$(q_G > 0) \wedge (\mathrm{G}_{\mathrm{frei}}^{\mathrm{Ein}}(s)) \wedge (q_G^n = 0) \rightarrow_{D \cdot (1 - p_{\mathrm{G}}^{\mathrm{HG}})} (q_G - 1) \wedge (e_G = 1) \wedge (q_G^n = 1) \tag{12}$$

und der Start eines Durchfahrprozesses über

$$(q_G > 0) \wedge (\mathrm{HG}_{\mathrm{frei}}^{\mathrm{Durch}}(s)) \wedge (q_G^n = 1) \rightarrow_{D \cdot p_{\mathrm{G}}^{\mathrm{HG}}} (q_G - 1) \wedge (hg_G = 1) \wedge (q_G^n = 0) \tag{13}$$

$$(q_G > 0) \wedge (\mathrm{HG}_{\mathrm{frei}}^{\mathrm{Durch}}(s)) \wedge (q_G^n = 1) \rightarrow_{D \cdot (1 - p_{\mathrm{G}}^{\mathrm{HG}})} (q_G - 1) \wedge (hg_G = 1) \wedge (q_G^n = 1) \tag{14}$$

modelliert.

Eine Durchfahrt ab GFK O ist dann beispielsweise mit den Übergängen

$$(hg_O = 1) \rightarrow_{\mu_{B,\mathrm{O}}} (hg_O = 2) \tag{15}$$

$$(hg_O = 2) \rightarrow_{\mu_{B,\mathrm{W}}} (hg_O = 0) \tag{16}$$

definiert, analog für GFK W.

Bei der Bedienung des Einfahrprozesses in die Gleisgruppe muss der Eingangsstrom erneut aufgeteilt werden, diesmal zwischen den beiden Ausfahrrichtungen. In den entspre-

chenden Übergängen für GFK O,

$$(e_O = 1) \wedge (k_{\text{Belegt}} < k_{GG}) \rightarrow_{\mu_{B,\text{O}} \cdot p_O^{\text{Wende}}} (e_O = 0) \wedge (tr_O + 1) \tag{17}$$

$$(e_O = 1) \wedge (k_{\text{Belegt}} < k_{GG}) \rightarrow_{\mu_{B,\text{O}} \cdot (1 - p_O^{\text{Wende}})} (e_O = 0) \wedge (tr_W + 1), \tag{18}$$

wird die Anzahl haltender Züge für wendende Züge in Richtung GFK O ($(tr_O + 1)$) und sonst in Richtung GFK W ($(tr_W + 1)$) erhöht.

Für den Halteprozess in der Gleisgruppe muss in den Übergängen

$$(tr_G > 0) \wedge (\text{G}_{\text{frei}}^{\text{Aus}}) \rightarrow_{\mu_{B,\text{G}} \cdot tr_G} (d_G = 1) \wedge (tr_G - 1) \tag{19}$$

$$(tr_G > 0) \wedge (\neg \text{G}_{\text{frei}}^{\text{Aus}}) \rightarrow_{\mu_{B,\text{G}} \cdot tr_G} (tr_G^w + 1) \wedge (tr_G - 1) \tag{20}$$

$$(tr_G^w > 0) \wedge (\text{G}_{\text{frei}}^{\text{Aus}}) \rightarrow_{D} (d_G = 1) \wedge (tr_G^w - 1) \tag{21}$$

der Belegungsstatus des Ausfahr-GFKs G beachtet werden. Falls er durch eine andere Forderung belegt ist, wird die Anzahl auf Ausfahrt wartender Züge erhöht (20) und die sonst direkte Ausfahrt (19) muss nach Freimeldung des GFKs erfolgen (21). Die mehrkanalige Modellierung der Gleisgruppe erlaubt eine gleichzeitige Bedienung von allen haltenden Zügen in (19) und (20).

Die Bedienung des Ausfahrtprozesses in einem GFK G erfolgt schließlich in

$$(d_G = 1) \rightarrow_{\mu_{B,\text{G}}} (d_G = 0) \tag{22}$$

erneut mit der verketteten Bedienrate $\mu_{B,\text{G}}$ des Ausfahr-GFKs.

C. Zusätzliche Informationen zum Beispielbahnhof

In Abbildung 5 ist der Gleisplan von *Aachen Hbf* abgebildet. Mit Hilfe der Fahrstraßen und Haltegleise aus den Fahrplandaten in Tabelle 2 konnten die Konflikte der einzelnen Zugfahrten bestimmt werden. Die Ankunfts- und Abfahrtszeiten für den Personenverkehr Aachen Hbf wurden aus Online-Plänen [5] für den 02.07.2025 in der betrachteten Stunde zwischen 07:00 Uhr - 08:00 Uhr gewonnen und die Angaben zum Güterverkehr wurden aus [3] übernommen.

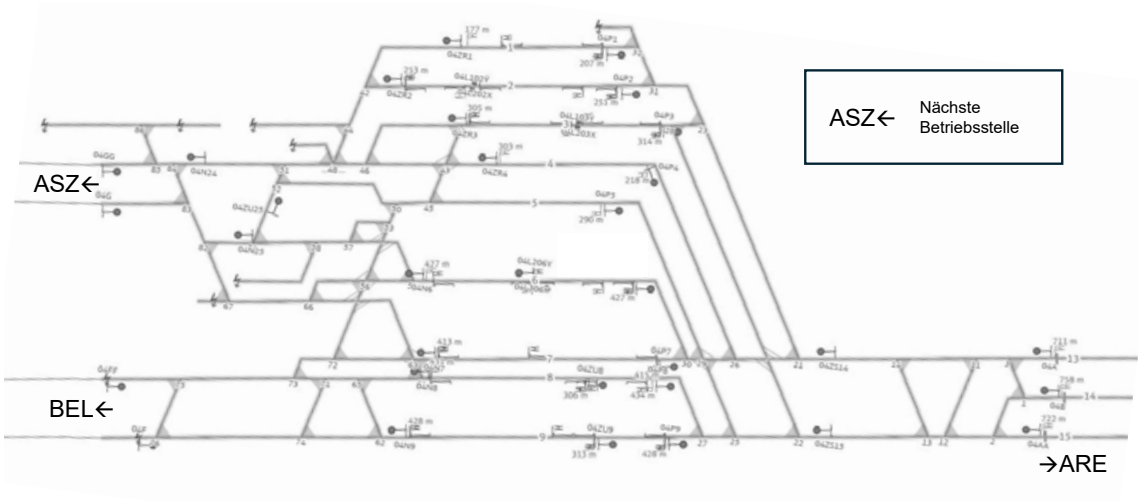


Abbildung 5: Infrastrukturplan von Aachen Hbf; modifiziert nach [2].

Tabelle 2: Fahrplandaten in Aachen Hbf zwischen 07:00 und 08:00 Uhr nach [5, 3]

Zug	Ankunft	Abfahrt	Halt [min]	Sperrzeit Einfahrt [min]	Sperrzeit Ausfahrt [min]	Fahrweg Ankunft	Fahrweg Abfahrt	Gleis	GFK Einfahrt	GFK Ausfahrt	Zugart
RB 30007	07:00	07:01	1	3	2	ASZ-8	8-ARE	8	GFK W	GFK O	NV
RB 30056	07:01	07:02	1	2	3	ARE-1	1-ASZ	1	GFK O	GFK W	NV
RE 26857	07:06	07:07	1	3	2	ASZ-8	8-ARE	8	GFK W	GFK O	NV
RE 26802	07:07	07:51	44	2	2	ARE-2	2-ARE	2	GFK O	GFK O	NV
RE 18911	07:12	07:17	5	3	3	ASZ-6	6-ASZ	6	GFK W	GFK W	NV
EST 9414	07:17	07:22	5	2	2	ARE-9	9-BEL	9	GFK O	GFK W	FV
S 41	07:22	07:37	15	2.5	3	BEL-8	8-BEL	8	GFK W	GFK W	NV
RB 10313	07:27	07:37	10	3	3	ASZ-6	6-ASZ	6	GFK W	GFK W	NV
RB 30509	07:30	07:31	1	3	2	ASZ-7	7-ARE	7	GFK W	GFK O	NV
RB 30554	07:31	07:32	1	2	3	ARE-1	1-ASZ	1	GFK O	GFK W	NV
ICE 11	07:36	07:39	3	2.5	2	BEL-9	9-ARE	9	GFK W	GFK O	FV
RE 26404	07:39	08:20	41	3	3	ASZ-7	7-ASZ	7	GFK W	GFK W	NV
RE 10902	07:41	08:18	37	2	2	ARE-3	3-ARE	3	GFK O	GFK O	NV
RE 18913	07:42	07:47	5	3	3	ASZ-6	6-ASZ	6	GFK W	GFK W	NV
NJ 40468	07:47	08:09	22	2	2.5	ARE-9	9-BEL	9	GFK O	GFK W	FV
GV 128-1	07:17	07:17	0	4	4	ASZ-5	5-ARE	5	GFK W	GFK O	GV
GV 128-2	07:14	07:14	0	4	4	ARE-4	4-ASZ	4	GFK O	GFK W	GV

Session 1
Sirikorn Satitrattanashewin, Jan Eisold, Beate Lux

Capacity and Quality Assessment of Track Groups in Railway Nodes

Satitrattanashewin, Sirikorn[1]; Eisold, Jan[1]; Lux, Beate[1]

[1]Institute of Railway Systems and Public Transport, Technische Universität Dresden

Summary

The nodes of the railway network play an essential role in the quality and efficiency of rail passenger and freight transport. Various methods and parameters can be used to assess the quality and capacity, especially in the track groups of these stations, but these are only comparable to a limited extent and, from today's perspective, only provide limited information. At the same time, mathematical optimisation approaches offer the opportunity to further develop the quality of the available methods and quality benchmarks.

Keywords: Railway Operation, Nodes, Capacity, Quality, Track Groups

1 Introduction

Rail transport plays a key role in passenger mobility and freight transportation – particularly in terms of its inherently low energy consumption and decentralized emissions. An efficient railway system relies on a balanced interplay between performance, quality, and economic efficiency. Assessing capacity and evaluating operational quality and cost-effectiveness are core tasks of railway operations research. Railway network nodes, as critical components of the infrastructure, significantly influence overall capacity. Particular attention is given to the track groups within these nodes, as they serve as stops, transfer points and stabling and handling facilities in passenger transport and play an important role in freight transport as essential components of the service networks in the form of freight stations and marshalling yards. Analysing the performance of such track groups is therefore crucial for identifying bottlenecks at an

early stage, optimising operational processes, and sustainably improving capacity utilization. Depending on the planning level, different methods for quality assessment and the dimensioning of track groups are available according to the specific research question and application context, including mathematical-analytical methods, timetable-based approaches, simulation, and optimisation approaches. Different parameters and quality benchmarks have been developed [1], [2], but these are only comparable to a limited extent [3] and, from today's perspective, only have limited informative value and detail.

Due to the limitations of existing methods, there is a need to develop and apply new approaches particularly using mathematical optimisation and data-driven analytics [4]. The choice of suitable quality indicators and performance benchmarks depends on the method used. For newly developed approaches, further research is required to define and validate appropriate quality indicators [5].

The aim of this contribution is to identify appropriate performance indicators and associated benchmarks for analysing the performance of track groups. The scientific contribution of this work lies in the analysis of existing and newly developed approaches through a qualitative and quantitative comparison of the different performance indicators. Based on this, indicators suitable for measuring operational quality are to be derived and further developed within the framework of mathematical optimisation. Subsequently, corresponding quality benchmarks need to be defined.

This paper is organised as follows: To give a short introduction into the relevant methods for investigating track groups, section 2 presents the UIC 406 methodology and section 3 provides an insight into the German DB Richtlinie 405. Section 4 summarises the established parameters and benchmarks. Section 5 reviews a new method, section 6 provides a qualitative and quantitative comparison of the methods and parameters. Finally, Section 7 concludes the main findings and an outlook on future research directions.

2 UIC 406 — Universal Method to Evaluate Capacity

The core of this approach is to calculate capacity consumption through timetable compression and to evaluate the number of possible train paths for a given line, node, or corridor. For track groups, the occupation of individual tracks is represented using a track occupation graph. Compression is performed separately for each individual track, in the same manner as on a line. It should be noted that track occupation begins with the occupation of the preceding route node and ends with the clearance of the adjacent route node. Thus, the capacity consumption of each track is largely independent of the others, making parallel use of the tracks potentially viable [2].

3 DB Richtlinie 405 - Railway Capacity in German Practices

Railway operations research studies in Germany, as guided by DB Richtlinie 405 (Track Capacity Guideline) [1], aim to analyse and evaluate both existing and anticipated states of the railway system as a basis for the optimal use of track capacity and the dimensioning of infrastructure. Three main methods are available: the mathematical-analytical method, the experimental method (simulation), and timetable studies (constructive method).

Mathematical-Analytical Method

The mathematical-analytical method is based on queuing theory. Given that the characteristics of track groups correspond to systems with multiple parallel service channels, track groups can be modeled as multi-channel queues. A principal advantage of this method is that it requires relatively less time and only rough estimates of input data compared to other methods, making it particularly suitable for long-term planning. Consequently, the mathematical-analytical method is commonly applied for the dimensioning of infrastructure. However, the applicability of this method is limited by the requirement that all tracks within a track group must be mutually substitutable; that is, every train must be able to use any track in the group without restriction. In practice, it is rare for all tracks to have identical characteristics, such as sufficient track length, adequately long platforms, appropriate electrification and voltage, and accessibility from and to lines via route nodes. If these criteria are not met, the track group must be divided into partial track groups consisting of tracks that satisfy only some of the required properties [1].

Simulation Method

Simulation methods are particularly suitable for the integrated analysis of route nodes and track groups, or of multiple nodes. Unlike the mathematical-analytical method, simulation allows the modelling of specific characteristics of individual tracks. Simulation methods are capable of generating results using parameters such as secondary delays, the probability that a train will have to wait before entering the track group, the optimal number of trains that can be accommodated, and the number of tracks required within the track group during simulated operational execution. Since this method analyses the infrastructure in an integrated manner, for example as nodes, these parameters can only be derived indirectly for track groups. In particular, the parameter of secondary delays is especially suitable for the evaluation of nodes. However, due to the multiple iterations required, simulation methods are currently not recommended for determining performance indicators of track groups, as they involve significantly greater effort compared to other methods [1].

Timetable Studies (Constructive Method)

This method is based exclusively on the analysis of time consumption, with the calculation of capacity consumption or the degree of occupancy based on the timetable being common practice. The constructive method is capable of generating a track occupation graph and determining parameters such as the number of tracks required within the track group. This method can also be applied for the dimensioning of infrastructure, provided that the service requirements are already clearly defined. A track occupation graph enables the optimisation of track group occupancy from both service and operational perspectives. However, the results obtained are always valid only for the predefined, specific boundary conditions. Due to the effort involved, the investigation of operational execution should be limited to a few critical cases [1]. The compression method according to UIC 406 can be classified as a type of timetable studies method in German practice.

4 Established Parameters and Benchmarks

In this section, key parameters and their corresponding benchmarks used for capacity and quality assessment—particularly focused on track groups in nodes—are described.

Degree of Occupancy (Occupancy Rate)

UIC 406 proposes a concatenated occupancy rate for track groups between 40% and 50% per track, with an additional time rate from 150% to 100%, and considers an acceptable quality of service to be achieved when overall capacity consumption reaches 100%. Since the usage of each track can be independent, the degree of occupancy can be calculated for individual tracks and aggregated for a group of tracks or the entire track area [2, pp. 28, 30]. In contrast, DB Richtlinie 405 currently provides reference values according to four quality levels (premium, optimal, risky, and insufficient). For track groups, however, there is not yet sufficient empirical data available. It should also be noted that these parameters do not provide statements regarding operational quality, as the procedure does not consider train priorities or knock-on and initial delays [3], [6, pp. 277-278].

Waiting Probability

The parameter “Waiting Probability” (also referred to as “Waiting Risk”), the probability that a train must wait in front of a track group, is described in DB Richtlinie 405 (see table 1). It can be calculated using mathematical-analytical methods. Using simulation methods, the waiting probability may also be determined directly by the tools. If this is not the case, DB Richtlinie 405 defines an alternative procedure: Additional fictitious tracks with the same properties as the real ones are virtually added to the track group, and

alternative, lowest-priority routes are assigned to these tracks for all trains. The simulation is then performed, and if any trains still use the lowest-priority fictitious track, further fictitious tracks are added until no trains use the last added one. The ratio of trains using the fictitious tracks to the total number of trains provides an estimate of the waiting probability, representing the proportion of trains that would have been queued in preceding track groups if the fictitious tracks had not been assumed (Module 0403 Pt. 5(8) and 7(6) in [1]).

Table 1: Quality Benchmark for Waiting Probability at DB InfraGO AG
Adapted from DB 405.0104 Pt. 20 [1], translated by the authors

zul. pw [%] for quality level	**premium**	**optimal**	**risky**	**insufficient**
		performance range (LB)		
Infrastructure Utilisation	Underutilisation	Appropriate utilisation	Very high utilisation	Excessive utilisation
Passenger traffic	≤ 1.0	2.5	≥ 3.5	≥ 5.0
Freight traffic	≤ 2.5	5.0	≥ 7.0	≥ 10.0
Mixed-Traffic	$\leq p_{Rz} \cdot 1.0 + p_{Gz} \cdot 2.5$	$p_{Rz} \cdot 2.5 + p_{Gz} \cdot 5.0$	$\geq p_{Rz} \cdot 3.5 + p_{Gz} \cdot 7.0$	$\geq p_{Rz} \cdot 5.0 + p_{Gz} \cdot 10.0$

p_{Rz} = Proportion of passenger traffic, p_{Gz} = Proportion of freight traffic
zul. pw [%] = Permissible waiting probability in percent

The necessary number of tracks is estimated in an iterative process, in which the number of tracks in the (partial) track groups is repeatedly adjusted until the waiting probability falls below the defined quality benchmark (see Table 1). The required number of tracks for each partial track group is determined by dimensioning the respective subgroups.

5 Mathematical Optimisation–Based Method for Track Group Assessment: RANGO

NEOMOBIL has developed RANGO, an optimisation-based decision support tool for the dimensioning of rail freight facilities, which employs mathematical optimisation techniques, formulated as a Mixed Integer Linear Programme (MILP), and is used in a framework called ‘Dimensioning Analysis’ (DimA) [7]. It supports the determination of the number and length of tracks required, the number of additional fictitious tracks required, track utilisation (expressed as the percentage of track occupation), and the percentage of unrealised track utilisations, depending on the given scenarios [7], [5]. With respect to

quality assessment, the results obtained from the optimisation calculations can be evaluated from different perspectives. Referring to DB Richtlinie 405, RANGO can estimate the waiting probability based on the percentage of unrealised track utilisations as the ratio of trains using fictitious tracks to the total number of trains [1]. However, it should be emphasised that the percentage of unrealised track utilisations determined this way is not conceptually identical to the waiting probability for a track group as calculated using analytical methods based on queuing theory. This must be considered when interpreting the results and transferring established quality benchmarks to this new procedure [5]. Furthermore, the results for track utilisation may also be assessed using established boundaries for occupancy rates specified in UIC 406 [2].

6 Qualitative and Quantitative Comparison of the different methods and parameters

As the different methods and tools use diverse approaches to investigates track groups, they use different input data, build up on diverse infrastructure and process modelling, provide different output and therefore different quality benchmarks must be used to interpret the results. To compare the methods, we selected three tools, analysed first the qualitative differences, and second the quantitative differences between them.

The GLEISE tool in LUKS®, a software by the company quattron [8], is a standard software used at DB InfraGO AG. Using a queueing-theoretic, mathematical–analytical approach, it applies the quasi-exact solution by Hertel [9] for the computational determination of waiting probabilities and the expected value of the queue in front of a single or multi-channel service point (Module 0105 Pt.5(13) in [1]).

The simulation programme SimBed (as detailed in the Appendix in [10]) is used in teaching and research at TU Dresden. It enables the modelling of a railway system as a service system and the simulation of its operational processes. Compared to analytical methods such as the GLEISE tool, SimBed allows the representation of more complex service system structures and the consideration of various characteristics of arrival flows and service processes. However, due to the abstraction inherent in service system models, the simulation still represents a simplified version of reality [10, p. 3].

In RANGO, tracks can be assigned individual characteristics (e.g., length or accessibility). The software estimates the percentage of unrealised track utilisations with respect to the number of available and additional fictitious tracks and—in line with DB Richtlinie 405—interprets the ratio of trains routed to fictitious tracks to the total number

of trains as an estimate of the waiting probability. It also reports the track occupancy rate for each track. It uses a mathematical optimisation model that minimises the number of tracks used and the occupation on additional tracks [11] using multiple optimisation to deal with various timetables [5].

The qualitative comparison of the methods using the described examples of the respective tools is shown in Table 2.

Table 2: Qualitative Comparision of the GLEISE Tool, SimBed Tool,and RANGO Tool Adapted from Table 5.1 in [12]; translated and modified by the authors

GLEISE tool in LUKS®	SimBed	RANGO
Underlying Method		
Mathematical-analytical method	Queuing theory-based simulation	Mathematical optimisation
Input Parameters		
- Expected Time of Arrival (ET_A) - Coefficient of variation of arrival intervals (VT_A) - Expected Time of Service (ET_B) - Coefficient of variation of service time (VT_B) - Investigation period in hours - Number of service channels (number of tracks in the group)	- Expected Time of Arrival (ET_A) - Coefficient of variation of arrival intervals (VT_A) - Expected Time of Service (ET_B) - Coefficient of variation of service time (VT_B) - Number of tracks in track groups	- Forecast traffic volume, such as incoming and outgoing trains, for various timetables generation - Train characteristics - Operational technologies and infrastructure constraints
Output Parameters		
- Degree of occupancy (occupancy rate): ratio of Expected Time of Service (ET_B) to Expected Time of Arrival (ET_A). - Degree of occupancy per track (average; equal across all tracks) - Waiting probability - Average queue length - Average waiting time - Total waiting time - Average waiting time of waiting trains (conditional mean) - Upper 95% confidence limit for the waiting time (superior value) - Upper 95% confidence limit for the queue length (superior value)	- Degree of occupancy (occupancy rate) for a track group - Waiting probability - Average waiting time across all trains (including standard deviation) - Average waiting time for waiting trains (including standard deviation) - Percentage of time with zero/one/two/more than two trains waiting - Average number of served trains - Average queue length	- Degree of occupancy for the entire track group (occupancy rate) - Number of tracks - Number of additional fictitious tracks - Percentage of occupation times per track - Cumulative percentage of all track-occupation times across the tracks (degree of fulfilment) - Percentage of unrealised track utilisations (degree of non-fulfilment) - Ratio of unrealised track utilisations to realised track utilisations
Modelling of infrastructure		
It is suitable for simple cases with a single arrival flow and requires that all tracks be mutually substitutable; only one network element can be represented at a time.	It extends this by allowing multiple arrival flows and partial track groups, but still models only one network element at a time.	It offers the highest flexibility and level of detail, supporting complex infrastructure features and interactions among multiple network elements, albeit at the cost of increased modelling and computational effort.
Modelling Accuracy		
Low accuracy: relies on limited input parameters, exact only in specific cases ($VT_A = VT_B = 1$), does not model adjacent areas	Moderate accuracy: accepts more detailed inputs (arrival distributions; multiple partial track groups/flows), but still does not model adjacent areas.	Very high accuracy: models detailed real-world processes (e.g., daily load profiles) and includes both microscopic and macroscopic constraints.

The most obvious finding from the comparison is that RANGO offers the highest modelling accuracy, while GLEISE and SimBed are limited by their input parameters and by the constraint that all tracks must be unique. Nevertheless, GLEISE is particularly well-suited for estimating homogeneous track groups due to its relatively low computational efforts and the use of established quality indicators. For complex facilities and questions, however, additional detailed analyses are necessary [7]. It is important to highlight that the results from the analytical method are calculated using equations and are reproducible, whereas SimBed employs synchronous simulation and RANGO multiple optimisation.

For investigating the quantitative differences of the three methods we used again the three respective tools we carried out three experimental series of varying complexity. In this process, the input parameters were systematically varied in order to capture the sensitivity of the key figures to changes in the input variables and to attribute differences in the key figures to specific causes. Some of these calculations were conducted within the framework of a student research project by Pionteck [12]. The following section presents the key experimental results; full details of the experiment can be found therein.

To enable a quantitative comparison, the experiments were conducted in sequence, starting with a simple track group and progressing to two more complex infrastructure track groups, each considering different distributions of arrival traffic flows and service times.

Experiment Series 01: Single Track Group

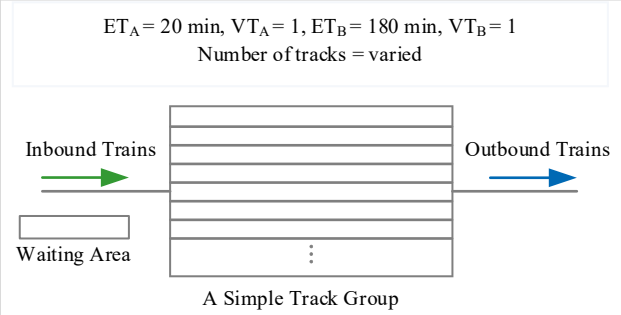


Figure 1: Schematic of the infrastructure for Experimental Series 01
Adapted from Figure 6.1.1 in [12]; translated and modified by the authors.

Experiment Series 01 was designed to investigate how the results differ under the same configuration across three distinct methods. Due to the constraints of the GLEISE tool, a scenario with one traffic flow and a simple track group—where all tracks are mutually substitutable—was simulated and analysed using all three methods (see Fig. 1). During the experiment, the number of tracks was varied and evaluated against established benchmarks and optimal threshold values. Inbound trains enter the track group with an

expected time of arrival (ET_A) of 20 minutes and an expected time of service (ET_B), or mean occupancy time, of 180 minutes. Both parameters are exponentially distributed ($VT_A = 1$; $VT_B = 1$). In both GLEISE tool and SimBed, the waiting area in front of the track group is assumed to be infinite. In contrast, RANGO assumes nine available tracks and assigns any trains that cannot be accommodated within these tracks to additional fictitious tracks, which possess the same properties as the first nine tracks.

The results of the experiments are summarised in Table 3. In this table, the results are compared across three distinct methods for two parameters: waiting probability and degree of occupancy. The waiting probability is classified as a quality parameter, whereas the degree of occupancy is an infrastructure-related parameter reflecting track utilisation. Comparing these two parameters provides valuable insights, as it allows the results to be interpreted from both a quality and an infrastructure utilisation perspective.

Table 3: Results for Experiment Series 01
authors' own illustration

Number of tracks	Waiting Probability			Degree of Occupancy		
	GLEIS	SimBed	RANGO	GLEIS	SimBed	RANGO
11	43.0 %	41.4 %	7.6 %	81.8 %	81.3 %	74.6 %
12	26.6 %	25.5 %	4.8 %	75.0 %	74.6 %	**70.5 %**
13	15.8 %	14.3 %	2.9 %	69.2 %	68.6 %	66.4 %
14	8.9 %	7.7 %	1.7 %	64.3 %	63.6 %	62.4 %
15	4.8 %	4.4 %	0.9 %	**60.2 %**	**59.5 %**	58.7 %
16	2.5 %	2.0 %	0.5 %	56.3 %	55.8 %	55.2 %
17	1.2 %	1.1 %	0.3 %	52.9 %	52.5 %	52.1 %
18	0.5 %	0.4 %	0.2 %	50.0 %	49.5 %	49.3 %
19	0.2 %	0.2 %	0.1 %	47.4 %	46.9 %	46.7 %
20	0.1 %	0.1 %	0.0 %	45.0 %	44.6 %	44.4 %

Waiting probabilities, as computed by each method, are highlighted in green for the value closest to—without exceeding—the 5% benchmark for freight traffic. The number of tracks required indicates the optimal number. For the GLEISE tool and SimBed, the results are similar: the optimal number of tracks is 15, with waiting probabilities of 4.8% and 4.4%, respectively. In contrast, RANGO reports a waiting probability of only 0.9% (highlighted in yellow) and determines the optimal number of tracks as 12, with a waiting probability of 4.8% at this level.

For the degree of occupancy or concatenated occupancy rate, values within the range proposed by UIC 406—maximum 50%—are highlighted, as this range is recommended as a minimum threshold for track utilisation based on the compression method. All three methods yield similar results in this regard, indicating that at least 18 to 20 tracks would be required to maintain the degree of occupancy within this minimum range.

When comparing the perspectives of quality and infrastructure utilisation, the required number of tracks can differ significantly, even when assessed using the same method. For example, according to RANGO, if the optimal number of tracks is determined based on the waiting probability, 12 tracks are sufficient, resulting in a waiting probability of 4.8%. However, at this level, the degree of occupancy reaches 70%, which is considerably higher than the minimum threshold. According to UIC 406, the recommended concatenated occupancy rate is between 40% and 50%, corresponding 18 to 20 tracks.

Additional experiments based on the configuration of Experiment Series 01 are presented in [12], in which the input parameters were systematically varied. In total, 16 experiments were conducted, enabling a more detailed and in-depth analysis.

Experiment Series 02: Three Track Groups with Two Arrival Traffic Flows

Experiment Series 02 increases the complexity of the infrastructure under investigation by introducing three parallel track groups and dividing the inbound traffic into two distinct arrival flows (see Fig. 2).

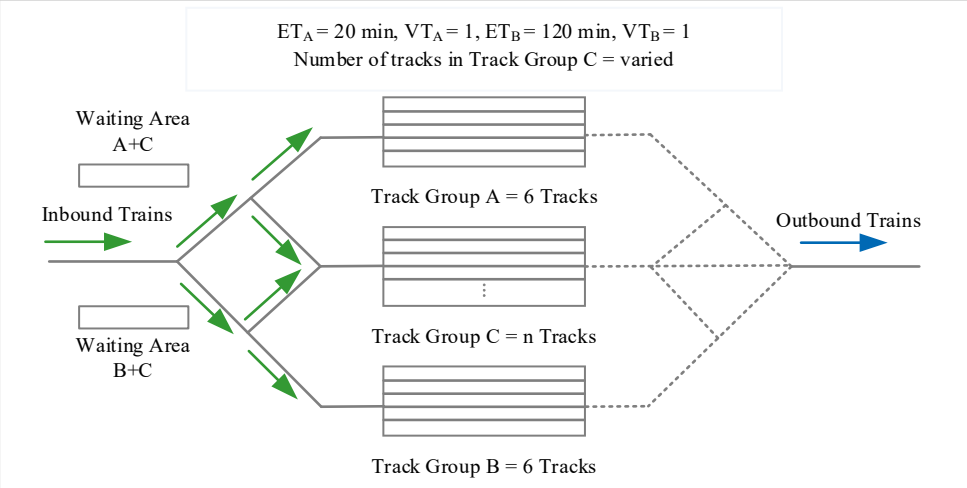


Figure 2: Schematic of the infrastructure for Experimental Series 02
Adapted from Figure 6.1.2 in [12]; translated and modified by the authors

This scenario is designed to reflect more realistic operational conditions, such as those found in larger railway nodes or marshalling yards, where multiple track groups serve different traffic flows or operational purposes. All tracks within each track group are

mutually substitutable; however, the tracks between different track groups are not mutually substitutable, as they are connected to different lines. Due to the constraints of the GLEISE tool, which supports only simple track groups, this scenario was simulated and analysed using SimBed and RANGO. Track Groups A and B each comprise six tracks, while the number of tracks in Track Group C was varied and evaluated against established benchmarks and optimal threshold values.

This schematic of the infrastructure for more complex railway nodes illustrates the level of abstraction inherent in service system models. Flows of inbound trains can be modelled in the simulation; however, the movement of outbound trains beyond the route nodes cannot be represented and is indicated by dashed lines, signifying that these aspects are either unknown or not considered at this stage. Each inbound train occupies a track for a certain period and then disappears from the model. In reality, exit routes must be considered with potential route exclusions between them. Additionally, trains from several directions often enter the same track group as multiple arrival traffic flows. There are no waiting areas in front of the track groups; instead, trains are located at other operating points on the approaching lines. Thus, the model represents a simplified version of reality.

In this experiment, inbound trains for each partial arrival flow enter the track groups with an expected time of arrival (ET_A) of 20 minutes and an expected time of service (ET_B) of 120 minutes. The first partial traffic flow enters either Track Group A or C, while the second enters either Track Group B or C.

In the SimBed tool, there are two waiting areas: one in front of Track Groups A+C and another in front of Track Groups B+C; both are assumed to be infinite. In the RANGO tool, no waiting areas are available. Instead, any trains that cannot be accommodated in Track Group A or Track Group B are assigned to Track Group C. The results of the experiments are summarised in Table 4.

For SimBed, the optimal total number of tracks is 19, with Track Group C comprising seven tracks. The waiting probabilities for waiting areas A+C and B+C are 4.8% and 4.5%, respectively, resulting in an average of 4.6% across both areas. In contrast, RANGO reports a waiting probability of only 1% (highlighted in yellow) and yields an optimal total of 16 tracks, with only four tracks in Track Group C and a waiting probability of 4.4%. In RANGO, the waiting probability is directly related to the number of additional fictitious tracks in Track Group C. It can be observed that the optimal number of tracks determined by RANGO is lower than that obtained by SimBed. The results are consistent with those of Series 01. It is important to note that the waiting probabilities are computed differently in each method.

Table 4: Results for Experiment Series 02
authors' own illustration

Total number of tracks	SimBed					RANGO	
	Number of tracks per track group		Degree of Occupancy	Average Waiting Probability	Waiting Probability	Degree of Occupancy	Waiting Probability
15	Group A	6	80.8 %	38.1 %	38.1 %	74.2 %	7.0 %
	Group C	3	75.1 %				
	Group B	6	81.1 %		38.0 %		
16	Group A	6	78.1 %	24.8 %	25.2 %	**71.5 %**	4.4 %
	Group C	4	65.4 %				
	Group B	6	77.7 %		24.4 %		
17	Group A	6	75.6 %	13.8 %	13.9 %	68.5 %	2.7 %
	Group C	5	55.6 %				
	Group B	6	75.6 %		13.7 %		
18	Group A	6	74.6 %	8.4 %	8.7 %	65.4 %	1.7 %
	Group C	6	48.8 %				
	Group B	6	74.4 %		8.2 %		
19	Group A	6	**74.3 %**	4.6 %	4.8 %	62.4 %	1.0 %
	Group C	7	**42.8 %**				
	Group B	6	**73.6 %**		4.5 %		
20	Group A	6	73.7 %	2.5 %	2.3 %	59.5 %	0.6 %
	Group C	8	38.0 %				
	Group B	6	73.7 %		2.6 %		

Additional experiments based on the configuration of Experiment Series 02 are presented in [12], in which four input parameters were systematically varied, as in Series 01. In total, another 16 experiments were conducted.

Experiment Series 03: Two Track Groups with Two Different Arrival Traffic Flows

Experiment Series 03 builds upon Series 02 by introducing different arrival traffic flows between the track groups. It is designed to investigate how the results differ when the track groups are assigned varying numbers of inbound trains and tracks, i.e., with an unsystematic distribution. Track Group A comprises at least five tracks and can be varied to include more than five tracks, while Track Group B comprises a fixed five tracks. A schematic of the modelled infrastructure is shown in Figure 3.

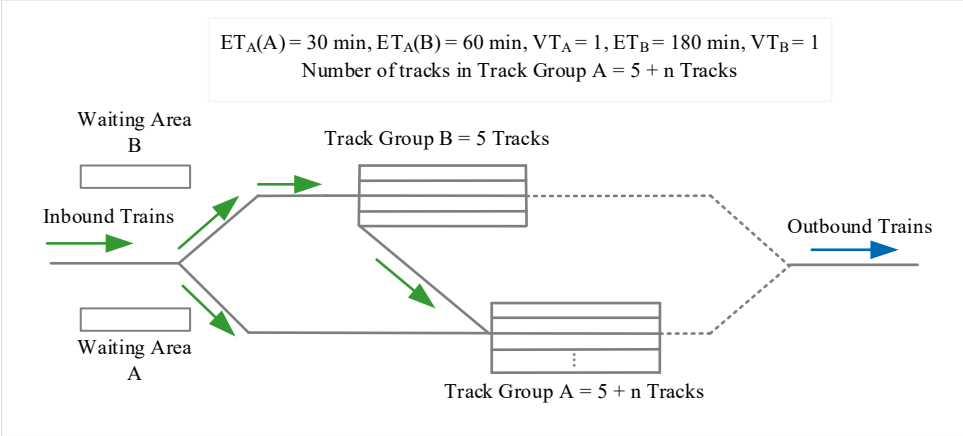


Figure 3: Schematic of the infrastructure for Experimental Series 03
Adapted from Figure 6.1.3 in [12]; translated and modified by the authors

In this experiment, inbound trains for each partial arrival flow enter the track groups with different expected times of arrival (ET_A). One arrival traffic flow can be served only by Track Group A, while the other can be served by either Track Group A or Track Group B. In other words, inbound trains from the traffic flow with an ET_A of 30 minutes can occupy only Track Group A. The other group of inbound trains, with an ET_A of 60 minutes, can occupy either Track Group A or Track Group B. Because Track Group A can be accessed by both arrival traffic flows, additional tracks are planned to be added to this group to ensure efficient utilisation. The expected time of service (ET_B) is 180 minutes for both track groups.

In the SimBed tool, there are two waiting areas: one in front of Track Group A and another in front of Track Groups B; both are assumed to be infinite. In the RANGO tool, as in previous experiments, no waiting areas are available. Instead, any trains that cannot be accommodated in Track Group A or Track Group B are assigned to additional fictitious tracks in Track Group A. The results of the experiments are summarised in Table 5. For SimBed, the optimal total number of tracks is 17, with Track Group A comprising twelve tracks. The waiting probabilities for waiting areas A and B are 3.2% and 0.6%,

respectively, resulting in an average of 2.3% across both areas. In contrast, RANGO yields an optimal total of fourteen tracks, with only nine tracks in Track Group A and a waiting probability of 3.4%. The waiting probability for Track Group A is consistently higher than that for Track Group B, as more inbound trains enter Track Group A. Across all series, the experimental results consistently indicate that RANGO identifies a lower optimal number of tracks than SimBed.

Table 5: Results for Experiment Series 03
authors' own illustration

Total number of tracks	SimBed					RANGO	
	Number of tracks per track group		Degree of Occupancy	Average Waiting Probability	Waiting Probability	Degree of Occupancy	Waiting Probability
11	Group A	6	108.9 %	73.9 %	99.6 %	68.6 %	15.6 %
	Group B	5	59.7 %		22.8 %		
12	Group A	7	103.3 %	50.1 %	68.0 %	66.8 %	9.8 %
	Group B	5	56.6 %		14.2 %		
13	Group A	8	95.6 %	30.0 %	40.9 %	64.3 %	5.9 %
	Group B	5	54.8 %		8.0 %		
14	Group A	9	89.1 %	16.8 %	22.9 %	**61.4 %**	3.4 %
	Group B	5	54.0 %		4.5 %		
15	Group A	10	83.5 %	9.5 %	13.0 %	58.2 %	1.8 %
	Group B	5	53.3 %		2.3 %		
16	Group A	11	78.7 %	5.0 %	6.8 %	55.1 %	0.9 %
	Group B	5	52.9 %		1.3 %		
17	Group A	12	**73.8 %**	2.3 %	3.2 %	52.1 %	0.4 %
	Group B	5	**53.0 %**		0.6 %		
18	Group A	13	69.7 %	1.0 %	1.3 %	49.3 %	0.0 %
	Group B	5	52.6 %		0.2 %		

Additional experiments based on the configuration of Experiment Series 03 are presented in [12], in which four input parameters were systematically varied for the various objectives of the experiments. In total, twenty-three experiments were conducted, enabling a more detailed and in-depth analysis.

7 Summary of the Comparison and Outlook

The qualitative analysis demonstrates that while the GLEISE tool is efficient for homogeneous and simple track groups, it is limited by its requirement for mutual substitutability of tracks and a restricted set of input parameters. SimBed extends the modelling capability to more complex scenarios, allowing for multiple arrival flows and partial track groups, but still represents a simplified version of reality. RANGO, based on mathematical optimisation, offers the highest flexibility and modelling accuracy, supporting complex infrastructure and operational scenarios.

Quantitatively, three experimental series were conducted. Across all series, the results were consistent: RANGO's optimisation approach yields a lower optimal number of tracks than GLEISE and SimBed for a given quality level. However, the conceptual differences in how waiting probability and track utilisation are calculated must be considered when interpreting the results. As shown in Table 3-5, the difference in the optimal number of tracks between the methods is remarkably consistent, with RANGO typically requiring three fewer tracks than the other approaches.

A possible explanation for this finding is that, in RANGO, inbound trains are accommodated on additional fictitious tracks. As a result, subsequent trains arriving when there are available tracks within the track groups can be serviced immediately. This reduces the likelihood of waiting compared to scenarios in which all trains must wait in sequence in a queue. Consequently, the waiting probability is lower in this case than in situations where every inbound train is required to queue according to its order of arrival. Therefore, it can be concluded that the percentage of unrealised track utilisations serves as an approximate estimate of the waiting probability, which tends to be lower.

Each experiment highlighted in this section serves as a representative example of the three series of experiments. In addition, the research project included a further fifty-two experiments in which input parameters were systematically varied. These additional experiments confirmed the sensitivity of the results to changes in arrival and service time distributions. Reflecting on the broader set of results, it would be valuable for future research to examine the deviation of RANGO's results—specifically, the percentage of unrealised track utilisations (degree of non-fulfilment)—from those of LUKS or SimBed, as reported by waiting probability, for example, by quantifying how much lower this percentage is. Such reference values could serve as benchmarks for future studies and practical applications.

Considering the parameter "Degree of Occupancy," which is evaluated according to the recommended values from UIC 406, the concatenated occupancy rate for track areas should be between 40% and 50%. It should be noted that these values are specifically designed for evaluation using the compression method. However, the values obtained from RANGO are derived from optimisation techniques, while those from GLEISE and SimBed are based on queuing theory; none of these methods employ the compression approach.

According to UIC 406, the concatenated occupancy time is calculated without any buffer between compressed train paths. Therefore, an additional time rate—ranging from 150% to 100%—must be added to the occupancy time to achieve a consumption rate of 100%, which is considered to meet the accepted quality of service. In this context, the recommended concatenated occupancy rate of 40% to 50% should be regarded as a minimum threshold, not as a target or upper bound for practical track utilisation.

The experimental results are consistent with this observation, as none of the experiments resulted in a degree of occupancy below 50%, and most scenarios required a higher utilisation rate to meet the defined quality standards. Since only a minimum threshold is currently defined, further research is needed to determine the target or upper bound for practical track utilisation in track groups.

Literature

[1] DB Netz AG / DB InfraGO AG, “Richtlinie 405 - Fahrwegkapazität,” 2022.

[2] International Union of Railways (UIC), “Code 406 - Capacity,” 2013.

[3] M. Bär, W. Lenze, J. Eisold and N. Nießen, “Benchmark von Qualitätsmaßstäben für Kapazitätsermittlungen,” in *Proceedings of the 2nd International Railway Symposium Aachen, pp. 513-525*, 2019.

[4] J. Eisold and O. Freitag, “Weiterentwicklung der eisenbahnbetriebswissenschaftlichen Verfahren für Zugbildungsbahnhöfe durch Nutzung der mathematischen Optimierung,” in *Proceedings of the 3rd International Railway Symposium Aachen, pp. 449-461*, 2021.

[5] J. Eisold, F. Maciejewicz and O. Freitag, "Entwicklung eines Dimensionierungsverfahrens für Zugbildungsbahnhöfe auf Basis der mathematischen Optimierung," in *Proceedings of the 4th International Railway Symposium Aachen, pp. 460-472*, 2023.

[6] N. Nießen, A. Oetting, A. Ćosović, R. D. Etzold, D. Menne, J. Schüttler, U. Steinborn, W. Weinhold, A. Wolters and F. Zwanziger, Planungs- und Betriebsmanagement für das System Bahn, 1. Auflage ed., Berlin: Bahn Fachverlag, 2023.

[7] O. Freitag, S. Speer, O. Konovalow and L. Meissner, "Entscheidungsunterstützung zur nachfragegerechten Dimensionierung von Anlagen des Schienengüterverkehrs," *Eisenbahntechnische Rundschau,* no. 07-08/2020, pp. 44-48, 2020.

[8] "LUKS – Leistungsuntersuchung Knoten und Strecken," quattron, [Online]. Available: https://www.quattron.com/leistungen/software-it-digital-data-management/luks/. [Accessed 01 07 2025].

[9] G. Hertel, Analytische Modellierung und formelmäßige Behandlung von Standard-Bedienungssystemen mit störanfälligen Kanälen, Dissertation HfV Dresden, 1986.

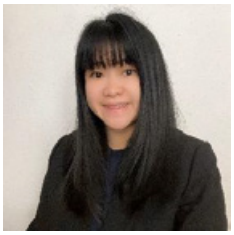
[10] J. Eisold, *Modellierung und Simulation von Bahnbetriebsprozessen: Kapitel III Simulation von Bahnbetriebsprozessen. Vorlesungsskript.,* Dresden: Technische Universität Dresden, 2025.

[11] O. Konovalow, B. Haack and S. Bäcker, "Dimensionierungsanalyse DimA: Verfahren zur Dimensionierung von Anlagen des Schienengüterverkehrs als Entscheidungsunterstützung für zukünftige Investitionen.," in *Proceedings of HEUREKA 2024*, Cologne, 2024.

[12] B. Pionteck, *Untersuchung der Kenngrößen und Qualitätsmaßstäbe für die Bemessung von Gleisgruppen,* Dresden, 2025.

Session 1
Sirikorn Satitrattanashewin, Jan Eisold, Beate Lux

Authors



Satitrattanashewin, Sirikorn

Sirikorn Satitrattanashewin studied Railway System Engineering (M.Sc.) at TU Dresden. She initially worked in software development and as an interlocking logic engineer at Bombardier Transportation Signal (Thailand), now known as Alstom. Since 2023, she has been a research assistant for self-learning tools and is working at the railway operations laboratory at the faculty of transport and traffic sciences, while pursuing her doctoral studies.



Eisold, Jan

Dr.-Ing. Jan Eisold studied transportation engineering at TU Dresden and has been working there since 2012 as a research associate at the Chair of Railway Operations. He successfully completed his doctorate in 2020. His work focuses on railway operations research, resilience, capacity, and rail freight transport, in which he is active in both teaching and research.



Lux, Beate

Dr.-Ing. Beate Lux studied traffic engineering at the TU Dresden. Initially, she worked in aircraft construction and in the project management of railway infrastructure projects. In 2002, she returned to TU Dresden and has been researching and teaching in the field of public transport operations. Since 2020, she is responsible for the railway operations laboratory.

IRSA 2025

SESSION 2

KAPAZITÄT II

RAILWAY CAPACITY II

Session 2
Harald Buschbacher

Targeted infrastructure investments for more efficient use of railway capacity by different market segments

Buschbacher, Harald [1]

[1]Schieneninfrastruktur-Dienstleistungsgesellschaft mbH

Keyworts: Railway capacity management, ETCS, ERTMS, integrated clock-face timetabling, open-access competition, modal shift

1 Introduction

In order to contribute to climate action in the upcoming years, the railway sector is expected to deliver more services and connections in order to allow for modal shift from more harmful modes of transport. In particular, new cross-border long-distance passenger trains have a high potential for greenhouse gas avoidance by modal shift from passenger aviation. Furthermore, the long-distance passenger segment of the railway market is growing as past steps of market liberalization acts make impact and more and more railway companies announce their plans to connect cities with their passenger trains.

Capacity requests for new train services, not expected during planning of rail infrastructure often leads to capacity conflicts with existing services. The current state of negotiations about the upcoming EU rail capacity regulation comprises a more structured process of capacity planning and allocation. However, there will be no legal basis to limit the capacity offered for long-distance passenger trains strictly to the intended numbers of such trains according to timetable and infrastructure concepts elaborated many years ago.

Unfortunately, on many railway lines it is far from trivial to reconcile additional long-distance passenger trains with other capacity demands: On mix-use railway lines with slower freight and local passenger train paths between the faster long-distance passenger train paths, additional fast long-distance passenger train paths usually require additional overtaking of slower trains. This causes not only longer running time for those trains, but also specifically longer travel time for passengers missing their interchange connections, as the increased running time does not fit in the integrated clock-face timetable anymore.

A conventional solution for additional capacity needs would be additional tracks along the whole line, requiring for significant sections tunneling because of limited available space on the surface next to the existing line. In order to contribute to climate action by solving rail capacity bottlenecks of the next 5-10 years, such huge investments, if not already ongoing, take too much time for implementation and they have relevant undesired climate effects during the construction phase. The aim of this paper is to identify solutions, contributing to a better balance between the benefits of additional long-distance services, disadvantages for freight and local passenger trains and costs for infrastructure improvements.

2 Method

Train paths of faster and slower trains get most close to each other next to stations that represent nodes of the integrated clock-face timetable or where overtaking between faster and slower trains takes place. In order to facilitate additional trains without significantly compromising travel or transport time of at least some trains, we must focus on the technically and operationally feasible headway time close to these stations.

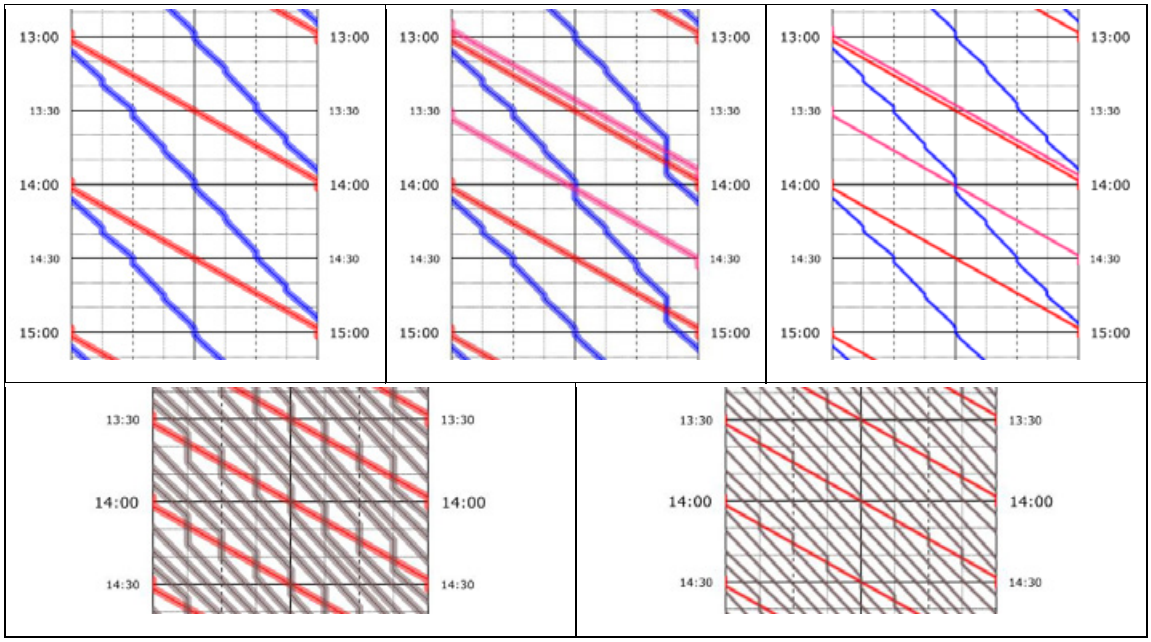


Figure 1: simplistic graphic timetable for mix-use railway lines with different speed per market segment, showing only one track of a double-track line. Blue: local passenger trains, red: regular-interval long-distance passenger trains, pink: additional long-distance passenger trains, brown: freight trains. Solid, narrow lines represent the actual trainpath, wider lines of lower opacity represent the trainpath including the half headway time before and the half headway time after.

The upper left corner of Figure 1 represents an example of line as a part of an integrated clock-face timetable with local and long-distance passenger trains. At each end of the line section, there is a node station with interchange connections between long-distance and local passenger trains, also to and from trains on branch lines. In the middle part of the upper row of Figure 1 additional long-distance passenger trains have been introduced into the timetable. Now the long-distance trains must overtake the local trains, increasing the travel time of the local trains to such extent that they arrive in the node after the intended node time. Interchange connections to other trains are broken and interchanging passengers experience a travel time increase of a whole service interval of the connecting service, usually between 30 and 120 minutes. With shorter headway times (slim lines in the right upper corner of Figure 1), it is possible to introduce additional long-distance trains without losing interchange connections between existing trains.

The lower part of Figure 1 illustrates the effect of different headway times on the mutual compatibility between freight trains and long-distance passenger trains: In case of longer headway times (left), there is capacity for three freight trains per 30 minutes and they have to stop at every third station in order to let a long-distance passenger train overtake. In case of shorter headway times (right), one more freight train can run every 30 minutes and freight trains must stop only at every fourth station for overtaking.

The reference case for the improvement approaches considered in this paper is a mix-use double-track railway line with ETCS-only operation and short block sections as a state-of-the-art train control system. For this infrastructure reference case, four timetable resp. operational reference cases were defined: overtaking procedures between long-distance passenger trains and local trains as well as between long-distance passenger trains and freight trains, both with or without a stop of the overtaking train in the respective station.

These four reference cases of overtaking are represented in Figure 2 to Figure 5. In these figures, the colour of the lines in the graphic timetable corresponds to the colour of the track in the line scheme above the graphic timetable. Solid lines represent freight trains, dashed lines represent long-distance passenger trains and dotted lines represent local passenger trains. The vertical axis means the time (minutes of an hour of regular-interval operation) and the horizontal axis means the distance (kilometres to the left and to the right from the station where the overtaking takes place). These figures show both directions and represent a symmetric timetable around the symmetry minute zero. The same legend applies to Figure 7 to Figure 13, Figure 17, Figure 18 and Figure 20 in Section 3 (Results).

Session 2
Harald Buschbacher

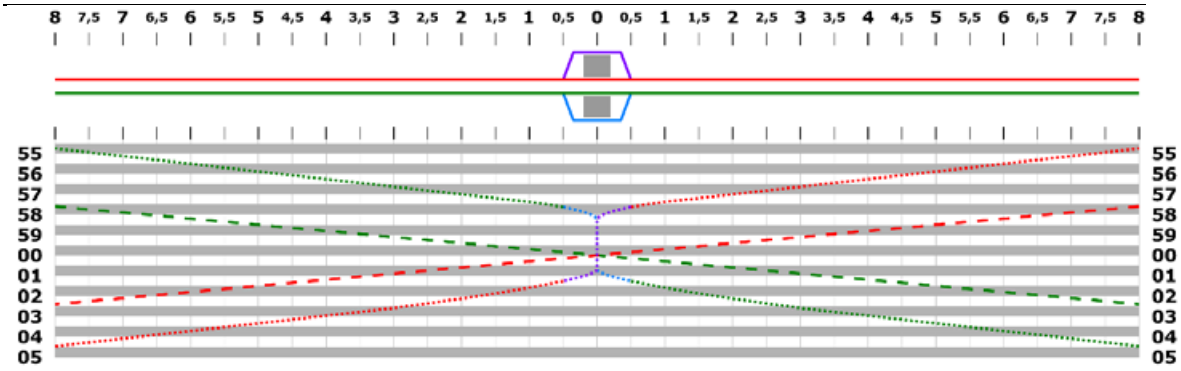


Figure 2: graphic timetable of an overtaking of a local passenger train by a long-distance passenger train passing through the overtaking station

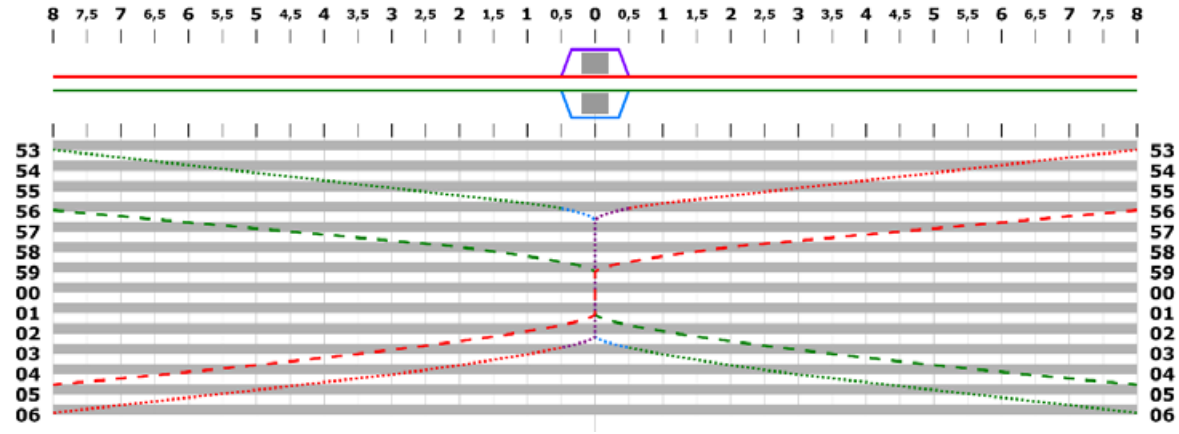


Figure 3: graphic timetable of an overtaking of a local passenger train by a long-distance passenger train stopping in the overtaking station

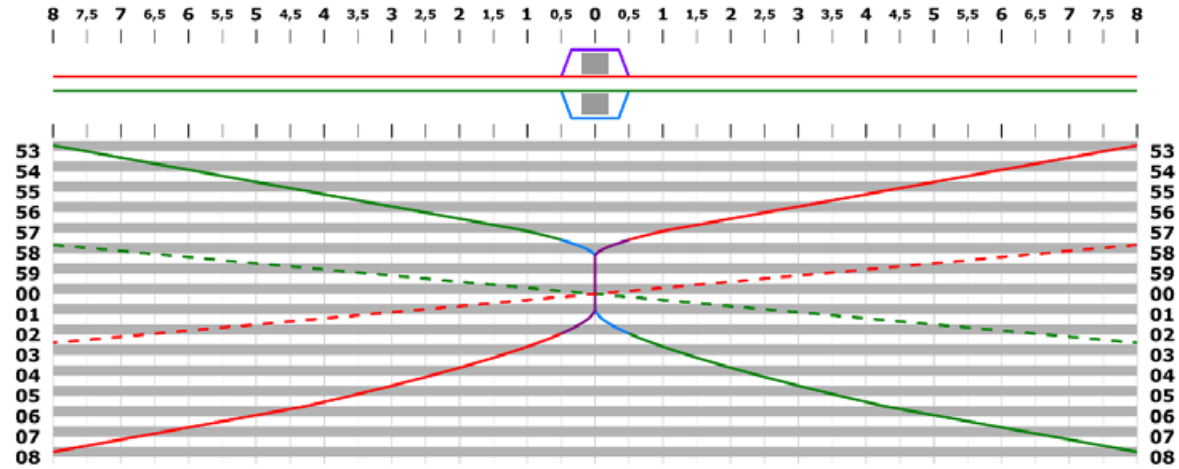


Figure 4: graphic timetable of an overtaking of a freight train by a long-distance passenger train passing through the overtaking station

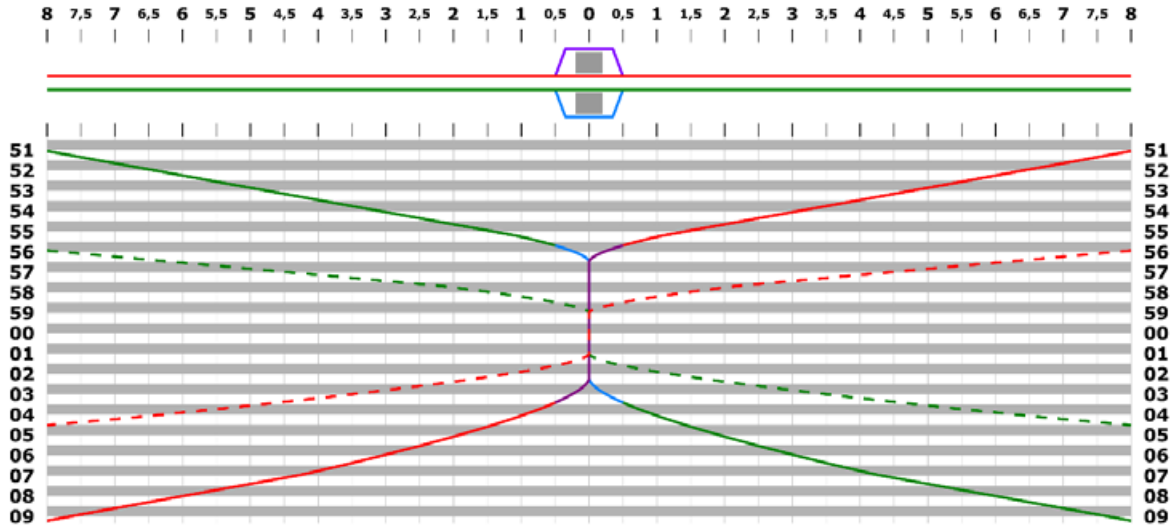


Figure 5: graphic timetable of an overtaking of a local passenger train by a long-distance passenger train stopping in the overtaking station

For calculations of acceleration, deceleration and headway for both the reference cases and the considered improvements, the following calculation assumptions were made:

- Speed level on mix-use railway lines:
  - Long-distance passenger trains: 200 km/h
  - Local passenger trains: 160 km/h
  - Freight trains: 100 km/h
- Train length:
  - Long-distance passenger trains: 400 m
  - Local passenger trains: 200 m
  - Freight trains: 700 m
- Concrete trainsets for the calculation of acceleration and deceleration:
  - Long-distance passenger trains: Loco-hauled train with two Siemens Eurosprinter ES64U4 locomotives and 14 carriages, total weight of all carriages 770 t, braking characteristics 150 R.
  - Local passenger trains: Siemens Desiro Mainline EMU, three-car-version, braking characteristics 150 R
  - Freight trains: Siemens Eurosprinter ES64U4 locomotive with 47 two-axle carriages of a total weight of 2115 t, braking characteristics 80 R.

  Acceleration and deceleration was calculated on a flat line using the FBS / iPLAN timetabling software [1].

- Braking curves: For the calculation of minimum headways between the train front and the end of authority the braking curve calculation tool of the European Railway Agency (ERA) [2] was used, applying the following parameters:
  - Braking characteristics: 150 P for passenger trains, 80 P for freight trains
  - Train length: 160 m for passenger trains, 700m for freight trains
  - Lambda braking model
  - 7,81 m distance between train front and ETCS antenna

  For the headway calculations, the "indication" braking curve was applied.
- Block section length: 200 m
- Length of switches: 150 m per switch
- Response time of signalling system and physical movement time of switches: 30 s
- Slip distance between regular stopping position of a passenger train at the platform and end of movement authority: 80 m

These calculation assumptions lead to a headway time at constant speed of about 1:45 minutes between two passenger trains and 2:20 minutes between two freight trains.

Following these calculation inputs, between two trains involved in an overtaking, headway time after departure is much shorter than before arrival. This difference stems from the braking distance the fast overtaking train has to pass before arrival, which is longer than the slip distance the overtaken train has to pass after departure. The other parts of the headway time are the same before arrival and after departure (see Figure 6).

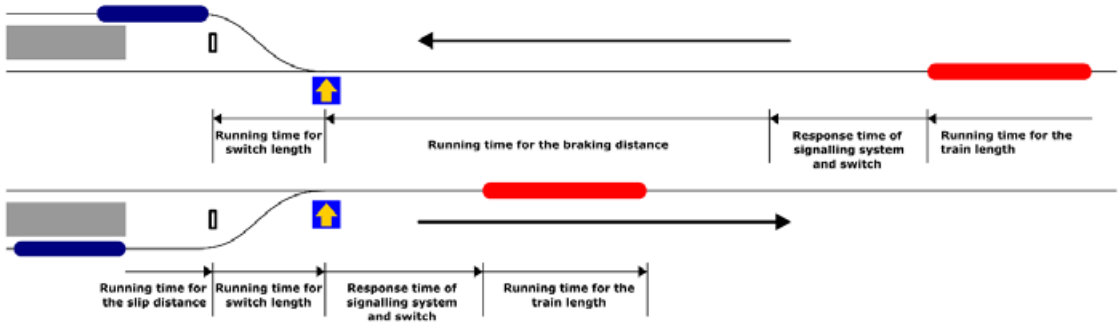


Figure 6: Composition of headway time before arrival and after departure of trains of different speed and stopping pattern under the conditions of ETCS only operation or another train control system with short blocks, but long calculated braking distance.

Under legacy train control systems like PZB, there is no such asymmetry between headway times before arrival and after departure, because first PZB is based on shorter braking distance assumptions (at same speed) and second, headway time after departure is determined by the feasible block section length that is much longer than under ETCS.

Following the principle of timetable symmetry in order to ensure the same interchange connections in both directions, the project focused on measures to shorten the headway time before arrival to an amount similar to the headway time after departure.

This paper presents only solution approaches that turned out to be promising. For further details on the method and less promising solution approaches, see the full report [3].

3 Results

3.1 Triple tracking for arriving trains

If overtaking between a local and a long-distance passenger train takes place without a stop of the long-distance train, a 6 km long triple track section before the station shortens the headway before arrival by about a minute, leading to nearly perfect symmetry.

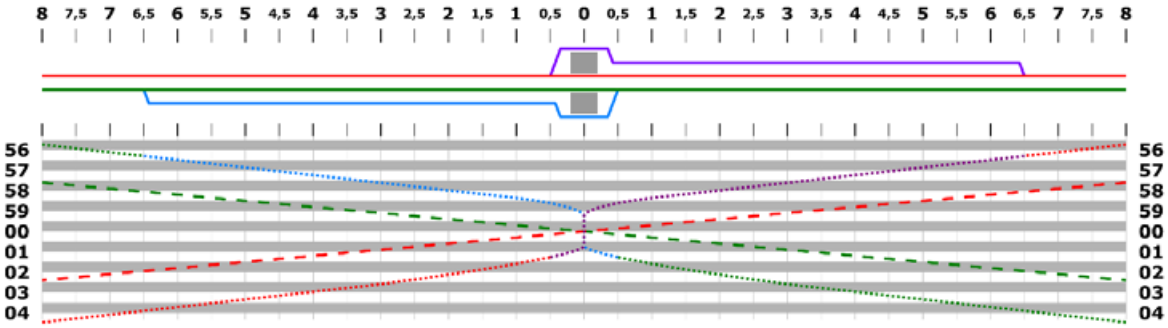


Figure 7: graphic timetable of an overtaking of a local train by a long-distance train passing through the overtaking station with a 6 km triple track section before arrival

If the long-distance passenger train also stops in the station where overtaking takes place, the headway reduction before arrival is similar, but slight asymmetry remains.

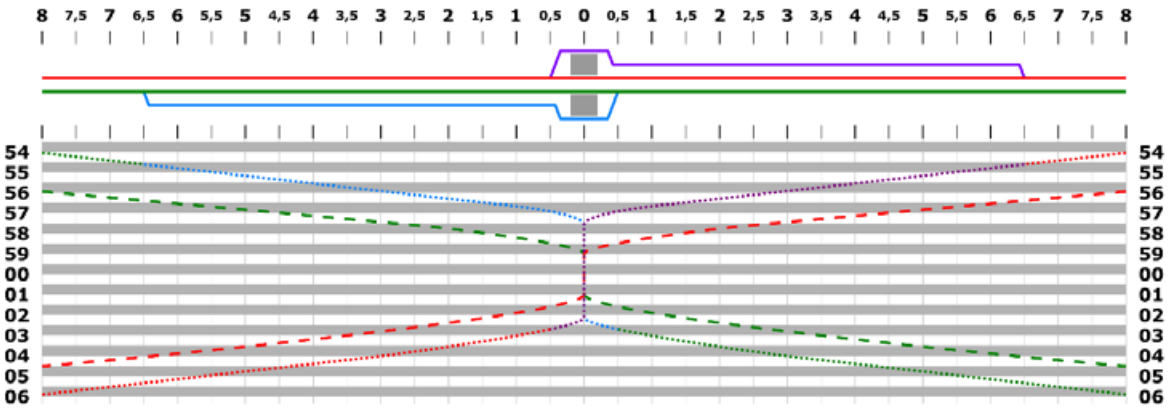


Figure 8: graphic timetable of an overtaking of a local train by a long-distance train stopping in the overtaking station with a 6 km triple track section before arrival

Because of more difference in speed and deceleration rate between freight and long-distance passenger trains (compared to local vs. long-distance passenger trains), triple tracking has more capacity effect when used for overtaking between freight and long-distance passenger trains: headway time before arrival is reduced by about two and a half minute.

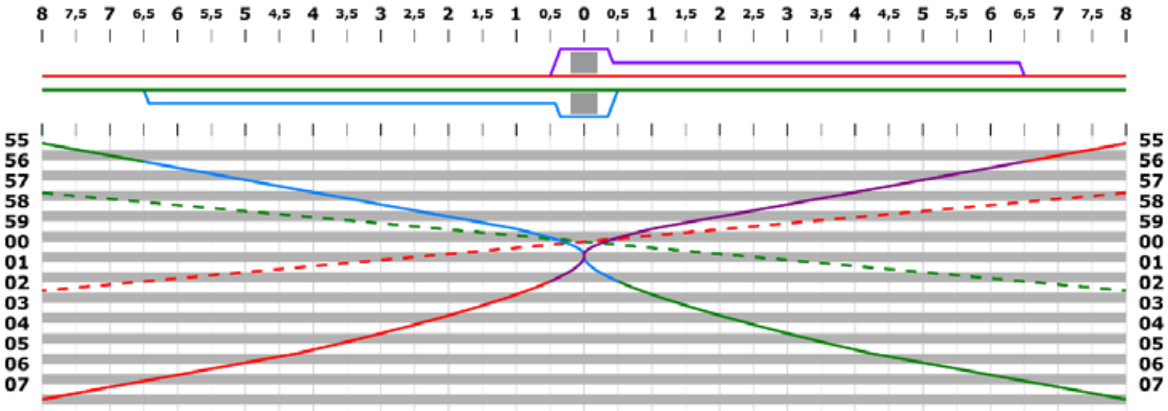


Figure 9: graphic timetable of an overtaking of a freight train by a long-distance passenger train passing through the overtaking station with use of a 6 km triple track section before arrival.

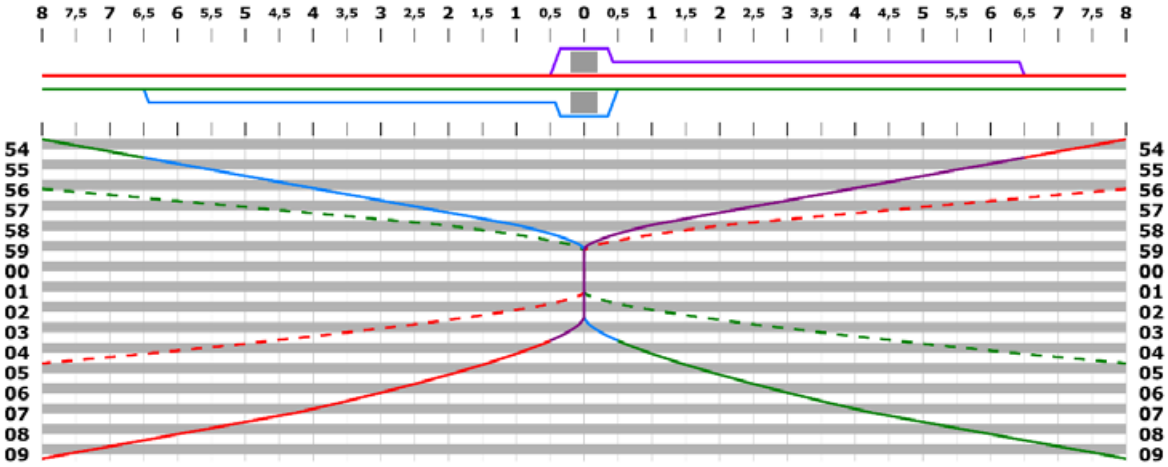


Figure 10: graphic timetable of an overtaking of a freight train by a long-distance passenger train stopping in the overtaking station with use of a 6 km triple track section before arrival.

A similar reduction in headway time by triple tracking can be achieved also for faster and slower passenger trains, if the triple track section comprises a stop of the local passenger trains: regardless whether the long-distance train stops in the station where the overtaking takes place or not, the headway before arrival is reduced by more than 90 seconds in addition to the effect of a triple track section without a stop. The shortening of the headway before arrival overcompensates the difference to the headway after departure

according to the reference case. This means, that additional measures would be useful to compensate the asymmetry, e.g. similar triple tracking on the opposite ends of the respective sections or asymmetric transfer times in nodes with interchange on the same platform or between different platforms depending on the direction of travel.

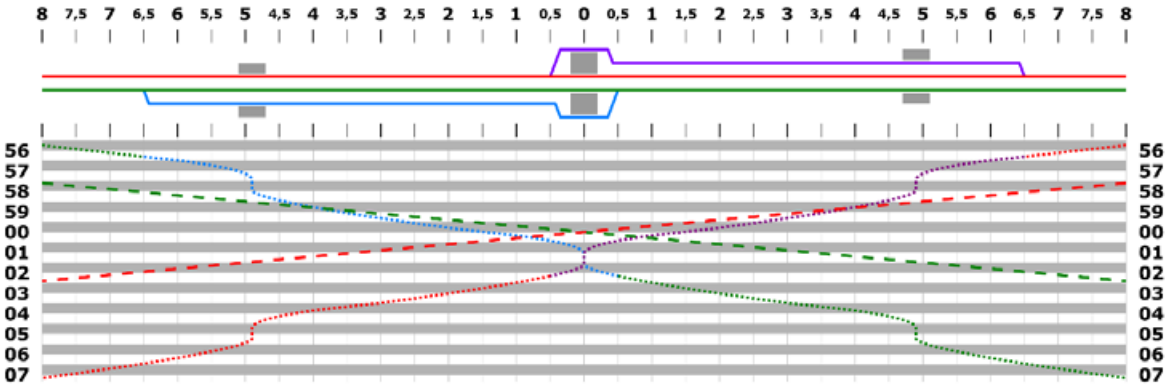


Figure 11: graphic timetable of an overtaking of a local passenger train by a long-distance passenger train passing through the overtaking station with use of a 6 km triple track section before arrival, with a stop of the local passenger train in the triple track section.

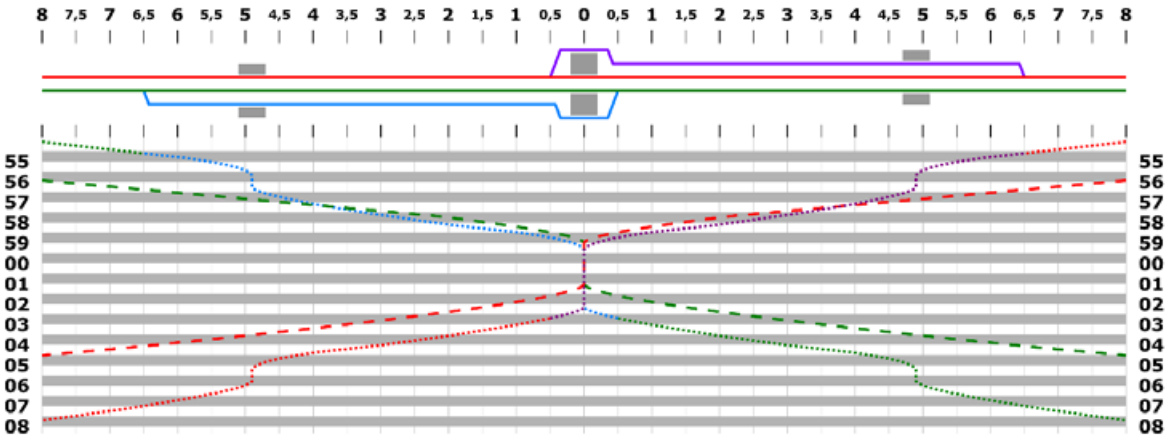


Figure 12: graphic timetable of an overtaking of a local passenger train by a long-distance passenger train stopping in the overtaking station with use of a 6 km triple track section before arrival, with a stop of the local passenger train in the triple track section.

3.2 Combinations of triple tracking and use of the opposite direction's track

Triple tracking is difficult in many locations due to land use conflicts. Therefore, it was investigated whether the use of the opposite direction's track would be a viable alternative to triple tracking for the resolution of capacity issues between faster and slower trains. Most of several examined variants would most probably have less positive effects on capacity in one direction than negative effects on capacity in the other direction, but one

variant is promising: Where triple tracking is impossible in the urban area close to the station where faster trains shall overtake slower ones, the slower trains could use a longer triple track section more distant to the station and then bridge the gap until the beginning of the station by using the opposite direction's track. In case of a 1,5 km double track section between a 4,5 km triple track section (with a stop for local trains) and the station, the headway time between arriving local and long-distance passenger trains is reduced by 2:38 minutes. At the same time, the theoretically possible headway of freight trains running constantly in the opposite direction becomes 3:16 minutes longer. At first sight, the positive effect in one direction is still smaller than the negative effect on the other direction. However, in the situation of faster and slower trains approaching close to each other a node station of the integrated clock-face timetable, headway time can be more decisive than between freight trains in the opposite direction, which might be anyway unfeasible at these concrete minutes because of other capacity conflicts along the line.

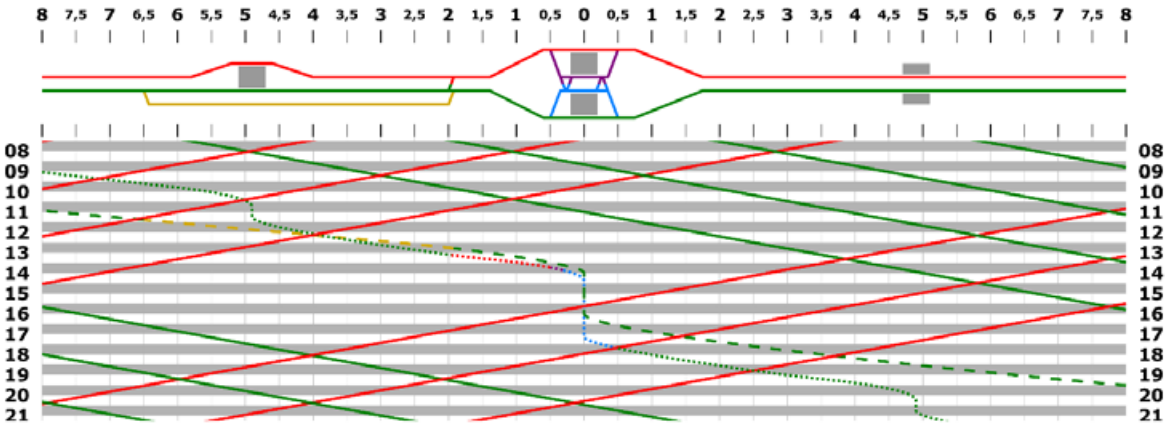


Figure 13: graphic timetable; trains running from left to right: overtaking of a local passenger train by a long-distance passenger train stopping in the overtaking station with use of a 4,5 km triple track section and 1,5 km of the opposite direction's track before arrival, with a stop of the local passenger train in the triple track section; trains running from right to left: freight trains running at constant speed with minimum possible headway between each other

For practical implementation of such a combination of triple tracking and use of the opposite direction's track, a suitable track configuration is decisive: The triple tracking becomes useless if the routes of the faster and the slower train cross each other at the end of the triple track section. A layout of the station and the triple track section with middle track(s) for slower trains and outer tracks for faster trains is favourable in terms of less expensive switches and more capacity efficient operation of train services that perform a turnaround after terminating in the respective station. On the other hand, such a station layout consumes more land than conventional stations with fastest tracks in the middle because of larger curve radii of the outer tracks.

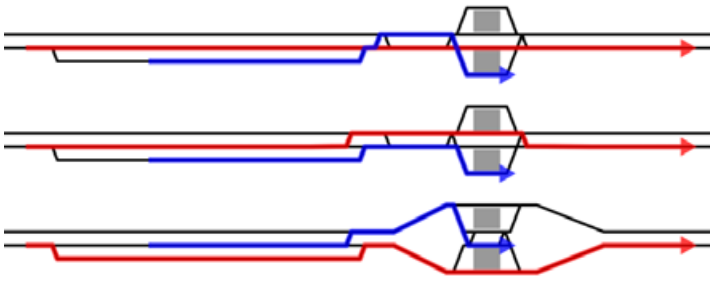


Figure 15: different layouts of triple track sections not directly attached to a station; blue: slower trains, red: faster trains; upper figure: track layout unsuitable for overtaking by combination of triple track section and use of the opposite direction's track; middle figure: track layout suitable for overtaking by combination of triple track section and use of the opposite direction's track, but requiring switches suitable for high speed in all directions; lower figure: track layout suitable for overtaking by combination of triple track section and use of the opposite direction's track, not requiring switches suitable for high speed in all directions



Figure 14: station layout with fast tracks in the middle, turnaround of a local train service with crossing routes between departing local train and passing long-distance train (blue arrow to the left and red arrow to the right)

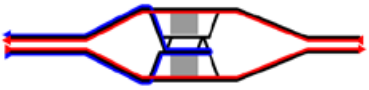


Figure 16: station layout with slow tracks in the middle, turnaround of a local train service (blue) without crossing routes

If the solution with partial triple tracking and partial use of the opposite direction's track according to Figure 13 is implemented together with an optimisation of the station layout according to Figure 16, eliminating the crossing routes in the context of turnarounds, the overall balance of shortenings and prolongations of headway times is positive compared to the reference case.

Another use case for triple tracking and station layouts following the principle of faster trains running always on the outer tracks is continuous triple tracking between two adjacent stations:

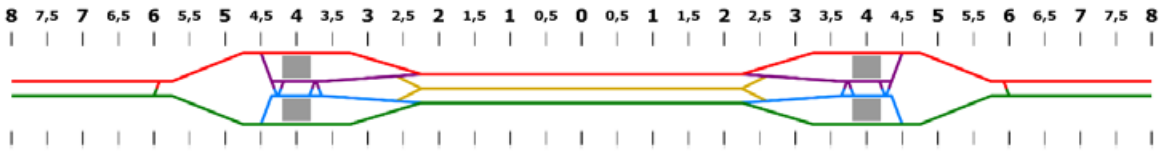


Figure 17: continuous triple track section between two stations to be used alternately for overtaking in both direction

Such a solution would allow for double use of the additional track, alternately for arrivals from the left side to the right station and from the right side to the left station. It would be rather suitable for overtaking between freight and long-distance passenger trains than between faster and slower passenger trains as overtaking in different stations per direction leads to asymmetry, undesired in timetables for passenger services.

3.3 Offset between passenger stop of long-distance trains and overtaking between freight and long-distance passenger trains

Capacity improvement without triple tracking is possible by optimised arrangement of stations where long-distance passenger trains stop and stations where these trains overtake freight trains: the train paths of a freight train and a fast passenger train can be accommodated closer to each other if the overtaking does not take place in a station where the long-distance passenger train stops, but at least the braking distance beyond that station. In this case, the fast long-distance train can brake later when approaching the stop without adjusting to the poor deceleration rate of the freight train.

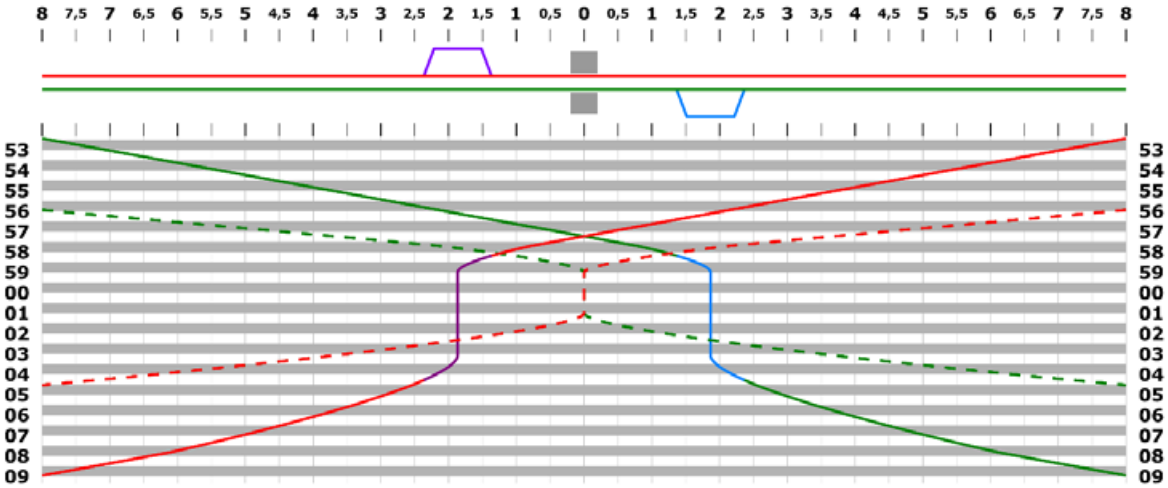


Figure 18: graphic timetable of an overtaking of a freight train by a long-distance passenger train with different stations for passenger stop and overtaking.

Such a solution is particularly useful in case of long sections without suitable locations for overtaking stations, e.g. long tunnels: Ideally, there is a stop for passenger trains immediately before and after the tunnel and a station for overtaking of freight trains slightly before and after these passenger stops.

3.4 Use of helper engines for freight trains within station limits

As running time difference between faster and slower trains reduce the usable capacity on a line section, better acceleration of freight trains would help to increase capacity on critical sections between stations where freight trains can be overtaken by long-distance passenger trains. If train operators deploy generally an additional engine to improve acceleration, they have to bear the costs of the second locomotive for the whole route, despite the need for better acceleration might be limited to a few sections. In these cases, it seems more efficient if the infrastructure manager provides a helper engine pushing the

freight train only for the very beginning of the acceleration phase and staying in the boundaries of the respective station without even reaching the switches at the end of the station before returning to its parking position.

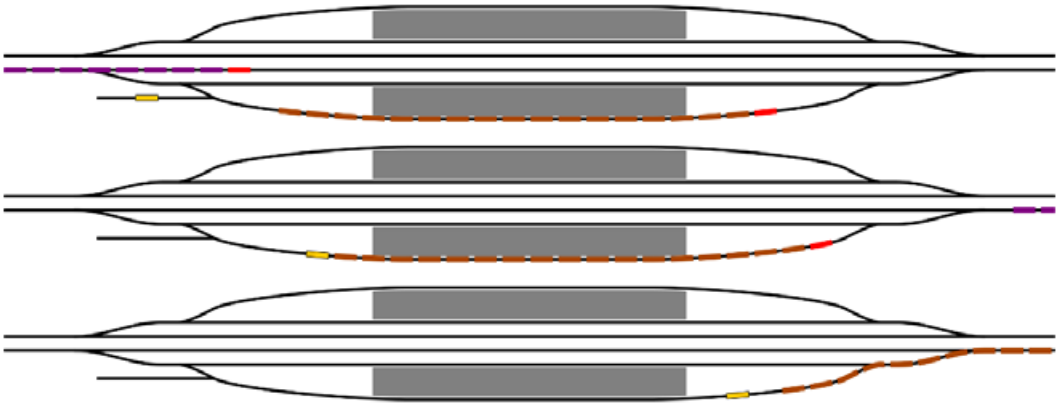


Figure 19: Helper engine (yellow), improving the acceleration of a freight train (brown cars, red locomotive) within station limits after overtaking by a long-distance passenger train (purple cars, red locomotive)

The effect of such use of a helper engine was examined by running time calculations for freight trains of 2000 t and 3000 t weight (without locomotive) and a Siemens Eurosprinter ES64U4 locomotive as a main engine. If the accelerating train is pushed over the first 600 m by a helper engine providing 260 kN additional tractive effort, it will reach the next station 41 seconds (2000 t train) resp. 59 seconds (3000 t train) earlier, allowing the next train to run accordingly earlier too. At the moment when the helper engine stops pushing, the speed of the train is 42 km/h (2000 t train) resp. 34 km/h (3000 t train).

3.5 Arrangement of additional long-distance passenger train paths in the transfer time between trains of the integrated clock-face timetable

With modern train control systems and short block sections, technically feasible headway times become significantly shorter than transfer times in nodes, in particular if transfer includes relevant walking times along a platform or from one platform to another one [4]. In other words, where interchange connections are foreseen between local and long-distance passenger trains running on the same line, the requirement of sufficient transfer time leads to a longer timespan between arrival of one train and departure of the other

train than the sum of technically feasible headway time plus stopping time of the long-distance train.

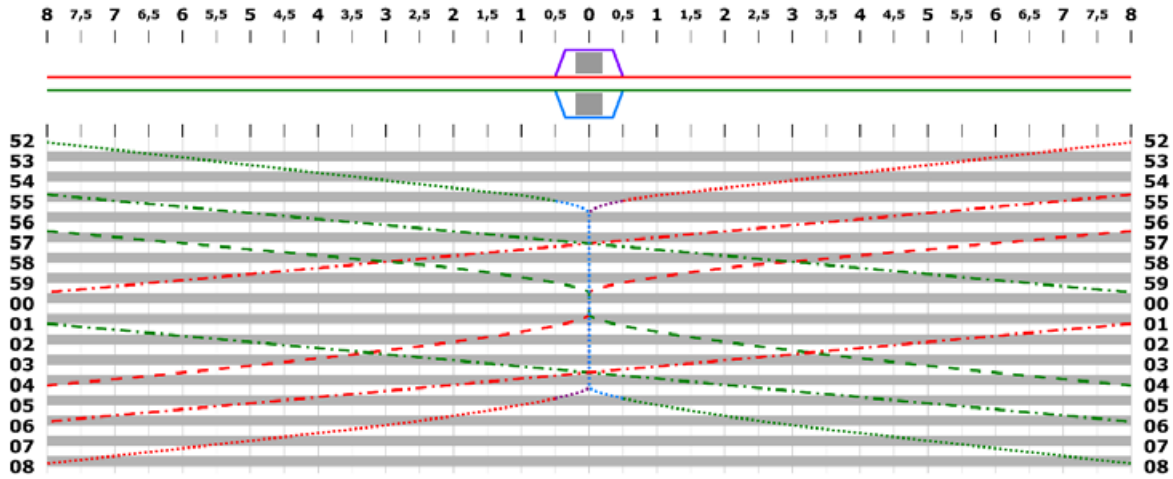


Figure 20: graphic timetable of an overtaking between a long-distance and a local passenger train in ETCS-only operation with additional long-distance trains passing during the transfer time between long-distance and regional trains. The dashed-dotted lines represent additional long-distance trains operated on top of the integrated regular-interval timetable.

As a result, the transfer time could be used to run an additional long-distance train that passes the respective station shortly before or after the stop of the regular long-distance train without a conflict with regional trains arriving before and departing after the stop of the long-distance train. Figure 20 represents such a solution based on the general calculation assumptions of this paper according to Section 2 (Method) and a one-minute-stop of the long-distance train, resulting in transfer times of 4:30 to 5:00 minutes between long-distance and local passenger trains. For stable and resilient operation, it seems more reasonable to run one additional train either before or after the regular-interval long-distance train stopping in the respective station than using both of these extra train paths. Whether the additional train shall take the earlier or the later option displayed in Figure 20 could be decided at short notice according to the current operational situation. In contrary to triple tracking, optimised spacing between stations for passenger stops and for overtaking, or provision of helper engines, this solution does not require the infrastructure manager to invest in additional physical assets compared to the reference cases described in Section 2 (Method). Nevertheless, it might require the introduction of new methods of capacity allocation and management to allow flexibility to run the additional train either before or after the regular-interval train.

4 Conclusions

4.1 Most promising solution approaches

In the decisive situation of slower and faster trains approaching a node or overtaking station tightly after each other, there are three ways to shorten headway times to a sufficient extent to facilitate one further long-distance passenger train path without compromising other market segments:

- Making freight trains and long-distance passenger trains run on different tracks over the last 6 km before arrival at the station. If continuous triple tracking is not possible over the whole distance, bridging of a short distance before the station by using the opposite direction's track should be considered in detail.
- The same solution for the situation with local and long-distance passenger trains approaching the station achieves the necessary reduction of headway time only if there is a stop of the local passenger train in the triple track section.
- Use of the transfer time between regular-interval long-distance trains and local trains for an additional long-distance train to pass the station without stopping.

The other solutions mentioned in this paper could help to facilitate additional train paths together with other measures or existing capacity margins.

Even if not used for regular operation, all the considered infrastructure investments contribute to punctuality, reliability and resilience.

4.2 Potential cost efficiency of targeted infrastructure investments

A rough indication for cost efficiency can be calculated as the number of additional train-kilometers per additional track-kilometers: In case of a 200 km/h mix-use double track line with an hourly interval of long-distance trains, a realistic distance between two nodes of the integrated clock-face timetable is about 60 km. 12 km of additional tracks (6 km triple tracking at either end of the section) would facilitate one additional long-distance train per hour and direction or 120 train-km per hour in total. In terms of infrastructure utilization, this is equivalent to a new double-track line used for 10 trains per hour and direction. This indicator can be even significantly higher, if such targeted infrastructure investments remove bottlenecks, facilitating additional trains over much longer distance. Furthermore, the intention of selective triple tracking is to achieve lower costs per additional track-km compared to entirely new lines or continuous quadruple tracking.

Literature

[1] iRFP e.K., Fahrplanbearbeitungssystem FBS (1.8.2) [Software], http://www.irfp.de.

[2] European Railway Agency (ERA), “Braking Curves Simulation Tool,” 09 01 2024. [Online]. Available: https://www.era.europa.eu/system/files/2022-11/era_braking_curves_simulation_tool_en.xls?t=1733133736.

[3] H. Buschbacher, "Minimalinvasive Lösungsansätze zur besseren trassentechnischen Verträglichkeit von integralem Taktfahrplan, taktüberlagerten Personenfernverkehrszügen sowie Güterzügen," 16 06 2025. [Online]. Available: https://www.schig.com/fileadmin/Media/Blogs/2025/Minimalinvasive_Loesungsansaetze_zur_besseren_trassentechnischen_Vertraeglichkeit.pdf. [Accessed 16 06 2025].

[4] ÖBB-Infrastruktur AG, "Mindestübergangszeiten (MÜZ) - Minimum Transfer Times," 25 03 2025. [Online]. Available: https://infrastruktur.oebb.at/geschaeftspartner/schienennetz/snnb/snnb-2026/snnb-2026-anhaenge/mindestuebergangszeiten.

[5] J. Pachl, Systemtechnik des Schienenverkehrs - Bahnbetrieb planen, steuern und sichern, 8. Auflage ed., Berlin, Heidelberg, New York: Springer-Verlag, 2016.

Author



Buschbacher, Harald

Harald Buschbacher, graduated in Spatial Planning/Transportation Planning works at SCHIG mbH, the competence centre for the Austrian railways, focusing on strategic passenger timetable planning according to the principle of integrated regular interval timetables. Harald is familiar with cross-border capacity issues through his participation in the advisory group of the EuroLink working group of RNE (RailNetEurope).

Session 2
Cédric Kekes, Franziska Theurich, Karl Nachtigall

Evaluating network capacity effects of ETCS L2 in a timetable-based capacity assessment model

Kekes, Cédric[1], Theurich, Franziska[1] and Nachtigall, Karl[1]

[1]Technische Universität Dresden

Abstract

We evaluate the impact of the European Train Control System (ETCS) on railway network capacity in Germany. Using a timetable-based assessment model, we analyse how different infrastructure equipment scenarios affect residual freight capacity. Applied to a test network in Northern Germany spanning Hamburg, Bremen and Hanover, the approach demonstrates how ETCS implementation affects network-wide freight capacity under realistic deployment scenarios. The research aims to provide insights for strategic ETCS deployment planning across railway networks, considering the challenges of mixed equipment operation during transition periods.

Keywords: railway networks, timetable-based capacity assessment, ETCS

1 Introduction

Railway networks face increasing traffic demands whilst approaching saturation. As infrastructure planning paradigms shift towards timetable-based design, evaluating residual capacity for freight trains under a fixed passenger timetable gains importance. Capacity gains are expected with the introduction of the European Train Control System (ETCS). In Germany, the National Implementation Plan (NIP) for ETCS [7] targets the network-wide conversion to ETCS L2 by 2043, requiring all vehicles to be equipped accordingly. However, the plan also identifies that current funding gaps jeopardise this timeline. Therefore,

we need to analyse equipment scenarios regarding both infrastructure and rolling stock. Whilst the effects of upgrading individual lines or corridors to ETCS have been studied extensively, network-wide approaches applied in practical contexts remain rare.

This paper's contribution is extending our timetable-based capacity assessment to quantify ETCS L2 effects on a network scale, accounting for both infrastructure upgrade locations and rolling stock capabilities, and to test the approach in a real-world network. We analyse scenarios comparing the status quo with ETCS L2 deployment in Northern Germany according to an implementation step of the NIP. The overall aim is to provide information to decision makers about potential capacity effects of upgrading parts of a railway network for the overall network. The remainder of this paper is structured as follows: in Section 2, we present related work on railway capacity and its dependencies with signalling and control technology. Afterwards, in Section 3, the timetable-based capacity assessment model used is shown and the incorporation of ETCS L2 into the model introduced. Subsequently, in Section 4, we present our case study in an initial application in Northern Germany and discuss the methodology's application, and conclude in Section 5 with insights for strategic ETCS deployment planning and an outlook on further work.

2 Related work

Railway capacity is a well-studied field of research and multiple approaches exist to assess capacity of railway infrastructure, often with a focus on a particular element type in the network (e.g. station, corridor). However, railway capacity depends on a variety of factors: the interdependencies between number of trains, homogeneity of travel times, average speed and operational stability [1]. Consequently, several indicators for capacity exist, including the throughput of the infrastructure (e.g. number of trains, passenger numbers, freight volume), the stability (e.g. overall delay on the infrastructure, average delay for each train), and the saturation of the system (e.g. infrastructure-occupation times) [15]. Each of these indicators can be influenced by the signalling and train control equipment in place; consequently, the signalling and train protection system used also has an impact on overall capacity [9]. This section will briefly present recent works regarding the interdependencies between capacity and signalisation.

2.1 ETCS capacity enhancement studies

The wide interests in the capacity enhancement effects of ETCS manifests in multiple publications on this topic. Mikulčić et al. [13] provide a comprehensive literature review, finding that ETCS L2 overlay on conventional signalling yields minor capacity improvements unless combined with new block division configurations. Their analysis reveals most studies focus on double-track suburban and urban lines using microscopic approaches. The studies are mostly microscopic in nature, highlighting the need for further research in a combined micro-, meso- and macroscopic level. Recent corridor-based studies demonstrate varying levels of ETCS capacity benefits. Cansu et al. [16] assess the capacity effects of upgrading legacy ATC2 signalling systems to ETCS Level 2, using a mixed-method approach combining analytical models for stations and microscopic simulation for lines. Their case study focusses on a 67 km railway corridor in Northern Italy. In his dissertation, Wagner [17] quantifies headway changes under different ETCS configurations with a focus on railway station capacity, demonstrating that ETCS L2 requires additional measures, such as shorter block sections and strategies maintaining line speed until the beginning of switch areas, to achieve capacity gains. Knutsen et al. [12] evaluate ETCS Hybrid Level 3 using RailSys simulation, finding modest headway improvements in two-train scenarios but minimal network-wide effects on capacity utilisation and punctuality, identifying infrastructure bottlenecks as key limiting factors.

2.2 Research outlook

Current literature reveals a gap between well-established corridor-based ETCS capacity studies and practical network-wide assessment approaches. Whilst individual line benefits are documented, system-wide effects during mixed equipment operation, traffic redistribution patterns, and phased implementation strategies remain largely unexplored. This research aims to address this gap by developing a comprehensive network-wide framework for assessing ETCS Level 2 capacity impacts under realistic deployment scenarios, extending beyond corridor-focused approaches to capture system-wide interdependencies critical for strategic implementation planning. The approach is based on the timetable-based capacity assessment model, which is presented in the next section.

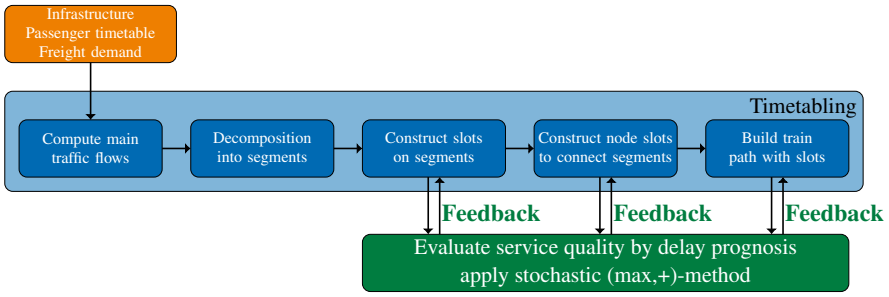


Figure 1: Schematic representation of the approach for capacity assessment

3 Framework and Methodology

The research uses the capacity assessment approach first presented in [3] and further developed in [10, 11, 4]. The approach assesses residual freight capacity in a railway network with a given passenger timetable, whilst ensuring the resulting combined schedule maintains robustness and meets specified delay performance criteria. The methodology enables comparative analysis by computing slots both with and without ETCS, allowing for direct evaluation of capacity improvements. In this section, we briefly present the timetable-based capacity assessment model before discussing how to include ETCS L2 to perform evaluations on its network capacity effects.

3.1 Timetable-based Capacity Assessment Model

The model is built on two pillars: Freight train scheduling as a mathematical optimisation problem and evaluating the robustness of the resulting timetable with a stochastic (max,+)-approach. As freight demand fluctuates, a slot-based train path allocation similar to that proposed in [14] is used. The timetabling module constructs slots for train paths based on infrastructure, a given passenger timetable, and rail freight demand, including the capacity utilisation within nodes. Meanwhile, the evaluation module provides feedback in the form of delay assessment through stochastic delay propagation using an event-driven and (max,+)-based framework akin to [8]. A schematic representation of the assessment approach is shown in Figure 1.

3.1.1 Input data

The model uses three main inputs stored in a database capable of accommodating different sources for scenario analysis:

- **Railway infrastructure data:** Modelled on a mesoscopic level as a graph, with arcs representing tracks and vertices representing stations or block signals. Thus, block sections of each railway line are considered, enabling us to calculate accurate headway times. As node capacity plays an important part in overall railway network capacity, nodes are modelled on a microscopic level as a separate station graph. Here, vertices represent all relevant points for railway operations, such as signals, switches, stop positions, whilst arcs represent the track sections connecting these points. On this microscopic station graph, possible routes in the node are determined to compute train paths through the node, considering all possible conflicts that can arise. Whilst the mesoscopic network including the timetable is available as a dataset for the 'Deutschlandtakt' [2], the microscopic station graphs are derived from OpenStreetMap data. More on this approach can be found in [11].
- **Passenger rail timetable:** We consider a conflict-free passenger rail timetable. The 'Deutschlandtakt' dataset provides us with information about the stations a train passes through, as well as its arrival and departure times at each station. Using this data, passenger train paths are computed.
- **Rail freight demand:** Represented by generating an origin-destination-matrix containing the number of freight demand from one station to another (e.g. in trains). It is estimated using statistical data and calibrated using data from rail traffic counts.

3.1.2 Timetabling module

The first step is to identify the routes freight trains will take through the railway network. Therefore, we first compute the theoretical capacity of the macroscopic network. For each arc, we determine the occupancy limit, i.e. how many additional freight trains could be scheduled by means of a heuristic. These capacity limits and the origin-destination matrix are the input for solving the multicommodity flow problem, using a branch-and-price approach. The resulting flows are then split into smaller segments on which uninterrupted train paths, called slots, are constructed. Splitting points are selected manually based on

criteria including availability of siding tracks for freight trains, allowing stops for breaks and connecting different slots to freight paths. Slots are composed of so-called speed profiles, describing usage and occupancy of the railway infrastructure elements whilst considering train dynamics. Within nodes, the slot creation process involves selecting appropriate routes on the microscopic station graph whilst considering operational constraints and potential conflicts. The slots subsequently serve as building blocks for the train path allocation model, which links up slots to form train paths through the network so that demand, safety regulations, and service quality standards are satisfied.

Speed profile generation: Speed profiles are calculated for different train categories, allowing consideration of different train dynamics as well as signalling and safety systems. Minimum headway times are computed between speed profiles sharing infrastructure, while distinguishing between stations and the open line. In railway nodes, speed profiles are computed on suitable paths within the node linking entries, exits and stop positions.

Slot construction and selection: A slot consists of speed profiles v_i with their respective start times t_i, forming the basis for computing freight train paths. The number of slots required per segment is based on multi-commodity-flow problem result. Let R_s be a set of freight train requests on a segment, defined by an origin o^r, a destination d^r and the computed (minimum) demand $d_{r,\min}$. For each request r, the aim is to compute a set of conflict-free slots S_r fulfilling the demand and fitting to the given passenger timetable. Slot quality is defined as $q_s = 1 - \frac{d_s}{d_s^*}$ where d_s is the planned travel time from origin to destination and d_s^* the minimum possible travel time. Thus, we ensure slots with shorter total duration receive higher quality scores than slots with long waiting times.

As different segments often share tracks, we define segment groups for which slots are computed jointly to ensure conflict-free usage across segments. For robustness, we add buffer times to the calculated headways between speed profiles, beginning with a target buffer time. If no feasible slots can be found, we reduce the buffer time step-wise down to the minimum headway time. Slots are initially generated using a heuristic algorithm, after which the solution is improved through column generation. The Slot Selection Problem maximises total slot quality while satisfying minimum demand $d_{r,\min}$ for each request$r \in R$. Conflict-free selection is guaranteed through constraints based on maximal cliques $C \in \mathscr{C}$, computed from pairwise slot conflicts. These clique constraints ensure that at most one slot per clique is selected, as all slots within a clique are in conflict with each other. The optimisation uses column generation with LP relaxation solved by Gurobi.

Slot allocation: With suitable slots computed, we allocate them to requests. The goal is to provide high-quality train paths for as many requests as possible, modelled as a path-based MILP. The model considers conflicts between paths, stop position capacity limitations, and robustness considerations based on delay propagation analysis. More on the slot allocation model can be found in [10]

3.1.3 Evaluation module

The evaluation module assesses the robustness of the computed slots and timetables using stochastic delay propagation through an event-driven (max,+)-based framework [9]. Train runs are modelled as a sequence of departure and arrival events related by running and dwelling activities. Stochastic delays are applied, and their effects computed using (max,+)-algebra to determine delay probabilities for each network event. This provides feedback to adjust slot selection, buffer times, or eliminate problematic slots. More on the evaluation module can be found in the corresponding publications [10, 4].

3.2 Incorporating ETCS L2 into the model

As the capacity assessment model described in subsection 3.1 was primarily developed for conventional block signalling, some adjustments to include ETCS L2 needed to be made. These adjustments primarily affect the infrastructure input, but also the speed profile generation and blocking time calculation. As the focus of this contribution is the evaluation of the network capacity effects of ETCS Level 2 according to Germany's National Implementation plan, we focused our modelling on the planned implementation of ETCS L2 with line side signalling (ETCS-L2mS) and without (ETCS-L2oS) in Germany. Given our mesoscopic modelling approach, we distinguish between mesoscopic railway network representation and the microscopic node network modelling:

Speed profile generation: As the speed profile generation also includes blocking time calculation, their creation process needs to be adjusted. Whilst the computation of running time and clearing time remains unchanged, the approach time is computed over the train's individual braking distance instead of the distance between the distant and the block signal. The individual braking distance is derived using train category-specific deceleration values and gradient information from OpenStreetMap.

Mesoscopic network modelling: Legacy block sections are divided into sub-blocks, ensuring compatibility for mixed-equipment scenarios. Their target length is set a variable parameter. For a given legacy block of length l and a target sub-block length l_{cab}, the number of sub-blocks is calculated as $k = l/l_{\text{cab}}$ and then rounded to the nearest integer, thus distributing the legacy block length evenly among k sub-blocks, with actual sub-block lengths adjusted accordingly to match the legacy block length exactly. Afterwards, speed profiles are computed for both signalling types, with ETCS profiles only used in the slot selection process when both infrastructure and train category are equipped accordingly.

Microscopic station modelling: For the microscopic station graph, the complex ETCS planning rules pose a greater challenge. For the capacity assessment, the location of balise groups and block markers needs to be approximated. This is due to the facts that balise group locations are important for release speed, overlap and flank protection, thus directly affecting headways and capacity. However, in practice, their location is restricted by infrastructure, e.g. switch areas and electrical section boundaries. Also, equipping a station with both ETCS and the legacy PZB system poses its own challenges: overlap lengths cannot be reduced and ETCS location markers cannot be placed within the overlap of the legacy signals [5]. To address these complications, some automatic planning approaches are proposed for infrastructure planning, e.g. [6]. However, it remains to be seen how these can be applied in practice. For our capacity assessment, we currently assume that existing signal locations are transformed into bidirectional ETCS location markers. Furthermore, the overlap calculation is changed, reducing their length and their release time for the station speed profile generation and headway calculation, reflecting ETCS braking capabilities. This captures some, but not all, capacity benefits of ETCS in stations. Further refinements to station speed profile generation and automated ETCS block subdivision within station areas are subject to ongoing research.

Rolling stock considerations: As laid out in Section 1, transitioning the railway network to ETCS L2 requires that not only infrastructure is upgraded, but also that railway vehicles are equipped with ETCS On-board-units. The NIP identified that in 2023, about 92 % of all railway vehicles in Germany lacked ETCS on-board equipment. For freight rail specifically, of around 5,200 vehicles, only 700 are equipped with ETCS [7]. The evaluation must therefore account for scenarios where infrastructure is upgraded whilst vehicles are not. This is modelled by introducing freight train categories with and without ETCS equipment and varying their share across scenarios.

4 Case Study

4.1 Scenario description

The test case network is located in Northern Germany, spanning between the cities of Hamburg, Bremen and Hanover, specifically between the freight yards of Maschen (south of Hamburg), Bremen and Seelze (west of Hanover), shown in Figure 2. It represents one of the first sections the NIP aims to convert to ETCS L2oS by 2028: Lines 1280 from Jesteburg (south of Maschen) to Buchholz and lines 2200 and 1283 between Buchholz and Bremen will be converted with legacy systems decommissioned. Thus, all trains running on these sections have to be equipped with ETCS. We compare these two scenarios:

- **Scenario 1:** Status quo with the legacy PZB and LZB system in place
- **Scenario 2:** ETCS L2oS on lines 1280 (Jesteburg – Buchholz) and 2200/1283 (Buchholz – Bremen), target ETCS block length: 300 meters

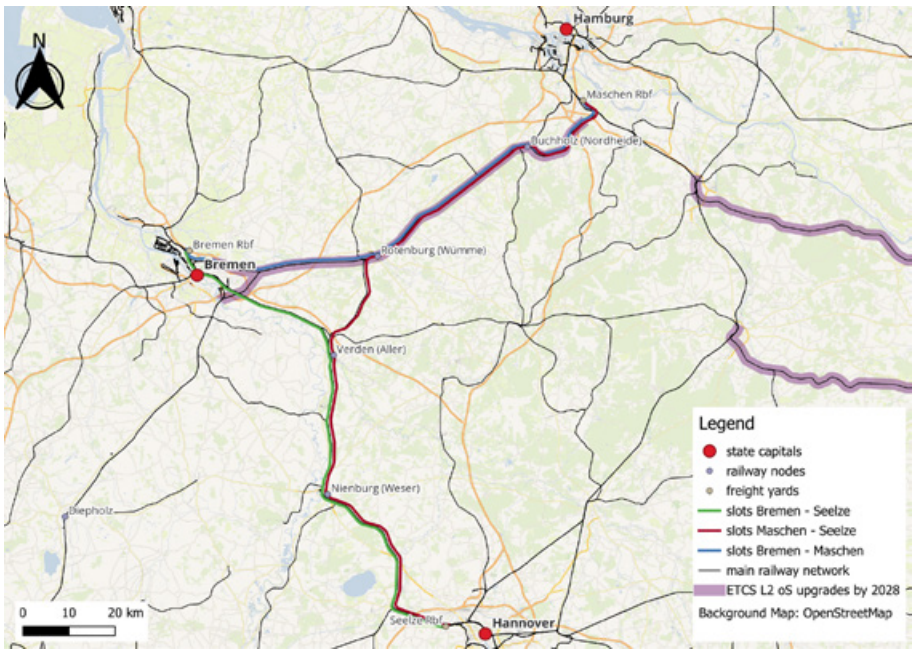


Figure 2: Test case network with ETCS implementation by 2028 [7]

The goal is to evaluate the capacity benefits in Scenario 2 for the origin-destination pairs Maschen – Seelze and Maschen – Bremen. With this focus, we manually select routes using the upgraded corridors. To capture network effects, we also include a route sharing track segments with the primary route between Verden and Seelze, resulting in the following requests with specified paths:

1. Bremen Rbf – Seelze Rbf (117 km)
2. Maschen Rbf – Seelze Rbf (167 km)
3. Maschen Rbf – Bremen Rbf (106 km)

These share infrastructure as follows: Maschen – Seelze and Maschen – Bremen share a route between Maschen Rbf and Rotenburg, using the double-tracked line 1280 towards Buchholz and the triple-track line 2200/1283 from there to Rotenburg, where they diverge. Request Maschen – Bremen continues towards Abzw. Utbremen, then leaves the main line via a single-track access line to reach Bremen Rbf. Request Maschen – Seelze continues from Rotenburg via a partially single-tracked line to Verden, where it meets the double-tracked line from Bremen. In Verden, requests from Bremen and Maschen converge and continue to Seelze on a double-tracked line.

Each request is assigned an identical minimum demand. Subsequently, conflict-free slots for the requests are computed for a two-hour time window, conflict-free to each other and the passenger timetable. The analysis proceeds in two phases: Initially, only request 2 (the longest route) is scheduled to determine whether the number of slots increases with ETCS upgrades. Subsequently, all three requests are scheduled simultaneously. In scenario 2, requests 2 and 3 require ETCS equipment, whilst request 1 operates exclusively on lines not equipped with ETCS. We assume all passenger trains are ETCS-equipped and compare the change in slots between the Scenarios 1 and 2. Train parameters used for calculating braking curves are given in Table 1, with standard freight trains (catTU) and ETCS-equipped variants (catTU_Cab). With the focus on strategic planning and the operating speed on the network not exceeding 200 km/h, a simplified conversion model is used. The evaluation module of the capacity assessment approach is not used, as the focus lies on the maximum number of slots that can be scheduled.

Table 1: Train categories used in the case study

ID	Description	v_{max} [m/s]	a_{acc} [m/s²]	a_{brake} [m/s²]	Length [m]	Brake setup [s]	ETCS
catA	FV 300 km/h	83.33	0.5	-1.0	400	0	yes
catB	FV 249 km/h	69.1	0.5	-1.0	400	0	yes
catD	FV 230 km/h	63.89	0.5	-1.0	400	0	yes
catF	FR 200 km/h	55.56	0.5	-1.0	250	0	yes
catH	FR 160 km/h light	44.44	0.8	-1.0	150	0	yes
catN	Local train	38.89	1.2	-1.0	200	0	yes
catTU	Standard freight	27.77	0.15	-0.4	700	45	no
catTU_Cab	Freight with ETCS	27.77	0.15	-0.4	700	45	yes

4.2 Results and Analysis

Preliminary findings indicated ETCS capacity gains do not always manifest, but depend on infrastructure characteristics, line speed and timetable constraints. The results demonstrate that they are also dependent on network configuration and demand patterns.

Table 2: Number of freight train paths for the scenarios

	Freight requests	# of slots (Scenario 1)	# of slots (Scenario 2)
Single-Route analysis	Maschen Rbf – Seelze Rbf	5	7
	Seelze Rbf – Maschen Rbf	5	6
Multi-Route analysis	Bremen Rbf – Seelze Rbf	5	5
	Seelze Rbf – Bremen Rbf	5	5
	Maschen Rbf – Seelze Rbf	2	3
	Seelze Rbf – Maschen Rbf	3	3
	Maschen Rbf – Bremen Rbf	4	4
	Bremen Rbf – Maschen Rbf	4	4

Single Route Analysis: To isolate the effects of ETCS from network-wide scheduling constraints, the first analysis focussed solely on the Maschen Rbf – Seelze Rbf route (167 km). This single-route scenario demonstrates ETCS L2 benefits, with Scenario 2 enabling 3 additional freight train paths compared to Scenario 1.

Multi-Route Analysis: When computing all three requests simultaneously, the comparison between Scenario 1 and Scenario 2 shows a minor improvement in the number of scheduled freight train paths: for the request Maschen Rbf – Seelze Rbf, an additional slot can be scheduled. This translates into 14 slots on the upgraded corridor between Rotenburg and Jesteburg, shared by the requests Maschen Rbf – Seelze Rbf and Maschen Rbf – Bremen Rbf. On the legacy corridor (Bremen –) Verden – Seelze, the total number of slots increases from 15 to 16 slots for both directions.

Therefore, the ETCS upgrades in Scenario 2 yield a small capacity benefit. This is further shown in the train graphs presented in Figures 4 (for Scenario 1) and 5 (for Scenario 2). Coloured in shades of green and blue are the freight train paths, coloured in shades of orange and red are the passenger trains. Comparing the two train graphs, the benefits of ETCS L2 can be seen: reduced headway times that enable tighter train path bundling. The additional slot for the request Maschen Rbf – Seelze Rbf is achieved by bundling this slot (departing around 00:25) with a passenger train and a slot for the request Maschen Rbf – Bremen Rbf. This enables scheduling the slot on the line Verden – Seelze, where it

meets traffic, passenger and freight, coming from Bremen. However, Figure 5 also shows unused gaps in the train graph on the Rotenburg–Buchholz section (e.g., around 00:18 from Maschen Rbf towards Rotenburg), yet these do not translate into additional slots for Maschen-originating requests. For the request Maschen Rbf – Bremen Rbf, the limiting factor is therefore not the main line between Rotenburg and Maschen. Given the network topology, the limiting factor may be the single-track acces line linking Bremen Rbf to the main line. This assumption is confirmed by the blocking time stairways for this request (see Figure 3): on the left-hand side, the single-track line used by the request is visible, with no gaps identifiable for the two-hour time-window. The Maschen – Seelze requests on the other hand are constrained both by the schedule on the Rotenburg – Buchholz section as well as the schedule on the Verden – Seelze section.

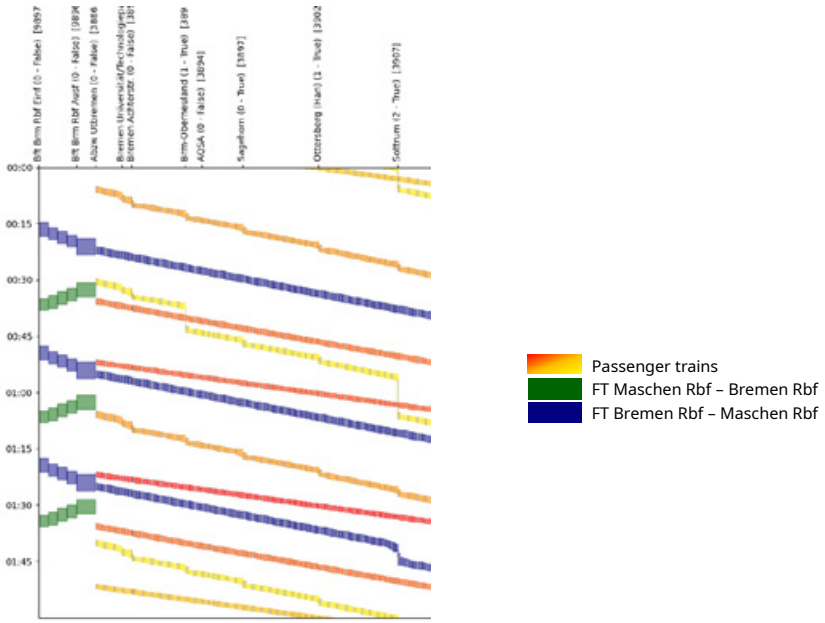


Figure 3: Block time stairway excerpt for the approach to Bremen Rbf, scenario 2

Analysis of Results: Comparing results between multi-route and single-route scenarios shows the interaction between ETCS benefits and network topology. In the single-route analysis, ETCS L2 implementation yielded 3 additional freight paths for the request Maschen – Seelze corridor by reducing headways on the upgraded segment. However, in the multi-route analysis, this benefit diminishes to just 1 additional path. These findings demonstrate that while ETCS successfully reduces headways on upgraded corridors, the realized capacity gains depend on whether downstream bottlenecks or converging traffic flows limit the network.

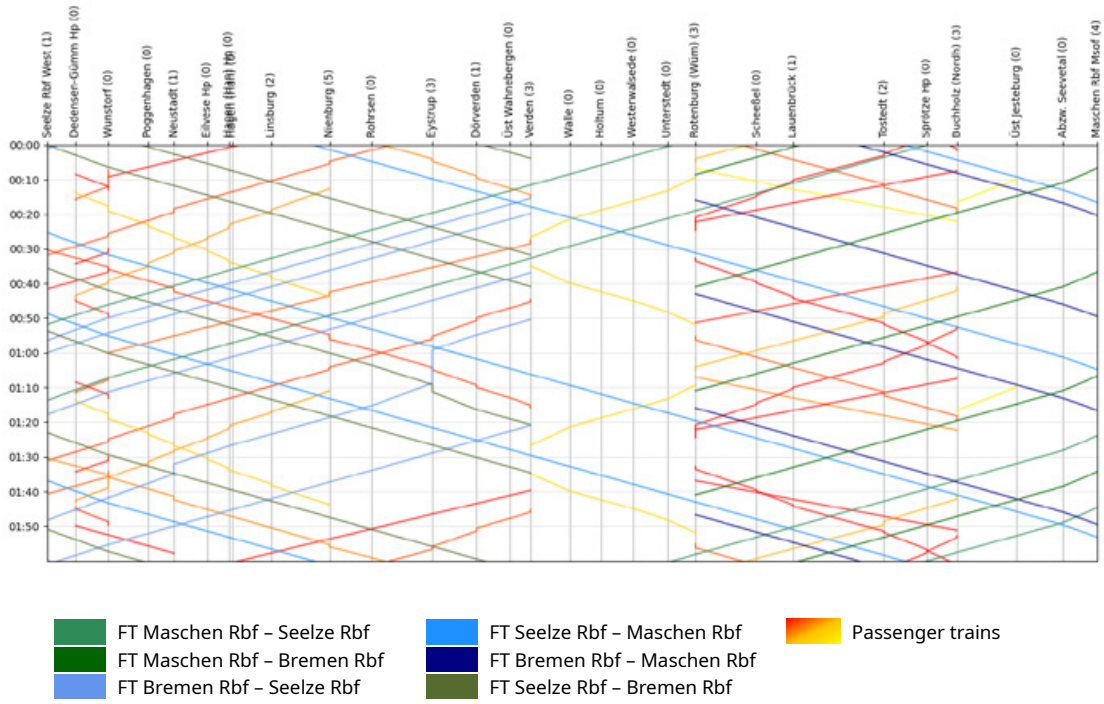


Figure 4: Train Graph Seelze – Maschen for Scenario 1, multi-route analysis

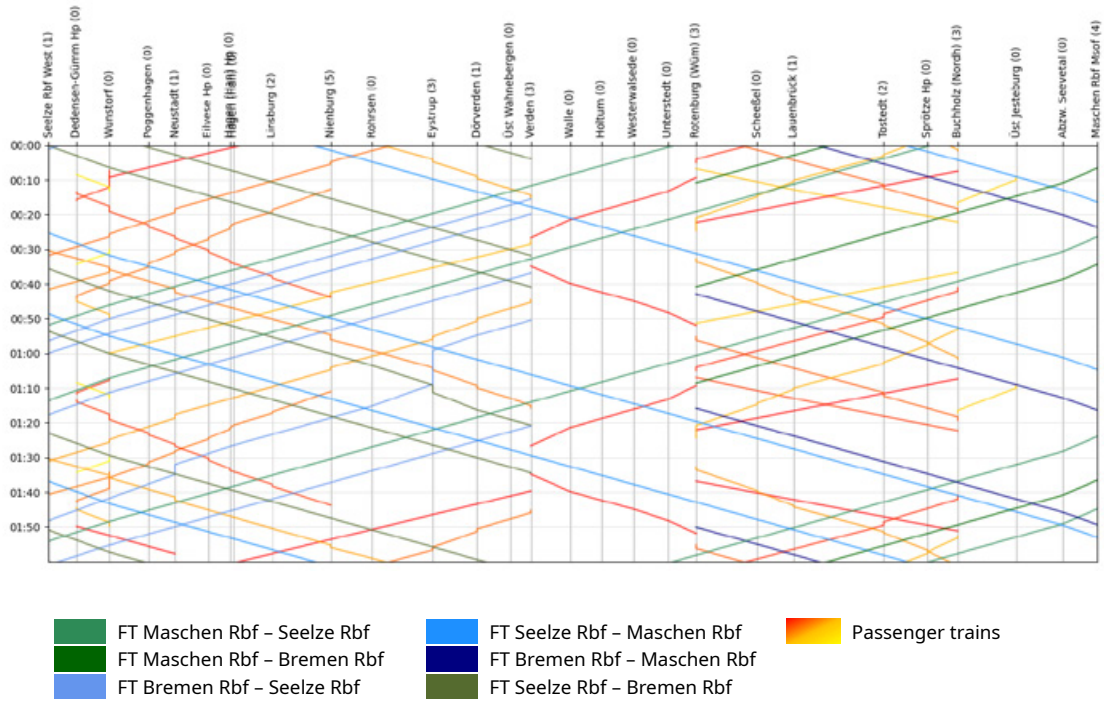


Figure 5: Train Graph Seelze – Maschen for Scenario 2, multi-route analysis

5 Conclusion and Future Work

This research introduces an approach towards evaluating ETCS L2 implementation effects on network-wide residual freight capacity across different phases of an implementation strategy. The combined meso- and microscopic approach allows for modelling of ETCS L2 close to a real-world implementation and evaluate capacity effects, considering both infrastructure and rolling stock constraints. The case study reveals insights about ETCS capacity benefits in realistic network scenarios. Whilst ETCS Level 2 yields a measurable improvement in the isolated corridor analysis, the multi-route scenario yields much smaller improvements due to network-level scheduling constraints and infrastructure bottlenecks. This emphasises the importance of coordinated network-wide approaches rather than isolated corridors when focussing on the capacity gains of ETCS.

However, the approach still has its limitations. The rolling stock categorisation, whilst comprehensive, may not fully represent the diversity of freight train characteristics and requires further sensitivity analysis. For our ETCS modelling, the approach using target sub-block lengths instead of optimised block lengths is simpler, but may underestimate capacity. For stations, the current modelling approach of transforming existing signal locations into ETCS location markers underestimates capacity gains in nodes. Both assumptions require sensitivity analysis in future studies. Future work will focus on modelling ETCS in railway nodes and incorporation into the network model, as already done for conventional signalling, so that capacity gains in railway nodes also may be quantified. Additionally, investigating different implementation sequences and infrastructure bottleneck mitigation strategies will provide insights for maximising ETCS capacity benefits.

Acknowledgment

This research is partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant number 512633826.

Literature

[1] *UIC Code 406 - Capacity*. 2013. – ISBN 978–2–7461–2158–4

[2] *Daten des Zielfahrplans Deutschlandtakt im Format railML TT*. Mobilithek.

Bundesministerium für Verkehr (BMV). https://mobilithek.info/offers/573181445239877632. Version: Mar 2023

[3] Comelli, S. ; Weik, N. ; Kekes, C. ; Theurich, F. ; Nachtigall, K.: Fahrplanbasierte Ermittlung der Kapazität von Eisenbahnnetzwerken unter Berücksichtigung der Betriebsqualität. In: *Eisenbahntechnische Rundschau* (2024), Nr. 6

[4] Comelli, S. ; Weik, N. ; Theurich, F. ; Kekes, C. ; Nachtigall, K.: Analysis of Approximation Schemes for Stochastic Delay Propagation in Railway Systems. In: *RailDresden 2025*, 2025

[5] Denissen, J. ; Flieger, M. ; Kümmling, M. ; Küpper, M. ; Wanstrath, S.: Optimierung der Blockteilung mit ETCS Level 2 im Digitalen Knoten Stuttgart. In: *SIGNAL + DRAHT* 113 (2021), August, Nr. 7+8

[6] Dillmann, S. ; Hähnle, R.: Automated Planning of ETCS Tracks. Version: 2019. doi.org/10.1007/978-3-030-18744-6_5. In: *Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification*. Springer International Publishing, 2019, 79–90

[7] Eisenbahn-Bundesamt: *Nationaler Umsetzungsplan ETCS*. 2024

[8] Goverde, R.M.P.: Railway timetable stability analysis using max-plus system theory. In: *Transportation Research Part B: Methodological* (2007), S. 179–201

[9] Goverde, R.M.P. ; Corman, F. ; D'Ariano, A.: Railway line capacity consumption of different railway signalling systems under scheduled and disturbed conditions. In: *Journal of Rail Transport Planning & Management* (2013), S. 78–94

[10] Kekes, C. ; Comelli, S. ; Theurich, F. ; Nachtigall, K. ; Weik, N.: Robust freight train scheduling by allocating pre-constructed slots. In: *Operations Research Proceedings 2024: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR)*, 2025

[11] Kekes, C. ; Theurich, F. ; Nachtigall, K. ; Comelli, S. ; Weik, N.: Railway nodes in an integrated timetable-based network capacity assessment model. In: *RailDresden 2025*, 2025

[12] Knutsen, D. ; Olsson, N.O.E. ; Fu, J.: Capacity evaluation of ERTMS/ETCS hybrid level 3 using simulation methods. In: *Journal of Rail Transport Planning & Management* (2024), 100444. doi.org/10.1016/j.jrtpm.2024.100444

[13] Mikulčić, M. ; Mlinarić, T. J.: Railway Capacity Enhancement with Modern Signalling Systems – A Literature Review. In: *Promet - Traffic & Transportation* 33 (2021), Februar, Nr. 1, 141–152. doi.org/10.7307/ptt.v33i1.3664

[14] NACHTIGALL, K.: Modelling and Solving a Train Path Assignment Model with Traffic Day Restriction. In: *Operations Research Proceedings 2015*, Springer International Publishing, 2017, 621–627

[15] SZYMULA, C. ; BEŠINOVIĆ, N. ; NACHTIGALL, K.: Quantifying periodic railway network capacity using petri nets and macroscopic fundamental diagram. In: *Transportation Research Part C: Emerging Technologies* 158 (2024), Januar, 104436. doi.org/10.1016/j.trc.2023.104436

[16] UZGIDIM CANSU, Ömür ; ATIEH, Kianinejadoshah ; STEFANO, Ricci: Influence of Signalling Systems on the Capacity of Railways by Lines and Nodes Assessment Methods. In: *Transportation Research Procedia* 69 (2023), 321–327. doi.org/10.1016/j.trpro.2023.02.178

[17] WAGNER, A.: *Influence of train and traffic control on railway station capacity*, TU Wien, Diss., 2024. doi.org/10.34726/HSS.2024.123221

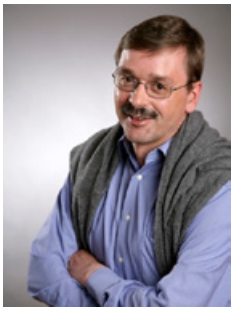
Authors



Kekes, Cédric
Dipl.-Ing.
Research assistant
Chair of Traffic Flow Sciences
TU Dresden



Theurich, Franziska
Dr. rer. nat
Postdoctoral research assistant
Chair of Traffic Flow Sciences
TU Dresden



Nachtigall, Karl
Prof. Dr. rer. nat.
Head of the Chair of Traffic Flow Sciences
TU Dresden

Session 2
Zheng Ning, Egidio Quaglietta, Mahnam Saeednia

Railway Infrastructure Capacity Evaluation of Modular Rail Pods under ETCS Level 2 Signalling System

Ning, Zheng[1], Quaglietta, Egidio[1] and Saeednia, Mahnam[1]

[1]Department of Transport and Planning, Delft University of Technology, Delft, the Netherlands

Abstract

Rail Pods are an emerging concept of modular self-propelled rail vehicle which can interchangeably move freight and transport for a more customer-oriented rail service. Pods are envisaged to operate on-demand with the possibility of forming platoons by either physically or virtually coupling at stations. In such a context, it is essential to quantify the actual service capacity of rail pod platooning, taking into account heterogeneous convoy structures and infrastructure constraints such as signalling rules and block occupation. This study extends UIC Code 406/blocking time theory to applied to Pod platoons, proposing a novel optimization model that integrates traction, cruising, and braking speed profiles alongside safe separation constraints. This approach enables coordinated optimization of cruising speeds across various platoon structures to minimize track capacity consumption. The model is applied to a case study considering the ETCS Level 2 signalling system. The results obtained for such a case study illustrate the ability of the proposed model to identify operational speed and composition of rail pods' platoons which lead to capacity effective use of the existing infrastructure. The proposed method provides potentials for a more flexible allocation of modular rail cars based on demand configuration.

Keywords: Capacity Evaluation; Pods4rail; Modular rail platoons; Optimized speed configuration; ETCS Level 2 signalling system

Session 2
Zheng Ning, Egidio Quaglietta, Mahnam Saeednia

1 Introduction

Delivering customer-oriented rail services requires greater system flexibility and seamless operations, prompting the development of innovative operational paradigms. Pods4Rail, an EU-funded project developing a next-generation modular and smart railway system, introduces autonomously driven, reconfigurable self-propelled rail cars (Pods) with freight or passenger capability. Operating on existing infrastructure, Pods are designed to function either independently or dynamically form platoons via coupling or decoupling mechanisms. This dynamic platooning capability enables the adjustment of their structure to meet variable transport demands, thereby enhancing overall system efficiency through increased flexibility (for more details, see Pods4Rail, 2024[7]).

Pods operations are subject to a dual constraint system. At the transport service level, arrival and departure times at key nodes must strictly adhere to predefined time windows. At the physical operational level, constraints include both inherent infrastructure parameters (e.g., maximum platoon length, speed and acceleration thresholds) and dynamic safety spacing requirements. These constraints ensure both the punctuality of transport services and the safety and stability of system operations.

Capacity evaluation is a key challenge when introducing novel transport modes compatible with existing infrastructure. UIC 406 analytical method is the most widely used capacity model providing solid methodological and comparative work on static railway capacity, using headway/braking curve formulas and UIC code validation based on given timetable [3, 9]. Non-timetabled operational concepts have also been proposed to enhance flexibility, executing dynamic decision-making through techniques like Reinforcement Learning [2].While the underlying capacity evaluation principles of these approaches align well with the operational logic of Pods, a critical limitation lies in the implicit assumption that all trains are homogeneous—characterized by standardized lengths, masses, and performance profiles. This assumption becomes inadequate in the context of Pod systems, which are characterized by variable platoon sizes leading to dynamic headways. As a result, conventional capacity indicators—such as the number of trains per hour—fail to accurately reflect the true utilization or throughput of such flexible and modular operations.

Rail transit is shifting from fixed consists to flexible platooning, driven by key technologies like Virtual Coupling (VC) [1, 8] and digital automatic coupling [6]. These technolo-

gies provide the conditions for Pods platoon operation. While the Pods4Rail project has explored conceptual architecture and demand-driven scheduling for Pods systems [4, 5], significant gaps remain in capacity performance analysis. Especially, parameter variations (e.g., platoon length changes), heterogeneous platoons, and how these unique Pod features integrate with existing infrastructure and ensure the safe separation.

This study addresses critical knowledge gaps by analyzing how platoon structures, and coordinated speed interactions affect capacity consumption under ETCS Level 2 system. It incorporates operational constraints related to signaling, train separation, speed transitions, and heterogeneous platoon characteristics, and establishes essential foundations for implementing Pods systems on existing infrastructure and provides essential inputs for scheduling studies. The paper is organized as follows. Section 2 presents the problem statement, introducing the concept of Pods, Pod platoons and their structure. Section 3 extends blocking time theory and the UIC 406 capacity model to develop a capacity occupancy model for Pod platoons. Based on this, a nonlinear optimization model is designed to determine the optimal coordinated cruising speeds for platoons, minimizing infrastructure occupancy while maintaining safety margins. Simulation results are presented in Section 4, and Section 5 concludes the paper.

2 Problem Statement

The modular transport unit system, referred to as Pods, consists of two core components (as shown in Figure 1): Transportation Units (TUs) and Carriers. TUs serve as modular, demand-responsive loading container units that can be autonomously decoupled from and reattached to Carriers. This design enables the seamless transfer of loaded TUs across different transport modes, including rail and road, thereby facilitating efficient intermodal operations. Carriers act as mobile service terminals, with their dispatching strategies adhering to the spatio-temporal pickup and delivery requirements of the TUs. Upon assembly at designated operating stations, TUs and a Carrier form a complete Pod.

As illustrated in Figure 2, rail corridors equipped with the Pods system must accommodate multiple heterogeneous platoons, each capable of dynamically adjusting its length through coupling and decoupling operations. A platoon traverses an ordered sequence of fixed block sections, where each block is released only after the physical tail of the platoon clears the section.

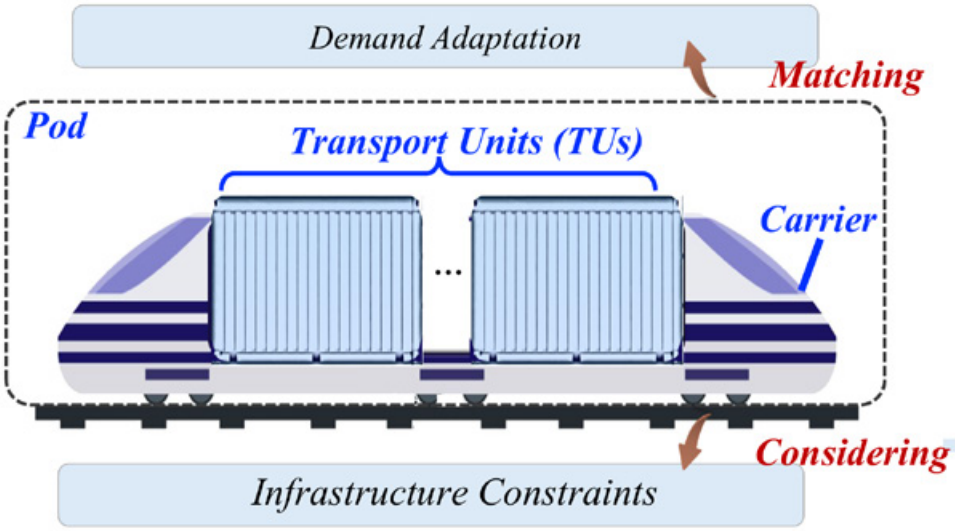


Figure 1: Pods architecture with Transportation Units (TUs) and Carriers

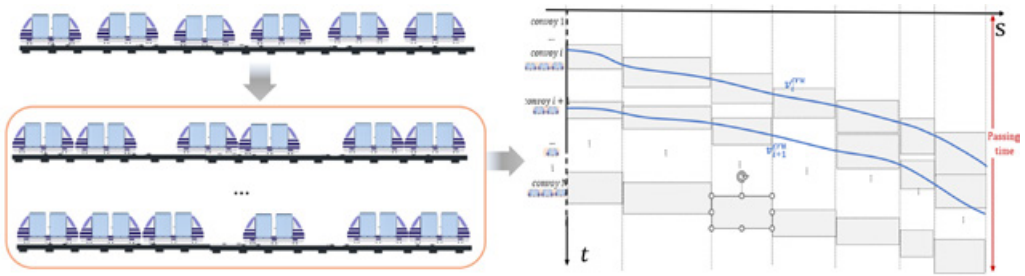


Figure 2: Block Occupancy Diagram for Heterogeneous Pod Platoon Traffic

Let $\mathscr{P} = \{P_1, P_2, \ldots, P_N\}$ denote the set of rail pod platoons involved in a demand request. This system $P_i \in \mathscr{P}$ is characterized by the following features:

- **Formation Type** $\mathscr{P}$: Defines the platoon structure, such as a single Pod, a short platoon, or a long platoon, which determines platoon operation attribute.
- **Speed Profile Conditions**: The motion of each platoon must comply with a predefined traction–cruise–braking trajectory. For a given platoon P_i, selecting an optimal cruising speed v_i is essential to minimize block occupancy across multiple blocks, even if theoretical optimal speeds might vary per block.
- **Platoon traffic Interaction**: Due to heterogeneous platoon characteristics, the optimal performance of heterogeneous Pod Platoons is interdependent, necessitating the calculation of coordinated cruising speeds.

Therefore, evaluating the performance of a given Pod Platoons Traffic requires minimizing infrastructure capacity consumption by coordinately optimizing the cruising speeds of all constituent platoons.

3 Model Formulation

3.1 Notions and Assumptions

The notations used in this paper are formally defined in Table1

The proposed method modeled under these fundamental assumptions, aligned with ETCS Level 2 signaling system operational specifications:

- Constant acceleration/deceleration profiles during traction (a_{tra}) and braking (a_{bra}) phases.
- Static platoon composition maintaining fixed pod count (n) throughout operational cycles
- All single pods are identical in terms of physical dimensions and mass. That is, each pod has the same length, weight, and performance characteristics.

3.2 Infrastructure occupation model

The blocking-time theory quantifies infrastructure occupation by defining the total duration a train renders a block section unusable goverde2013railway, bevsinovic2017microscopic, wang2020evaluating. However, this framework assumes fixed train lengths and a dispatching pattern, which limits its applicability to dynamically formed and reconfigurable pod platoons. Based on this, we extend the basic model to account for platoon composition with variable lengths and internal spacing. Fig. 3 illustrates the spatio-temporal blocking profiles for a pod platoon operating under two typical scenarios: (a) cruising and (b) station approach and departure. In each subfigure, the top part presents the speed profile across segmented infrastructure blocks, while the bottom part visualizes the corresponding block occupation timeline.

For cruising segments (Fig. 3 (a)), the block occupation time for a pod platoon is gov-

Table 1: Symbol Definitions with Classification

Symbol	Definition	Unit
Parameters (System Constants)		
L_{tlen}	Single Pod vehicle length	m
S_{vc}	Virtual coupling safety spacing	m
a_{bra}	Maximum braking deceleration	m/s²
a_{tra}	Maximum traction acceleration	m/s²
v_{sat}	Station speed limit	m/s
L_{sec}	Track section signaling length	m
L_{pla}	Station approach protection zone length	m
L_{over}	Platform throat overlap length	m
L_j^{total}	Cumulative position before block j (sum of preceding blocks)	m
L_{sat}	Station cumulative position k	m
T_r	Route establishment time	s
T_{free}	Block release delay time	s
T_D	Mandatory dwell time at platform	s
n_i	Number of Pods in platoon i	-
T_D	Mandatory dwell time at platform	s
t_{couple}	Setting time for forming platoon	s
α_j	Block type indicator (0:station,1:track)	-
Intermediate Variables (Calculated)		
$T^{\text{sat}}(n)$	Station blocking time for n-Pod platoon	s
T_{ap}	Approach braking time ($v_i^{\text{cru}}/a_{\text{bra}}$)	s
T_{pass}	Block traversal time	s
T_{sm}	Safety margin time behind last Pod	s
T_{tlen}	Train length passage time	s
T_{over}	Station throat clearance time	s
$v_{i,j}^{\text{in}}$	Platoon i entry speed at block j	m/s
$v_{i,j}^{\text{out}}$	Platoon i exit speed at block j	m/s
L_{platoon}	Total platoon length ($nL_{\text{tlen}} + (n-1)S_{vc}$)	m
d_{bra}	Braking distance ($(v_i^{\text{cru}})^2/(2a_{\text{bra}})$)	m
$t_{N,M}$	Total block occupation time for last platoon	s
$T_{i,j}^{\text{pass}}$	Platoon i passing time at block j	s
Decision Variables (Optimizable)		
$h_{\text{min}}^{(i-1,j)}$	Minimum headway between platoon $i-1$ and i at block j	s
v_i^{cru}	Optimized cruising speed of platoon i	m/s

erned by:

$$T^{\text{line}}(n) = T_r + \frac{v_i^{\text{cru}}}{a_{\text{bra}}} + \frac{L_{\text{sec}} + nL_{\text{lenP}} + (n-1)S_{intra} + L_{\text{sm}}}{v_i^{\text{cru}}} + T_{\text{clear}} \tag{1}$$

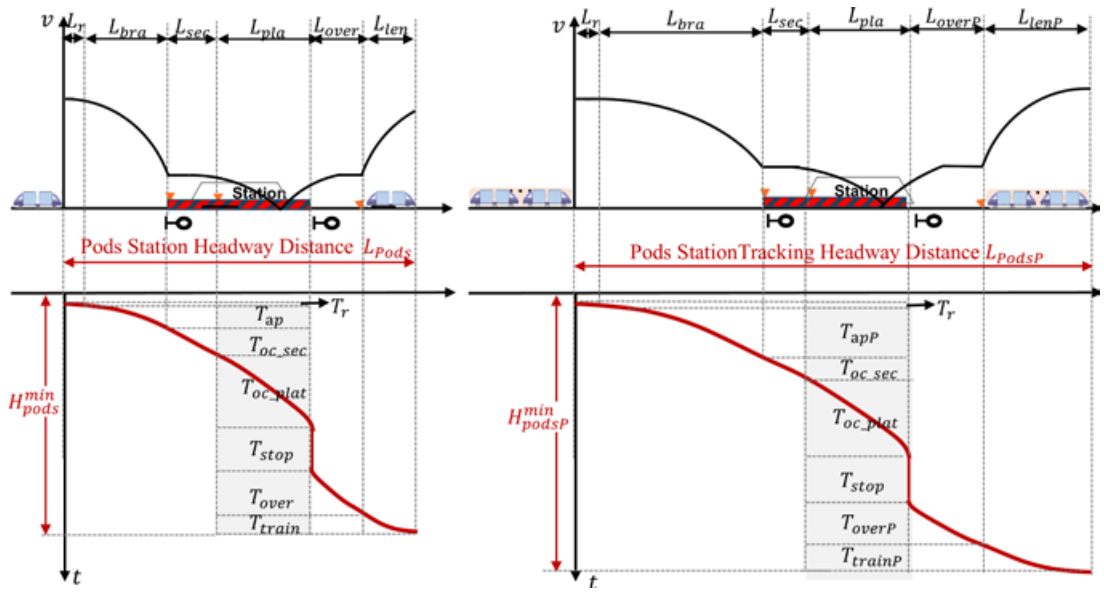


Figure 3: Block Occupancy Diagram for Heterogeneous pod Platoon Traffic

Here, nL_{lenP} represents the total physical length of the pod platoon, while $(n-1)S_{intra}$ denotes the safety separations between adjacent pods in the platoon. In contrast, other components, such as reaction time T_r, approaching time, and release delay, are governed by the real-time coordination system and are considered independent of the platoon size.

This formulation indicates that the primary length-dependent terms are nL_{lenP} and $(n-1)S_{intra}$, which directly prolong the time required for complete clearance of a block. Accordingly, the model captures how larger platoon formations lead to extended occupation, critical for capacity estimation and control design.

For station segments (Fig. 3 (b)), additional components must be considered, including dwell time and constrained acceleration profiles. The blocking time of a pods platoon is:

$$T^{\text{sat}}(n) = T_r + \frac{v_i^{\text{cru}}}{a_{bra}} + \frac{L_{pla} - \frac{(v_i^{\text{cru}})^2}{2a_{bra}}}{v_i^{\text{cru}}} + T_D + [T_{lapP}(n) + T_{tlenP}(n)] + T_{rel} \tag{2}$$

When generalized to an n-pod platoon, the time required to fully clear the station throat increases proportionally to the platoon length. To ensure safe operations and avoid conflicts, the station overlap length must be correspondingly extended. Given the allowable station speed and the overlap length, the rear-end passing time of the pod platoon through

the station and overlap sections is computed as follows:

$$T_{\text{lapP}}(n) + T_{\text{lenP}}(n) = \begin{cases} \sqrt{\dfrac{2\left(nL_{\text{len}} + (n-1)S_{\text{intra}} + L_{\text{over}}\right)}{a_{\text{tra}}}}, & v_i^{\text{cru}} \geq Vcrit \\ \dfrac{2a_{\text{tra}}\left(nL_{\text{len}} + (n-1)S_{\text{intra}} + L_{\text{over}}\right) + \left(v_i^{\text{cru}}\right)^2}{2a_{\text{tra}}v_i^{\text{cru}}}, & v_i^{\text{cru}} < Vcrit \end{cases} \tag{3}$$

where $Vcrit$ denotes the critical velocity defined by:

$$Vcrit = \sqrt{2a_{\text{tra}}\left(nL_{\text{len}} + (n-1)S_{\text{intra}} + L_{\text{over}}\right)}.$$

Compared to cruising segments, station areas exhibit higher sensitivity to platoon length due to the complexity of station throat sections and overlap track areas. Capturing this effect through an extended formulation is crucial to accurately evaluate infrastructure capacity under operational and signalling constraints.

3.3 Capacity Optimization Model

As illustrated in the previous section, the capacity consumed by a given pods' service schedule is measured by the total time the infrastructure is occupied. Thus, a capacity-effective schedule aims to minimize such an infrastructure occupation. Assuming a fixed start time for the rail pods' service, this translates into reducing the time at which the last platoon in a schedule cycle clears the considered infrastructure. Given a section composed of M blocks, the objective is to calculate the capacity consumption of N pod platoons, where each platoon consists of n pods. The objective function is defined as:

$$\min\ T_{\text{pass}} = t_{N,M} - t_{1,1} \tag{4}$$

where, the total passing time $t_{N,M}$ is computed through an iterative process over platoon movements and block transitions.

- For each block j, the earliest time at which platoon i can enter is determined by the minimum headway from the preceding platoon:

$$t_{i,j} = t_{i-1,j} + h_{\text{min}}^{(i-1,j)} \tag{5}$$

- The passing time of a platoon from consecutive blocks is determined as::

$$t_{i,j} = t_{i,j-1} + T^{\text{pass}}_{i,j-1} \tag{6}$$

The passage time $T^{\text{pass}}_{i,j}$ of platoon i through block j is important for evaluating infrastructure capacity, especially under the Pods system with heterogeneous platoons. Modeling the movement of these mixed-structure platoons is essential, accounting for variations in formation length, distinct acceleration/deceleration profiles, and the time cost associated with coupling process.

$$T^{\text{pass}}_{i,j} = \begin{cases} (n_i - 1)t_{\text{couple}} + T_{\text{stop}} + f_{\text{sta}}(l_j), & \text{Station operation} \\ f(v_{\text{in}}, v_{\text{out}}, l_j), & \text{Acceleration/Deceleration} \\ \dfrac{l_j}{v^{\text{cru}}_i}, & \text{Cruising} \\ \dfrac{l_j}{\max(|v_{\text{in}}|, 0.1)}, & \text{Other cases} \end{cases} \tag{7}$$

where

$$v^{\text{out}}_{i,j} = \begin{cases} v_{\text{tar}}, & \text{if } v_{\text{out}} \text{ reached} \\ \sqrt{|v^2_{\text{in}} \pm 2al_j|}, & \text{otherwise} \end{cases}$$

The Pods platoons operation process is governed by the following constraints defined for all trains $i \in \mathscr{I}$ and block segments $j \in \mathscr{J}$:

$$v^{\text{in}}_{i,j+1} = v^{\text{out}}_{i,j}, \qquad \forall i \in \mathscr{I},\ j \in \{1, \ldots, J-1\} \tag{8}$$

$$t_{1,1} = 0 \tag{9}$$

$$V_{\min} \le v^{\text{in}}_{i,j},\ v^{\text{out}}_{i,j} \le V_{\max}, \qquad \forall (i,j) \in \mathscr{I} \times \mathscr{J} \tag{10}$$

$$v^{\text{cru}}_i \ge v_{\text{sat}}, \qquad \forall i \in \mathscr{I} \tag{11}$$

$$h^{(i-1,j)}_{\min} \ge T^{\text{block}}_{i,j}, \qquad \forall i \in \mathscr{I} \setminus \{1\},\ j \in \mathscr{J} \tag{12}$$

Speed continuity requires that platoon i's output speed at block j equals its input speed at block $j+1$, as shown in Eq. (8). Eq. (9) indicates that the lead platoon departs from the first block at time zero, establishing a temporal reference. The speeds are constrained within the operational limits $[V_{\min}, V_{\max}]$, as shown in Eq. (10), and the cruising speed v^{cru}_i must be no less than the station limit v_{sat}, according to Eq. (11).

The minimum time headway $h_{\min}^{(i-1,j)}$ is imposed to maintain safe separation by ensuring that it exceeds the block occupation time of the preceding platoon, as stated in Eq. (12). The term $T_{i-1,j}^{\text{block}}$ represents the blocking time of platoon $i-1$ at block j.

The non-linear coupling between vehicle speed, time consumption, and headway constraints creates nonconvex solution space. This complexity hinders the direct application of standard optimization solvers such as CPLEX or Gurobi. However, observing that the objective function's complexity is contrasted by a relatively small number of constraints, we apply a hybrid Interior Point - Augmented Lagrangian strategy to solve the model.

4 Analysing capacity impacts of rail pods' platoon formation and operational speed

Simulation cases are provided in this section considering the ETCS level 2 signalling system to illustrate the effectiveness of the proposed approaches.Six Pods with varying platoon structures were selected to verify the proposed model. Other related parameters can be approximated in table 2.

Table 2: Simulation Parameters for Case Study

Structural		**Dynamic**		**Operational**	
Symbol	Value (Unit)	Symbol	Value (Unit)	Symbol	Value (Unit)
M	6 (-)	a_{bra}	1.0 (m/s²)	v_{sat}	20 (m/s)
L_{tlen}	100 (m)	a_{tra}	0.8 (m/s²)	$v_{\max}$	60 (m/s)
S_{vc}	30 (m)			$v_{\min}$	20 (m/s)
Geometric		**Control**		**Time**	
L_{sec}	[1000,1500×4,1000] (m)	T_r	4 (s)	T_{stop}	30 (s)
L_{pla}	100 (m)	T_{rel}	3 (s)	t_{couple}	120 (s)
L_{over}	50 (m)	α_j	[0,1,1,1,1,0] (-)		
L_{sm}	200 (m)				

4.1 Analysis of the relation capacity-cruising speed for uniform pod platoons

First, consider a homogeneous platoon configuration, where each platoon consists of the same number of Pods, resulting in equal platoon lengths. Under a uniformly block section length configuration, the minimum tracking time headway equals the block occupation time, leading to a line headway defined as $H^{\text{block}} = T^{\text{block}}$. The corresponding line capacity is derived from the UIC standard as follows:

$$C_{\text{block}} = \frac{3600}{H^{\text{block}}} \tag{13}$$

Based on this formulation, analysis is conducted on the relationship between speed, tracking interval, and capacity within the station blocking section under homogeneous platoon flow conditions. As shown in Figure 4, the approach area of the station imposes additional restrictions on the separation of the train due to the need for precise stopping and a longer blockage time.

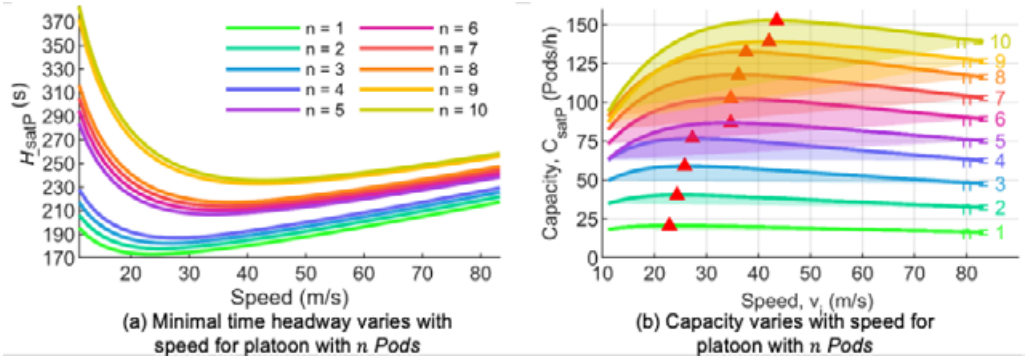


Figure 4: Optimal Speed Matching under Maximum Capacity and Headway

Results indicate that the station capacity is not directly proportional to speed or platoon length. The station capacity is much smaller than the section capacity, so when trains enter a section, there is some space to adjust the platoon formation through speed variation. When a train cluster forms platoons of different configurations, there will be varying supply rates within the time window for pods transportation.

4.2 Capacity analysis of heterogeneous pod platoons for optimised cruising speeds

To further evaluate the impact of heterogeneous platoon configurations on infrastructure capacity, we consider all possible platoon arrangements composed of 6 Pods under a fixed infrastructure setup. By modeling all integer partitions of 6 Pods, a total of 32 distinct platoon structures are generated, ranging from a single long platoon to multiple shorter, evenly distributed platoons. Figure 5 illustrates the infrastructure occupation time corresponding to each configuration. The X-axis represents the total track occupation time, while the Y-axis indexes the specific platoon structures.

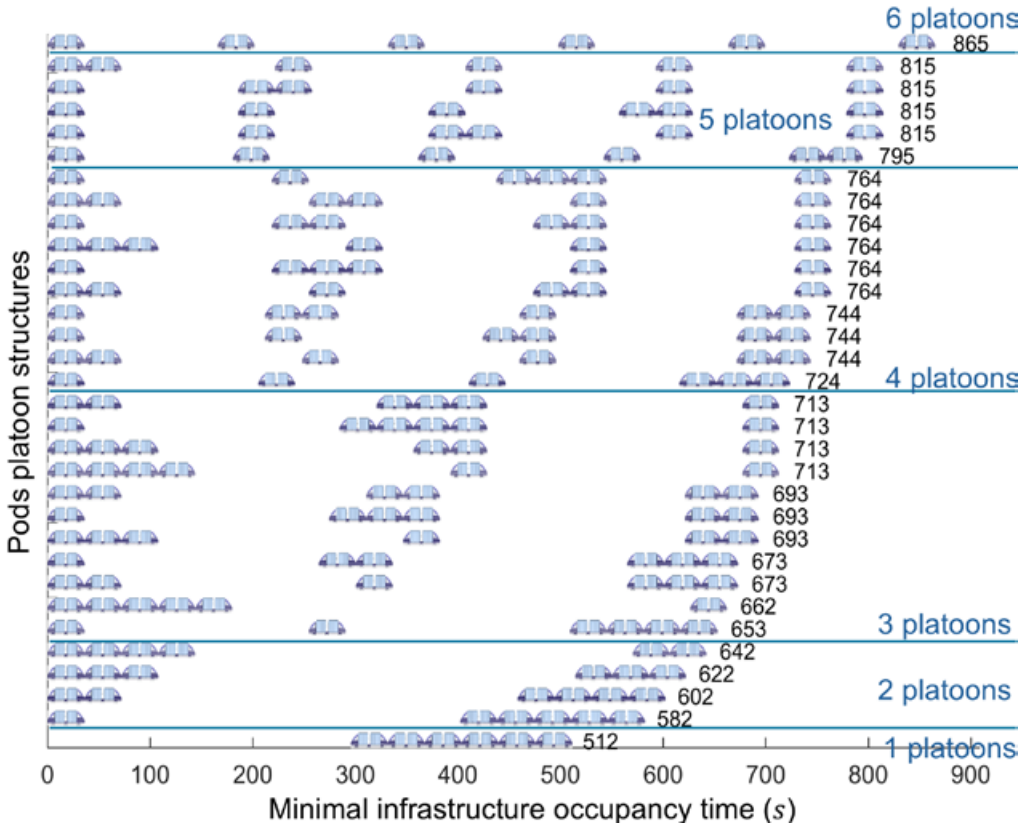


Figure 5: Minimum infrastructure occupation time for heterogeneous platoons

The comparison reveals that platoon structure has a significant influence on infrastructure consumption. The optimal platoon configuration can improve system efficiency by up to 40.8%. These findings highlight the importance of structural design in platoon scheduling and provide a quantitative basis for future optimization in dynamic platoon management

strategies. Results suggest that, in the absence of platform length constraints, forming larger platoons can significantly enhance line throughput and infrastructure utilization.

Even with the same number of platoons using the Pods set, variations in platoon length configuration can lead to significant differences in infrastructure occupation. For instance, a `2-3-1` structure and a `1-1-4` structure both consist of three platoons, where numbers indicate the number of Pods included in each platoon, but their temporal occupation across the blocking sections differ due to the order and size distribution of the platoons.

To investigate this effect, Figure 6 presents the block time occupation and optimal cruise speed configurations under `2-3-1` and `1-1-4` structures. The horizontal axis denotes the cumulative length of the blocks, while the rectangular blocks represent the minimum safe separation time required by each block. The upper part of Figure 9 shows the safety separation and speed trajectory of the block, and the lower part shows its coordinated speed configuration. The curved lines represent the coordinated speed profiles computed for each platoon. Each group of lines corresponds to one platoon, and the number of lines in each group reflects the number of Pods within that platoon.

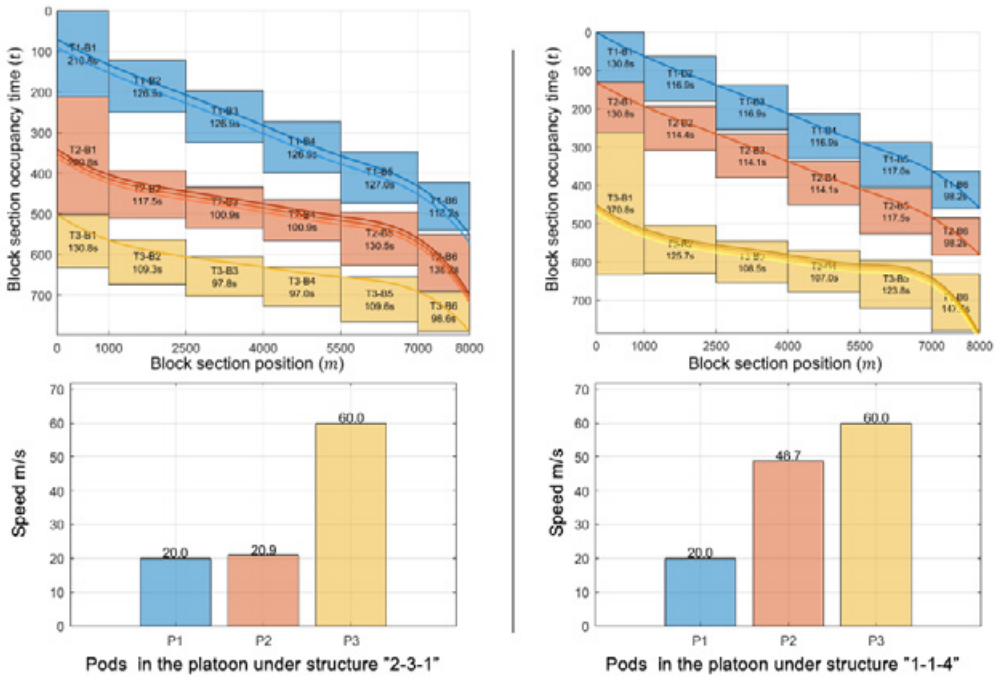


Figure 6: Minimum occupation time for heterogeneous platoons with optimal speed

The simulation results indicate that optimal velocity configurations can minimize block occupation time under different platoon layouts. The departure block, which includes

additional time for platoon formation, often becomes the bottleneck in overall capacity consumption. This visualization highlights how structural layout and coordinated velocity planning jointly impact infrastructure usage, reinforcing the need for formation and speed co-optimization in platoon scheduling.

Interestingly, the optimal solutions do not require all trains to operate at maximum speed. Instead, a trend emerges where leading platoons are assigned relatively higher speeds, while trailing platoons operate at reduced velocities. This "fast-front, slow-rear" coordination effectively minimizes tracking gaps and optimizes infrastructure utilization. Moreover, the optimal cruise speed is not linearly proportional to platoon size. Longer platoons do not necessarily result in higher optimal speeds, highlighting the importance of joint optimization between structural layout and speed planning for capacity enhancement.

Further analysis was conducted to investigate the impact of platoon coupling time on infrastructure occupation and optimal platoon structure. The results, summarized in Table 3, indicate a near-linear increase in total occupation time with increasing coupling time.

Table 3: Impact of Platoon Coupling Time on Infrastructure Occupation

CouplingTime (s)	0	30	60	90	120	150	180	210
Occupation (s)	249	399	549	699	849	902	902	902
Structure	6	6	6	6	6	[1,1,1,1,1,1]		

When the coupling time exceeds a threshold (e.g., 150 seconds), the optimizer consistently selects the most uniform platoon structure (e.g., [1,1,1,1,1,1]), suggesting that structure-based optimization becomes ineffective. This is because the extended coupling time dominates the total occupation time, rendering any benefit from structural variation negligible. These findings imply that excessive coupling durations can negate the benefits of platoon-based structural planning. Therefore, the operational time window for platoon formation acts as a limiting factor in capacity optimization. This provides a quantitative reference for setting pick-up and delivery time constraints in Pods systems.

5 Conclusion

The Pods system offers higher flexibility for Customer-oriented rail services and seamless operations. The scheduling time windows must be carefully configured by considering in-

frastructure, especially under the safety constraints imposed by signaling systems. The dynamic nature of Pods—featuring variable platoon lengths and dispatch timing—renders conventional capacity assessment metrics unsuitable. To address this, we extend the classic blocking time theory and the UIC 406 methodology to support Pods platoons. A novel optimization model is proposed to minimize infrastructure occupation time. The model constructs a mapping between convoy structure, speed profile, and capacity consumption, enabling joint optimization of speed allocation. The model incorporates detailed operational rules and piecewise velocity profiles reflecting the traction–cruise–brake process. A hybrid Interior Point–Augmented Lagrangian method is employed to solve the resulting nonlinear optimization problem. Simulation results under ETCS level 2 signaling system settings show that infrastructure consumption is not linearly correlated with either platoon size or speed. The proposed coordinated structure-speed optimization can enhance capacity utilization by up to 40.8%. Optimal solutions typically exhibit a "fast-front, slow-rear" velocity pattern. Notably, operating all trains at maximum speed is neither necessary nor optimal.

Furthermore, the analysis reveals that platoon coupling time has a critical impact on capacity efficiency and optimal structure selection. When coupling time exceeds a threshold, the benefits of platoon-based optimization diminish. It will create feedback constraints on pickup and delivery windows in Pods deployment. The proposed modeling framework provides a foundation for analyzing the interactions between Pods system features and infrastructure constraints. Future work will extend the model to incorporate non-uniform mass distribution among Pods and adapt it to moving block signaling environments. This study provides a quantitative foundation for the planning and design of Pods systems.

Acknowledgment

This project has received funding from the Europe's rail Joint Undertaking (JU) under GA No. 101121853. It is supported by the Europe's Rail JU and its members funded by the European Union. Views and opinions expressed are those of author(s) only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Literature

[1] FLAMMINI, Francesco ; MARRONE, Stefano ; NARDONE, Roberto ; PETRILLO, Alberto ; SANTINI, Stefania ; VITTORINI, Valeria: Towards railway virtual coupling. In: *2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)* IEEE, 2018, S. 1–6

[2] KHADILKAR, Harshad: A scalable reinforcement learning algorithm for scheduling railway lines. In: *IEEE Transactions on Intelligent Transportation Systems* 20 (2018), Nr. 2, S. 727–736

[3] LANDEX, Alex: Evaluation of Railway Networks with Single Track Operation Using the UIC 406 Capacity Method. In: *Networks and Spatial Economics* 9 (2009), 7-23. https://api.semanticscholar.org/CorpusID:108512316

[4] LIAO, Ximeng ; HAN, Jihee ; MAHNAM, Saeednia ; PAZ MARTINEZ, Aaron: Unlocking the Potentials of Modularity in Railways, a Heuristic Framework for Pods Scheduling. In: *27th IEEE International Conference on Intelligent Transportation Systems, ITSC 2024* IEEE-Institute of Electrical and Electronics Engineers, 2024

[5] LIAO, Ximeng ; HAN, Jihee ; SAEEDNIA, Mahnam: Modular Vehicle Routing on Railways: Opportunities for Intermodality. In: *2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)* IEEE, 2024, S. 572–577

[6] NOLD, Michael ; CORMAN, Francesco: Dynamic train unit coupling and decoupling at cruising speed: Systematic classification, operational potentials, and research agenda. In: *Journal of Rail Transport Planning & Management* 18 (2021), S. 100241

[7] PODS4RAIL CONSORTIUM: *Pods4Rail Project Homepage*. https://pods4rail.eu/, 2025. – Accessed: 2025-04-23

[8] QUAGLIETTA, Egidio ; WANG, Meng ; GOVERDE, Rob M.: A multi-state train-following model for the analysis of virtual coupling railway operations. In: *Journal of Rail Transport Planning & Management* 15 (2020), S. 100195

[9] WIDYASTUTI, H ; BUDHI, WS: Railway capacity analysis using Indonesian method and UIC code 405 method. In: *IOP Conference Series: Materials Science and Engineering* Bd. 930 IOP Publishing, 2020, S. 012059

IRSA 2025

SESSION 3

BAHNSTEIGE

RAILWAY PLATFORMS

Session 3
Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

Rahmenbedingungen für eine Implementierung von Bahnsteigbarrieren und Zugangskontrollen im deutschen Vollbahnsystem

Kliem, Maximilian[1], Lenze, Wiebke[1], Pfeifer, Andreas[1], Nießen, Nils[1],
Horn, Florian[2], Jendritzki, Felix[2],
Bertl, Melanie[3], Jordanski, Christina[3], Koch, Susanne[3]

[1]Verkehrswirtschaftliches Institut der RWTH Aachen University
[2]DB InfraGO AG, Geschäftsbereich Personenbahnhöfe
[3]Fachbereich Wirtschaft und Recht der Frankfurt University of Applied Sciences

Zusammenfassung

Sicherungsmaßnahmen, wie bspw. Bahnsteigbarrieren oder Zugangskontrollen, können die Sicherheit und das subjektive Sicherheitsempfinden erhöhen. Während solche Maßnahmen bereits weltweit vorrangig auf Metro- und Stadtbahnsysteme zum Einsatz kommen, ist ihre Anwendung im deutschen Vollbahnsystem bisher nicht wissenschaftlich untersucht. Daher wird für diesen Anwendungsbereich die Eignung verschiedener Sicherungsmaßnahmen aus technisch-betrieblicher, wirtschaftlicher sowie rechtlicher Sicht geprüft. Darüber hinaus wird mithilfe von Personenflusssimulationen sowie einer praktischen Erprobung der Nachweis erbracht, ob die Umsetzung der Sicherungsmaßnahmen einzeln oder in Kombination unter den bestehenden baulichen und eisenbahnbetrieblichen Rahmenbedingungen im deutschen Vollbahnsystem möglich und zielführend ist.

Keywords: Bahnsteigbarrieren, Gefahrenabwehr, Kritische Infrastruktur, Prävention, Reisendensicherheit, Zugangskontrollen

Session 3
Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

1 Einleitung

In Bahnhöfen ereignen sich immer wieder Unfälle oder Delikte, bei denen Menschen bspw. auf die Bahngleise bzw. zwischen Zug und Bahnsteig stürzen oder von einem durchfahrenden Zug erfasst werden [1]. Um die Gefahr solcher Gleisunfälle zu reduzieren und dadurch die Sicherheit sowie das Sicherheitsempfinden der Reisenden nachhaltig zu verbessern, wird im Rahmen dieses Beitrags eine Implementierung von zusätzlichen Sicherungsmaßnahmen (Bahnsteigbarrieren und Zugangskontrollen) im deutschen Vollbahnsystem geprüft. Im Rahmen der Untersuchung werden verschiedene Systemvarianten der Sicherungsmaßnahmen identifiziert und anschließend technisch-betriebliche, wirtschaftliche sowie rechtliche Faktoren zur Machbarkeit analysiert und bewertet. Mithilfe von Personenflusssimulationen sowie einer praktischen Erprobung wird der Nachweis erbracht, ob die Umsetzung der Sicherungsmaßnahmen einzeln oder in Kombination unter den bestehenden baulichen und eisenbahnbetrieblichen Rahmenbedingungen im deutschen Vollbahnsystem möglich und zielführend ist.

2 Subjektive Sicherheit an Bahnsteigen

Während objektive Sicherheit auf systematisch gemessenen Risiken basiert, bspw. Kriminalitäts- und Unfallstatistiken, beruht subjektive Sicherheit auf der Situationsdefinition durch die Personen, nach individuellen Einschätzungen und Vorstellungen [2]. Das Sicherheitsempfinden spiegelt in der Regel nicht die objektive Gefährdungslage wider [3]. Sicherheitsmaßnahmen können in unterschiedlichem Maße auf je objektive und subjektive Sicherheit wirken [4]. In nationalen sowie internationalen Studien wurden vielfältige Einflussfaktoren auf das Sicherheitsempfinden an Bahnhöfen gefunden (vgl. [3], [5]). Dabei wird der Schwerpunkt überwiegend auf Security-Aspekte gelegt. Neben personenbezogenen Merkmalen (bspw. Geschlecht) sind Bahnhofsdesign (bspw. offene Bauweise, Ausleuchtung) und Bahnhofszustand von Relevanz [6]. Ebenso werden Faktoren, die sich auf die persönliche Umwelt beziehen, wie Personendichte, Präsenz von Personal und Anwesenheit bestimmter, risikoassoziierter Personengruppen berücksichtigt [7]. Effekte von Zugangskontrollen und Bahnsteigbarrieren auf die subjektive Sicherheit sind der Studienlage nach bisher wenig erforscht.

Für die Messung von subjektiver Sicherheit im ÖPV werden in Studien unterschiedliche Operationalisierungen angewandt, auch jeweils davon ausgehend, welche Definition von subjektiver Sicherheit zugrunde gelegt wird und welche Ebene zum Gegenstand der Untersuchung gemacht werden soll (siehe u. a. [8]). Auf Grundlage der Studienergebnisse

erfolgte die Entwicklung eines Erhebungsinstrumentes zur Erfassung des subjektiven Sicherheitsempfindens bei Reisenden am Bahnsteig. Der Fragebogen ist prinzipiell, durch Anpassungen, in verschiedenen Situationen für Reisende des Bahnverkehrs einsetzbar. Schwerpunkt bei der Entwicklung des Erhebungsinstrumentes war die möglichst umfassende Abbildung des subjektiven Sicherheitsempfindens durch die Mitaufnahme von Fragen, die verschiedene Bereiche abbilden. Neben einer Frage zum allgemeinen Sicherheitsgefühl („*Wie sicher haben Sie sich am Bahnsteig gefühlt?*“)[1], erfolgt in dem Fragebogen deshalb auch die Untersuchung acht verschiedener Gefährdungsbereiche, die am Bahnsteig relevant sein könnten.

Die Auswahl der Gefährdungsbereiche erfolgte in Anlehnung an die „SKiD“-Studie des Bundeskriminalamts zur Sicherheit und Kriminalität in Deutschland [9] bei der das Sicherheitsgefühl in Bezug auf verschiedene Deliktarten abgefragt wird, und wurde ergänzt durch weitere Bereiche, die das Sicherheitsempfinden in der Situation am Bahnsteig abbilden sollen.

Daraus ergaben sich die folgenden acht Bereiche, zu denen jeweils eine Frage[2] im Erhebungsinstrument enthalten ist:

- Sexuelle Belästigung
- Persönlicher physischer Angriff
- Diebstahl
- Terroranschlag
- Gesundheitliche Gefährdung
- Absichtliches ins Gleis gestoßen werden
- Unabsichtliches ins Gleis fallen (Unfall)
- Fallen zwischen Zug und Bahnsteigkante (Unfall)

Zudem beinhaltet das Erhebungsinstrument Fragen zu möglichen Einflussfaktoren, die auf Basis der Studienlage und Relevanz für Barrieren und Zugangsbeschränkungen ausgewählt wurden, wie u. a. Personenanzahl, Tag/Nacht, Alter, Bildungsabschluss, Art des Bahnsteigs, Wartezeit, Geschlecht, mit/ohne Begleitung.

[1] Genauer Wortlaut der Fragestellung: Nun möchten wir etwas darüber erfahren, wie sicher oder unsicher Sie sich am Bahnsteig gefühlt haben. Bitte geben Sie auf der Skala von -3 (= Sehr unsicher) bis 3 (=Sehr sicher) den Wert an, der Ihr Sicherheitsempfinden am besten widerspiegelt. Wie sicher haben Sie sich am Bahnsteig gefühlt?

[2] Beispiel für Fragestellung im Erhebungsinstrument für den Bereich „Diebstahl“: Bitte geben Sie auf der Skala von -3 (= Sehr stark beunruhigt) bis 3 (= Gar nicht beunruhigt) einen Wert an, der Ihr Empfinden am besten widerspiegelt. Inwieweit fühlten Sie sich beunruhigt, dass Ihnen etwas gestohlen wird?

Nach einem qualitativen Pretest und darauf beruhenden Anpassungen, erfolgte die Testung des Erhebungsinstrumentes im Rahmen eines quantitativen Pretests mit etwa 140 Teilnehmenden. Die Auswertung der nicht repräsentativen Daten gibt unter anderem Hinweis darauf, dass die subjektive Sicherheit in Bezug auf bestimmte Gefährdungsbereiche unterschiedlich hoch ausfällt (siehe Abbildung 1). Deshalb erfolgt auch bei Testung der Maßnahmen das getrennte Abfragen der einzelnen Gefährdungsbereiche.

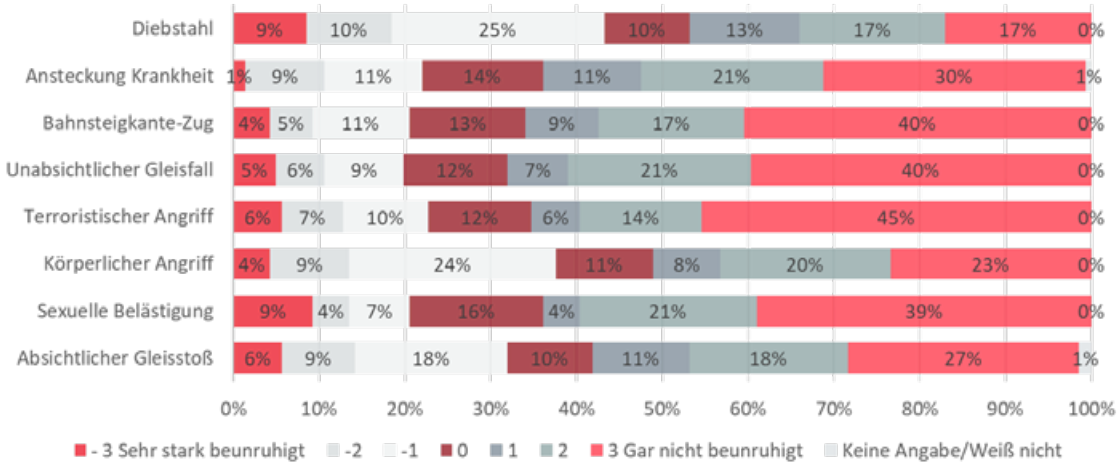


Abbildung 1: Übersicht über Beunruhigung durch verschiedene Deliktarten/Bereiche, keine repräsentative Befragung, N=141. Hinweis: Die Befragten wurden zu ihrem letzten Aufenthalt an einem Bahnsteig befragt. [Eigene Darstellung]

Anwendung fand der Fragebogen, mit Anpassungen für die konkrete Situation, unter anderem bei der Durchführung der praktischen Erprobung im September 2025.

3 Varianten von Bahnsteigbarrieren und Zugangskontrollen

Sowohl Bahnsteigbarrieren als auch Zugangskontrollsysteme kommen weltweit, vorrangig bei Metro- und Schnellbahnsystemen, bereits erfolgreich zum Einsatz. Um eine möglichst fundierte Beurteilung der Sicherungsmaßnahmen vorzubereiten, sind verschiedene Systemvarianten sowohl der Bahnsteigbarrieren als auch der Zugangskontrollen zu identifizieren. Je nach Ausprägungsform ergeben sich unterschiedliche Wechselwirkungen mit den Rahmenbedingungen des deutschen Vollbahnnetzes, welche eine separate Bewertung der Varianten in Kapitel 4 erforderlich machen.

Insgesamt sechs Varianten von Bahnsteigbarrieren wurden in die Untersuchung aufgenommen. Dazu zählen die gebräuchlichen Platform Screen Doors (PSD), Platform Safety Gates (PSG) oder Sicherheitsbarrieren mit Lücken (SB-L), aber auch neuartige Lösungen

wie Variable Pitch Platform Screen Doors (VP-PSD), Verticle Platform Screen Doors (VPSD) oder Rope Platform Screen Doors (RPSD). Von den genannten Varianten kommen bisher nur SB-L im deutschen Vollbahnnetz auf der Schnellfahrstrecke Hamburg – Berlin zum Einsatz.

Grundsätzlich lassen sich diese Varianten in Abhängigkeit ihrer Bewegungsrichtung unterscheiden. So können sie entweder vertikal (PSD, PSG und VP-PSD) oder horizontal (RSD und VPSD) öffnen oder wie die SB-L unbeweglich sein. Zudem ergeben sich Unterschiede hinsichtlich der Durchlässigkeit der Bauformen und somit der spezifischen Schutzwirkung, der Höhe sowie der Öffnungspositionen.

Für die Zugangskontrollen werden im Rahmen der Untersuchung ebenfalls verschiedene Varianten betrachtet. Grundsätzlich werden sogenannte Drehsperren und Schwingtüren unterschieden. Die Wahl einer der beiden Formen hat aber nur geringe Auswirkungen auf technische oder betriebliche Faktoren. Für die weiteren Analysen ist vielmehr die Aufstellposition innerhalb des Bahnhofes relevant. So werden bei Zugangskontrollen verschiedene Schutzradien abgegrenzt (vgl.

Abbildung 2). Ein „gesicherter Bahnsteig“ bspw. sieht Kontrollen direkt vor dem Betreten des Bahnsteiges vor. Ein „gesicherter Bahnsteigbereich“ auf der anderen Seite umfasst mehrere Bahnsteige und sieht eine Kontrolle, zum Beispiel vor dem Betreten einer Verteilinfrastruktur wie einem Personentunnel vor. Eine weitere Möglichkeit ist die Absicherung des gesamten Bahnhofs mithilfe von Zugangskontrollen.

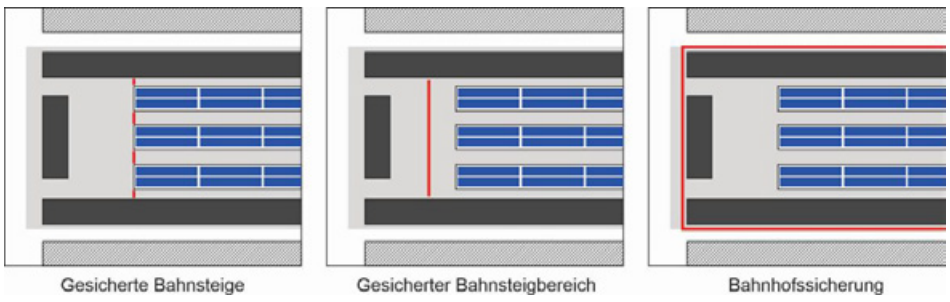


Abbildung 2: Schutzradien von Zugangskontrollen [Eigene Darstellung nach [10]]

4 Machbarkeitsstudie von Sicherungsmaßnahmen

Die Anwendung der vorgestellten Sicherungsmaßnahmen ist derzeit vorrangig auf Metro- und Stadtbahnsysteme fokussiert [11]. Um eine Übertragbarkeit der Sicherungsmaßnahmen von diesen geschlossenen, unabhängigen Schienenbahnsystemen auf Vollbahnsysteme in Deutschland bewerten zu können, wurde eine Untersuchung der technisch-betrieblichen, rechtlichen und wirtschaftlichen Rahmenbedingungen durchgeführt.

Session 3
Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

4.1 Technische Sicht

Zur Beurteilung der technisch-betrieblichen Umsetzbarkeit kann ein umfangreicher Kriterienkatalog herangezogen werden, der unterschiedliche Aspekte sowie Kenngrößen des Systems Bahn abbildet und auf Basis verschiedener Regelwerke und Zusammenhänge entwickelt wurde [12]. Bezüglich der technischen Umsetzbarkeit wurden einerseits die konkreten Schutzwirkungen hinsichtlich bestimmter Gefahrenpotenziale analysiert, um den Nutzen der betrachteten Sicherungsmaßnahmen zu bewerten, aber auch ein Abgleich mit den Anforderungen aus den baulichen Gegebenheiten durchgeführt. Geprägt durch das hohe Alter der Eisenbahninfrastruktur in Deutschland existieren besonders viele Anforderungen, welche bei Implementierung zusätzlicher Sicherungsmaßnahmen erfüllt werden müssen. Dabei gibt es Anforderungen, welche sich vorrangig auf Bahnsteigbarrieren und andere, welche sich vorrangig auf Zugangskontrollen auswirken. So wirken sich etwa die Bahnsteigkrümmung oder der Lichtraum vorrangig bei der Wahl der Bahnsteigbarriere aus, nicht aber bei der Wahl der Zugangskontrolle oder des Schutzradius. Die berücksichtigten Kriterien wurden, wie in Abbildung 3 dargestellt, in insgesamt 8 Gruppen unterteilt und anschließend für jede Systemvariante getrennt bewertet.

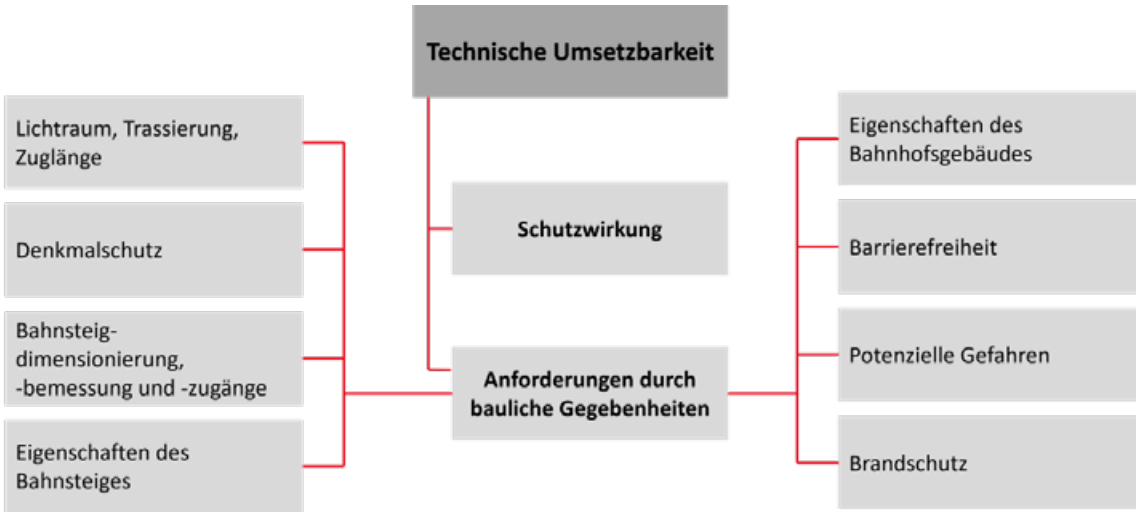


Abbildung 3: Bewertungskriterien der technischen Machbarkeitsstudie [Eigene Darstellung]

In der durchgeführten Machbarkeitsstudie hat sich gezeigt, dass potenzielle Konflikte zwischen den technischen Rahmenbedingungen und den Sicherungsmaßnahmen nach dem derzeitigen Untersuchungsstand nur in geringem Umfang zu erwarten sind. Es besteht kein grundsätzlicher Ausschluss einer bestimmten Variante, jedoch ergeben sich Anwendungsszenarien, welche die Wahl bestimmter Sicherungsmaßnahmen bevorzugen

lassen. Es wurde festgestellt, dass die Sicherheitsbarrieren mit Lücken (SB-L) hinsichtlich der relevanten Gefahrenpotenziale nur wenige Schutzziele erfüllen.

Bei der Implementierung von Sicherheitsmaßnahmen in einem konkreten Bahnhof müssen stets die individuellen Randbedingungen berücksichtigt werden, um die am besten geeignete Ausführungsform auszuwählen.

4.2 Betriebliche Sicht

Ähnlich der technischen Umsetzbarkeit wurden auch zur Beurteilung der betrieblichen Umsetzbarkeit verschiedene Kriterien definiert und im Anschluss bewertet. Die Auswirkungen auf den eigentlichen Bahnbetrieb wurden ebenso berücksichtigt, wie auch Auswirkungen auf die Passagierbewegungen sowie eine Bewertung der Anlagenresilienz bzw. von Störungsfolgen (siehe Abbildung 4).



Abbildung 4: Bewertungskriterien der betrieblichen Machbarkeitsstudie [Eigene Darstellung]

Die Auswirkungen auf den Bahnbetrieb fallen grundsätzlich gering aus. Einfahr- und Durchfahrgeschwindigkeit sollten durch die Errichtung von Bahnsteigbarrieren nicht beeinträchtigt werden. In manchen Fällen erlauben Bahnsteigbarrieren durch den Schutz der Reisenden auf dem Bahnsteig bspw. vor der Sogwirkung vorbeifahrender Züge sogar höhere Geschwindigkeiten.

Alle Bahnsteigbarrieren können im Vergleich mit dem Ausgangszustand den Fahrgastwechsel in mindestens gleicher Güte ermöglichen, wenn nicht sogar eine Verkürzung der erforderlichen Zeiten herbeiführen. Nur ausgewählte Bahnsteigbarrieren können bei einem Fahrzeugpark mit unterschiedlichen Türpositionen eingesetzt werden. Hier sind vor allem die neuartigen Varianten RSD, VPSD und VP-PSD zu bevorzugen, da sie bei variablen Türabständen eines heterogenen Fahrzeugparks, wie es in Deutschland der Fall ist, eingesetzt werden können. Die Wahl der Bahnsteigbarriere wird somit maßgeblich durch den verwendeten Fahrzeugpark beeinflusst. Auf Strecken mit einem heterogenen Fahr-

zeugpark sind vertikal öffnende Systeme zu bevorzugen. Eine homogene Fahrzeugauswahl begünstigt hingegen die Verwendung der etablierten horizontal öffnenden Bahnsteigbarrieren.

Bei der Analyse der betrieblichen Auswirkungen von Zugangskontrollen wurde der Schutzradius „gesicherter Bahnsteig“ als nachteilig bewertet. Durch die bahnsteigscharfen Kontrollen werden Umsteigevorgänge, die einen Bahnsteigwechsel erfordern, erheblich verlängert und so die Reisezeiten u. U. negativ beeinflussen. Die Ausgestaltung von Zugangskontrollen wird maßgeblich von bahnhofsspezifischen Faktoren (insbesondere dem Bahnhofslayout) beeinflusst.

Aus betrieblicher Perspektive ergeben sich somit für den Regelbetrieb keine wesentlichen Hindernisse, welche einer Implementierung von Bahnsteigbarrieren und Zugangskontrollen nach aktuellem Kenntnisstand im Allgemeinen behindern würden. Es könnten aber Faktoren herausgearbeitet werden, die bestimmte Ausführungsvarianten in bestimmten Szenarien bevorzugen lassen.

4.3 Rechtliche und wirtschaftliche Sicht

Hinsichtlich der rechtlichen Umsetzbarkeit der Sicherungsmaßnahmen wurden diverse Richtlinien, Verordnungen, Gesetze und Normen begutachtet. Anhand der ausgewerteten Dokumente scheinen keine grundsätzlichen Einschränkungen zu bestehen, die der Implementierung von Zugangskontrollen und Bahnsteigbarrieren entgegenstehen. Vielmehr sind gem. Art. 27 der Verordnung (EU) 2021/782 über die Rechte und Pflichten der Fahrgäste im Eisenbahnverkehr geeignete Maßnahmen zu ergreifen, um die persönliche Sicherheit der Fahrgäste in den Bahnhöfen zu gewährleisten. Eine sichere und ordnungsgemäße Betriebsdurchführung bedarf zudem einer ausreichenden Berücksichtigung u. a. von Aspekten der baulichen Sicherheit, des Brandschutzes und der Barrierefreiheit [13].

Für die Bewertung der wirtschaftlichen Umsetzbarkeit wurden die Kostenfaktoren und -treiber für den gesamten Lebenszyklus der Anlagen identifiziert. Anschließend wurden die Effekte verschiedener Einflussparameter, wie bspw. das Personenaufkommen oder die Bahnhofsgestaltung, auf die Kostenkomponenten betrachtet. Auch wenn es dabei nicht möglich ist, die Kosten für die Implementierung im gesamten deutschen Vollbahnnetz zu benennen, muss davon ausgegangen werden, dass die Kosten insgesamt erheblich ausfallen werden. Sowohl die Errichtung als auch der Betrieb und die Instandhaltung der Anlagen werden mit umfangreichen Aufwendungen verbunden sein.

5 Simulation der Sicherungsmaßnahmen

Mit Hilfe der Personenflusssimulation „CAST Pedestrian“ wurde eine realitätsnahe Abbildung des Betriebsgeschehens sowie eine Beurteilung der Wechselwirkungen der einzelnen Sicherungsmaßnahmen mit gegebenen Bahnhofs- und Bahnsteiglayouts untersucht [14, 15]. Um die Übertragbarkeit der Simulationsergebnisse auf einen möglichst großen Teil des deutschen Vollbahnnetzes zu ermöglichen, wurden für die Simulation drei verschiedene Beispielbahnhöfe in unterschiedlichen Größenordnungen ausgewählt (siehe Tabelle 1).

Tabelle 1: Ausgewählte Beispielbahnhöfe

Großer Bahnhof	Mittelgroßer Bahnhof	Kleiner Bahnhof
Köln Hbf	Aachen Hbf	Düsseldorf Benrath

Die Auswahl der relevanten Beispielbahnhöfe wurde zudem von weiteren Faktoren beeinflusst. Es wurde u. a. angestrebt, möglichst viele verschiedene Zugangsmöglichkeiten und Bahnsteigformen abzudecken. So verfügen alle ausgewählten Bahnhöfe sowohl über reguläre Treppenzugänge als auch über Fahrstühle zur Nutzung durch mobilitätseingeschränkte Reisende. In Köln Hbf wurden zudem Fahrtreppen simuliert, sodass alle relevanten Zugangsmöglichkeiten zum Bahnsteig abgebildet werden. Verschiedene Bahnsteigformen konnten ebenfalls über die unterschiedlichen Beispielbahnhöfe abgedeckt werden. So bestehen Unterschiede hinsichtlich der Krümmung sowie der Platzverhältnisse an Auf- und Abgängen.

Für die Simulation wurden weitere Randbedingungen festgelegt. So wurden beispielsweise Störfallszenarien entworfen, welche den Ausfall einer Bahnsteigbarriere, den Ausfall einer Zugangskontrolle sowie potenzielle Steigerungen der Passagieraufkommen umfassen. Auf diese Weise können sowohl eventbezogene Nachfragespitzen als auch zukünftige Nachfragesteigerungen abgebildet werden.

Zur eigentlichen Modellierung der Bahnhofsinfrastruktur wurden verschiedene Rahmenbedingungen und Besonderheiten der Prozesse im Bahnhof. Basis der Modellierung ist eine Aufstellung aller relevanten personenindividuellen Prozessschritte, welche ein Reisender beim Ein-, Aus- bzw. Umstieg durchläuft (bspw. Ticketkauf, Einkaufen, Warten am Bahnsteig). Es ergibt sich in der Simulation für jeden Reisenden eine individuelle „Aufgabenliste“, die je nach zuvor festgelegten Charakteristika in ihrer zeitlichen Dauer

variiert. Zudem bestand ein besonderer Fokus auf der realitätsnahen Abbildung des Fahrgastwechsels. Der Ein- und Ausstiegsprozess wurde derart modelliert, dass Passagiere im Regelfall zuerst den Zug verlassen und anschließend auf dem Bahnsteig wartende Passagiere den Zug betreten. Auch die Verteilung der Reisenden auf dem Bahnsteig sowie die Verteilung der Reisenden auf verschiedene Zugtüren wurde möglichst realitätsnah nachgebildet.

Bei der Modellierung der Sicherungsmaßnahmen wurde der Abstand zwischen der Bahnsteigkante und der Bahnsteigbarriere basierend auf Herstellerangaben mit 50 cm angesetzt. Eine Reduzierung dieses Abstands ist bei einer tatsächlichen Implementierung ggf. unter Berücksichtigung der durch das Lichtraumprofil und den Betrieb gegebenen Rahmenbedingungen und voraussichtlich nur im Einzelfall möglich.

Die Modellierung der Zugangskontrollen wurde initial für alle in

Abbildung 2 gezeigten Varianten vorgenommen. Im Rahmen der Simulation wurde jedoch der „gesicherte Bahnsteig“ als Testfall verworfen, da sich in allen betrachten Bahnhöfen ein nicht marktgerechtes Platzangebot in den Unterführungen oder auf dem Bahnsteig ergab. Aus diesem Grund wurde für die weiteren Simulationen lediglich der Schutzradius „gesicherter Bahnsteigbereich“ betrachtet. Die Variante „Bahnhofssicherung“ wird nicht simuliert, da grundsätzlich ein möglichst kleiner Schutzradius angestrebt wird.

Nach Festlegung der verschiedenen Basisparameter und Modellierung der spezifischen Prozesse im Bahnhof, erfolgte die Durchführung der Simulation. Für die vier definierten Testfälle wurden dabei in einem statistisch abgesicherten Umfang Simulationsdurchläufe zur Beschreibung des Systemverhaltens durchgeführt, sodass eine Auswertung zielführend möglich war.

Die Auswertung der Simulation umfasst zwei wesentliche Bestandteile. So wurden sowohl Kenngrößen identifiziert, quantitativ analysiert und mit Referenzwerten abgeglichen als auch qualitative Analysen von 3D-Visualisierungen der Passagierbewegungen vorgenommen. Hierbei besteht ein starker Bezug zu den Herangehensweisen aus dem Flughafenwesen. Insbesondere die von der International Air Transport Association (IATA) verwendeten Level-of-Service im „Airport Development Reference Manual“ [16] wurden für die Simulationsauswertungen adaptiert. Diese definieren im Zusammenhang mit verschiedenen Serviceeinrichtungen und Infrastrukturbestandteilen die zugehörigen Referenzwerte, welche zur Sicherstellung einer ausreichenden Servicequalität zu erreichen sind. So hat jeder Reisende einen spezifischen Platzbedarf, welcher abhängig von seiner Position im Gebäude abzugrenzen ist. Ist dieser Platzbedarf eingehalten, kann sowohl von einem ausreichenden Komfort für den Reisenden als auch von ausreichenden

Flussverhalten der Passagierströme ausgegangen werden. Weiteren wesentlichen Einfluss auf den Komfort der Reisenden haben zudem Wartezeiten, welche sich für die Reisenden infolge der Sicherungsmaßnahmen ergeben. Diese sind ebenfalls auf bestimmte Grenzwerte zu beschränken (vgl. [16]). Es existieren Grenzwerte beispielsweise für Wartezeiten bei der Gepäckausgabe oder der Sicherheitskontrolle. Für Ticketkontrollen, welche an Flughäfen oft als Teil der Sicherheitskontrolle vorgenommen werden, sind durch die IATA keine spezifischen Referenzwerte vorgesehen, da sie für den Zeitbedarf an der Sicherheitskontrolle eine untergeordnete Rolle spielen. Daher wurde im Rahmen der Untersuchung die Adaption anderer Referenzwerte notwendig. Die nach dem „Airport Development Reference Manual" akzeptablen Wartezeiten für Sicherheitskontrollen wurden anhand der regulären Bedienzeiten für Sicherheitskontrollen und der Bedienzeiten für Zugangskontrollen skaliert. Es ergab sich eine zulässige Wartezeit im Belastungsfall von 60 Sekunden. Insgesamt konnten somit vier wesentliche Messgrößen identifiziert und zulässige Referenzwerte festgelegt werden:

- Platzangebot auf dem Bahnsteig allgemein [m²/Pers.]
- Platzangebot auf dem Bahnsteig in kritischen Bereichen [m²/Pers.]
- Platzangebot an der Sicherheitskontrolle [m²/Pers.]
- Wartezeit an den Zugangskontrollen [s]

Die Simulationsauswertung zeigt, dass alle untersuchten Kenngrößen eine marktgerechte Servicequalität (mindestens „Optimum" nach Einordnung der IATA) ergeben. Auch unter Anwendung von Störfallszenarien kann zumeist eine marktgerechte Servicequalität festgestellt werden. So kann der Ausfall einzelner Anlagen durch die unterstellte Dimensionierung kompensiert werden. Bei signifikanten Steigerungen der Passagieraufkommen ist die Anzahl der Zugangskontrollen zu skalieren. In den simulierten Beispielbahnhöfen ist eine solche Erhöhung der Anzahl an Zugangskontrollen mit den vorhandenen Platzverhältnissen für den Schutzradius „gesicherter Bahnsteigbereich" jeweils vereinbar. Negative Wechselwirkungen zwischen Bahnsteigbarrieren und Zugangskontrollen fallen während der Analyse der Ergebnisse nicht auf. Die Simulation hat darüber hinaus gezeigt, dass sich der Testfall „gesicherter Bahnsteig" aufgrund beengter Platzverhältnisse in Durchgangsbahnhöfen i. d. R. nicht marktgerecht implementieren lässt.

6 Praktische Erprobung

Zur anschließenden Überprüfung und Validierung der Simulationsergebnisse sowie vorrangig zur weiteren Erfassung der Auswirkungen der Sicherungsmaßnahmen auf das subjektive Sicherheitsempfinden der Reisenden, welches im Rahmen der durchgeführten Simulation nicht abbildbar war, werden die Sicherungsmaßnahmen einer praktischen Erprobung unterzogen. Hierzu wurde ein umfangreicher Test mit circa 400 Proband:innen an einem Bahnsteig im Bahnhof „Berlin Olympiastadion“ durchgeführt. Zur vollständigen Erfassung der Auswirkungen wurden dabei unterschiedliche Nutzergruppen in die Erprobung eingebunden. Verschiedene Szenarien wurden mithilfe von Mockups realitätsnah nachgestellt und mittels Videotechnik aufgezeichnet. Dabei fand ebenfalls der entwickelte Fragebogen zur Erfassung des subjektiven Sicherheitsempfindens, angepasst an die konkrete Simulationssituation, erneut Anwendung.

7 Fazit

Die bisherige Analyse zeigt, dass Bahnsteigbarrieren und Zugangskontrollen unter Beachtung anwendungsspezifischer Rahmenbedingungen (insb. des Bahnhofslayouts und der Homogenität des Fahrzeugparks) grundsätzlich im deutschen Vollbahnnetz integrierbar sind. Die wirtschaftlichen Aufwände einer Implementierung sind jedoch erheblich und mit dem resultierenden Nutzen abzuwägen. Aufbauend auf den erzielten Erkenntnissen werden im Anschluss an die praktische Erprobung Handlungsempfehlungen für Sicherungsmaßnahmen an Bahnhöfen aufgestellt. Ziel ist es dabei, Migrationspfade für eine kurzfristige Umsetzung zu definieren.

Die Projektbearbeitung erfolgt im Zuge der Bekanntmachung „Anwender – Innovativ: Forschung für die zivile Sicherheit II“ des BMFTR im Rahmen des Programms „Forschung für die zivile Sicherheit“ der Bundesregierung.

Literatur

[1] hib 913/2023, Bahnverkehr-Störungen durch Personen im Gleisbett, Berlin: Deutscher Bundestag, Parlamentsnachrichten, 2023.

[2] B. Frevel, Innere Sicherheit. Elemente der Politik, Wiesbaden: Springer VS, 2018.

[3] L. Gerhold, K. Dorner, E. Brandes und J. Hartmann, „Subjektives Sicherheitsempfinden und subjektive Sicherheitswahrnehmung," in *Sicherheitsempfinden, Sicherheitskommunikation und Sicherheitsmaßnahmen. Ergebnisse aus dem Forschungsverbund WiSima*, Berlin, 2020, pp. 9-58.

[4] V. Ceccato, C. Sundling und G. Gliori, „What makes a railway station safe and for whom? The impact of transit environments on passengers' victimisation and safety perceptions," *European Transport Research Review,* Bd. 16, 2024.

[5] C. Sundling und V. Ceccato, „The impact of rail-based stations on passengers' safety perceptions. A systematic review of international evidence," *Transportation Research Part F: Traffic Psychology and Behaviour,* Bd. 86, pp. 99-120, 2022.

[6] V. Ceccato, N. Gaudelet und G. Graf, „Crime and safety in transit environments: Asystematic review of the English and the French literature," *Public Transport,* Bd. 14, pp. 105-153, 2022.

[7] N. Schlüter, J.-P. Nicklas und P. Winzer, „Unsicherheiten am Bahnhof – Sicherheitswahrnehmungen im öffentlichen Personennahverkehr," in *Innere Sicherheit nach 9/11*, Wiesbaden, Springer Fachmedien, 2016, pp. 149-164.

[8] S. Kemme, J. Bendler und J. Struck, „Subjektive Sicherheit – Ein Überblick über Konzeptualisierungen, Forschungsstand und Impulse für die Prävention. In Deutscher Präventionstag (Hrsg.), Sicherheit im Wandel. Expertisen zum 29. Deutschen Präventionstag (S. 50-93)," 2024.

[9] C. Birkel, D. Church, A. Erdmann, A. Hager und N. Leitgöb-Guzy, „Sicherheit und Kriminalität in Deutschland – SKiD 2020," Bundeskriminalamt, 2020, Wiesbaden.

[10] N. Baron und N. L. Bot, „Railway station boarding controls: issues and limits. Performing security to secure," 2020.

[11] A. Okada und T. Ohisma, „Assessing the potential of half-height platform screen doors to prevent personal injury accidents: Evidence from the Tokyo metropolitan area railway network," *IATSS Research,* Bd. 49, pp. 81-92, 2025.

[12] DB InfraGO AG, Richtlinie 813, 2022.

[13] DIN Deutsches Institut für Normung e.V., DIN 18040 - 1 "Barrierefreies Bauen - Planungsgrundlagen - Teil 1: Öffentlich zugängliche Gebäude", Berlin: Beut Verlag GmbH, 2023.

[14] E. Bosina, Meeder, Mark und U. Weidmann, „Pedestrian flows on railway platforms," in *17th Swiss Transport Research Conference*, Monte Veritá / Ascona, 2017.

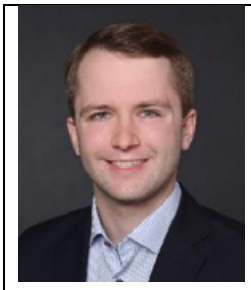
[15] Y. Ahn, T. Kowada, H. Tsukaguchi und U. Vandebona, „Estimation of Passenger Flow for Planning and Management of Railway," in *World Conference on Transport Research - WCTR 2016*, Shanghai, 2016.

[16] International Air Transport Association (IATA), Airport Design Reference Manual, 10th Edition, Montreal - Geneva, 2016.

Session 3

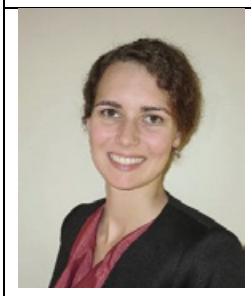
Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

Autoren



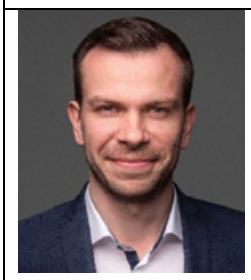
Kliem, Maximilian

Maximilian Kliem ist seit 2022 wissenschaftlicher Mitarbeiter am Verkehrswissenschaftlichen Institut der RWTH Aachen im Forschungsbereich Eisenbahnbetriebswissenschaft. Zuvor absolvierte er ein Studium des Bauingenieurwesens an der RWTH Aachen mit der Vertiefungsrichtung Verkehrswesen.



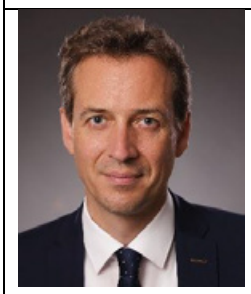
Lenze, Wiebke

Dr.-Ing. Wiebke Lenze ist seit 2017 wissenschaftliche Mitarbeiterin am Verkehrswissenschaftlichen Institut der RWTH Aachen im Bereich der Eisenbahnbetriebswissenschaft. Das Studium des Bauingenieurwesens mit der Vertiefungsrichtung Verkehrswesen absolvierte sie an der RWTH Aachen und der NTNU in Norwegen.



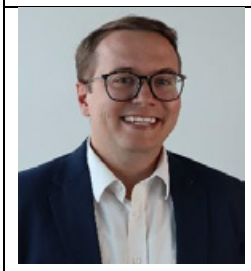
Pfeifer, Andreas

Dr.-Ing. Andreas Pfeifer ist seit 2021 Oberingenieur am Verkehrswissenschaftlichen Institut der RWTH Aachen. Nach dem Studium des Wirtschaftsingenieurwesens mit der Fachrichtung Bauingenieurwesen an der RWTH Aachen promovierte er dort 2019. Von 2014 bis 2020 war er als wissenschaftlicher Mitarbeiter am Verkehrswissenschaftlichen Institut tätig.



Nießen, Nils

Univ.-Prof. Dr.-Ing. Nils Nießen leitet seit 2013 das Verkehrswissenschaftliche Institut der RWTH Aachen. Nach dem Studium des Bauingenieurwesens und der Wirtschaftsgeographie an der RWTH Aachen promovierte er dort 2008. Anschließend war er Projektingenieur bei HaCon in Hannover und Geschäftsführer der VIA Consulting & Development GmbH.

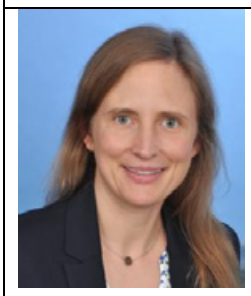
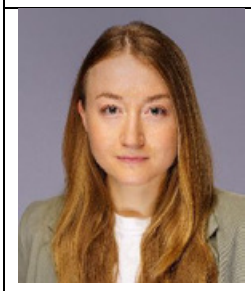
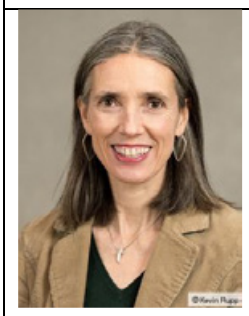


Horn, Florian

Florian Horn ist seit 2024 Projektleiter des Forschungsvorhabens Sicherheitsbahnhof (Sicherheitsforschung), das zusammen mit der Bundespolizei im Auftrag des BMIs und BMVs neue Ansätze für die Sicherheit am Bahnhof finden soll. Nach dem Studium des Security Managements an der TH Brandenburg war er in mehreren Funktionen des (europäischen) Sicherheitsmanagements tätig.

Session 3

Maximilian Kliem, Wiebke Lenze, Andreas Pfeifer, Nils Nießen, Florian Horn, Felix Jendritzki, Melanie Bertl, Christina Jordanski, Susanne Koch

	Jendritzki, Felix
	Felix Jendritzki ist Mitarbeiter der DB InfraGO AG im Geschäftsbereich Personenbahnhöfe und arbeitet dort seit 2022 mit am Forschungsvorhaben Sicherheitsbahnhof.
	**Bertl, Melanie**
	Melanie Bertl ist seit 2024 wissenschaftliche Mitarbeiterin an der Frankfurter University of Applied Sciences an der School of Personal Development and Education (ScoPE) sowie am Fachbereich Wirtschaft und Recht. Sie studierte Soziologie mit Schwerpunkt auf Methoden der empirischen Sozialforschung an der Universität Bielefeld und an der Philipps-Universität Marburg.
	**Jordanski, Christina**
	Christina Jordanski ist seit 2024 wissenschaftliche Mitarbeiterin an der Frankfurt University of Applied Sciences am Fachbereich Wirtschaft und Recht. Das Studium der Soziologie mit Schwerpunkt auf Methoden der empirischen Sozialforschung absolvierte sie an der Universität Bielefeld und der Universität zu Köln.
	**Koch, Susanne**
	Prof. Dr. Susanne Koch ist seit 2008 Professorin für Allgemeine Betriebswirtschaftslehre und Logistik an der Frankfurt University of Applied Sciences und geschäftsführende Direktorin der School of Personal Development and Education. Nach dem Bergbaustudium an der Technischen Universität Clausthal und der Promotion an der Universität Bremen war sie in der Beratung und für Lufthansa Systems tätig.

Session 3
Sebastian Herwartz-Polster, Aaron Paz Martinez, Johannes Pagenkopf, Marc Sparsbrod, Felix Heizler

Der Höhenunterschied zwischen Fahrzeug und Bahnsteigkante. Erfassung der Einstiegssituationen im deutschen SPNV.

Herwartz-Polster, Sebastian[1], Paz Martinez, Aaron[1], Pagenkopf, Johannes[1], Sparsbrod, Marc[2], Heizler, Felix[2]

[1]Deutsches Zentrum für Luft und Raumfahrt (DLR) – Institut für Fahrzeugkonzepte
[2]Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt (DZSF)

Zusammenfassung

Selbstständig bedienbare fahrzeugseitige Einstiegshilfen können helfen, mobilitätseingeschränkten Reisenden den Ein- und Ausstieg flexibel an allen Verkehrsstationen ohne Unterstützung von Bahnpersonal zu ermöglichen. Um entsprechende Anwendungsfälle abzuleiten, ist es notwendig, Art und Ausprägung der höheninkompatiblen Einstiegssituationen im Schienenpersonennahverkehr zu erfassen, zu quantifizieren und zu bewerten. Um dies aussagekräftig und belastbar für ganz Deutschland zu tun, wurden Daten zu Bahnsteighöhen, Fahrzeugeinstiegshöhen und Fahrplänen akquiriert, miteinander verschnitten und ausgewertet. Das Ergebnis ist eine flächendeckende und fahrtengenaue Analyse der Höhenunterschiede zwischen Fahrzeug und Bahnsteigkante im deutschen SPNV.

Keywords: Barrierefreiheit, selbstständige Bedienbarkeit, Einstiegshilfe, Einstiegshöhe, Bahnsteighöhe, Einstiegssituationen, Kompatibilitätskarte, SPNV

Sebastian Herwartz-Polster, Aaron Paz Martinez, Johannes Pagenkopf, Marc Sparsbrod, Felix Heizler

1 Einleitung

Gemäß Artikel 9 der UN-Behindertenrechtskonvention (UN-BRK) (Vereinte Nationen, 2006) sind öffentliche Transportmittel für alle Menschen zugänglich zu gestalten. Dabei ist eine größtmögliche Unabhängigkeit aller Fahrgäste zu gewährleisten (Art. 20 UN-BRK). Eine zentrale Barriere stellt der häufig auftretende Höhenunterschied zwischen Bahnsteigen und Fahrzeugen dar. Das ist hauptsächlich auf die historisch gewachsenen, uneinheitlichen Bahnsteighöhen in Deutschland (von < 380 mm bis zu 1030 mm) und sich daraus je nach Gebiet und Einsatzzweck ergebenden unterschiedlichen Zielhöhen (550 mm und 760 mm (Kieffer, Ernst, & Jasper-Ottenhus, 2014)) zurückzuführen, die in Deutschland selbst innerhalb einzelner SPNV-Netze bzw. Linien variieren. Zudem variieren Einstiegshöhen, auch bei Fahrzeugen einer Fahrzeugplattform – abhängig von SPNV-Verkehrsnetz und Einsatzort. In der Folge ergeben sich häufig keine niveaugleichen Ein- und Ausstiege. Als niveaugleicher Einstieg gemäß TSI PRM (Europäische Kommission, 2014) gilt eine Einstiegssituation dann, wenn der Höhenunterschied zwischen Bahnsteigkante und Fahrzeugeinstieg maximal 50 mm, und der horizontale Spalt maximal 75 mm beträgt. Fahrzeugseitige Einstiegshilfen sind in Schienenfahrzeugen bereits als Hublifte oder Rampensysteme im Einsatz, die bisher umgesetzten Systeme erfordern jedoch stets die Bedienung durch eingewiesenes Personal der Eisenbahnverkehrsunternehmen. Der Einsatz selbstständig bedienbarer fahrzeugseitiger Einstiegshilfen kann dazu beitragen, mobilitätseingeschränkten Reisenden den Ein- und Ausstieg ohne weitere Unterstützung durch Personal oder Begleitpersonen zu ermöglichen. Um die technischen Anforderungen an solche Einstiegshilfen zu formulieren und entsprechende Anwendungsfälle abzuleiten, ist es erforderlich, Art und Ausprägung der höheninkompatiblen Einstiegssituationen im Schienenpersonennahverkehr quantitativ zu erfassen.

In der durch das DZSF beauftragten Studie „Selbständige Bedienbarkeit fahrzeugseitiger Einstiegshilfen“ werden Anforderungen an solche Einstiegshilfen definiert. In diesem Beitrag werden die Ergebnisse einer quantitativen Erfassung der Höhenunterschiede zwischen Fahrzeug und Bahnsteigkante im deutschen SPNV vorgestellt. Dies liefert Aufschluss über die generellen Bedarfspotenziale für fahrzeugseitige Einstiegshilfen

2 Methodik

Die Studie verfolgt einen datengetriebenen Ansatz um die Höhenunterschiede zwischen Fahrzeugeinstieg und Bahnsteighöhe (Höhendifferenz Bahnsteigkante zu

Schienenoberkante) in Deutschland zu erfassen. Dieser Abschnitt beschreibt die verwendeten Datenquellen und das Vorgehen um die Daten nutzbar zu machen und zu analysieren.

Für die Erfassung der Bahnsteighöhen wurde das Infrastrukturkataster (ISK-V) des Eisenbahnbundesamtes (Eisenbahn-Bundesamt, 2023) herangezogen. Die in Deutschland eingesetzten SPNV-Fahrzeuge und deren Netz- bzw. Linienzugehörigkeit wurden recherchiert und als Datensatz erstellt (eine genauere Beschreibung erfolgt weiter unten). Um die tatsächlichen Höhenunterschiede (an welchem Gleis hält das Fahrzeug) sowie deren Häufigkeit (wie oft hält das Fahrzeug an diesem Gleis) zu erfassen, wurden Bahnsteighöhen und Fahrzeugeinstiegshöhen anhand des deutschlandweiten Sollfahrplans (Delfi e.V., 2025) verknüpft. Diese Datensätze haben keine gemeinsamen Identifier, sind inkonsistent aufgebaut und weisen Datenlücken auf. Sie sind also zunächst inkompatibel. Um dem zu begegnen, wurden algorithmische Lösungen entwickelt. Abbildung 1 stellt die Prozessierung und Verknüpfung der Datensätze schematisch dar.

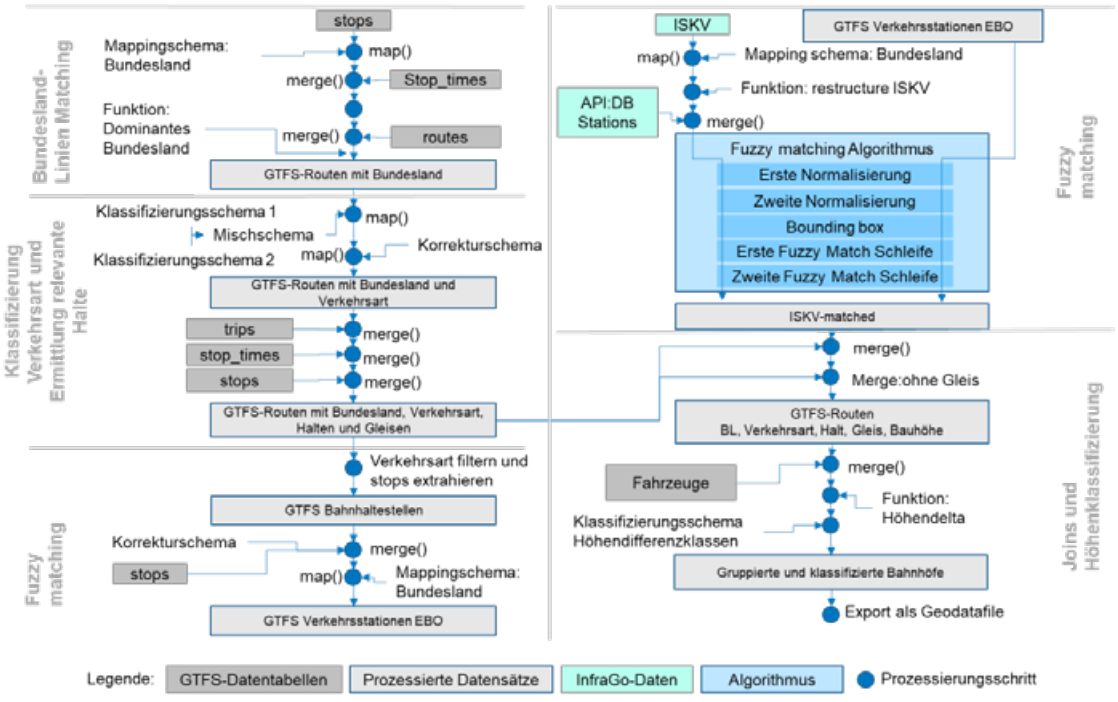


Abbildung 1: Schema der Verknüpfung der Datensätze für Fahrzeuge und Bahnsteighöhen mithilfe des Sollfahrplandatensatzes.

Beispielsweise ist die Kartierung der Verkehrsart im Sollfahrplan unzuverlässig (es gibt konkurrierende Kartierungsschemata in dem Datensatz, Schienenersatzverkehre sind mal als Busverkehre kartiert, mal als Schienenverkehr etc.). Daher wurde zunächst für alle

Halte die euklidische Entfernung zum Schienennetz ermittelt. Dafür wurden Schienennetzgeometrien aus OpenStreetMap (OpenStreetMap-Beitragende, 2025) verwendet. Anschließend wurden alle vorkommenden Routen mit deren Halten und den entsprechenden Distanzen zum Schienennetz ausgestattet. So konnten alle Linien ausgewählt werden, die vollständig am Schienennetz liegen. Dann wurde für jede Linie berechnet, in welchem Bundesland der Großteil der angefahrenen Verkehrsstationen liegt, um Linien mit Bundesländern zu attribuieren. Dies ist notwendig, da Linienbezeichnungen im Regionalverkehr in Deutschland nicht einheitlich sind.

Um die Bahnsteighöhen im ISK-V-Datensatz mit dem Fahrplan zu verbinden, mussten die uneinheitlichen Bezeichnungen der Verkehrsstationen harmonisiert werden. Daher wurde ein mehrstufiger Zuordnungsprozess entwickelt, der sowohl textuelle Ähnlichkeit als auch geografische Nähe berücksichtigt. Zunächst wurden die Stationsnamen beider Datensätze durch eine Normalisierungsfunktion bereinigt und eine geografische „Bounding Box“ mit einem Radius von zwei Kilometern berechnet. Anschließend erfolgte ein schrittweises, iteratives „Fuzzy Matching“. So wurden für jede SPNV-Linie im Fahrzeugdatensatz die passenden Fahrten aus dem Fahrplan, die Halte an den Bahnsteigen sowie die entsprechenden Bahnsteighöhen in einem Datensatz zugeordnet. Für jede Haltestelle pro Linie wurde die Häufigkeit der Halte pro Woche erfasst und für jeden Halt ein neuer Datenpunkt eingefügt.

Um die Fahrzeugeinstiegshöhen zu berücksichtigen, wurde ein Fahrzeugdatensatz erstellt, welcher Einstiegshöhen und Einstiegshilfen der im Einsatz befindlichen SPNV-Fahrzeuge beinhaltet. Im Folgenden wird dieser Datensatz spezifisch beschrieben. Der SPNV umfasst eine Vielzahl unterschiedlicher Fahrzeugtypen mit einer breiten Variation von Einstiegshöhen. Die Fahrzeuge stammen aus unterschiedlichen Jahren mit entsprechend verschiedenen Zulassungsbedingungen. Zudem unterscheiden sich regelmäßig die Fahrzeuganforderungen bei den netzspezifischen Bestellungen. So kommt es vor, dass Fahrzeuge einer Baureihe mit unterschiedlichen Einstiegshöhen ausgeliefert wurden und werden. Erschwerend kommt hinzu, dass nicht jedes Fahrzeug über seine Lebensdauer hinweg in dem Netz eingesetzt wird, für das es bestellt wurde. Da kein zentrales und umfassendes Kataster der Einstiegshöhen und eingesetzten fahrzeugseitigen Einstiegshilfen existiert, wurde eine umfassende Recherche von SPNV-Fahrzeugen und deren Einstieghöhen netz- bzw. linienspezifisch durchgeführt. Die Recherche erfolgte in zwei Phasen.

Zunächst erfolgte die Zuordnung von Fahrzeugen zu den Linien bzw. Netzen und den jeweiligen Verkehrsverträgen. Die Webseite „wettbewerb.spnv-deutschland.de“

(Rosenthal, 2009 - 2025) bietet eine nach Jahren sortierte detaillierte Auflistung vergangener und geplanter SPNV-Verkehrsverträge. Aus dieser Quelle wurden die SPNV-Netze für jedes Jahr im Zeitraum 2007-2029 entnommen, bereinigt, bei Bedarf um die Linienbezeichnung und das jeweilige Bundesland ergänzt und in eine konsolidierte Tabelle mit Daten nach Bundesland, Beginn und Ende der Laufzeit des Verkehrsvertrags, Eisenbahnverkehrsunternehmen, Linie, Streckenverlauf, sowie eingesetztem Fahrzeug zusammengeführt. Dabei wurden nur Verkehrsverträge berücksichtigt, deren Laufzeit nach 2024 endet.

Darauf folgte die Erhebung von Informationen zu Einstiegshöhen und Einstiegshilfen für jedes Fahrzeug in den jeweiligen Linien bzw. Netzen, basierend auf der zuvor vorgenommenen Zuordnung. Hierbei wurden Primärquellen von Fahrzeugherstellern und EVU verwendet (Siemens Mobility GmbH, 2024; Stadler Rail AG, 2016; DB Fernverkehr AG, 2021; DB Fernverkehr AG, 2025; Niederbarnimer Eisenbahn-AG, kein Datum). Die Nutzung von Internetarchiven über die Wayback Machine war von großer Bedeutung. Sekundärquellen wie Wikipedia und Foren wie „mobil.nrw“, „oepnv-info“, „hellertal.startbilder“ und „nahverkehr-franken“ (Wikipedia, o.J.; Krämer, 2025; Kollorz, 2024; Dittmer, 2016; Schwarz, 2022) kamen ebenfalls zum Einsatz.

Für 97 % der erfassten SPNV-Fahrzeuge konnten die Daten zur Einstiegshöhe dokumentiert werden, davon 70% mit Datensicherheit (DS) 1 (sehr sicher), 14 % mit DS 2 (sicher, wie z.B. bei Fahrzeugen des gleichen Modells innerhalb eines Netzes, in dem Daten zum jeweiligen Fahrzeugmodell einer Linie vorliegen) und 1,5 % mit DS 3 (relativ sicher, wie bei der Verwendung eines Zuordnungsschlüssels für bestimmte Doppelstockwagons). Bei 11,8 % der erfassten Linien waren die Fahrzeuge zum Abschluss der Analyse (April 2025) nicht zugewiesen.

Bei der Erfassung der fahrzeugseitigen Einstiegshilfen konnten für rund 44 % der erfassten Linien entsprechende Systeme nachgewiesen werden. Für die recherchierten Datenblätter von Fahrzeugen, die keine Angaben zum Einsatz von Einstiegshilfen enthalten, wird angenommen, dass in diesen Fahrzeugen keine entsprechenden Systeme verbaut sind. Dieser Fall betrifft etwa 28 % der Linien. Insgesamt konnten für 72 % der dokumentierten Fahrzeuge direkte Informationen zum Einbau oder Nicht-Einbau von Einstiegshilfen ermittelt werden. Der resultierende Datensatz beschreibt den Großteil des deutschen SPNV-Fuhrparks.

Aus der Verknüpfung der Datensätze der Bahnsteighöhen und der Fahrzeuge mittels des Fahrplans ergibt sich ein Datensatz der Einstiegssituationen in Deutschland. Dieser Datensatz hat die in Tabelle 1 dargestellte Struktur.

Tabelle 1: Schematische Darstellung des Ergebnisdatensatzes

Bundesland	SPNV-Linie	Bahnhof	Gleis	Bahnsteighöhe	Fahrzeug	Fahrzeug-einstiegshöhe
1	RE1	1	1	550 mm	BR1	600 mm
1	RE1	2	2	760 mm	BR1	600 mm
…	…	…	…	…	…	…
16	RE2	1	11	760 mm	BR2	800 mm
16	RE2	2	21	380 mm	BR2	800 mm

Um Implikationen und Muster in den Daten zu erkennen wurde das erstellte Datenprodukt mit Karten und Häufigkeitsstatistiken analysiert. Final wurden Anwendungsfälle abgeleitet. Die Ergebnisse dieser Analyse werden im folgenden Kapitel vorgestellt.

3 Höhenkompatibilität Bahnsteig-Fahrzeug

Das Ergebnis der angewandten Methodik ist eine konsistente und umfassende Datenbasis zur Kompatibilität zwischen Bahnsteigkante und Fahrzeugeinstieg in Deutschland. Anhand dieser Daten lassen sich liniengenau die Bahnsteighöhen quantifizieren, wie in Abbildung 2 am Beispiel Berlin-Brandenburg gezeigt werden kann. Hier sind alle SPNV-Linien im Betrachtungsraum dargestellt mit Ausnahmen der Linien, die zum Großteil im grenzüberschreitenden Ausland liegen bzw. bei denen zum Zeitpunkt der Datenerhebung Schienenersatzverkehre aktiv waren. Die grüne Einfärbung indiziert die Anzahl der Ein- und Ausstiege pro Tag. Von den 37 betrachteten SPNV-Linien haben nur fünf Linien eine einheitliche Bahnsteighöhe (in der Abbildung markiert durch einen roten Rahmen).

Die Grafik zeigt, wie häufig unterschiedliche Bahnsteighöhen auf einer Linie vorkommen und dass es auch zu vergleichsweisen großen Höhenunterschieden kommt. So werden beispielweise auf vier Linien Bahnsteige mit der Höhe 960 mm und Bahnsteige mit einer Höhe kleiner als 380 mm angefahren (RB24, RB27, RE4 und RE8) und auf 19 Linien Bahnsteighöhen von 760 mm und Bahnsteige mit Bahnsteighöhen kleiner 380 mm. Der Höhenunterschied zwischen den Bahnsteigen ist dort größer als 380 mm.

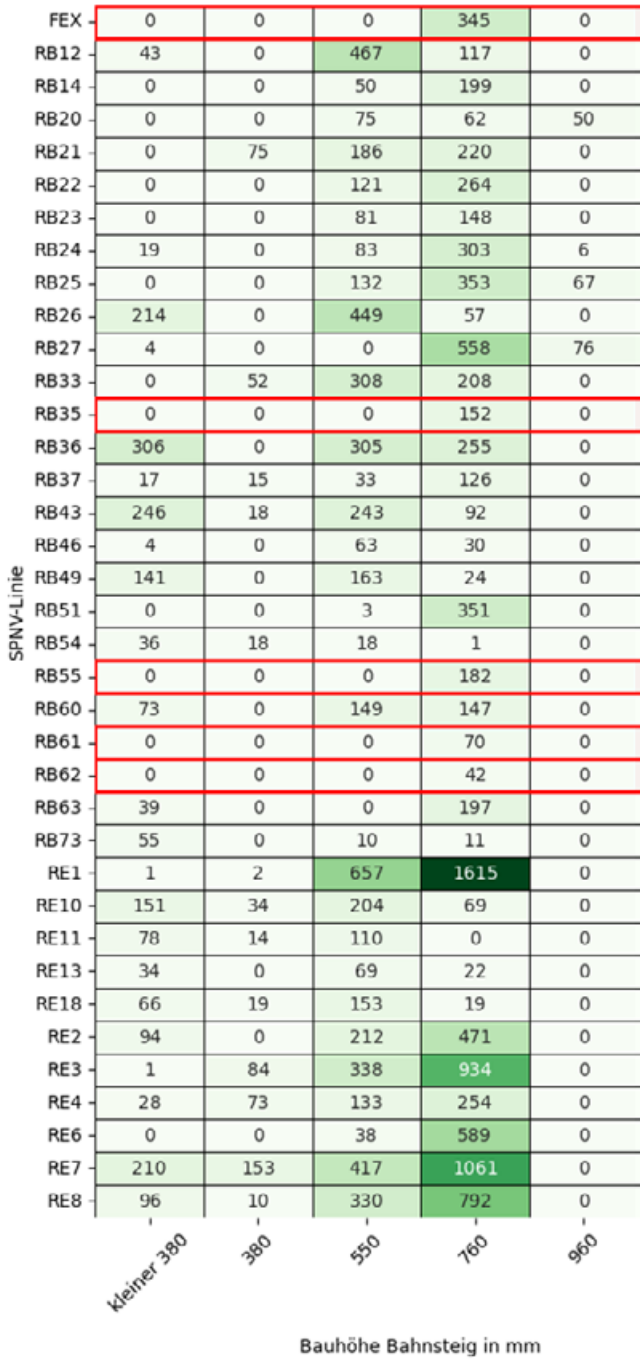


Abbildung 2: Anzahl von Einstiegssituationen pro Tag entlang von SPNV-Linien in Berlin und Brandenburg

Durch den Abgleich der Bahnsteighöhen mit den Einstiegshöhen der Fahrzeuge lässt sich zudem der Anteil niveaugleicher Ein- und Ausstiegssituationen für ganz Deutschland beschreiben (Tabelle 2). Fahrzeugeinstiegshöhen sind häufig nicht zu den

Bahnsteighöhen kompatibel. So machen Ein- und Ausstiege an Bahnsteigen mit einer Bauhöhe von 550 mm 24,9 % aller Ein- und Ausstiegssituationen aus, von denen 84,9 % niveaugleich sind. Häufiger kommen Ein- und Ausstiege an Bahnsteigen mit 760 mm vor (37,8 %). Diese sind jedoch nur zu 64,3 % niveaugleich. In der Gesamtheit beträgt der Anteil der niveaugleichen Ein- und Ausstiege ca. 54 %, wobei für 6,6 % der Gesamtheit der Ein- und Ausstiegssituationen keine Aussage getroffen werden kann, da die Bahnsteighöhe aufgrund von Datenlücken in den Quelldatensätzen nicht bekannt ist.

Tabelle 2: Anteil niveaugleicher Einstiege im deutschen SPNV

Bahnsteighöhe [mm]	Anteil an allen Einstiegen [%]	Davon niveaugleich [%]
< 380	7,1	0,0
380	10,2	1,4
550	24,9	84,9
760	37,8	64,3
960*	13,2	59,9

** Metropol-S-Bahn Netze Berlin, Hamburg, München ausgeschlossen.*

Einen Eindruck der räumlichen Verteilung der Höhenkompatibilität vermittelt Abbildung 3. Die Kompatibilitätskarte zeigt die Häufigkeit von niveaugleichen bzw. nicht-niveaugleichen Einstiegssituationen pro Verkehrsstation. Verkehrsstationen können eine Vielzahl von Einstiegssituationen aufweisen (verschiedene Bahnsteighöhen eines Bahnhofs und Einsatz verschiedener Fahrzeuge). Die Farbcodierung gibt an, wie groß der Anteil der höhenkompatiblen Ein- und Ausstiege an der jeweiligen Verkehrsstation ist.

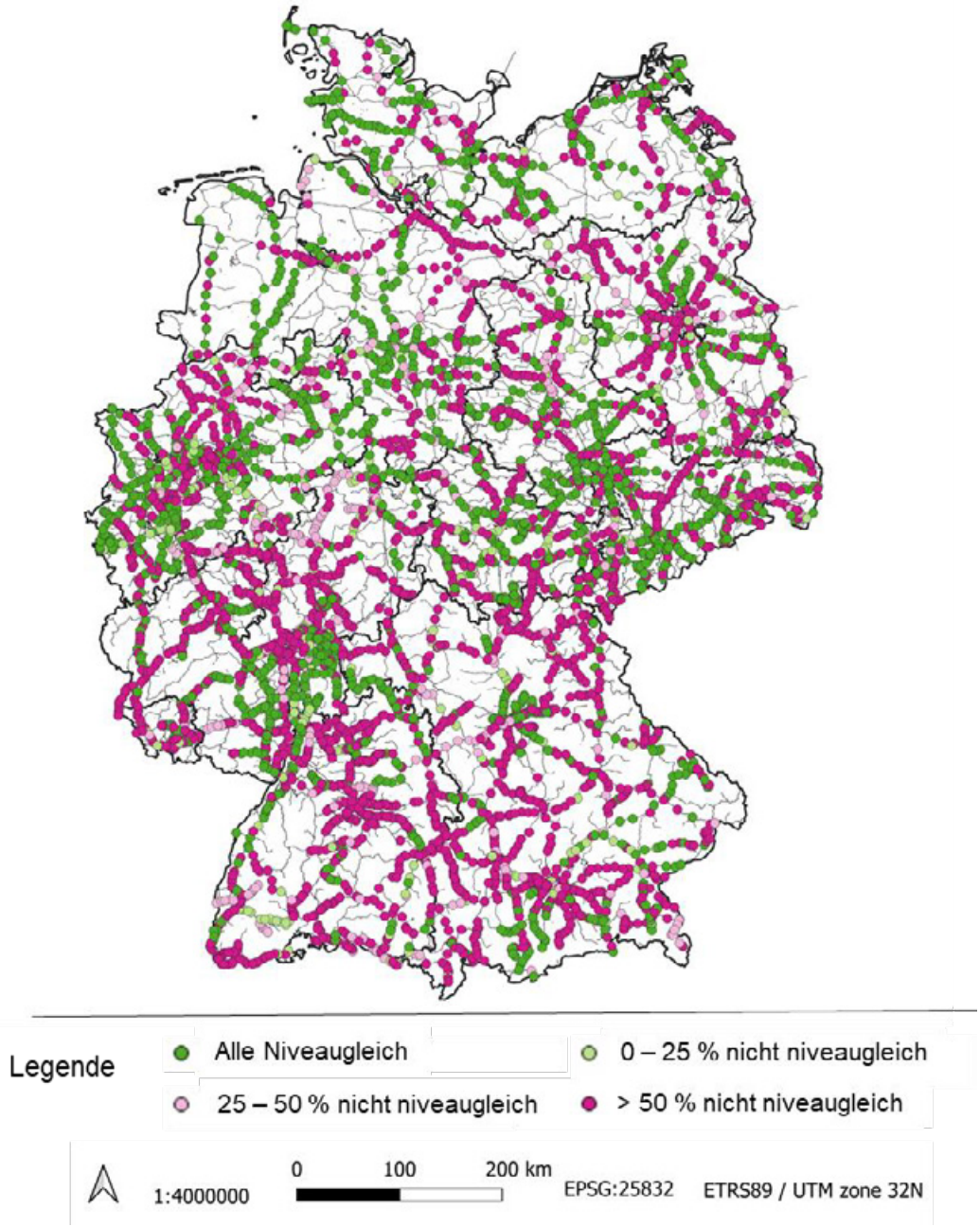


Abbildung 3: SPNV-Kompatibilitätskarte: Anteil der niveaugleichen Ein- und Ausstiege an den Verkehrsstationen.

Abbildung 4 zeigt die Häufigkeitsverteilung der Höhenunterschiede zwischen Fahrzeugeinstieg und Bahnsteigkante. Positive Vorzeichen bedeuten, dass die Fahrzeugeinstiegshöhe über der Bahnsteighöhe liegt (Einstieg nach oben). Negative Vorzeichen bedeuten, dass die Fahrzeugeinstiegshöhe kleiner als die Bahnsteighöhe ist. Die Rotfärbung markiert Höhendifferenzen, die entsprechend TSI-PRM die maximale

Stufenhöhe abbilden, also +230 mm und -160 mm. Diese Grenzen dienen als Richtwert zur Einordnung der Höhenunterschiede von EAS. Höhenunterschiede, die über diesem Grenzwert liegen, können Rollstuhlfahrende zumeist nur mit einem Hublift überwinden (Kieffer, Ernst, & Jasper-Ottenhus, 2014).

Die größte Gruppe von Höhenunterschieden ist mit 53,6 % die mit Ein- und Ausstiegen zwischen einschließlich -50 mm und +50 mm. Die Gruppe zwischen einschließlich -160 mm und -60 mm macht rund 8 % aus, die zwischen 60 mm und 230 mm rund 9,3 %. Die größeren Höhenunterschiede treten anteilig bei rund 22,5 % der EAS auf (ca. 5 % kleiner als -160 mm und 17,5 % größer als 230 mm).

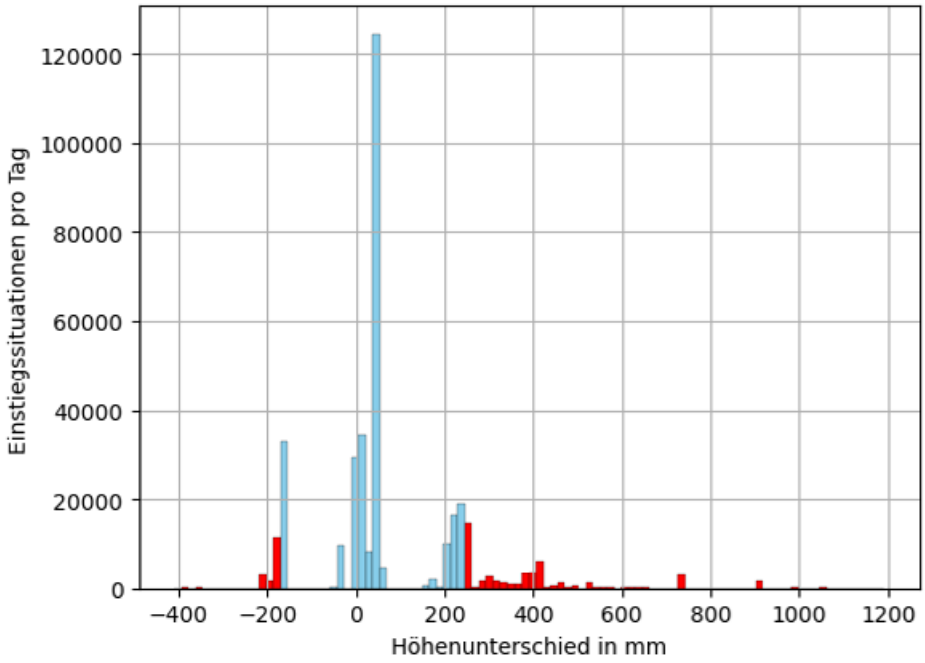


Abbildung 4: Häufigkeit und Ausprägung der Höhenunterschiede zwischen Bahnsteigkante und Fahrzeugeinstieg.

Abbildung 5 zeigt Einstiegssituationen von häufig im deutschen SPNV eingesetzten Baureihen. Dabei wird der Höhenunterschied zwischen Bahnsteig und Fahrzeugeinstieg auch hier entsprechend der TSI-PRM klassifiziert (±50 mm sowie +230 mm/-160 mm).

Fahrzeuge wie der Siemens Desiro HC und der Alstom Coradia Max mit ihren Kombinationen aus Einzel- und Doppelstockwagen und bis zu drei verschiedenen Einstiegshöhen im selben Zugverband bieten in etwa 79 % der Einstiegssituationen niveaugleiche Einstiege an. Etwa 8,7 % mehr könnten ggf. mit geeigneten Rampen

überwunden werden, und für die restlichen 12 % wäre der Einsatz von Hubliften oder Bahnsteiganpassungen notwendig.

Bei einstöckigen Fahrzeugen im SPNV variiert der Anteil der niveaugleichen Einstiege zwischen rund 50 % beim Alstom Lint 51 und 81 und rund 63 % beim Stadler FLIRT. In 49 % bzw. 32 % der Einstiegssituationen müssten umfassendere Maßnahmen ergriffen werden, um Barrierefreiheit zu erreichen, wie zum Beispiel die Nachrüstung mit Hubliften, die Anpassung der Bahnsteighöhe oder der Einsatz anderer Fahrzeuge mit niveaugleichen Ein- und Ausstiegen.

Auffällig ist der Stadler GTW, bei dem 48 % der EAS potenziell mit einem Rampeneinsatz barrierefrei werden, vorausgesetzt, dass eine Nachrüstung baulich möglich ist. Auch der Siemens Desiro Classic und der Stadler RegioShuttle RS1 hätten diese Möglichkeit in 17% bzw. 23 % der Fälle. Auch zu beachten sind Extremfälle wie die DB-Baureihe 612 (Neigetechnik-Dieseltriebzüge), die bauartbedingt eine Einstiegshöhe von 1.290 mm aufweist. Für den Einstieg in dieses Hochflurfahrzeug müssen - wie auch im Fernverkehr üblich - Stufen überwunden werden. Hier können ggf. Hublifte verwendet werden.

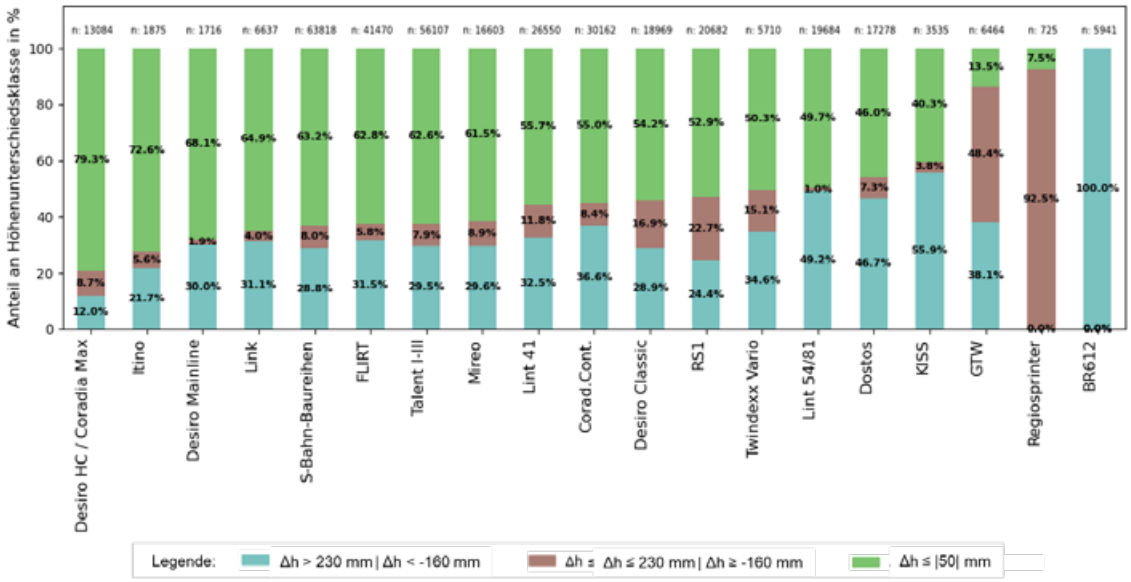


Abbildung 5: Verteilung der Einstiegssituationen in Deutschland für ausgewählte SPNV-Fahrzeuge. n = Anzahl der erfassten Einstiegssituationen pro Tag.

4 Diskussion

Aus der quantitativen Analyse des Höhenunterschieds zwischen Fahrzeug und Bahnsteigkante wird ersichtlich, dass es im deutschen Eisenbahnnetz zahlreiche

Höhenunterschiede gibt und ein großer Anteil der SPNV-Linien keinen durchgehend niveaugleichen Ein- und Ausstieg aufweist. Aus den Daten lassen sich, unter den zuvor getroffenen Annahmen der überwindbaren Höhendifferenzen mit Rampe und Hublift, Anwendungsfälle für verschiedene Systeme von Einstiegshilfen ableiten (Abbildung 6). Ein Anwendungsfall beschreibt die Höhenunterschiede zwischen Fahrzeugeinstieg und Bahnsteigkante und deren Ausprägung entlang von SPNV-Linien. Die prozentuale Häufigkeit der Anwendungsfälle (Abbildung 6 rechts) entspricht dem Anteil der SPNV-Linien mit den beschriebenen Höhenunterschieden.

Anwendungsfall	Höhendifferenz zw. Fahrzeug und Bahnsteig
1 – Geringe Höhenunterschiede	$\|\Delta h\| \leq 50$ mm
2 – Homogen moderate Höhenunterschiede	-160 mm $\leq \Delta h \leq 230$ mm Bei durchgängig gleicher Höhendifferenz
3 – Moderate Höhenunterschiede	-160 mm $\leq \Delta h \leq 230$ mm
4 – Durchgängig große Höhenunter-schiede	$(\Delta h > +230$ mm$) \mid (\Delta h < -160$ mm$)$
5 – Große und geringe Höhenunter-schiede	$(\Delta h > +230$ mm$) \mid (\Delta h < -160$ mm$)$ & $(\|\Delta h\| \leq 50$ mm$)$
6 - Große und moderate Höhenunterschiede	$\|\Delta h\| > 50$ mm
7 – Alle Höhenunterschiede	Δh = alle vorkommenden Höhenunterschiede

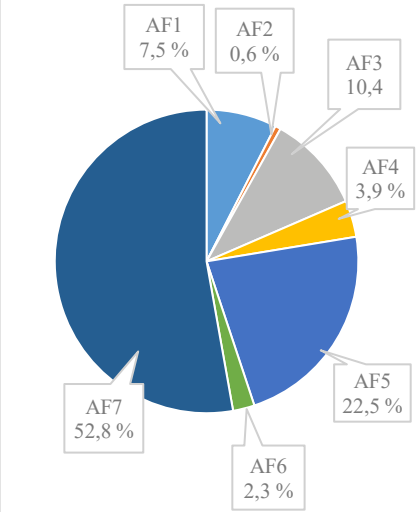


Abbildung 6 Anwendungsfälle verschiedener Einstiegshilfen und Häufigkeit ihres Auftretens in Deutschland

Nur wenige Fahrzeuge kommen mit horizontalen automatischen Spaltüberbrückungen aus, da SPNV-Linien selten ausschließlich niveaugleiche Einstiege aufweisen (Anwendungsfall 1). Um mobilitätseingeschränkten Fahrgästen eine selbstständige Mobilität im deutschen Schienenpersonenverkehr zu garantieren, ist der Einsatz von selbstständig bedienbaren Rampen und/oder Hubliften bei einem Großteil der Fahrzeuge notwendig (Anwendungsfälle 2-7). Auf Linien, die anteilig mindestens einige nicht-

niveaugleiche Ein- und Ausstiegssituationen aufweisen, ist der Einsatz fahrzeugseitiger Einstiegshilfen dabei naheliegend.

Zu diskutieren sind die hier zu Grunde gelegten Höhendifferenzen sowohl für einen niveaugleichen Einstieg (± 50 mm), als auch für den Einsatz von Rampen (-160 mm bis +230 mm) hinsichtlich der tatsächlichen Zumutbarkeit für mobilitätseingeschränkte Fahrgäste. Die Anpassung dieser Grenzwerte ist Bestandteil einer späteren Phase des Forschungsprojekts und erfordert je nach Ergebnis eine Anpassung der Schwellenwerte für die quantitative Erhebung der Anwendungsfälle der Systeme Rampe und Hublift.

Zudem sei darauf hingewiesen, dass sich die dargestellten Ergebnisse überwiegend auf die verfügbaren Infrastrukturdaten der DB InfraGO beziehen. Diese machen einen erheblichen Anteil am deutschen Eisenbahnnetz aus und belegen bereits die Notwendigkeit für selbstständig bedienbare fahrzeugseitige Einstiegshilfen.

Auch bei der zukünftigen Beschaffung von Fahrzeugen ist zu berücksichtigen, dass trotz Plänen zur Harmonisierung der Bahnsteighöhen nicht unmittelbar mit einer einheitlichen Bahnsteighöhe von 760 mm gerechnet werden kann (DB Fernverkehr AG, 2021) (DB Fernverkehr AG, 2025). Umbauten mit 550 mm Bahnsteighöhe als Zwischenstufe sind und werden umgesetzt. Ein niveaugleicher Einstieg muss also in vielen Fällen an mehreren möglicherweise unterschiedlich hohen Einstiegen gewährleistet sein.

5 Ausblick

Aus der Analyse der derzeitigen Einstiegssituationen und deren Inkompatibilitätsgraden ergibt sich klarer Handlungsbedarf. Erforderliche Maßnahmen umfassen die Harmonisierung der Bahnsteighöhen, die Nachrüstung bestehender Fahrzeuge mit geeigneten Einstiegshilfen wie Rampen oder Hubliften sowie im Rahmen von Neubeschaffungen den Einsatz von Fahrzeugen ohne Höhendifferenz zu den Bahnsteigen. Fahrzeuge und Bahnsteige mit mehreren Einstiegs- und/oder Bahnsteighöhen können hier eine wichtige Rolle spielen.

Die vom DZSF beauftragte Studie, in der sich diese Analyse einbettet, konzentriert sich auf selbstständig bedienbare fahrzeugseitige Einstiegshilfen, die es mobilitätseingeschränkten Menschen ermöglichen sollen, ohne fremde Hilfe ein- und auszusteigen. In den nächsten Schritten werden in dem Projekt der aktuelle Stand der Technik fahrzeugseitiger Einstiegshilfen untersucht sowie eine kollaborative Anforderungserhebung an selbständig bedienbare fahrzeugseitige Einstiegshilfen mit verschiedenen Interessengruppen durchgeführt.

Literatur

DB Fernverkehr AG. (2021). *Das 4. Programm zur Barrierefreiheit der Deutschen Bahn AG.*

DB Fernverkehr AG. (2025). *5. Programm zur Barrierefreiheit der Deutschen Bahn AG.* Detusche Bahn AG.

Delfi e.V. (2025). Deutschlandweite Sollfahrplandaten (GTFS). Abgerufen am 31.03.2025 von https://www.opendata-oepnv.de

Dittmer, C. (2016). *deutsche-reisezugwagen.de*. Abgerufen am 30.01.2025 von https://www.deutsche-reisezugwagen.de/wagendaten/785-bdnrzf-dabpza/

Eisenbahn-Bundesamt. (2023). ISK-V - Infrastrukturkataster für die Veröffentlichung. Abgerufen am 30.01.2025 von https://www.eba.bund.de/DE/Themen/Finanzierung/LuFV/ISK-V/isk-v_node.html

Europäische Kommission. (2014). Anwendbare Technische Spezifikation für die Interoperabilität EU 1300/2014 - Zugänglichkeit für Menschen mit Behinderung und Menschen mit eingeschränkter Mobilität - (PRM).

Kieffer, E., Ernst, J., & Jasper-Ottenhus, C. (05 2014). Das Bahnsteighöhenkonzept der DB AG. *Eisenbahntechnische Rundschau*(2014 Nr.5), S. 36-42.

Kollorz, W. (2024). *nahverkehr-franken.de*. Abgerufen am 30.1.2025 von https://www.nahverkehr-franken.de/rbahn/dosto-techdat.html

Krämer, D. (2025). *oepnv-info.de*. Abgerufen am 30.01.2025 von https://www.oepnv-info.de/reisen/bahnreisen/informationen-zur-barrierefreiheit-und-ausstattung-von-zuegen/fahrzeuginformationen-nahverkehr

Niederbarnimer Eisenbahn-AG. (kein Datum). *neb.de*. Abgerufen am 30.01.2025 von https://www.neb.de/unternehmen/fahrzeuge/regio-shuttle-rs1/

OpenStreetMap-Beitragende. (2025). OpenStreetMap.

Rosenthal, K. (2009 - 2025). *wettbewerb.spnv-deutschland.de*. Abgerufen am 30.01.2025 von https://www.wettbewerb.spnv-deutschland.de/

Schwarz, A. (2022). *hellertal.startbilder.de*. Abgerufen am 30.01.2025 von https://hellertal.startbilder.de/bild/deutschland~triebzuege~br-650-stadler-regio-shuttle-rs1/794843/zwei-gekuppelte-stadler-regio-shuttle-rs1-bzw.html

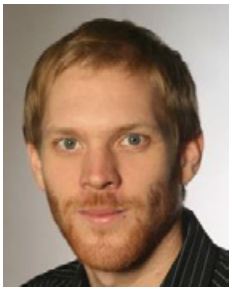
Siemens Mobility GmbH. (2024). *bwegt.de*. Abgerufen am 30.01.2025 von https://www.bwegt.de/fileadmin/media/PDF/Pressemitteilungen/2024/Anlage_Datenblatt-Mireo-B-Ortenau_08.04.2024.pdf

Stadler Rail AG. (Mai 2016). Abgerufen am 30.01.2025 von https://web.archive.org/web/20160524074030/http://www.stadlerrail.com/de/fahrzeuge/gelenktriebwagen-gtw

Vereinte Nationen. (2006). Übereinkommen über die Rechte von Menschen mit Behinderungen (UN-BRK). Übersetzung der Bundesrepublik Deutschland. New York.

Wikipedia. (o.J.). Abgerufen am 30.01.2025 von https://de.wikipedia.org/wiki/Bombardier_Talent_2

Autoren



Herwartz-Polster, Sebastian

Sebastian Herwartz-Polster, M.Sc. ist Geograf und ist beim DLR-Institut für Fahrzeugkonzepte als Wissenschaftlicher Mitarbeiter in der Abteilung Fahrzeugsysteme und Technologiebewertung. tätig. Der Schwerpunkt seiner Arbeit liegt im Bereich der Infrastrukturbewertung und in der Bewertung von Schienenfahrzeugen mit alternativen Antrieben.



Paz Martinez, Aaron

Dipl.-Ing. Paz Martinez ist Ingenieur für Maschinenbau und ist beim DLR-Institut für Fahrzeugkonzepte als Wissenschaftlicher Mitarbeiter in der Abteilung Fahrzeugsysteme und Technologiebewertung tätig. Der Schwerpunkt seiner Arbeit liegt im Bereich neuartige Schienenfahrzeugkonzepte.



Pagenkopf, Johannes

Johannes Pagenkopf, M.Sc., ist Wirtschaftsingenieur und ist beim DLR-Institut für Fahrzeugkonzepte als Leiter einer Forschungsgruppe in der Abteilung Fahrzeugsysteme und Technologiebewertung tätig. Der Schwerpunkt seiner Arbeit liegt im Bereich neuer Technologien für Schienenfahrzeuge.



Sparsbrod, Marc

Dipl.-Ing. Marc Sparsbrod ist Ingenieur für Schienenfahrzeugtechnik und arbeitet als wissenschaftlicher Referent und Doktorand am Deutschen Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt (DZSF). Er betreut den Themenbereich Barrierefreiheit.



Heizler, Felix

Felix Heizler, M.Sc., ist Verkehrsingenieur und arbeitet als wissenschaftlicher Referent am Deutschen Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt (DZSF). Er betreut die Themenbereiche ÖPNV und Automatisierung.

IRSA 2025

SESSION 4

GÜTERVERKEHR

FREIGHT TRAFFIC

Session 4
Manfred Enning, Bernd Schmidt

Baukasten für einen zukunftsfähigen Schienengüterverkehr in der Fläche

Enning, Manfred und Schmidt, Bernd

FH Aachen University of Applied Sciences, Aachen

Zusammenfassung

Die derzeitige Entwicklung des Einzelwagenverkehrs erfordert schnelle Antworten auf viele Fragen der Logistik, des Betriebs aber auch der personalsparenden Technik. Aktuell ist die Diskussion stark auf Technik der Digitalen Automatischen Kupplung (DAK) und deren bahnbetrieblichen Nutzen fokussiert. Im Beitrag wird der Einfluss der DAK im Zusammenspiel mit weiteren heute verfügbaren Technologien auf die Effizienz des Schienengüterverkehrs diskutiert. Dabei werden Szenarien betrachtet, die neben dem Hauptlauf vor allem auch die betrieblichen Nebenprozesse und die Schnittstelle zur Kundenlogistik umfassen. Hieraus werden Anforderungen an die weitere technische Entwicklung abgeleitet.

Keywords: Schienengüterverkehr, Logistik, Digitale Automatische Kupplung

1 Einleitung

Der Schienengüterverkehr (SGV) steht unter massivem betriebswirtschaftlichen und operativen Druck. Mit einem Anteil am Umsatz der Logistikbranche von ca. 2 % [8] mangelt es auch am für technische und betriebliche Innovationen erforderlichen Ertrag. Die für DB Cargo geforderte „Schwarze Null“ Ende 2026 [18] lässt zumindest für den Einzelwagenverkehr eher einen Kahlschlag als eine Ausweitung des Volumens erwarten, so dass der Modalsplit-Anteil des SGV von aktuell ca. 19 % [11] tendenziell abnehmen wird, mit allen Konsequenzen für die CO_2-Bilanz des Verkehrs.

Hoffnungsträger könnte die Domäne sein, die die Bahn in der Vergangenheit sehr gut beherrschte, nämlich die europaweite Vernetzung von Herstellern hochwertiger Kaufmannsgüter mit ihren Lieferanten und ihren Kunden. Gerade in diesem Segment hat die Güterbahn aber in den letzten Jahrzehnten am meisten Zustimmung verloren, weil sie die für die Verkettung notwendige Zuverlässigkeit nicht mehr bieten konnte und sie aufgrund veralteter Technik und Prozesse zu kompliziert, unflexibel und teuer geworden ist.

Die Digitale Automatische Kupplung DAK [3] ist als Retter des Einzelwagenverkehrs in aller Munde, weil sie ein Fülle von Problemen lösen kann. Sie schafft aber in der Übergangszeit eine Menge neuer Probleme und kann ihre Wirkung im Einzelwagensystem erst am Ende einer längeren Migrationsphase erbringen. Gleichzeitig verteuert sie jeden Wagen erheblich, so dass eine Wirtschaftlichkeit bei heutigen Quelle/Senke-Beziehungen und heutigen betrieblichen Prozessen über zentrale Zugbildungsanlagen nicht gesichert ist.

2 Bausteine eines SGV-Systems für die Fläche

2.1 Allgemeines und Beispielszenario

Ein schlichter Austausch der Schraubenkupplung gegen die Digitale Automatische Kupplung im klassischen Einzelwagensystem ist aus Sicht der Autoren nicht ausreichend, um die Flächenbedienung im Schienengüterverkehr wettbewerbsfähig zu machen. Auch die mit der DAK verbundenen Möglichkeiten zur weitergehenden Digitalisierung[1] der Zugbildungsprozesse (Abbildung 1) und der Zugfahrten reichen nicht aus, wenn Be- und Entladung sowie Bedienfahrten in der althergebrachten Weise erfolgen und Transporte über die wenigen Großknoten im Einzelwagennetz geroutet werden [2].

Im Folgenden wird versucht, aus einer ganzheitlichen Betrachtung von Transportprozessen abzuleiten, wie ein neues Produktionssystem im SGV als Ersatz des Einzelwagensystems gestaltet sein müsste und welche Anforderungen an die technischen Fähigkeiten der Güterwagen sich daraus ergeben.

Dazu wird die „Reise“ eines gedeckten zweiachsigen Güterwagens (z.B. Bauart Hbbillns 305) in einem möglichen Umlauf mit den Stationen gemäß Tabelle 1 betrachtet. Als Beispiel dient eine Logistik-Kette im Bereich der Lebensmittelindustrie. Ein Unternehmen

[1] Das Foto zeigt eine DAK Stufe 5 mit Pushbutton-Entriegelung, Strom- und Datenkupplung. „DAK Erprobung bei WLE“ am 10.07.2025

Session 4
Manfred Enning, Bernd Schmidt



Abbildung 1: Digitale Automatische Kupplung

Tabelle 1: Wagenumlauf-Stationen

Nr	**Prozessschritt**
1	Beladen/ Entladen, Schnittstelle zu automatischen Ladesystemen
2	Wagenbewegung im Anschluss, Selbstorganisation der Wagen
3	Wagenaustausch Zustellung/ Abholung, Briefkastenbedienung
4	Sortierung nach Destinationen, Flachrangieren mit angetriebenen Wagen
5	Wartung, Strategie für „High-Tech“ im Güterwagen

A stellt aus landwirtschaftlichen Grundstoffen ein Vorprodukt her, welches in einem Unternehmen B zusammen mit vielen weiteren Zutaten zu einem handelsfertig verpackten Lebensmittel verarbeitet wird. Die kleinste betrachtete Gebindegröße ist die Europalette (EPAL, 1200 x 800 mm) und aus Gründen der Vereinfachung wird unterstellt, dass Kühlung auf den Transportwegen nicht erforderlich ist.

Unternehmen A bezieht seine Rohstoffe von landwirtschaftlichen Betrieben in seinem näheren Umfeld. Zusatzstoffe für die Herstellung und Leerbehälter erhält es als palettiertes Transportgut zum Teil über die Straße. Unternehmen B erhält seine Rohstoffe über Straße und Schiene. Fertigprodukte werden über die Schiene in entfernt liegende Warenverteilzentren und über die Straße an Supermärkte in der Umgebung ausgeliefert.

2.2 Schnittstelle zur kundenseitigen Logistik

Beide Unternehmen benötigen Straßen- und Schienenanbindung. Die dominante Transportart „Straße“ definiert in aller Regel die innerbetriebliche Logistik. Hier hat sich eine Zeilenstruktur mit vielen nebeneinanderliegenden Lkw-Laderampen etabliert. Die bis heute häufigste Art der Beladung von Lkw-Sattelaufliegern erfolgt mit Gabelstaplern, wobei jeweils 1 bis 3 Europaletten aufgenommen werden. Für eine vollständige Beladung mit 33 Europaletten werden üblicherweise ca. 40 Minuten benötigt.

In manchen Bereichen etablieren sich gerade Automatic Truck Loading Systems (ATLS). Bei diesen werden die Paletten innerhalb der Logistikhalle bereits in der richtigen Anordnung für den Lkw-Versand (11 Reihen mit je 3 Paletten längs bzw. 16 Reihen mit je 2 Paletten quer) gestellt. Nach dem Andocken des Lkw an die Laderampe erfolgt die eigentliche Verladung mit ca. 6 m/min. also in 2-3 Minuten. ATLS aber auch Lösungen mittels Fahrerlosen Transportsystemen (FTS) verbreiten sich gerade schnell und das Marktvolumen wächst rasant[16]. Sie verschaffen dem Straßengüterverkehr einen gewaltigen Rationalisierungsschub. Anwendungen finden sich heute vorwiegend bei Shuttle-Verkehren (z.B. Direktverkehr zwischen Brauerei und einem entfernten regionalen Auslieferungslager in der Getränkeindustrie).

Genau solche Relationen können ein prädestiniertes Anwendungsfeld für einen leistungsstarken Quelle/Senke-Schienenverkehr sein. Wenn hier die Beladung aber in herkömmlicher Weise mit Gabelstaplern von der Seite erfolgen müsste, wäre die Schiene gegenüber

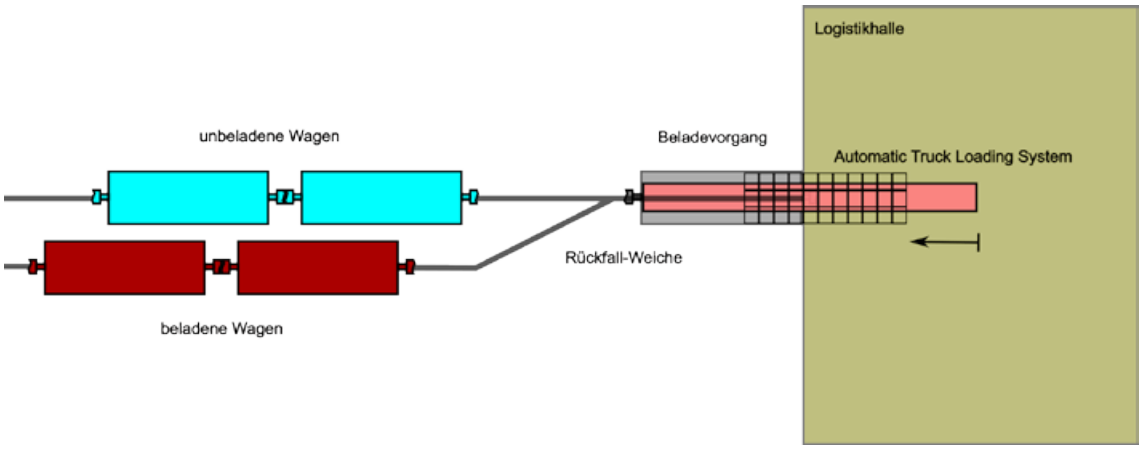


Abbildung 2: Kopframpenbedienung mit ATLS

der rationalisierten Straße chancenlos. Der oben genannte gedeckte Güterwagen Hbbillns 305 ist hinsichtlich der Ladefläche exakt auf den Lkw-Sattelauflieger abgestimmt. Ideal wäre, wenn der Güterwagen mit einem existierenden ATLS über die Stirnseite beladen werden könnte. Aber auch spezielle Beladesysteme für die Seiten- oder Stirnseitenbeladung vierachsiger gedeckte Güterwagen (66 EPAL Kapazität) können Wettbewerbsvorteile für die Schiene bieten.

Um einen Güterwagen für die moderne automatisierte Logistik auszurüsten, muss er über fahrzeugseitige Komponenten, wie z.B. Steuerungen, Fördersysteme, automatische Verriegelungs- und Ladungssicherungssysteme etc. verfügen. Diese müssen mit Strom versorgt und gesteuert werden. Wenn sich eine Beladung über die Stirnseite als vorteilhaft erweisen sollte, müssen Güterwagen baulich so gestaltet werden, dass sie an eine Laderampe andocken können und sie müssen idealerweise in einem automatisierten Prozess an der Rampe ausgetauscht werden (Abbildung 2).

2.3 Wagenbewegung im Anschluss

Es zeigt sich schnell, dass die Automatisierung des Wagenaustauschs für eine Kopframpenbedienung weder mit Lokomotiven noch mit Rangierrobotern oder Seilzuganlagen sinnvoll gestaltet werden kann. Immer wären An- und Abkuppelvorgänge notwendig. Selbst mit der ferngesteuert entkuppelbaren DAK wäre dies ein umständlicher und - wegen der Notwendigkeit einer Mindestgeschwindigkeit zum Kuppeln - sehr lautstarker Prozess. Wenn die Wagen aber die Fähigkeit hätten, sich selbst mit geringer Geschwindigkeit zu bewegen und

zwischen den Bewegungsphasen sicher zu stehen, wäre der Aufwand zur Automatisierung überschaubar. Aber auch für viele andere Aufgaben im Bereich einer Werksbahn oder eines Industrieanschlusses wäre ein kleiner (abkuppelbarer) Rangierantrieb in Verbindung mit einer automatischen Parkbremse vorteilhaft.

Für solche Rangierantriebe, die entweder gleichzeitig die Funktion einer Bremse übernehmen oder das pneumatische System des Wagens zur sicheren Bremsung nutzen sind unterschiedlichste Konzepte entwickelt worden. Serienreif – wenn auch vergleichsweise teuer – sind Systeme der US-amerikanischen Hersteller Parallel Systems[14] und Intramotev[9]. Nevomo[12] bietet in Europa ein in die Infrastruktur integriertes System an und auch das DLR entwickelt derzeit angetriebene Güterwagen.

Antriebe und Bremsen können unter Nutzung konventioneller Techniken manuell ferngesteuert betrieben werden. Für eine weitere Automatisierung der Prozesse zwischen Logistikhalle und Gleisanschluss bieten sich autonome, agentenbasierte Konzepte an. Jeder Wagen kennt seine „Mission“, z.B. zu einer bestimmten Zeit am ATLS zu sein und Ladung aufnehmen, dann zum Anschlussgleis fahren und sich in die Wagengruppe einzureihen. Dabei organisiert er sich im Sinne des „Internet der Dinge“ [15] einzeln oder in Gruppen selbständig und nutzt Komponenten der Infrastruktur, wie z.B. Weichen.

Hindernisse im Fahrweg müssen sicher erkannt werden. Bordgestütze Systeme, z.B. Lidar können den Fahrweg überwachen, sind aber recht aufwändig und während der späteren Fahrt im Zugverband nutzlos. Eine Lösung, bei der der Gefahrenraum von einem Tower aus oder mittels Drohnen überwacht wird, kann eine wirtschaftliche Alternative sein.

2.4 Zustellung/ Abholung

Die Zahl der Gleisanschlüsse ist von fast 12.000 im Jahr der Bahnreform 1994 auf gut 2.300 im Jahr 2021 gesunken. Zu den Gründen gehört neben der geringer werdenden Akzeptanz bei Verladern auch ein sehr hoher bahnbetrieblicher Aufwand für die Bedienung von Gleisanschlüssen, insbesondere von Gleisanschlüssen „an der freien Strecke“, die nicht direkt an einen Bahnhof angebunden sind. In sehr vielen Fällen erfolgt die Bedienung in der Form von „Sperrfahrten“. Sperrfahrten sind Hin- und Rückfahrten auf einem für „Zugfahrten“[2] gesperrten Gleis mit geringer Geschwindigkeit. Sie kosten in erheblichen Maße

[2]Zugfahrten erfolgen immer zwischen zwei Bahnhöfen

Trassenkapazität bei einer (in tkm gemessenen) sehr geringen verkehrlichen Leistung [13]. Der Wegfall von Bedienfahrten ermöglicht in vielen Fällen überhaupt erst einen stabilen Takt im Personenverkehr ohne kosten- und zeitaufwändige Ausbauten der Infrastrutkur und lag damit oft im Interesse der Eisenbahninfrastrukturunternehmen (EIU) und der Eisenbahnverkehrsunternehmen (EVU) im Personenverkehr.

Starke Quellen und Senken des Güterverkehrs (Produzierende Unternehmen, Logistikzentren) entstanden in den letzten Jahrzehnten vorwiegend „auf der grünen Wiese“ mit kurzer Straßenanbindung an die nächste Autobahnauffahrt. Bahnanschlüsse und die Lage des nächsten Bahnhofes verloren an Bedeutung. Wenn man Verkehre für die Schiene zurückerobern will, muss man Anschlüsse an der freien Strecke reaktivieren oder neu bauen. Weil sie meistens aus der Projektierung der Stellwerke herausgefallen sind, erfordert dies einen immensen Aufwand für Änderungen an Stellwerken und Zulassungsverfahren.

Im Rahmen des Projekts „Güterwagen 4.0“ wurden auch Konzepte für eine Anschlussbedienung untersucht, die betrieblich einem Stopp einer Zugfahrt auf der freien Strecke entspricht (Abbildung 3). Zentrales Merkmal ist, dass unmittelbar neben dem Streckengleis liegende „Briefkästen“ ohne Fahrten in die Verästelungen des Anschlusses bedient werden. Es gibt hier keine „Rangierabteilung“; der die Wagen bringende oder holende Zug bleibt auf der Strecke und setzt auch nicht zurück.

Vor und nach der Anschlussbedienung bilden die intelligenten „Güterwagen 4.0“ einen auf Zugintegrität und Bremswirkung geprüften Zug, der mit der Streckengeschwindigkeit z.B. zwischen zwei Takten des SPNV verkehrt [5]. Die Anschlussweiche muss in einem solchem Konzept nicht in die Stellwerksprojektierung einbezogen werden, weil in jeder Phase der Anschlussbedienung die Strecke aus der Sicht des Stellwerks als „belegt“ erscheint und Fahrten ins besetzte Gleis technisch ausgeschlossen sind.

Es ist lediglich ein zusätzlicher Achszählschienenkontakt für das Anschlussgleis zu ergänzen[13]. Dieses erfordert in der Regel keine Projektierungsänderung an der Stellwerkslogik, insbesondere nicht bei Stellwerken, die über getrennte Achszählanlagen (z. B. Motorzählwerke) verfügen. Auch die ordnungsgemäße Weichenlage der Anschlussweiche wird über diesen Kontakt signaltechnisch sicher abgebildet (Befahrmeldung). Allerdings muss das derzeitige Vorschriftenwerk für Zugfahrten angepaßt werden, um ein „Flügeln“ außerhalb eines Bahnhofs zu ermöglichen. Derartige Vorgehensweisen sind aber aus der Vergangenheit und aus dem Ausland bekannt und wurden regelmäßig bei gemischten Zügen

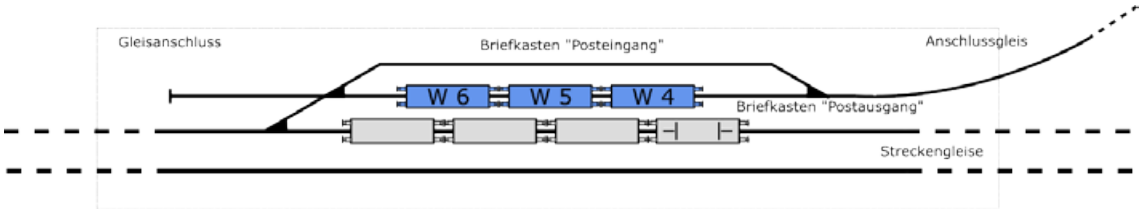


Abbildung 3: Gleisanschlusbedienung mit „Güterwagen 4.0“. Briefkastenbedienung

eingesetzt.

2.5 Sortierung nach Destinationen

Das bestehende Einzelwagensystem befindet sich leider in einer Abwärtsspirale. Solange das Wagenaufkommen groß und Personal im Rangierdienst preiswert verfügbar war, konnte ein engmaschiges Netz mit Netzzugangspunkten und Zugbildungsanlagen aufrecht erhalten werden. Mit dem Wegfall von Gleisanschlüssen fielen auch Satellitenbahnhöfe weg und das schrumpfende Wagenaufkommen zwang in Verbindung mit hohen Personalkosten zur Konzentration der Zugbildungsaufgaben auf wenige hoch automatisierte Anlagen. Auch heute ist es noch möglich, im Einzelwagenverkehr direkt zu fahren. Die allermeisten Relationen werden aber über die Rangierbahnhöfe geroutet mit der Folge mitunter großer Umwege. Damit schrumpft die wirtschaftliche Wettbewerbsfähigkeit aber auch - was gerne verschwiegen wird - der Umweltvorteil.

Im Einzelwagenverkehr frei werdende Kapazitäten (Fahrplantrassen, Gleise in Zugbildungsanlagen, Personal) bilden die ideale Grundlage für den Aufbau eines neuen Systems für die Flächenbedienung. Die automatische Kupplung von Strom und Daten ermöglicht Elementarfunktionen wie Erkennung der Zugreihung, Bremsprobe und Zugintegritätsprüfung [7]. Damit können Wagenumstellungen in Zukunft überall erfolgen, ohne dass Einrichtungen oder besonders geschultes Personal verfügbar sein muss. Bottleneck ist aber die Erzeugung der für das Trennen und Verbinden notwendigen Wagenbewegung. Der klassische Ablaufberg ist eine intelligente und energiesparende Methode, weil er nur ein einziges angetriebenes Fahrzeug (Abdrücklokomotive) benötigt, um viele Wagen in koordinierter Art und Weise zu bewegen.

Mit der Digitalen Automatische Kupplung mit fernsteuerbarer Entkupplungsfunktion (DAK Stufe 5) reicht theoretisch eine Weiche und zwei Gleise entsprechender Länge, um

jeden beliebigen Wagenaustausch durch Vor-/ Rückwärtsbewegungen („Flachrangieren“) durchzuführen. Bei geeigneter Ausstattung der Güterwagen ist zwischen den Bewegungen kein Verlassen des Führerstands notwendig. Angetriebene Güterwagen können auch hier einen Beitrag leisten, um solche Prozesse noch flüssiger und energiesparender durchzuführen. Es kann eine neues System entstehen, welches die Vorteile des heutigen Systems durch Bündelung der Einzelwagen auf den Hauptachsen zu langen Zügen mit einer hohen Anpassungsfähigkeit an wechselnde Verkehrsrelationen kombiniert.

2.6 Wartungsstrategie für „High-Tech“-Güterwagen

Ein Güterwagen, wie er in diesem Beitrag spezifiziert wird, ist konventionell in dem Sinne, dass er im Hauptlauf in einen konventionellen Zug mit indirekt wirkender Bremse eingestellt wird. Es macht aber wenig Sinn, ihn auch bezüglich betrieblicher Wartung wie einen normalen Güterwagen zu behandeln. Seine speziellen Fähigkeiten spielt er nur in Verbindung mit ebenso speziellen Systemumgebungen aus; er wird also - im Gegensatz zu den meisten Güterwagen im EWV, die nur im Lastlauf individuell disponiert werden und danach wieder in der „Reserve“ verschwinden - immer unter vollständiger Disposition betrieben werden. Eine Flotte solcher intelligenter, aktiver Güterwagen sollte bezüglich Wartung eher wie eine Flotte von Triebwagen im Personenverkehr behandelt werden. Außerdem bieten sich gute Möglichkeiten, von einem zeit- oder laufstreckenbezogenen Wartungsregime auf eine zustandsbasierte und prädiktive Wartung zu wechseln.

3 Güterwagen für ein neues System der Flächenbedienung

3.1 Offene Schnittstellen

Entscheidend für die Akzeptanz des Schienengüterverkehrs ist nicht, dass der Güterwagen funktioniert und sicher ist. Davon wird ausgegangen. Entscheidend ist, dass er sich heute in die logistischen Prozesse des Kunden einfügt und er so offen und so flexibel gestaltet ist, dass er auch mit zukünftigen, heute noch unbekannten Herausforderungen kompatibel ist.

Von zentraler Bedeutung hierfür sind offen definierte Schnittstellen und Normen. Je höher die Sicherheitsanforderungen, umso bequemer ist das „System aus einem Guss“, bei dem ein Hersteller/ Systemhaus alle Gewerke selber entwickelt und herstellt und nur die

Außenschnittstellen bezüglich der Sicherheit nachzuweisen sind. Unter dieser „Denke“ leidet der Bereich der streckenseitigen Leit- und Sicherungstechnik bis heute. Erst durch die EULYNX-Schnittstellen [1] wird die monolythische Struktur aufgebrochen und es beginnt unter der Überschrift DSTW ein Wettbewerb zu entstehen, der zu neuen kreativen Lösungen führt.

Bei der Digitalen Automatischen Kupplung ist der offene Ansatz teilweise gelungen. In einem Auswahlprozess hat sich der von Voith eingebrachte Scharfenberg Typ 10 Kupplungskopf durchgesetzt und alle Konkurrenten werden aufgrund der TSI und der Normung in der Lage sein, mechanisch, elektrisch und datentechnisch kompatible Kupplungen herzustellen. Im Bereich der digitalen Funktionen innerhalb eines Fahrzeugs hört aber derzeit die Kooperation auf. Hier wird aktuell davon ausgegangen, dass OEMs komplette elektronische Güterwagenausrüstungen auf den Markt bringen ohne standardisierte Schnittstellen im Inneren. Dieser Weg mag zu vergleichsweise schnellen Lösungen führen. Er ist aber langfristig betrachtet teuer und kann zu einem Innovationshemmnis werden.

Ein Problem, das jeglicher Innovation am Güterwagen im Wege stand, ist die fehlende Stromversorgung. Durch die DAK ist Abhilfe in Sicht. Strom kann durch den Zug durchgeleitet und in den Wagen in Akkumulatoren gespeichert werden. Aber auch mit lokalen Maßnahmen, wie dem Einbau von Achsgeneratoren, kann problemlos Strom z.B. für den Betrieb von Speicherprogrammierbaren Steuerungen, Sensoren und Aktoren bereit gestellt werden. Bislang hatte aber jeder Hersteller für jede Anwendung eigene Ideen zur Stromversorgung.

Im Jahr 2018 ist im Verein Deutscher Ingenieure eine Sektorinitiative entstanden, die zumindest für eine sehr wichtige Schnittstelle des Systems Güterwagen einen weitgehend akzeptierten Vorschlag gemacht hat. Mit der VDI-Richtlinie 5905 [17] steht heute eine Anerkannte Regel der Technik zur Verfügung, die Vereinbarungen zur Systemspannung (24 Volt) eines Güterwagenbordnetzes ebenso wie zu garantierende Leistungsparameter und Steckerformen beinhaltet.

Insbesondere die Vernetzung mit Komponenten der Ladestelle (siehe oben ATLS) wird erheblich erleichtert, wenn auf dem Wagen eine klassische 24 V Stromversorgung für Speicherprogrammierbare Steuerungen, Sensoren und Aktoren vorausgesetzt werden darf.

3.2 Fähigkeiten

Aus den oben genannten Szenarien lassen sich einige Forderungen an den Güterwagen der Zukunft ableiten, die im Folgenden ohne den Anspruch auf Vollständigkeit aufgelistet werden.

- **Standardisierte Stromversorgung** mit definierter Spannung, garantierten Leistungsparametern und genormten elektromechanischen Schnittstellen (siehe oben)
- **Bordrechner/ Lokale Intelligenz**: Das Internet der Dinge erfordert die Allgegenwärtigkeit von Fähigkeiten zur Datenverarbeitung und zur Kommunikation („Cyber-physical System", „Ubiquitous Computing"). Die Fähigkeiten müssen unabhängig davon sein, ob der Wagen Bestandteil einer Gruppe ist und ob eine Lokomotive zur Verfügung steht.
- **„Schwarmintelligenz"**: Viele Prozesse erfordern, dass Wagen miteinander kommunizieren. Zweck kann die Vermeidung von Begegnungen innerhalb der Gleise einer Anschlussbahn sein. Zweck kann auch die Aggregation von Daten für die Optimierung von Rangierprozessen sein[3].
- **Aktorik**: Im Bahnbetrieb der Zukunft ist keine menschliche Arbeitskraft verfügbar. Also müssen alle Handhabungen am Wagen durch elektromechanische Aktorik ersetzt werden. Im konventionellen Güterwagenbetrieb betrifft dies
  - *Parkbremse*: Ein Betrieb mit Hemmschuhen ist nicht automatisierbar. Jeder Güterwagen benötigt eine steuerbare Festhaltebremse (Federspeicher, motorbetriebene Handbremse)
  - *Luftabsperrhähne*: Auch Wagen mit Automatikkupplung werden über Luftabsperrhähne verfügen, die perspektivisch auch ferngesteuert/ automatisch betätigt werden müssen[4]
  - *Adaptionen* der Luftbremse: Güterwagen verfügen in der Regel über zwei verschiedene Bremsdynamikeinstellungen (G und P) und die Bremse muss meist an das Ladungsgewicht angepasst werden (Laststellung). Aktuell sind dies handbetätigte Hebel, die perspektivisch ferngesteuert/ automatisch betätigt

[3] Beispiel einer Aggregation: „Die Wagengruppe hat 12 Achsen, wiegt 165 Tonnen, hat 121 Tonnen Bremsgewicht, ist luftseitig gekuppelt und am Ende der Wagengruppe abgeschlossen". Mit dieser Information kann ein Lrf die Wagengruppe für eine Rangierfahrt ankuppeln, ohne sie kontrollieren zu müssen

[4] Die Automatische Kupplung kann grundsätzlich die Hauptlaufleitung beim Entkuppeln schließen. Damit wäre aber der wichtige Fail-Safe-Mechanismus der selbsttätigen Zugbremse ausgehebelt.

werden müssen[5].

  - *Zugschlusssignal*: Solange als Rückfallbetrieb für die Freimeldung eines Gleises der Zugschluss beobachtet werden muss, muss auch der Güterwagen der Zukunft ein Zugschlusssignal tragen. Um flexible Änderungen der Zugkonfiguration bei der „Briefkastenbedienung“ realisieren zu können, muss der Zugschluss z.B. als elektrische Signalleuchte einstellbar sein.

- **Offenes Betriebssystem**: Die Software des Güterwagens muss aus einem sicherheitskritischen Teil bestehen, der beliebig um nicht sicherheitskritische Anwendungsprogramme erweitert werden kann. Ähnliche Herausforderungen sind im Bereich der Automatisierungstechnik mit Echtzeitbetriebssystemen bewältigt werden. An der FH Aachen ist ein rudimentäres Güterwagenbetriebssystem „WagonOS“ entwickelt worden. Aktuell unterstützt es nur die prototypische Realisierung von Vorgängen am stehenden Zug, wie z.B. die „Assistierte Bremsprobe“ [4].
- **Einzelachsantrieb für Rangierbewegungen**: Für eine weitgehende Automatisierung führt in vielen Anwendungsszenarien an der autarken Bewegungsmöglichkeit der Wagen kein Weg vorbei.

4 Zusammenfassung

Schon in den vergangenen 20 Jahren sind maßgebliche Innovationen an Güterwagen in den Betrieb integriert worden. So sind Telematik-Grundfunktionen (Lokalisierung, Stoßüberwachung, ...) heute allgemein verfügbar. Mit der Digitalen Automatischen Kupplung wird ein gewaltiger Schritt gemacht, der alle Vorgänge der Zugbildung und des Betriebs von Güterzügen auf ein bisher ungekanntes technologisches Niveau hebt. Dieser Schritt ist extrem teuer und verlangt der Branche in der Übergangszeit hohe Entbehrungen ab. Um so wichtiger, dass man den Blick nicht auf den Bahnbetrieb verengt.

Nach dem letzten veröffentlichten „DAC General Master Plan 02“ [6] beginnt die Deployment-Phase im Jahr 2029 und geht bis in die Mitte der 30er Jahre. Alles unter der Voraussetzung, dass die Finanzierung gesichert wird. Aktueller Stand ist aber, dass noch nicht einmal die Herstellung und der Betrieb der 100 so genannten „Pioneer-Trains“ gesichert sind, mit denen die „final validation“ vor der TSI-Revision durchgeführt werden soll. Kunden des

[5] Es gibt Bestrebungen, ep-Bremse und Gleitschutz bei Güterwagen einzuführen. Diese würden die Umstellungen obsolet machen

Schienengüterverkehrs können nicht so lange auf Verbesserungen warten und überall in Europa entstehen gerade „Inseln“ aus technisch aufgerüsteten Wagenparks. Ganz aktuell: der „iWagon“ von VTG UK in Großbritannien [10]. Gemeinsam ist allen eine lokale Stromversorgung mit Achsgeneratoren. Hersteller und Betreiber sollten diese Entwicklungen aufgreifen und eine gemeinsame Sektorinitiative starten, die die technische Entwicklung aktiver und kooperativer Wagen und deren Standardisierung voranbringt. Inselbetriebe aus hoch produktiven Wagenparks können zu gegebener Zeit mit vergleichsweise geringen Kosten auf die DAK umgestellt werden.

Erst wenn die Güterbahn sich von den klassischen Bandagen löst und primär darauf schaut, was der Kunde braucht (und der Wettbewerber vielleicht schon bietet), kann sie sich zu dem entwickeln, was eigentlich immer ihre Aufgabe war: zum logistischen Rückgrat der Wirtschaft.

Literatur

[1] *A Consortium Dedicated To Shaping The Future Of Railway Signalling In Europe*. URL: https://eulynx.eu/ (besucht am 08. 07. 2025).

[2] Bundesnetzagentur. *Sondererhebung Einzelwagenverkehr Marktstrukturen und Wirtschaftlichkeit*. Sep. 2022. URL: https://data.bundesnetzagentur.de/Bundesnetzagentur/SharedDocs/Downloads/DE/Sachgebiete/Eisenbahn/Unternehmen_Institutionen/Veroeffentlichungen/Marktuntersuchungen/Sondererhebung/ewv_bericht.pdf.

[3] *Digitale Automatische Kupplung im Schienengüterverkehr*. URL: https://www.dac4.eu/ (besucht am 10. 07. 2025).

[4] Manfred Enning und Rudolf Hilgers. „Die Assistierte Bremsprobe als Brücke zur Vollautomatisierung des Schienengüterverkehrs“. In: *International Railway Symposium Aachen, Tagungsband* (2023).

[5] Manfred Enning, Bernd Dietrich Schmidt und Daniela Wilbring. „Auf dem Weg zur autonomen Anschlussbedienung“. In: *Privatbahn Magazin* 3 (2019).

[6] *EU DAC DELIVERY PROGRAMME*. URL: https://rail-research.europa.eu/rail-projects/eu-dac-delivery-programme/ (besucht am 16. 11. 2025).

[7] *FP5 – TRANS4M-R, Transforming Europe's Rail Freight*. URL: https://rail-research.europa.eu/rail-projects/fp5-trans4m-r/ (besucht am 10. 07. 2025).

[8] Markus Hecht. *Beitrag des Schienengüterverkehrs zur Erreichung der Klimaschutzziele*. 2016. URL: https://die-gueterbahnen.com/assets/files/news/pdf/beitrag-des-schienengueterverkehrs-zur-erreichung-der-klimaschutzziele.pdf.

[9] *Intramotev. Rail. Reborn. Ready to Roll*. URL: https://intramotev.com/ (besucht am 10. 07. 2025).

[10] *iWagon - The UK's First Digital Freight Wagon*. URL: https://uk.vtg.com/products-and-services/iwagon (besucht am 16. 11. 2025).

[11] *Marktanteile: Der Erfolgskurs der Güterbahnen gerät ins Stocken*. URL: https://www.allianz-pro-schiene.de/themen/gueterverkehr/marktanteile/ (besucht am 10. 07. 2025).

[12] „Nevomo and Captrain launch globally first MagRail Booster: an automated rail service in a German steel factory". In: (17. Dez. 2024). URL: https://www.nevomo.tech/en/nevomo-and-captrain-launch-globally-first-magrail-booster-automated-rail-service-german-steel-factory/ (besucht am 10. 07. 2025).

[13] Nils Nießen. *Gleisanschlüsse, technisch betrachtet*. 16. März 2022. URL: https://www.eba.bund.de/DE/Veroeffentlichungen/Vortraege/FET_2022/fet_2022_node.html.

[14] *Parallel Systems. Powering Autonomous Rail*. URL: https://www.moveparallel.com/ (besucht am 10. 07. 2025).

[15] Raphael Pfaff, Bernd Dietrich Schmidt und Manfred Enning. „Towards inclusion of the freight rail system in the industrial internet of things – Wagon 4.0". In: *Stephenson Conference, London* (März 2017).

[16] DataBridge Market Research. *Globaler Markt für automatisierte LKW-Beladesysteme – Branchentrends und Prognose bis 2029*. URL: https://www.databridgemarketresearch.com/de/reports/global-automated-truck-loading-system-market?srsltid=AfmBOopOjmfVnydmTF5FtpKfMejjL-5Pp9DVuJWGhnU9oH-p8qnhrUvj.

[17] Verein deutscher Ingenieure. *VDI 5905 Blatt 1 - Schnittstellen aktiver, kooperierender Güterwagen - Stromversorgung*. [Online; Stand 07.06.2021]. 2021. URL: https://www.vdi.de/richtlinien/details/vdi-5905-blatt-1-schnittst%20ellen-aktiver-kooperierender-gueterwagen-stromversorgung.

[18] „Vorgabe für DB Cargo: Von –468 Mio. EUR auf Null in zwei Jahren". In: *Rail Business* 27 (2025), S. 5.

Session 4
Manfred Enning, Bernd Schmidt

Autoren



Enning, Manfred

FH Aachen, Lehrgebiet Bahnsystemtechnik im Fachbereich Maschinenbau und Mechatronik. Mitglied im Fachbeirat Bahntechnik des VDI und im Wissenschaftlichen Beirat des TIS. Forschungsschwerpunkt: Aktive und vernetzte Güterwagen.

enning@fh-aachen.de



Schmidt, Bernd

FH Aachen, Lehrgebiet Bahnantriebe im Fachbereich Maschinenbau und Mechatronik. Zuvor war er für ein großes Beratungsunternehmen im Bahnbereich tätig. Seine Forschungsschwerpunkte sind die Energieeffizienz von Antriebssystemen, die Reduzierung von Verschleiß sowie die Betriebseffizienz auf der letzten Meile.

b.schmidt@fh-aachen.de

Session 4
Kurt Fuchs

DAK: Quantensprung für den Schienengüterverkehr oder technologisch Sackgasse

Fuchs, Kurt

Rechtanwalt

Zusammenfassung

Die Schraubenkupplung im Schienenverkehr in Europa ist ein Anachronismus. In fast allen anderen Regionen der Welt sind längst automatische Mittelpufferkupplungen Standard bei den Eisenbahnen. Die Nachteile der Schraubenkupplung sind offensichtlich: Personalintensiver, zeitaufwendiger Betrieb, schwere und gefährliche körperliche Arbeit beim Kuppeln und geringe Zuglasten. Nun soll die Digitale Automatische Kupplung (DAK) all das ändern: Sie soll nicht nur das manuelle mechanische Kuppeln entbehrlich machen, sondern auch Stromversorgung und Datenübertragung in Güterzügen gewährleisten. Die damit verbundenen Vorteile für den Schienengüterverkehr sind ebenso offensichtlich wie die Nachteile der Schraubenkupplung. Vor allem die Zugvollständigkeitskontrolle und die Einführung der EP-Bremse würden wesentliche Verbesserungen bringen. So könnten ortsfeste Gleisfreimeldeanlagen entfallen und ETCS Level 3 flächendeckend eingeführt werden. Die elektropneumatische-Bremse (EP-Bremse) würde höhere Geschwindigkeiten und/oder höhere Zuglasten und eine automatische Bremsprobe ermöglichen und den Verschleiß der Bremsen und Radsätze deutlich verringern. Sensoren, zum Beispiel zur Überwachung der Radsätze oder der Ladung wären möglich, bei ausreichender elektrischer Leistung auch Aktuatoren zur ferngesteuerten Be- und Entladung eines Zuges.

Über die Nachteile der DAK, die zumindest zum Teil ebenso offenkundig sind, wird dagegen wenig bis gar nicht diskutiert. Es fängt an bei den Kosten, die sich nicht auf die Kupplung und deren Einbau beschränken, sondern auch die Installation der entsprechenden Sensoren und der notwendigen Verkabelung in den Güterwaggons umfassen. Ein Güterwaggon würde sich dadurch um 50 bis 100 % in der Herstellung verteuern, vom ohnehin margenschwachen Schienengüterverkehr (SGV) nicht zu

finanzieren. Darüber hinaus erreicht die DAK nicht einmal ansatzweise die für eine geordnete Abwicklung des Güterverkehrs auf der Schiene erforderliche Zuverlässigkeit beim Kuppeln. Ohne eine solche wäre der SGV aber insgesamt in Frage gestellt. Auch ist die DAK nicht abwärtskompatibel zur Schraubenkupplung, was eine auch nur halbwegs sinnvolle Einführungsstrategie nahezu unmöglich macht und die Disposition von Güterwaggons über Jahre erheblich verkomplizieren würde. Schließlich würde eine entsprechende Ausrüstung von Güterwaggons den Wartungsaufwand für diese Fahrzeuggattung vervielfachen, umgekehrt aber die Zahl der geeigneten Werkstätten dramatisch reduzieren. Viel spricht also dafür, daß die DAK eine technologische Sackgasse darstellt. Ihre Einführung könnte den Schienengüterverkehr in seiner Existenz bedrohen.

Keywords: DAK, Zukunft des SGV

1 Einleitung

Mit der Schraubenkupplung verfügen die Eisenbahnen in Europa über einen museumsreifen Anachronismus. Diesen zu beseitigen ist allerdings aufgrund der schieren Zahl der betroffenen Fahrzeuge, immerhin mehrere Hundertausend, extrem aufwendig und kostspielig. Entsprechend ist der erste Anlauf mit der UIC-Kupplung, der in den 1960er Jahren gestartet wurde, schließlich Anfang der 1980er Jahre endgültig gescheitert. Nun soll mit der DAK ein neuer Anlauf genommen werden, mit dem diesmal auch noch alle erdenklichen anderen Probleme des Schienengüterverkehrs gelöst werden sollen.

2 Vorteile

Die Vorteile der DAK sind für jeden, der sich mit Schienengüterverkehr einigermaßen auskennt, offensichtlich.

2.1 Automatische Kupplung

Eine automatische Kupplung würde die Abwicklung des Schienengüterverkehrs erheblich vereinfachen, vor allem den Einzelwagenverkehr mit seinen häufigen Rangiervorgängen. Die körperlich anstrengende und gefährliche Arbeit beim Kuppeln der Güterwagen würde entfallen. Allein dadurch könnte ein erheblicher Produktivitätsgewinn erreicht werden. In Kombination mit der Stromversorgung im Güterzug und der Datenübertragung wäre auch ein automatisches Entkuppeln ohne

zusätzliche Infrastrukturelemente möglich, allerdings nur bei mit einem Triebfahrzeug verbundenen Waggons.

Auch erlaubt eine Mittelpufferkupplung gegenüber den Schraubenkupplungen die Übertragung größerer Zug- und Schubkräfte und dadurch erheblich größere Zuglasten. Schließlich führt eine Mittelpufferkupplung gegenüber den bisher verwendeten Schraubenkupplungen zu einem signifikant geringeren Verschleiß an Rädern und Schienen.[1]

2.2 Zugvollständigkeitskontrolle

Stromversorgung und Datenübertragung im Güterzug würden es sehr einfach machen, eine Zugvollständigkeitskontrolle im Zug selbst zu implementieren. Dadurch wäre der flächendenkende Einsatz von ETCS Level 3 möglich. Infrastruktureinrichtungen zur Zugvollständigkeitskontrolle, wie Achszähler und Gleisstromkreise, die störanfällig und wartungsaufwendig sein, könnten dann vollständig entfallen, was die Kosten der Eisenbahninfrastruktur signifikant verringern würde.

2.3 Elektronpneumatische Bremse

Mit Stromversorgung und Datenübertragung im Güterzug wäre auch die Einführung einer elektropneumatischen Bremse möglich. Diese erlaubt gegenüber der reinen Druckluftbremse das gleichzeitige verzögerungslose anlegen aller Bremsen in einem Zug und würde so das Brems- und Fahrverhalten der Güterzüge nachhaltig verbessern, damit höhere Geschwindigkeiten oder größere Zuglasten ermöglichen und den Verschleiß an Bremsen, Rädern und Schienen reduzieren. Auch die Belastung der Kupplungen während der Fahrt würde sinken.

2.4 Sensorik

Schließlich würde Stromversorgung und Datenübertragung allerlei Sensorik im Güterzug ermöglichen, vor allem Sensoren zur Überwachung von Bremsen und Radsätzen, sowie der Radsatzlager ferner Sensoren zur Ladungsüberwachung und natürlich zur Überwachung der Kupplung selbst.

[1] Stuhr: Untersuchung von Einsatzszenarien einer automatischen Mittelpufferkupplung, Berlin, 2013

3 Nachteile

Während die Vorzüge der DAK auf der Hand liegen, sind die Nachteile einer flächendeckenden Einführung dieser Kupplung weniger offensichtlich, dafür aber nicht weniger gravierend. Abgesehen von den horrenden Kosten, die mit 15.000 bis 20.000 Euro pro Waggon angegeben werden, spielen sie in der Diskussion über die Einführung der DAK aber kaum eine Rolle.

3.1 Kupplung

Zunächst erscheint schon die Wahl der Scharfenbergkupplung als mechanischer Teil fraglich. Dieser Teil der Kupplung ist entscheidend für ein zuverlässiges Kuppeln der Güterwagen. Eine Zuverlässigkeit von 99 % bei den Kupplungsvorgängen dürfte als absolute Untergrenze für einen funktionierenden Güterverkehr anzusehen sein, also ein fehlerhafter Kupplungsversuch auf 100 Kupplungsvorgänge, was einer Störung bei etwa jedem vierten Zug entsprechen würde. Bisher wurde diese Zuverlässigkeit, soweit bekannt, im Probebetrieb nicht erreicht. Da der bisher durchgeführte Probebetrieb noch nicht die rauen Bedingungen des Regelbetriebs erreicht, müßte die Zuverlässigkeit hier noch deutlich darüber liegen. Die langjährigen Erfahrungen mit diesem Kupplungstyp im Personenverkehr lassen es zweifelhaft erscheinen, ob mit der Scharfenbergkupplung eine für den Schienengüterverkehr ausreichende Zuverlässigkeit erreicht werden kann. Der Kupplungstyp wird bisher nirgendwo im Güterverkehr verwendet.

Ein weiterer großer Nachteil dieses Kupplungstyps ist, daß er nicht abwärtskompatibel zur Schraubenkupplung ist. Dies würde die Organisation des Schienengüterverkehrs für die Dauer der Umrüstung auf diesen Kupplungstyp erheblich verkomplizieren und eine vergleichsweise kurze Einführungsphase erfordern, um den Schienengüterverkehr nicht zu stark zu beeinträchtigen. Dies wiederum würde aber zu anderen erheblichen Problemen bezüglich der Produktions- und Werkstattkapazitäten für die Herstellung und den Einbau der neuen Kupplung führen und dadurch auch die Kosten gegenüber einer längeren Übergangsphase noch einmal erhöhen.

3.2 Leistung

Soweit bekannt, ist für die Stromversorgung über die DAK 230V/400V Dreiphasen-Wechselstrom mit einer Gesamtleistung für den Zug im einstelligen Kilowattbereich vorgesehen. Hierfür müßten zunächst relativ große Leitungsquerschnitte im Zug vorgesehen werden, damit bei 750 m Zuglänge die Verluste nicht zu groß sind und im

letzten Waggon auch noch Strom fließt. Dabei dürften für die Übertragungsverluste im Zug nicht so sehr die in den Waggons verlegten Leitungen, sondern die Kupplungen maßgeblich sein. Sind die Kontakte in den Kupplungen nicht richtig sauber oder die Kontaktflächen relativ klein, dürften hier die größten Übertragungsverluste entstehen. Unter den rauen Bedingungen des SGV scheint es äußerst zweifelhaft, daß es gelingt, mit vertretbarem Aufwand die Kontakte in den Kupplungen dauerhaft hinreichend sauber und auch sonst intakt zu halten. Hierbei ist zu berücksichtigen, daß durch eine Störung der Stromübertragung auch nur in einer einzigen Kupplung die Zugvollständigkeitskontrolle bereits ausfallen und die Weiterfahrt des Zugs damit unmöglich gemacht würde.

Die relativ geringe elektrische Gesamtleistung erlaubt faktisch nur den Betrieb von Sensorik im Zug, nicht aber von Aktuatoren, zum Beispiel zum automatischen Be- und Entladen oder zur Versorgung von Kühlcontainern mit Strom. Allenfalls können Steuerelemente für pneumatische Aktuatoren betrieben werden.

Es gibt auch Überlegungen, die Güterwaggons mit Batterien auszurüsten. Dies ist jedoch nur dann sinnvoll, wenn Sensorik und Steuerelemente über Funk angesteuert werden können. Sind die Waggons mit einem Triebfahrzeug verbunden, ist keine Batterie erforderlich. Sind sie ohne Triebfahrzeuge abgestellt, ergeben Batterien nur dann einen Sinn, wenn Sensorik und Steuerelemente über Funk angesteuert werden können.

3.3 Datenübertragung

Dies führt zum Thema Datenübertragung im Zug. In der DAK ist ein mechanisches Element zur Kupplung der Datenleitungen im Zug vorgesehen. Dies führt zur selben Problematik der mechanischen Verbindung dieser Leitungen in den Kupplungen. Die entsprechenden Kontakte müssen hinreichend sauber und auch sonst intakt sein, um eine zuverlässige Datenübertragung zu gewährleisten. Auch hier scheint es sehr zweifelhaft, dies unter den rauen Bedingungen des Schienengüterverkehrs mit vertretbarem Aufwand dauerhaft zu gewährleisten. Werden die Waggons mit Batterien und einer Übertragungsmöglichkeit per Funk ausgerüstet, könnte auf mechanisch zu kuppelnde Datenleitungen ganz verzichtet werden und die Datenübertragung grundsätzlich per Funk erfolgen.

Um die Sensorik in den Güterwaggons sinnvoll nutzen zu können, muß vorab Art und Umfang der Sensorik und das Format der zu übertragenden Daten festgelegt werden.

Nur so wäre weiterhin ein freizügiger Einsatz aller Güterwaggons wie bisher möglich. Sicherheitsrelevante Daten, insbesondere bezüglich der Zugvollständigkeitskontrolle und der Bremsen müßten redundant übertragen werden können. Auch hier muß vorab festgelegt werden, wie die Redundanz ausgestaltet wird und was bei Abweichungen zwischen den redundanten Daten oder der Störung eines von mehreren Übertragungskanälen geschehen soll.

Sollen Änderungen an der Sensorik oder der Art oder des Umfangs der übertragenen Daten erfolgen, müßten diese auf allen Güterwaggons und Lokomotiven nachträglich implementiert werden.

3.4 Wartung

Nahezu keine Beachtung hat bisher gefunden, welche Auswirkungen die Einführung der DAK auf die Wartung von Güterwaggons hätte. Werkstätten hierfür sind derzeit überwiegend sehr einfach ausgestattet, insbesondere in den entlegeneren Teilen Europas. Eine niedrige zweistellige Anzahl unterschiedlicher Werkzeuge, sowie eine einstellige Anzahl unterschiedlicher Ersatzteile genügt für fast jede Reparatur eines Güterwaggons, was in vielen Fällen auch eine mobile, ortsungebundene Instandhaltung ermöglicht. Für die Instandhaltung oder auch nur den Tausch elektrischer und elektronischer Komponenten ist kaum eine dieser Werkstätten ausgestattet. Die Einführung der DAK würde also nicht nur erhebliche Investitionen in die Fahrzeuge erfordern, sondern auch in die Werkstätten, und zwar nicht nur in deren Ausstattung, sondern auch in die Ausbildung des dort tätigen Personals. Es ist zu erwarten, daß die Zahl der Werkstätten für Güterwaggons signifikant abnehmen würde.

Demgegenüber würde der Wartungsaufwand erheblich steigen. Bei Schienenfahrzeugen für den Personenverkehr verursachen Elektrik und Elektronik in den Fahrzeugen einen erheblichen Anteil des Wartungs- und Instandhaltungsaufwands. Eine Verdopplung des Wartungsaufwands bei Güterwaggons durch die Einführung der DAK dürfte die absolute Untergrenze sein. Auch eine Verdrei- oder –vierfachung ist je nach Gesamtausstattung der einzelnen Waggons nicht unplausibel. Eine deutlich höhere Nachfrage nach Wartungsleistungen würde also auch ein erheblich verknapptes Angebot treffen. Die Verfügbarkeit des vorhandenen Wagenparks dürfte deutlich sinken.

3.5 Kosten

Den wesentlichen Knackpunkt an der Einführung der DAK stellen aber die damit verbundenen Kosten dar. Die Angaben hierzu schwanken, dürften aber insgesamt zu niedrig angesetzt sein. Konsistent sind die Angaben pro Güterwagen in Höhe von ca. 20.000 € für die Umrüstung. Daraus folgt bei 500.000 umzurüstenden Fahrzeugen eine Gesamtsumme von über 10 Mrd. Euro für die Umrüstung, eine Summe, welche der margenschwache Schienengüterverkehr auf keine Fall aufbringen kann. Dabei dürfte dieser Betrag die absolute Untergrenze der Kosten sein. Nicht berücksichtigt ist dabei, daß für die notwendige schnelle Umrüstung des gesamten Fahrzeugparks Fertigungskapazitäten für die schnelle Herstellung der notwendigen Komponenten und Werkstattkapazitäten für die Umrüstung aufgebaut werden müßten, die anschließend ganz überwiegend nicht mehr benötigt würden. Alternativ könnten die Komponenten über einen längeren Zeitraum auf Vorrat produziert werden. Dies würde die Einführung jedoch deutlich verzögern und wäre mit erheblichen Lagerkosten über lange Zeiträume verbunden. Das Problem der Kapazitäten für die eigentliche Umrüstung bliebe aber auch dann bestehen. Ob die Mitgliedstaaten oder die EU die notwendigen Mittel aufbringen werden als Förderung des Schienengütervkehrs, ist zur Zeit völlig offen. Auch die Kosten für die Investitionen in die Werkstätten sind dabei noch nicht berücksichtigt.

Auch wenn eine möglichst schnelle Umrüstung angestrebt würden, wären wegen der schieren Zahl der umzurüstenden Fahrzeuge immer noch einige Jahre hierfür erforderlich. In dieser Zeit würde die Disposition der Güterwaggons erheblich verkompliziert, da die DAK nicht abwärtskompatibel ist, gemischte Züge also nicht möglich wären. Auch dies stellt einen erheblichen Kostenfaktor dar.

Eine staatliche Starthilfe würde das Problem, daß die Einführung der DAK auch die Preise für neue Güterwaggons deutlich verteuern würde, bei einfachen Waggons um mindestens 50 %, je nach Ausstattung eventuell auch deutlich mehr, nicht lösen. Auch dies kann ganze Teile des Schienengüterverkehrs wirtschaftlich grundsätzlich in Frage stellen. Schon heute sind erhebliche Teile des Schienengüterverkehrs chronisch defizitär. Eine deutliche Steigerung bei den Kosten für das Rollmaterial würde vielen Verkehren wirtschaftlich endgültig den Garaus machen.

Zu den einmaligen Kosten für Umrüstung und Beschaffung kommen aufgrund des erheblich höheren Wartungsaufwands noch deutlich höhere Instandhaltungskosten.

Auch die geringere Verfügbarkeit der Waggons würde die Kosten weiter nach oben treiben.

Ob die erhofften Produktivitätsgewinne, wenn sie sich denn realisieren lassen, all diese Kostensteigerungen kompensieren können, erscheint zumindest zweifelhaft.

4 Schlußfolgerungen

Das Projekt DAK kämpft, wie dargestellt, noch mit vielen Problemen, von denen einige, zum Beispiel das Thema Werkstätten, noch gar nicht Teil der Diskussion sind. Eine Betriebsreife der Kupplung ist derzeit nicht absehbar. Noch weniger ist klar, wie deren Einführung nach Erreichen der Betriebsreife organisiert und finanziert werden soll. Das Problem dabei ist: Der Ersatz der Schraubenkupplung durch eine zuverlässige automatische Kupplung ist längst überfällig. In wenigen Jahren wird es schlicht unmöglich sein, ausreichend Personal für die körperlich anstrengende und gefährliche Handarbeit beim Kuppeln der Güterzüge zu finden. Daß mit der DAK auch Stromversorgung und Datenübertragung im Güterzug eingeführt werden sollen, erhöht die Komplexität des Projekts und dessen Kosten erheblich und gefährdet so den erfolgreichen Abschluß

Sinnvoll wäre daher, zunächst nur das eigentliche Kupplungsproblem zu lösen, idealerweise mit einer Kupplungsausführung, die bereits bewährt, am Markt verfügbar und vor allem abwärtskompatibel ist. Die von der Firma Faiveley Transport entwickelte C-AKv zum Beispiel würde diese Bedingungen erfüllen. Nach Klärung der Finanzierungsfrage könnte eine Umrüstung auf diesen Kupplungstyp kurzfristig beginnen und wegen der Abwärtskompatibilität über einen sinnvollen Zeitraum gestreckt umgesetzt werden, der die oben geschilderten Probleme bei einer flächendeckenden Einführung der DAK vermeiden würde.

Eine Stromversorgung in den Güterzügen könnte mit der C-AKv bei Bedarf später nachgerüstet werden. Bezüglich der Datenübertragung spricht dagegen für Güterzüge viel für eine funkbasierte Lösung. So ließen sich parallel zur Umrüstung der Kupplungen viele Funktionen kurzfristig und relativ preiswert dort umsetzen, wo sie den größten Nutzen bringen, ohne daß alle Güterwaggons mit den entsprechenden Systemen ausgestattet werden müßten. Auch für die Zugvollständigkeitskontrolle ließe sich sicher eine funkbasierte Lösung finden. Die ist vielleicht technisch nicht so elegant, wie die über die DAK anvisierte Lösung, dafür aber wahrscheinlich auch relativ kurzfristig und deutlich preiswerter umzusetzen.

Abschließend erscheint die DAK also zumindest zur Zeit als technologische Sackgasse. Und viel spricht dafür, daß der Schienengüterverkehr keine Zeit hat, auf eine endgültige Klärung des Sinns oder Unsinns dieser vermeintlichen Innovation zu warten.



Session 4
Kurt Fuchs

Autor



Fuchs, Kurt

Dr. Kurt Fuchs studierte zunächst Maschinenbau mit Schwerpunkt Schienenfahrzeugtechnik an der RWTH Aachen und anschließend Jura an der Universität Münster. Er ist seit 2001 als Rechtsanwalt in Köln zugelassen und hat 2013 zum Thema „Netzzugang im Schienenverkehr“ promoviert. Er hält regelmäßig Vorträge und veröffentlicht Aufsätze zu rechtlichen und technischen Themen rund um den Schienenverkehr.

IRSA 2025

SESSION 5

ZUSTANDSÜBERWACHUNG I
CONDITION MONITORING I

Session 5
Mirco Janßen, Sven Jenne, Martin Möhring,
Pascal Gil, Henning Seeger, Christian Schindler

Entwicklung eines Zustandsüberwachungssystems für Losradachsen

Janßen, Mirco[1], Jenne, Sven[1], Möhring, Martin[1]

Gil, Pascal[2], Seeger, Henning[2], Schindler, Christian[2]

[1]Gutehoffnungshütte Radsatz GmbH, GHH-BONATRANS Group
[2]Institut für Schienenfahrzeuge, RWTH Aachen University

Zusammenfassung

Die Zustandsüberwachung von Losradachsen bietet Potenzial zur Steigerung der Sicherheit und Verfügbarkeit im Schienenverkehr. Durch die Analyse von Schwingungssignalen lassen sich Schäden an Wälzlagern und Laufflächen der Räder frühzeitig erkennen. Im Rahmen eines Projekts wurde ein bestehender Einzelachsprüfstand in Zusammenarbeit mit der RWTH Aachen erfolgreich umgebaut und auf die Anforderungen von Losradachsen angepasst. Die Prüfstandsuntersuchungen bestätigten die Wirksamkeit der eingesetzten Analyseverfahren zur Zustandsdiagnose. Dies bildet die Grundlage für die Entwicklung eines Prototyps, der künftig im Flottenbetrieb getestet werden soll.

Keywords: Zustandsüberwachung, Diagnose, Predictive Maintenance, Losradachsen, Wälzlager

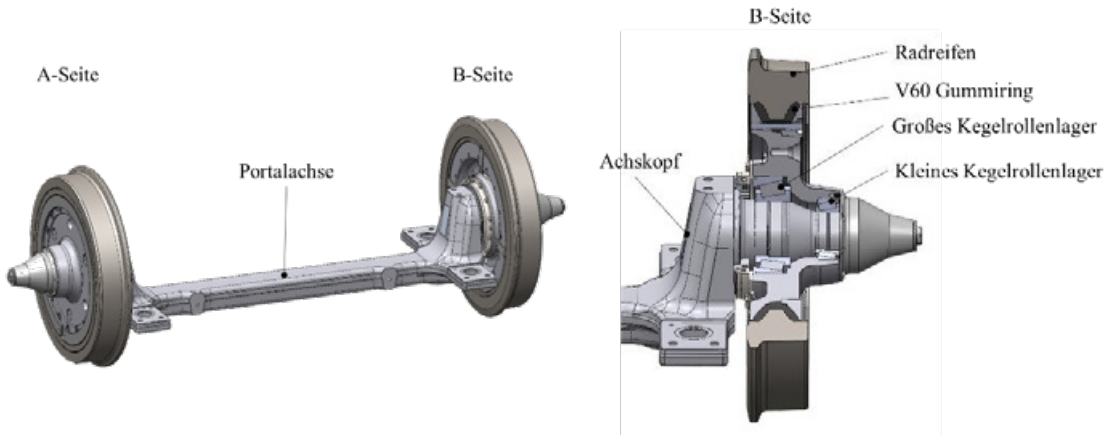
1 Einleitung

Die Mobilität auf der Schiene ist ein unverzichtbarer Bestandteil moderner Verkehrssysteme. Sicherheit, Verfügbarkeit und Wirtschaftlichkeit sind entscheidende Faktoren für den reibungslosen Betrieb. Eine der größten Herausforderungen für Betreiber ist die Zuverlässigkeit der Fahrwerke, insbesondere der Radsätze und Losradachsen, die im Betrieb hohen Belastungen ausgesetzt sind.

Session 5
Mirco Janßen, Sven Jenne, Martin Möhring,
Pascal Gil, Henning Seeger, Christian Schindler

Traditionell erfolgen Wartung und Instandhaltung nach festen Zeitintervallen oder Erfahrungswerten. Dieses Vorgehen birgt jedoch zwei wesentliche Nachteile: Entweder werden Bauteile zu früh ausgetauscht, was unnötige Kosten verursacht, oder es kommt vorzeitig zu unerwarteten Defekten, die zu Betriebsunterbrechungen und hohen Reparaturkosten führen. Hier setzt die Zustandsüberwachung (Condition Monitoring) an, um den Zustand der Fahrzeugkomponenten in Echtzeit zu überwachen. Im Fokus dieses Projekts stehen Losradachsen, welche im Folgenden näher vorgestellt werden.

Wie in Abbildung 1 zu sehen ist, ermöglichen Losradachsen (LRA), durch die Verwendung einer gekröpften Portalachse, die Konstruktion von Straßenbahnen mit niedriger Fußbodenhöhe um 300 mm (Niederflurstraßenbahnen).



a) b)

Abbildung 1:a) LRA und b) Lagerung der Räder

Im Gegensatz zu klassischen Radsätzen, bei denen die gummigefederten Räder starr über eine durchgehende Achswelle miteinander verbunden sind, sind die Räder bei einer LRA voneinander entkoppelt und einzeln über Wälzlager drehbar auf der gekröpften Portalachse gelagert (siehe Abbildung 1 b)). Marktreife Systeme zur Zustandsüberwachung dieser Bauart – insbesondere von Wälzlagern und Radreifen – existieren derzeit noch nicht. Die GHH-Radsatz GmbH verfolgt daher ein eigenes Entwicklungskonzept zur Zustandsüberwachung von Losradachsen.

2 Entwicklungskonzept

Das Entwicklungskonzept zur Zustandsüberwachung von LRA ist in Abbildung 2 schematisch dargestellt. Der aktuelle Fortschritt ist durch die vertikale Linie indiziert.

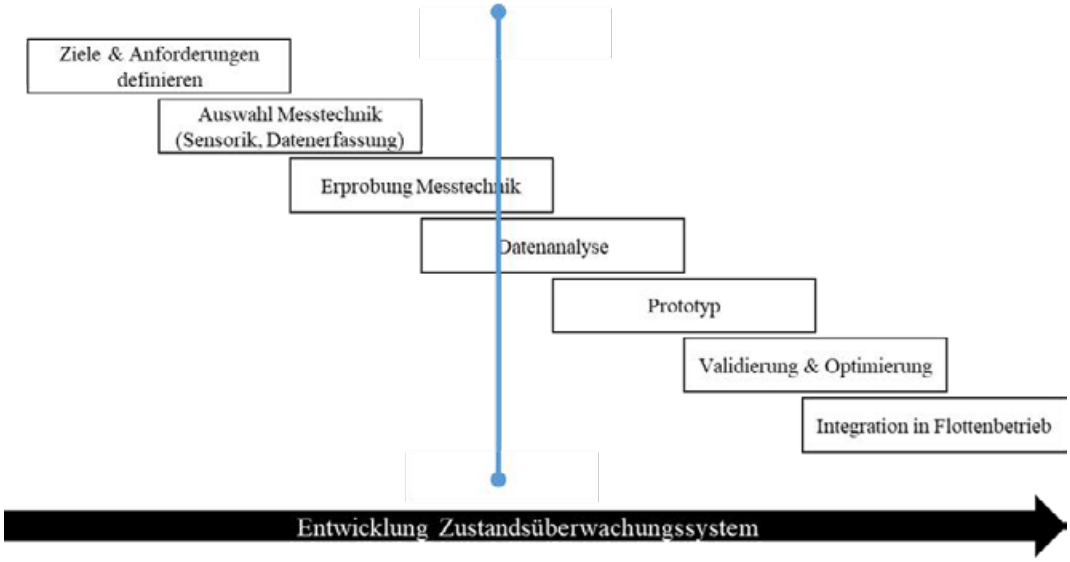


Abbildung 2: Entwicklungskonzept GHH-Radsatz für Zustandsüberwachungssystem LRA

Zunächst werden **Ziele und Anforderungen** definiert – insbesondere, welche Komponenten überwacht und welche technischen Rahmenbedingungen berücksichtigt werden sollen. Der Fokus liegt auf den **Radreifen** und **Wälzlagern**. Ob eine gleichzeitige Erfassung und Bewertung der Infrastruktur möglich ist, wird ebenfalls untersucht.

Es folgt die **Auswahl geeigneter Messtechnik,** also Sensoren und Methoden zur Datenerfassung. Anschließend erfolgt deren **praktische Erprobung** anhand von:

- Betriebsmessfahrten (realitätsnahe Bedingungen)
- Prüfstandsuntersuchungen (kontrollierte, wiederholbare Zustände)

Parallel dazu beginnt die **Datenanalyse** zur Identifikation möglicher Fehlermerkmale. Darauf basierend wird eine weitere LRA mit Messtechnik ausgestattet und bei einem Kundenfahrzeug im realen Betrieb getestet. Nach erfolgreicher Validierung und Optimierung wird die Technologie in das GHH-Produktportfolio integriert, sodass Betreiber sie zur langfristigen Zustandsüberwachung und vorausschauenden Wartung in ihren Flotten einsetzen können.

Im Rahmen dieser Veröffentlichung werden die Ergebnisse der Prüfstandsuntersuchungen vorgestellt.

3 Prüfstandsuntersuchungen

In diesem Kapitel wird zunächst die Modernisierung bzw. der Umbau eines bestehenden Einzelachsprüfstands des Instituts für Schienenfahrzeuge (IFS) der RWTH Aachen und im Anschluss das durchgeführte Versuchsprogramm vorgestellt.

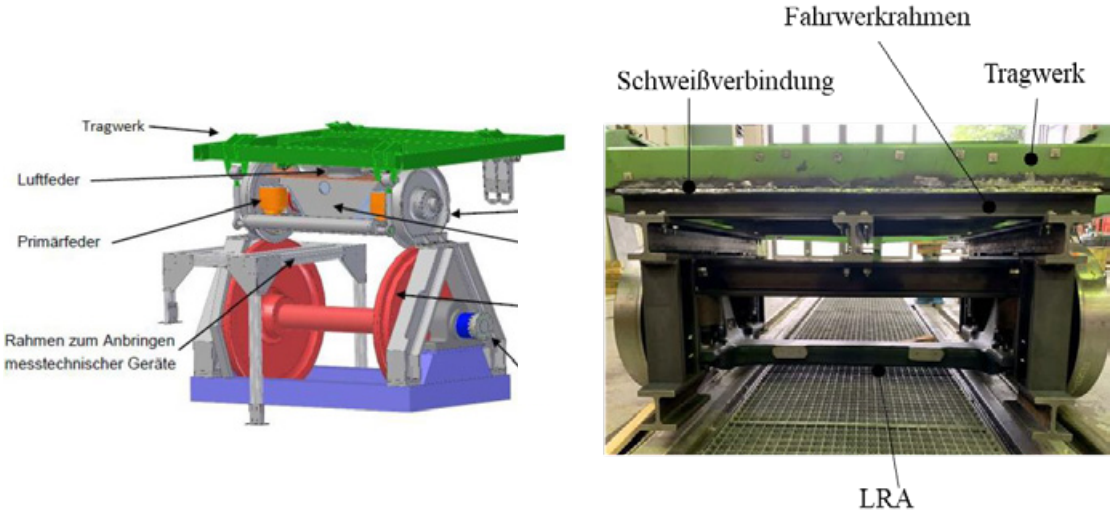
Session 5
Mirco Janßen, Sven Jenne, Martin Möhring,
Pascal Gil, Henning Seeger, Christian Schindler

3.1 Prüfstandsumbau

In Zusammenarbeit mit dem IFS wurde ein vorhandener Einzelachsprüfstand umfassend modernisiert und die mechanische Integration einer Losradachse ermöglicht. Die wesentlichen technischen und funktionalen Anforderungen seitens GHH an den Prüfstand umfassen hierbei:

- Mechanische Integration einer Losradachse
- Möglichkeit zum schnellen Ein- und Ausbau der Achse
- Einstellen verschiedener Achslasten (maximale Last: 12,5 t)
- Geschwindigkeiten der Losradachse zwischen 0 – 60 km/h
- Möglichkeit zur Integration verschiedener LRA-Modelle durch geringe Adaption der Aufnahmehalterung des Prüfstandes
- Es müssen keine Querkräfte mit dem Prüfstand realisiert werden.

Ausgehend von dem bestehenden Einzelachsprüfstand (siehe Abbildung 3 a)) wurde in einem Konstruktionsprozess nach VDI 2221 [1] die Entwicklung und Konstruktion des neuen Prüfstandes unter Berücksichtigung der o.g. Anforderungen umgesetzt.



a) b)

Abbildung 3: a) bestehender Einzelachsprüfstand und b) Angeschweißter Fahrwerkrahmen für LRA an bestehenden Tragwerksrahmen

Für die mechanische Integration wurde ein neuer Fahrwerkrahmen aus Stahlträgern konstruiert. Die Schnittstelle des Fahrwerkrahmens zur LRA bilden Töpfe zur Aufnahme der LRA-Primärfedern. Die Einbindung der originalen Primärfedern reduziert einerseits die übertragenen ungefederten Massen und führt darüber hinaus zu einem realistischeren Versuchsaufbau. Der Fahrwerkrahmen wird mit dem bereits vorhandenem Tragwerk

durch Schweißnähte stoffschlüssig verbunden (siehe Abbildung 3 b)). Durch den modularen Aufbau des Fahrwerkrahmens können in Zukunft mit nur geringfügigen Anpassungen weitere LRA-Typen aufgenommen werden.

Im Zuge der Modernisierung wurde die Elektronik vollständig erneuert. Der Antrieb erfolgt über einen Drehstrom-Asynchronmotor, der aufgrund der hohen Drehmomentanforderungen an ein Schneckengetriebe angeflanscht ist. Die Regelung des Motors erfolgt über einen Frequenzumrichter, der seine Befehle über einen Datenbus von einem Steuerungsrechner erhält. An letzterem können sämtliche Betriebsparameter wie beispielsweise der Motorstrom oder die durch einen Drehzahlgeber an der Getriebeausgangswelle erfasste Drehzahl angezeigt und überwacht werden.

Das Einstellen unterschiedlicher Achslasten erfolgt über Stahlplatten (je 2,5 t), welche auf dem Tragwerk befestigt werden können (siehe Abbildung 4 a)). Der Prüfstand ermöglicht in seiner derzeitigen Konfiguration die Aufnahme und den Betrieb einer Losradachse mit Geschwindigkeiten von bis zu 60 km/h bei einer maximalen vertikalen Achslast von 12,5 t.



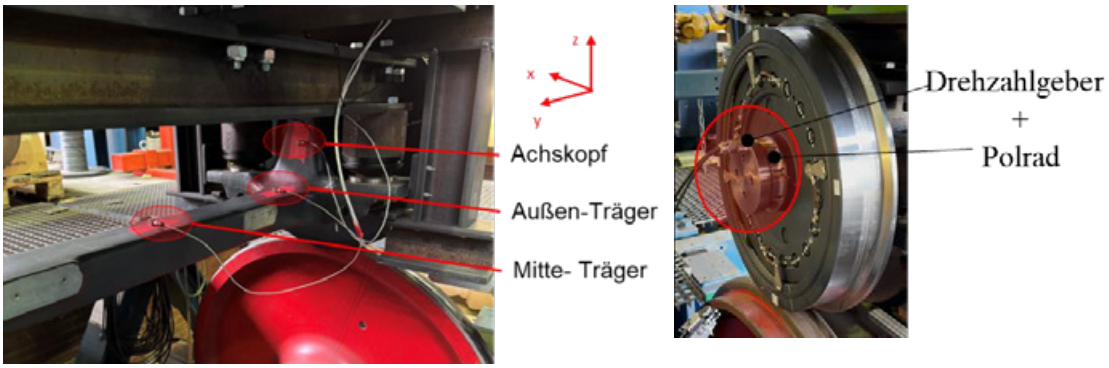
a) b)

Abbildung 4: a) Ursprungskonfiguration des Einzelachs-Rollprüfstandes b) aktuelle Konfiguration des LRA-Prüfstandes

Test-LRA und Prüfstand sind in Abbildung 4 b) zu sehen. Die Getriebeausgangswelle des Antriebs ist über eine Kupplung mit einem Treibradsatz (rot) verbunden. Durch den Rollkontakt werden die Räder der LRA in Rotation versetzt. Für den gewählten LRA-Typ nimmt das Durchmesserverhältnis von Treibrad zu LRA-Rad (Neuzustand ⌀ 600 mm) etwa einen Wert von 2 an.

3.2 Versuchsprogramm und -durchführung

Verwendet wurde eine neue Test-LRA, baugleich zu LRAs, die in Straßenbahnen des Typs Siemens Avenio eingesetzt werden. Die verbauten Kegelrollenlager (2 pro Seite) und Radreifen (Typ GHH® *V60*) entsprechen dem Neuzustand. Diese Konfiguration der LRA dient als Referenzzustand. Wie in Abbildung 5 a) dargestellt, wurden zur Erfassung der auftretenden Beschleunigungen während des Betriebs hochauflösende piezoelektrische Beschleunigungssensoren verwendet und an verschiedenen Positionen der LRA appliziert.



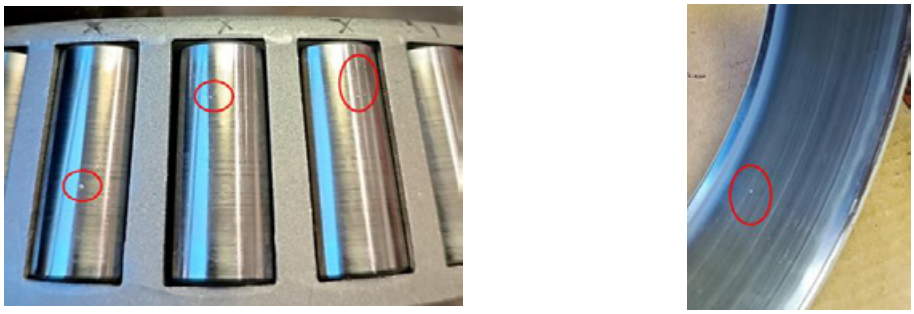
a) b)

Abbildung 5: a) Position der Beschleunigungssensoren und b) Drehzahlerfassung

Zusätzlich zu den Beschleunigungssignalen wird das Drehzahlsignal mittels eines Drehzahlgebers der Firma Lenord&Bauer hochfrequent über ein speziell für diesen Prüfstand und die Test-LRA konstruiertes Polrad (65 Zähne, Modul 2) erfasst. Die Ermittlung der Geschwindigkeit erfolgt aus Drehzahlsignal und Laufkreisdurchmesser des LRA-Rades.

Die Sensorik bestehend aus Beschleunigungsaufnehmern und Drehzahlgeber wird an das Datenerfassungssystem CronosFlex der Firma imc angeschlossen. Die Abtastraten wurden auf 50 kHz festgelegt.

Ein Lagerwechsel (nur auf Seite B der Test-LRA) ermöglichte den Vergleich von neuen und gebrauchten Lagerzuständen (Laufleistung ca. 500.000 km). Bei den gebrauchten Wälzlagern, die speziell für den Versuch ausgewählt wurden, waren sowohl am kleinen als auch beim großen Kegelrollenlager (siehe Abbildung 1 b)) leichte Vorschäden in Form von beginnenden Materialausbrüchen (Pittings) aufgrund von Materialermüdung an mehreren Wälzkörpern (Abbildung 6 a)) und den Außenringen (Abbildung 6 b)) zu detektieren.



a) b)

Abbildung 6: a) Materialausbrüche an Wälzkörpern für kleines Kegelrollenlager und b) am Außenring des großen Kegelrollenlagers

Im Rahmen dieser Veröffentlichung wird nur eine Auswahl der erprobten Versuchsreihen vorgestellt. Eine Übersicht dieser ausgewählten Versuche ist in Tabelle 1 zu sehen.

Tabelle 1: Versuchsreihen für Test-LRA-Zustand „Neu“ und „Gebraucht“

Versuchsreihe	Last	Soll-Geschwindigkeit	Lagerzustand	Radreifenzustand
V001	10t	20 km/h	Neu	Neu
V002		30 km/h		
V003		40 km/h		
V004		50 km/h		
V005		20 km/h	Gebraucht	
V006		30 km/h		
V007		40 km/h		
V008		50 km/h		

Bei jeder Versuchsreihe wurde die Soll-Geschwindigkeit angefahren, für ca. 3 min gehalten (konstanter Betrieb) und die Daten aufgezeichnet.

4 Datenanalyse

In diesem Kapitel wird zunächst die relevante Auswertungsmethodik gemäß VDI 3832 [2] und die Ordnungsanalyse als Auswertungsverfahren vorgestellt. Im Anschluss werden die Ergebnisse der Analyse präsentiert und im Hinblick auf ihre Eignung zur Bewertung des Lagerzustandes evaluiert.

4.1 Auswertungsmethodik

Bei der Analyse der aufgezeichneten Beschleunigungs- und Drehzahlsignale wurde sich stark an der VDI 3832 (April 2023) *Körperschallmessungen zur Zustandsbeurteilung von Wälzlagern in Maschinen und Anlagen* orientiert [2]. Auch wenn diese Richtlinie vorwiegend für Wälzlagerungen stationärer Anlagen und Maschinen ausgelegt ist, sind die fundamentalen Auswertungshintergründe auch für dynamische Anwendungen übertragbar,

sofern zusätzliche Schwingungs- und Körperschalleinwirkungen berücksichtigt werden. In der VDI 3832 wird die Erstellung eines sogenannten Hüllkurvenspektrums beschrieben, welches ein bewährtes Mittel zur zuverlässigen, frequenzabhängigen Analyse von Wälzlagerschäden darstellt. Schon kleinste Fehlstellen auf Laufbahnen oder Wälzkörpern erzeugen beim Überrollen kurze, hochfrequente Impulse, welche durch ein sogenanntes Hüllkurvenspektrum sichtbar gemacht werden können. Der grundsätzliche Ablauf zur Erstellung des Hüllkurvenspektrums ist in Abbildung 7 schematisch dargestellt.

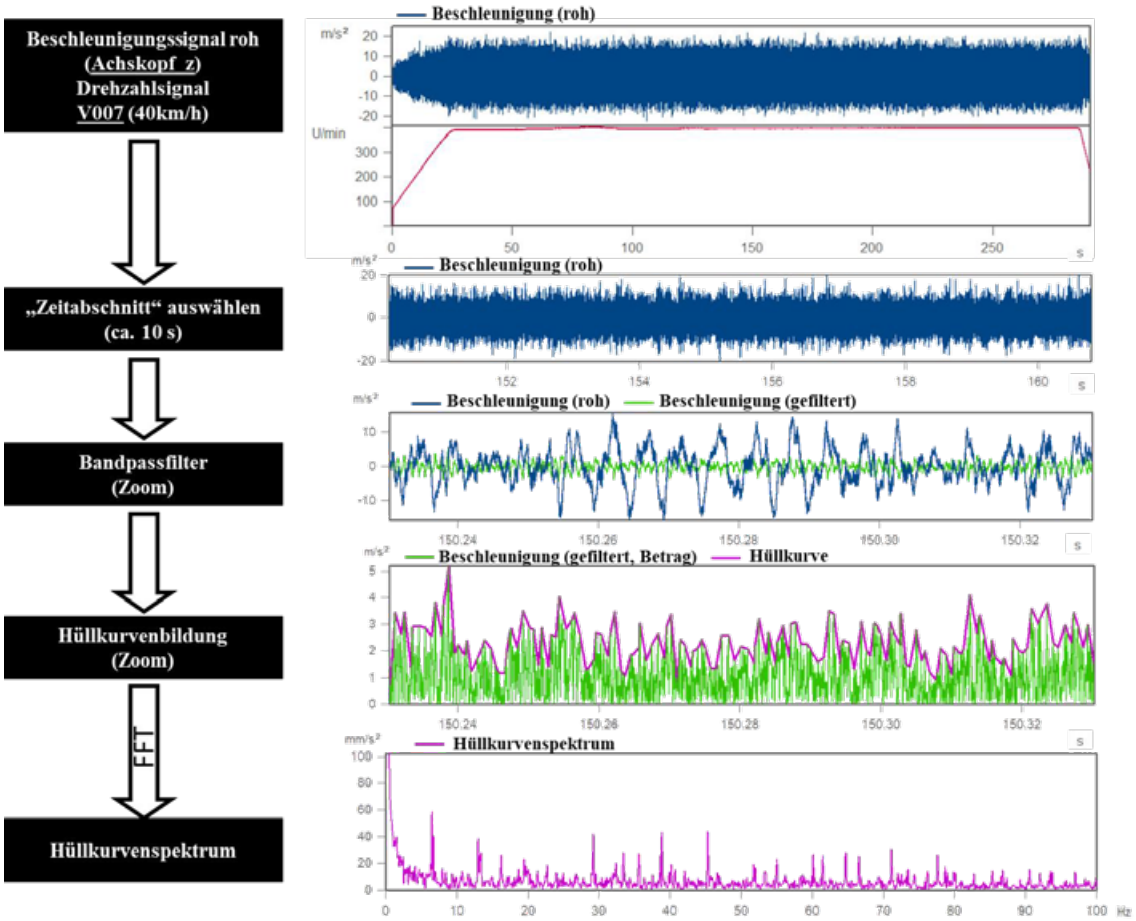


Abbildung 7: Schematische Darstellung der Erstellung des Hüllkurvenspektrums in Anlehnung an [2]

Voraussetzung für diese Art der Analyse ist eine konstante Drehzahl, sodass Anfahr- und Auslaufphasen der Versuchsreihe nicht in die Auswertung einbezogen werden. Aus dem aufgezeichneten Rohsignal wird ein Abschnitt mit konstanter Drehzahl und einer Länge von etwa 10 Sekunden extrahiert und anschließend einer Bandpassfilterung unterzogen. Im nächsten Schritt wird der Betrag des gefilterten Signals berechnet und eine Hüllkurve darübergelegt. Abschließend erfolgt eine klassische Fast-Fourier-Transformation (FFT) der Hüllkurve, um die relevanten Frequenzanteile des zugrundeliegenden Beschleunigungssignals sichtbar zu machen

Zusätzlich zur Analyse der auftretenden Frequenzspektren wurde eine Ordnungsanalyse durchgeführt. Dabei handelt es sich um ein Auswertungsverfahren, das insbesondere bei drehzahlvariablen Maschinen Anwendung findet. Im Gegensatz zur klassischen FFT, die das Spektrum in festen Frequenzintervallen (Hz) darstellt, analysiert die Ordnungsanalyse das Signal in Bezug auf die Rotordrehfrequenz. Das bedeutet, dass die Frequenzachse nicht in Hertz, sondern in sogenannten Ordnungen angegeben wird – also als Vielfache der Rotordrehfrequenz.

Durch diese Darstellung lassen sich drehzahlabhängige Signalkomponenten besser identifizieren, da sie unabhängig von der momentanen Drehzahl stets bei derselben Ordnung erscheinen [3].

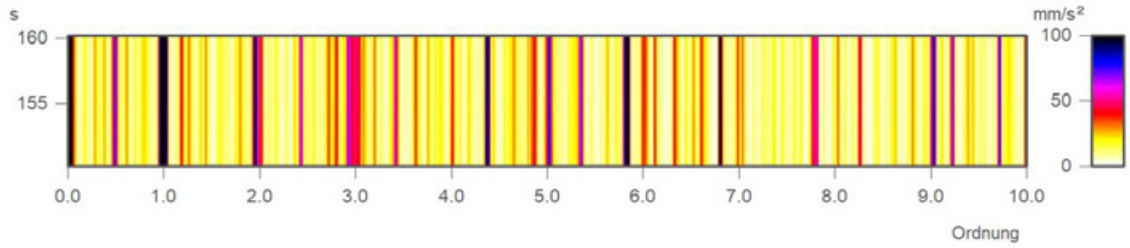


Abbildung 8:Ordnungsanalyse von V007_Achskopf_z für den gleichen Zeitabschnitt wie in Abbildung 7

Im folgenden Abschnitt werden nun die Ergebnisse vorgestellt und analysiert.

4.2 Ergebnisanalyse

Für die Zustandsüberwachung von Radreifen und insbesondere von Wälzlagern ist es von entscheidender Bedeutung, die relevanten Frequenzen zu kennen, die den jeweiligen Komponenten bei einer Radumdrehung zugeordnet werden können.

Bei Flachstellen auf dem Radreifen lässt sich die relevante Radfrequenz direkt aus dem Drehzahlsignal ableiten (U/min → U/s → Hz). Im Fall von Wälzlagern ist insbesondere die sogenannte charakteristische Überrollfrequenz von Interesse. Wenn ein Wälzkörper beispielsweise über eine Schadstelle rollt, führt dies zu stoßartigen Anregungen, die sich in Form von Körperschallsignalen ausbreiten. Diese Anregungen treten periodisch auf – mit einer Frequenz, die sich aus der spezifischen Kinematik des Lagers ergibt.

Die Überrollfrequenz wird maßgeblich durch die Lagergeometrie sowie durch die Lastverteilung im Lager beeinflusst. Sie ist aus diagnostischer Sicht besonders relevant, da Schäden an Laufbahnen oder Wälzkörpern typischerweise genau bei dieser Frequenz oder ihren Vielfachen Schwingungssignale verursachen. Die Berechnung der charakteristischen Frequenzen für Käfig, Wälzkörper, Innen- und Außenring kann der VDI-Richtlinie 3832 entnommen werden.

In folgender Abbildung 9 wird exemplarisch ein Vergleich zwischen den Versuchsreihen V004 und V008 (jeweils bei 10 t Achslast und 50 km/h) dargestellt, und zwar für die Konfigurationen mit neuem sowie gebrauchtem Kegelrollenlager. Als Auswerteposition wurde der Achskopf gewählt, wobei die vertikale Richtung als relevante Auswertungsrichtung betrachtet wurde.

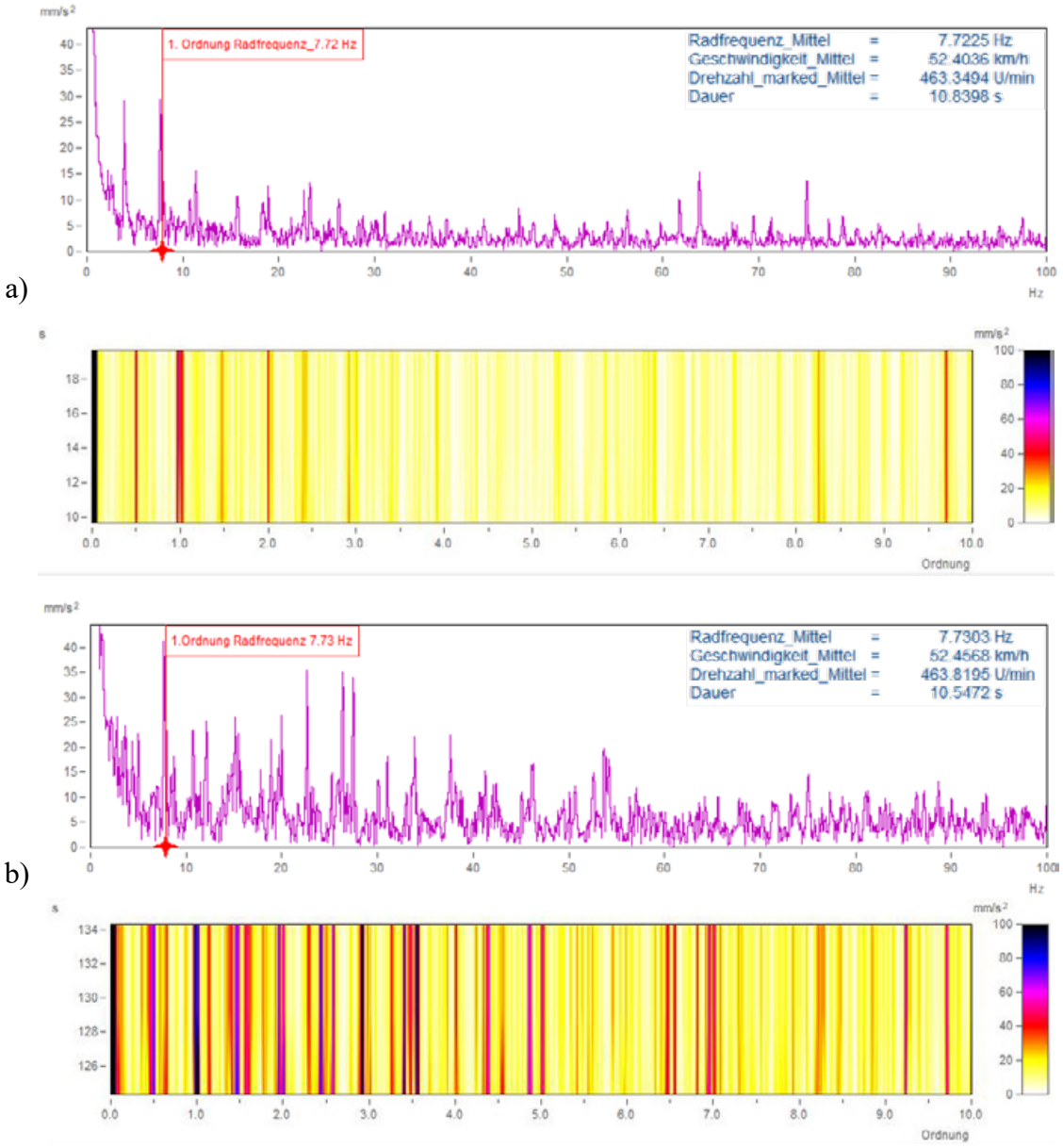


Abbildung 9: Gegenüberstellung von a) V004 und b) V008 für Achskopf_z

Aus platztechnischen Gründen können nicht alle Ergebnisse visuell dargestellt werden. Die nachfolgend beschriebenen Erkenntnisse gelten jedoch ebenfalls für die übrigen Versuchsreihen.

Bei der Analyse des Hüllkurvenspektrums zeigt sich, dass die Ordnungen der Radfrequenz (wobei die 1. Ordnung der Radfrequenz selbst entspricht) im Spektrum klar identifizierbar sind. Dies lässt den Rückschluss zu, dass Änderungen in der Amplitude der Radfrequenz als Indikator für den Zustand des Radreifens herangezogen werden können.

Darüber hinaus ist im Hüllkurvenspektrum ersichtlich, dass bei der Konfiguration mit gebrauchtem Lager im Vergleich zur neuen Ausführung mehrere zusätzliche Peaks bei unterschiedlichen Frequenzen auftreten. Dies deutet auf zusätzliche Stoßimpulse im System hin, die vermutlich auf Materialausbrüche an Wälzkörpern und Außenringen zurückzuführen sind.

Die Ergebnisse der Ordnungsanalyse bestätigen diese Beobachtungen. Insbesondere bei der Versuchsreihe V008 konnten deutlich mehr Anregungen im Beschleunigungssignal identifiziert werden als bei V004. Dies wird durch die intensivere Einfärbung der vertikalen Linien bei gleicher Skalierung und identischen Versuchsbedingungen besonders deutlich.

Somit konnte nachgewiesen werden, dass sowohl mittels Hüllkurvenspektrum als auch mittels Ordnungsanalyse Unterschiede im Beschleunigungssignal erkennbar sind.

5 Fazit & Ausblick

Die durchgeführten Prüfstandsuntersuchungen belegen, dass mit Hilfe von Hüllkurven- und Ordnungsanalyse der Zustand von Wälzlagern und Radreifen zuverlässig beurteilt werden kann. Besonders bei gebrauchten Lagern zeigen sich charakteristische Anregungen im Schwingungssignal, die auf beginnende Schäden hinweisen – ein klarer Nachweis für das Potenzial der Zustandsüberwachung.

Ein zentraler Erfolgsfaktor ist der umfassende Umbau des Einzelachsprüfstands. Die mechanische Integration der Losradachse, die flexible Anpassbarkeit für verschiedene Achstypen sowie die Erneuerung der Antriebs- und Steuerungstechnik haben die Grundlage für reproduzierbare Tests geschaffen. Der Prüfstand bildet damit einen wichtigen Baustein für die Entwicklung eines serienreifen Zustandsüberwachungssystems für Losradachsen.

Im nächsten Schritt sollen Betriebsmessfahrten unter realen Einsatzbedingungen durchgeführt werden, um die Praxistauglichkeit und Robustheit der bisher entwickelten Analyseverfahren zu überprüfen. Dabei sollen insbesondere die Einflüsse von Umgebungsbedingungen, Fahrdynamik und variabler Last auf die Signalqualität und Diagnosegüte untersucht werden.

Parallel dazu sind weitere Untersuchungen auf dem Prüfstand geplant, bei denen Kegelrollenlager mit definierten Schädigungsgraden (z. B. leichte, mittlere und starke Laufbahnbeschädigungen) eingesetzt werden. Ziel ist es, die Sensitivität und Selektivität der Diagnoseverfahren weiter zu verfeinern und Grenzwerte für den Lagerzustand zu definieren.

Nach erfolgreichem Abschluss dieser Untersuchungen soll ein Prototyp eines Zustandsüberwachungssystems entwickelt werden, der in eine ausgewählte Fahrzeugflotte integriert wird. Im Rahmen eines mehrmonatigen Flottenbetriebs soll das System umfassend getestet und hinsichtlich Langzeitstabilität, Detektionssicherheit und Wartungsfreundlichkeit evaluiert werden.

Autoren



Janßen, Mirco Dr.-Ing.
Gartenstr. 40, 46145 Oberhausen, Germany
Position: Forschung und Entwicklung / Research and Development



Jenne, Sven Dr.-Ing.
Gartenstr. 40, 46145 Oberhausen, Germany
Position: Direktor Konstruktion & Entwicklung / Director Engineering & Development



Möhring, Martin Dipl.-Ing.
Gartenstr. 40, 46145 Oberhausen, Germany
Position: Leitung Erprobung und Validierung / Manager Testing and Validation



Gil, Pascal M.Sc.
Institut für Schienenfahrzeuge, RWTH Aachen University
Seffenter Weg 8, 52074 Aachen
Position: Wissenschaftlicher Mitarbeiter



Seeger, Henning M.Sc.
Institut für Schienenfahrzeuge, RWTH Aachen University
Seffenter Weg 8, 52074 Aachen
Position: Absolvent Bahnsystemingenieurwesen



Schindler, Christian Univ.-Prof. Dr.-Ing.
Institut für Schienenfahrzeuge, RWTH Aachen University
Seffenter Weg 8, 52074 Aachen
Position: ehem. Lehrstuhlinhaber und Institutsdirektor

Literatur

[1] VDI 2221 Blatt 1: Entwicklung technischer Produkte und Systeme - Modell der Produktentwicklung, 2019

[2] VDI 3832: Körperschallmessungen zur Zustandsbeurteilung von Wälzlagern in Maschinen und Anlagen, 2013

[3] Franke, Dieter:Wälzlagerdiagnose an Maschinensätzen, Diagnose und Überwachung von Wälzlagerfehlern und –schäden, 1.Aufl., Dresden, SpringerVieweg, 2022

Session 5
Matthias Zelinka, Martin Joch, Alexander Schirrer, Stefan Jakubek

Effizientes EN 50318-konformes Simulationsmodell der Interaktion zwischen Pantographen und Oberleitung

Zelinka, Matthias BSc; Dipl.-Ing. Dr. Joch, Martin
Priv.-Doz. Dr. Schirrer, Alexander; Univ. Prof. Dr. Jakubek, Stefan

PJ Messtechnik GmbH
Institut für Mechanik und Mechatronik, TU Wien

Zusammenfassung

Dieser Beitrag präsentiert eine neuartige, schlankes, vollständig in MATLAB implementiertes Finite-Elemente-Modell zur Simulation der Interaktion zwischen Pantographen, Oberleitung und Deckenstromschiene. Der Fokus liegt auf hoher Recheneffizienz und Robustheit. Es lassen sich kritische Betriebszustände, wie Dachstromabnehmerabstände, systematisch identifizieren und bewerten.

Die Oberleitung zeichnet sich durch eine höchst komplexe Dynamik aus. Es breiten sich Verformungswellen in jeder Komponente (Fahrdraht, Tragseil und Y-Beiseile) mit unterschiedliche Ausbreitungsgeschwindigkeiten aus. Aufgrund der sehr geringen Dämpfung der Oberleitung bleiben mechanische Schwingungen zudem über lange Zeiträume bestehen. Solche schwach gedämpften Systeme neigen dazu, unerwünschte Resonanzeffekte zu erzeugen, welche die Stabilität der Pantograph-Oberleitungs-Interaktion gefährden können.

Neben klassischen Oberleitungssystemen mit Fahrdraht und Tragseil kann das Modell auch auf steife Systeme, wie etwa der Deckenstromschiene, übertragen werden. Diese Systeme weisen deutlich andere dynamische Eigenschaften auf, insbesondere eine wesentlich höhere Biegesteifigkeit.

Ergänzend zur der Simulation nach EN 50318 [1] werden Stromabnehmermessungen nach EN 50317 [2] durchgeführt, um die Wechselwirkung zwischen Pantographen und Oberleitung messtechnisch zu erfassen. Die Firma PJ Messtechnik ist ein ISO/IEC

17025 akkreditiertes Prüfinstitut, das Messfahrten zur Zulassung von Zügen durchführt. Der Messprozess umfasst den Aufbau und die Kalibrierung des Messsystems, die Erfassung der Kontaktkraft zwischen Schleifleiste und Fahrdraht sowie eine detaillierte Datenanalyse zur Bewertung der dynamischen Interaktion.

Keywords: Pantograph-Oberleitungs-Interaktion, Deckenstromschiene, FEM-Modell, OVE EN 50318, OVE EN 50317, Mehrfachtraktion, MATLAB-Simulation, Messtechnik

1 Einleitung

Die dynamische Interaktion zwischen Pantographen und Oberleitungssystemen bestimmt maßgeblich die Betriebssicherheit und den Verschleiß elektrifizierter Bahnsysteme. Da reale Messfahrten ressourcenintensiv und nur begrenzt wiederholbar sind, gewinnen simulationsgestützte Analysen zunehmend an Bedeutung.

Die Analyse solcher komplexen Systeme erfordert präzise Simulationswerkzeuge, die reale Bedingungen möglichst genau abbilden. Die EN 50318:2023 [1] liefert hierfür standardisierte Vorgaben zur Validierung numerischer Modelle. Besonders Anhang B der Norm bietet eine Grundlage zur Modellierung realer Oberleitungsabschnitte und ermöglicht damit den Übergang von idealisierten zu praxisnahen Simulationen.

Der vorliegende Beitrag greift diese Thematik auf und stellt ein neuartiges, in MATLAB implementiertes Finite-Elemente-Modell vor, das für die Analyse der Interaktion zwischen Pantographen und Oberleitung konzipiert ist und auf Recheneffizienz optimiert ist.

Aktuelle Normen und Simulationsstandards berücksichtigen Deckenstromschienen bislang nicht. Der vorgestellte Simulationsansatz erfasst Vorspann- und Steifigkeitsterme und ermöglicht somit auch eine präzise, recheneffiziente und robuste Bewertung dieser steifen Komponenten.

2 Technische Grundlagen und mechanisches Modell der Oberleitung

Die Untersuchung der Interaktion zwischen Pantographen und Oberleitungen bei Mehrfachtraktion ist von zentraler Bedeutung, um den Verschleiß des Fahrdrahts und der Schleifleisten des Pantographen zu identifizieren. Es kann bei Kontaktverlust der Pantographen mit der Oberleitung zur Bildung eines Lichtbogens kommen, welcher zu einem lokalen Schmelzen und Verdampfen des Materials führen kann. Ziel der Simulationsmethodik ist es, das reale mechanische System der Oberleitung in

Simulationen nachzubilden und beispielsweise hinsichtlich Kontaktverlust und Verschleiß kritische Dachstromabnehmerabstände zu identifizieren.

2.1 Technische Grundlagen der Oberleitung

Aus der Differentialgleichung (2.1) einer gespannten Saite ergibt sich die Wellenausbreitungsgeschwindigkeit v_{Welle} zu Gleichung (2.2). Diese ist eine physikalische Grenze für die Betriebsgeschwindigkeit eines Zuges. Nähert sich die Zuggeschwindigkeit der Wellenausbreitungsgeschwindigkeit, wachsen die Schwingungsamplituden rasch in unzulässige Größen und es kommt zu Kontaktverlust oder gar Oberleitungsschäden. In der Praxis bestätigt sich dieses Phänomen bei Hochgeschwindigkeitsfahrten. Bei Fahrten mit $v_{Zug} = v_{Welle}$ kann kein kontinuierlicher Kontakt zwischen Pantographen und Fahrdraht aufrechterhalten werden. Aus praktischen Erfahrungen geht hervor, dass die Wellenausbreitungsgeschwindigkeit mindestens das 1,4 bis 1,5 fache der Fahrtgeschwindigkeit betragen soll. Die Wellenausbreitung ergibt sich aus der Zugkraft T, der Dichte ρ und der Querschnittsfläche A des Fahrdrahts. [3]

$$\frac{\partial^2 w}{\partial x^2} - \frac{T}{\rho A}\frac{\partial^2 w}{\partial t^2} = 0 \quad (2.1)$$

$$v_{Welle} = \sqrt{\frac{T}{\rho A}} \quad (2.2)$$

2.2 Mechanisches Modell

Die Pantograph-Oberleitungs-Interaktion zeichnet sich durch eine höchst komplexe Dynamik aus, die zahlreiche Herausforderungen für die Modellierung und Simulation mit sich bringt. Folgende Aspekte charakterisieren das mechanische System:

Zusammenwirken verschiedener Komponenten

Die Oberleitung besteht aus einer Vielzahl an Komponenten, welche unterschiedliche mechanische Eigenschaften haben. Den größten Einfluss auf die dynamische Bewegung haben Fahrdraht, Tragseil, Hänger, Seitenhalter und Y-Beiseile. In der Oberleitung breiten sich mechanische Wellen aus, welche in jeder Komponente unterschiedliche Ausbreitungsgeschwindigkeiten haben.

Schwache Dämpfung

Aufgrund der sehr geringen Dämpfung der Oberleitung bleiben mechanische Schwingungen über längere Zeiträume bestehen. Solche schwach gedämpften Systeme

neigen dazu, unerwünschte Resonanzeffekte zu erzeugen, welche die Stabilität der Pantograph-Oberleitungs-Interaktion gefährden.

Wellenausbreitung und Reflexionen

Die durch die Interaktion angeregten mechanischen Wellen breiten sich über weite Distanzen entlang der Oberleitung aus. Sie werden an Maststützpunkten, Hängern und Y-Aufhängungen teils reflektiert und bilden in Summe komplexe Überlagerungseffekte, welche die Kontaktkraft zwischen Pantographen und Fahrdraht maßgeblich beeinflussen.

Großes Rechengebiet

Um nicht von Randeffekten und Einschwingvorgängen des Rechengebietes beeinflusst zu werden, muss die Systemgrenze des Oberleitungsmodells ausreichend groß gefasst werden. Der Einfluss von am Rand des Rechengebietes reflektierten Wellen auf den Auswerteabschnitt muss vernachlässigbar gering sein. Aufgrund der hohen Wellenausbreitungsgeschwindigkeit und der Reflexion dieser Wellen ist es notwendig, bereits bei kurzen Simulationsdauern ein großes Rechengebiet zu modellieren. Zusätzlich können auch sog. absorbierende Randbedingungen verwendet werden. [4] [5]

Numerische Steifigkeit und Nichtlinearitäten

Die Dynamik der Interaktion von Pantographen und Oberleitung ist numerisch steif und zeigt nichtlineares Verhalten. Numerische Steifigkeit bedeutet hier, dass das Verhältnis des größten zum kleinsten Eigenwert groß ist und kleine Zeitschritte in der Simulationsrechnung erforderlich macht. Nichtlinearitäten treten beispielsweise beim Durchknicken der Hänger auf und erfordern auch hier eine feine Zeitauflösung der Simulation.

3 Simulation Interaktion Pantographen und Oberleitung

3.1 Simulationsmodell

Die Implementierung erfolgt in MATLAB. Im Gegensatz zu dedizierten Finite-Elemente-Programmen ermöglicht dieser Ansatz die weitreichende Flexibilität bei der Modellbildung und Anpassung der Simulationsmethodik. Das Simulationsmodell basiert auf einer örtlich diskreten Finite-Elemente-Beschreibung der Oberleitung, welche über Hermit'sche Balkenelemente formuliert wird. Ein Element verbindet zwei Knoten mit jeweils zwei Freiheitsgraden, der vertikalen Verschiebung und der Verdrehung. Die

Position in Elementarkoordinaten entspricht ξ und in physikalischen Koordinaten x. Ein Element hat die spezifische Länge $2a$. [6, pp. 111-117]

$$\xi = \frac{x}{a} \tag{3.1}$$

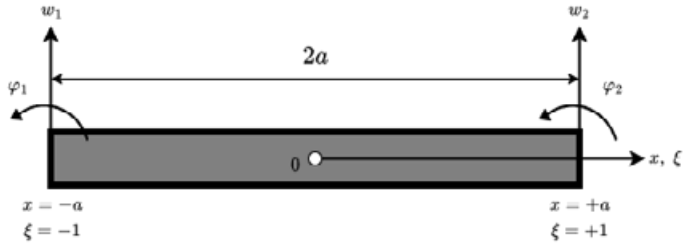


Abbildung 1: Balkenelement und seine lokalen Koordinatensysteme

Die Dynamik des Pantographen wird über ein entsprechendes Mehrkörpersystem modelliert und im Kontaktpunkt mit der Oberleitung gekoppelt. Die Simulation wird mit dem Newmark-Beta-Integrationsverfahren gelöst. [6, pp. 75-76]

Das Simulationsmodell der Interaktion zwischen Pantographen und Oberleitung basiert auf einer partiellen Differentialgleichung (pDGL) 4. Ordnung. (3.2) Die Differentialgleichung (2.1) wird um den Steifigkeitsterm $EI\frac{\partial^4 w}{\partial x^4}$ und einer Kraftdichte $f(x,t)$ erweitert. Die Kraftdichte ist von Position und Zeit abhängig und modelliert die Kontaktkräfte der Pantographen, die Hängerkräfte, die Seitenhalterkräfte und Gewichtskräfte. [7]

$$\rho\, A\frac{\partial^2 w}{\partial x^2} + EI\frac{\partial^2 w}{\partial x^2} - T\frac{\partial^2 w}{\partial t^2} = f(x,t) \tag{3.2}$$

Um die Systemmatrizen zu erhalten, muss die diskretisierte pDGL in die schwache Form überführt werden. Nach dem Aufbau der globalen Systemmatrizen aus den Elementmatrizen [6] ergibt sich ein gewöhnliches Differentialgleichungssystem 2. Ordnung, siehe Gleichung (3.5).

$$\boldsymbol{M}\ddot{\boldsymbol{q}} + \boldsymbol{D}(\boldsymbol{q})\dot{\boldsymbol{q}} + \boldsymbol{K}(\boldsymbol{q}) = \boldsymbol{F}(\boldsymbol{q}) \tag{3.3}$$

Die Dämpfung wird mittels Rayleigh-Dämpfung laut OVE EN 50318 [1] realisiert. Die globale Dämpfungsmatrix $\boldsymbol{D}$, Steifigkeitsmatrix $\boldsymbol{K}$ sowie der Kraftvektor $\boldsymbol{F}$ sind von dem Lagevektor $\boldsymbol{q}$ abhängig und müssen in jedem Zeitschritt aktualisiert werden. Es wird das Durchknicken der Hänger bei Druckbelastung, der Kontakt der Pantographen mit dem Fahrdraht und die Seitenhalterdynamik berücksichtigt.

Die FEM-Oberleitung wird mit einer steifen Feder für Zugbelastung mit dem Pantographen gekoppelt. [8] Die Pantographen werden als Mehrkörpersystem (MKS) laut OVE EN 50318 [1] modelliert. [9]

Im Folgenden sind drei Spannen der Hochgeschwindigkeits-Wechselstrom-Oberleitung mit Y-Aufhängung dargestellt. In Abbildung 2 ist die Seitenansicht und in Abbildung 3 ist die Draufsicht zu sehen.

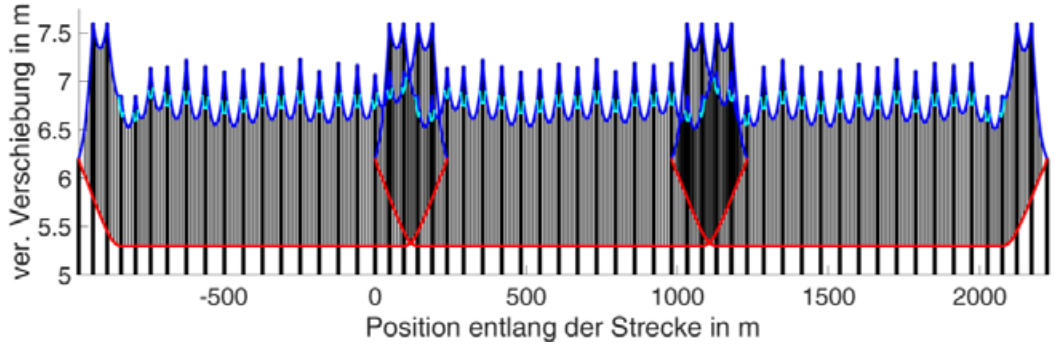


Abbildung 2: Seitenansicht der Oberleitung

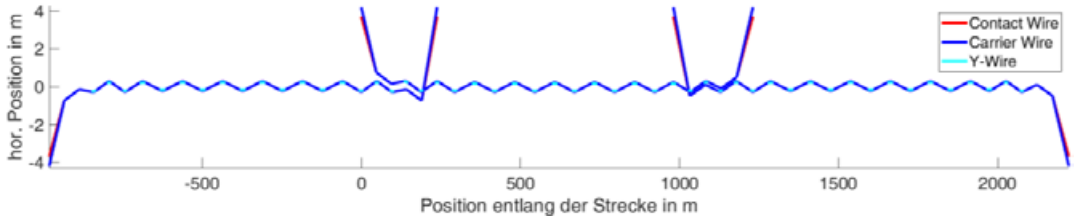


Abbildung 3: Draufsicht der Oberleitung

3.2 Konvergenzanalyse

Die Wahl der räumlichen Diskretisierung Δx und der Zeitschrittweite Δt ist entscheidend, um mit minimalem Rechenaufwand eine ausreichend genaue Lösung zu gewährleisten. Eine zu grobe Diskretisierung führt zu fehlerhaften Simulationsergebnissen, während eine zu feine Diskretisierung die Rechenzeit unnötig erhöht. Daher ist eine Konvergenzanalyse erforderlich, um ein optimales Gleichgewicht zwischen Genauigkeit und Rechenaufwand zu finden. Als Fehlermaß zur Bewertung der Genauigkeit der Simulation wird der relative Fehler $E_{sim,rel}$ herangezogen. Dieses beschreibt den Quotienten zwischen dem Fehler der Simulation gegenüber der Referenzlösung (E_{sim}) und dem Fehlermaß der Referenzlösung (E_{ref}).

$$E_{sim,rel} = \frac{E_{sim}}{E_{ref}} \cdot 100 \tag{3.4}$$

Das Fehlermaß der Referenzlösung ist als die Summe der quadratischen Abweichungen der Referenzkontaktkraft von dessen Mittelwert definiert. Das Signal wird an N äquidistanten Zeitschritten ausgewertet.

$$E_{ref} = \sum_{i=1}^{N} (F_{ref,i} - \bar{F}_{ref})^2 \tag{3.5}$$

Der Fehler der Simulation gegenüber der Referenzlösung ist die Summe der quadratischen Abweichungen der Kontaktkraft der Referenzlösung über N äquidistante Zeitschritte.

$$E_{sim} = \sum_{i=1}^{N} (F_{sim,i} - F_{ref,i})^2 \tag{3.6}$$

Die Konvergenzanalyse der Kontaktkraft bezieht sich auf den nachlaufenden Dachstromabnehmers. Es wird die Hochgeschwindigkeits-Wechselstrom-Oberleitung mit Y-Aufhängung nach EN 50318: Anhang B.2 [1] analysiert. Der Pantographenabstand beträgt 192m und die Zuggeschwindigkeit 302 km/h. Die Simulationsdauer ist 2.5s. Es wird jeweils die Elementgröße einer Komponente (Tragseil, Fahrdraht oder Y-Beiseil) variiert und die anderen Komponentenlängen mit 0.5m konstant gehalten. Die Referenzlösung ist eine Elementgröße Δx von 0.25m für alle Oberleitungskomponenten.

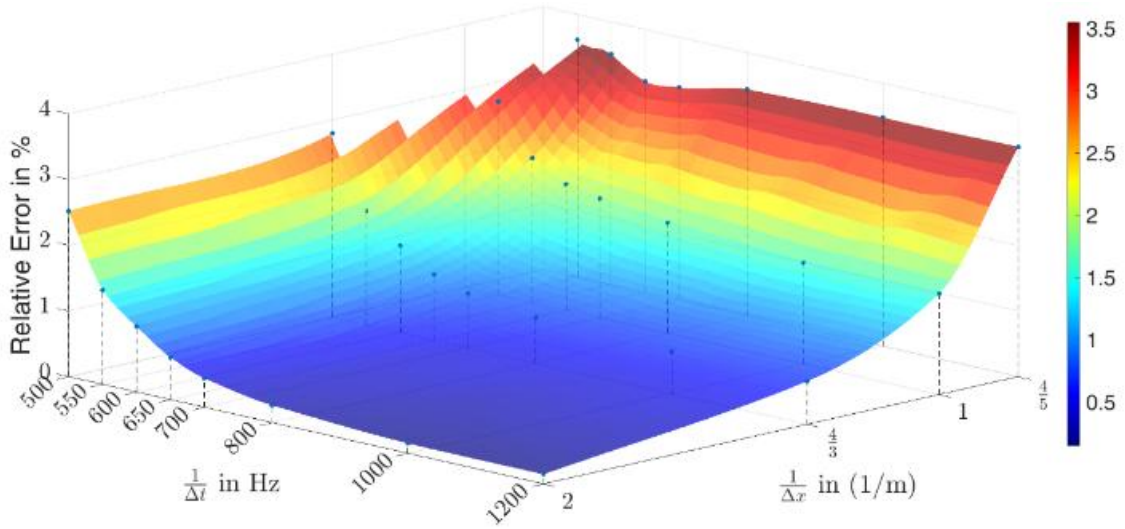


Abbildung 4: Konvergenz Fahrdraht

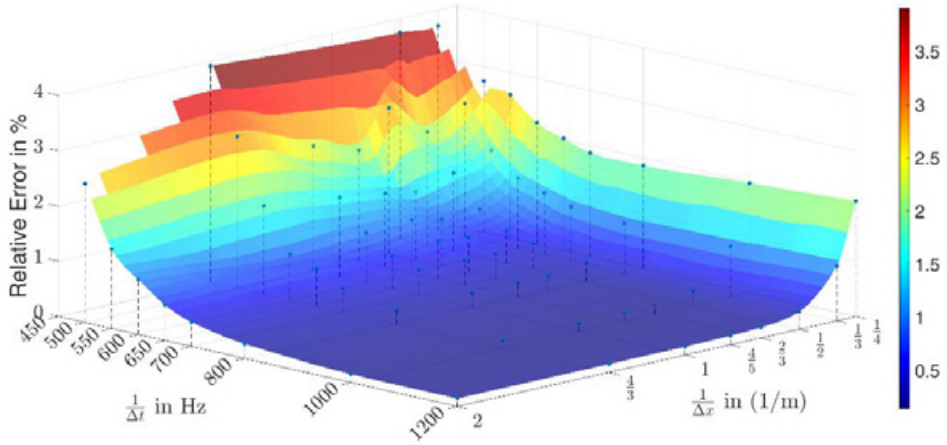


Abbildung 5: Konvergenz Tragseil

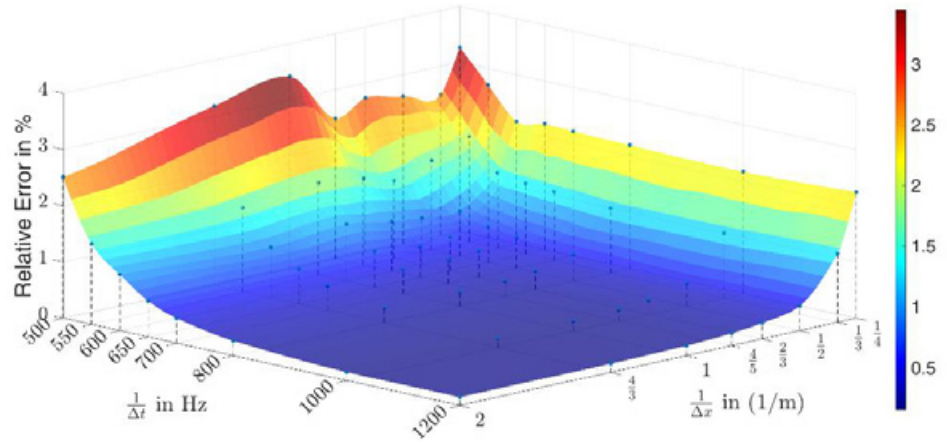


Abbildung 6: Konvergenz Y-Beiseil

Rechenaufwand

Neben der Genauigkeit spielt auch der Rechenaufwand eine entscheidende Rolle. Die Abhängigkeit des Rechenaufwands von der Diskretisierung zeigt sich durch:

$$t_{CPU} \propto \frac{1}{\Delta N} \cdot \frac{1}{\Delta t} \tag{3.7}$$

ΔN entspricht der Anzahl der Elemente.

3.3 Validierung

Die Validierung der Hochgeschwindigkeits-Wechselstrom-Oberleitung mit Y-Aufhängung mit gemessenen Daten wird nach EN 50318: Anhang B.2 [1] durchgeführt.

Die Kontaktkraft wird entsprechend den Vorgaben der EN 50317 mithilfe eines Filters sechster Ordnung geglättet.

Es werden drei Sektionen mit einer Gesamtlänge von 3204,25 m simuliert. Der Auswerteabschnitt ist der Bereich der mittleren Sektion und entspricht der Simulationszeit zwischen 5- und 22 s.

Es wird die Validierung mit folgenden Auflösungen durchgeführt. Es ist anzumerken, dass um Klemmen und Aufhängungen feiner diskretisiert wird als die maximale Elementauflösung. Alle Simulationen zeigen Ergebnisse, welche in den von der Norm vorgesehenen Intervallen liegen. Der relative Fehler $E_{sim,rel}$ bezieht sich auf die ungefilterte Kontaktkraft des nachlaufenden Pantographen und wird nach Formel (3.6) berechnet.

Tabelle 1: Elementauflösung, Zeitschrittweite, Rechenzeit und rel. Fehler

	Fahrdraht	Tragseil	Y-Beiseil	Zeitschrittweite	Rechenzeit	$E_{sim,rel}$
Variante 1	1 m	4 m	4 m	1/600 s	71 s	53.33 %
Variante 2	0.75 m	2 m	2 m	1/650 s	107 s	8.13 %
Variante 3	0.5 m	1.5 m	1.5 m	1/800 s	207 s	2.60 %
Variante 4	0.25 m	0.5 m	0.5 m	1/1200 s	669 s	0.05 %
Referenz	0.125 m	0.25 m	0.25 m	1/2000 s	1943 s	-

Im Folgenden wird die gefilterte Kontaktkraft der Variante 4 dargestellt:

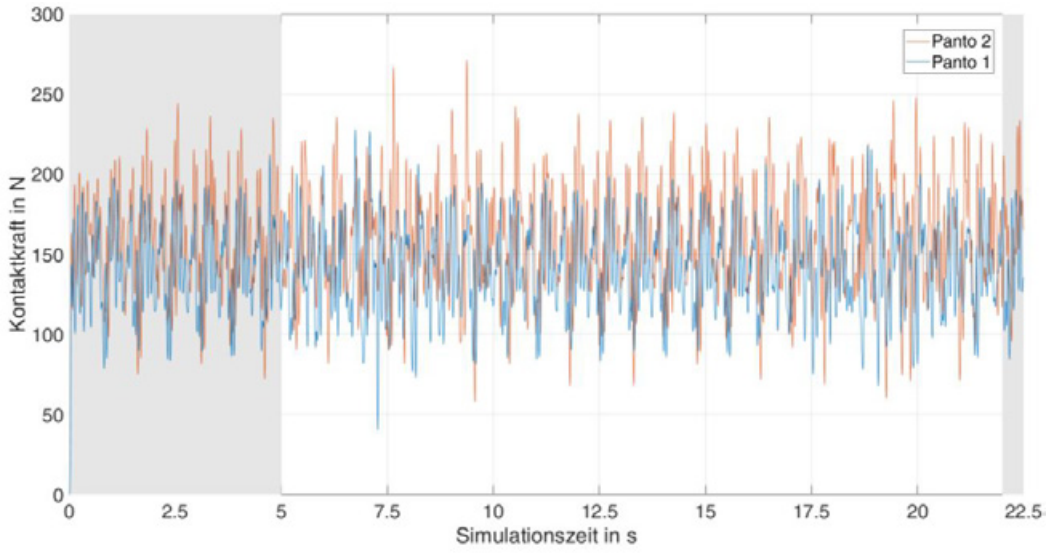


Abbildung 7: Kontaktkraft der Pantographen der Variante 4

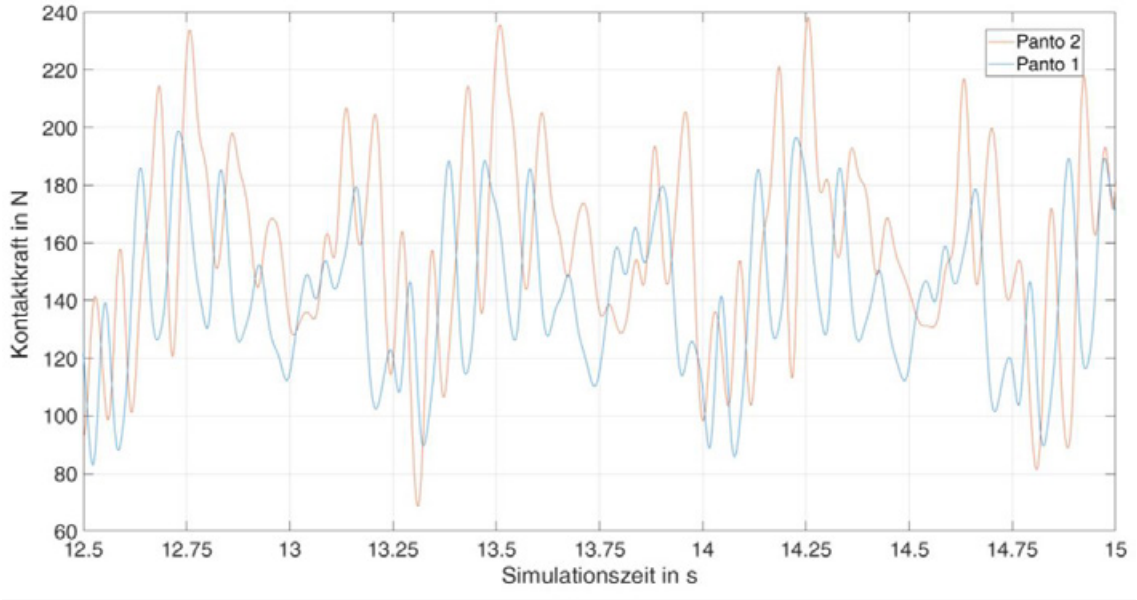


Abbildung 8 : Teilausschnitt der Kontaktkraft der Variante 4

Statische Validierung

Zum Vergleich der Simulationsdaten mit den statischen Werten werden die Verschiebung des Fahrdrahts, die Hängerlängen und die Elastizität verglichen. Die Elastizität entspricht dem Anhub geteilt durch die am Fahrdraht wirkende Kraft.

Für die statische Validierung werden wie in der Norm EN 50318: B.2.4 [1] vorgesehen die Stützpunkte und Hängerpositionen von Spannenfeld S14 bis S15 (419.02 – 544.98 m) betrachtet. Die Diskretisierung entspricht Variante 4.

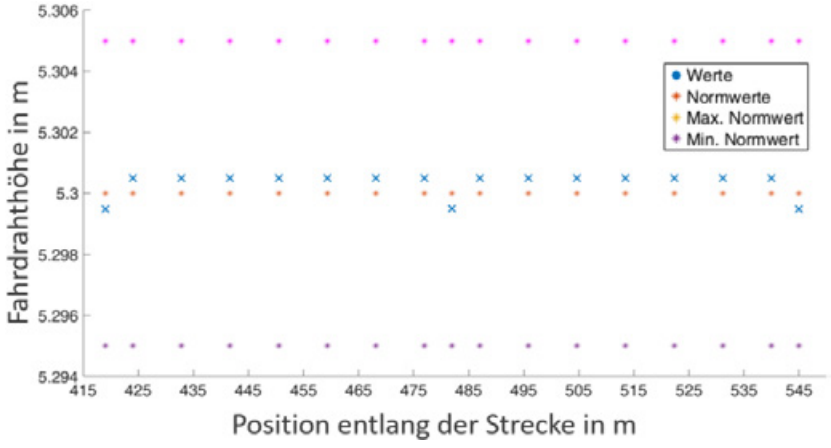


Abbildung 9: Fahrdraht Position

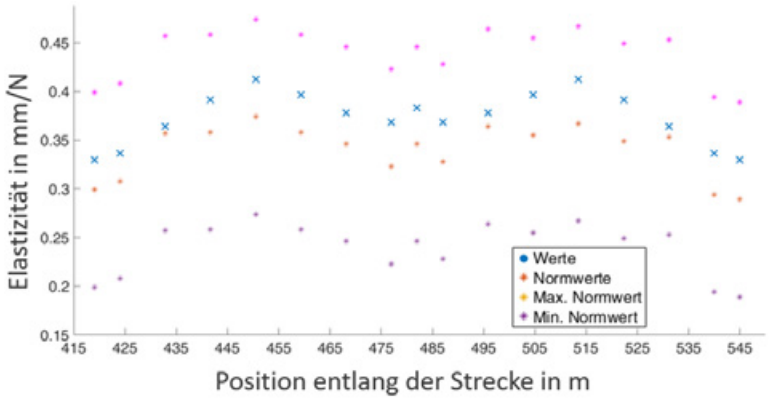


Abbildung 10 : Elastizität der Oberleitung

Dynamische Validierung

Für die Validierung des dynamischen Zusammenwirkens sind die Simulationsergebnisse mit den Messdaten aus EN 50318: B.2.5 zu vergleichen. Es folgt der Vergleich der simulierten Kontaktkraft mit den Normintervallen. Die Zahlenwerte sind in SI-Einheiten angegeben.

```
Checking Pantograph 1:
_____________________________________________________________________________
In Norm OVE EN 50318 interval:
   Yes   |    No   |
-----------------------------------------------------------------------------
    x    |         | meanValue: Value: 141.607 is in Norm interval [140.5, 145.5]
    x    |         | stdValue: Value: 26.6431 is in Norm interval [21.6, 32.4]
    x    |         | stdValue_FFT: Value: 26.6431 is in Norm interval [21.04, 31.56]
    x    |         | stdValue_5Hz: Value: 18.4596 is in Norm interval [15.84, 23.76]
    x    |         | stdValue_5Hz_20Hz: Value: 19.0315 is in Norm interval [13.84, 20.76]

Checking Pantograph 2:
_____________________________________________________________________________
In Norm OVE EN 50318 interval:
   Yes   |    No   |
-----------------------------------------------------------------------------
    x    |         | meanValue: Value: 159.227 is in Norm interval [158.5, 163.5]
    x    |         | stdValue: Value: 32.1675 is in Norm interval [26.08, 39.12]
    x    |         | stdValue_FFT: Value: 32.1675 is in Norm interval [25.28, 37.92]
    x    |         | stdValue_5Hz: Value: 22.5357 is in Norm interval [18.16, 27.24]
    x    |         | stdValue_5Hz_20Hz: Value: 22.6846 is in Norm interval [17.6, 26.4]
```

Abbildung 11 : Ausgabe Matlab Vergleich Kontaktkraft mit EN 50318

Des Weiteren wird in der EN 50318: B.2.5 der Anhub an Stützpunkt M13 und an Stützpunkt M14 vorgegeben.

Gemäß EN 50318 [1] muss der Anhub an den Stützpunkten M13 und M14 innerhalb von +20 mm bzw. -10 mm des Normwertes liegen.

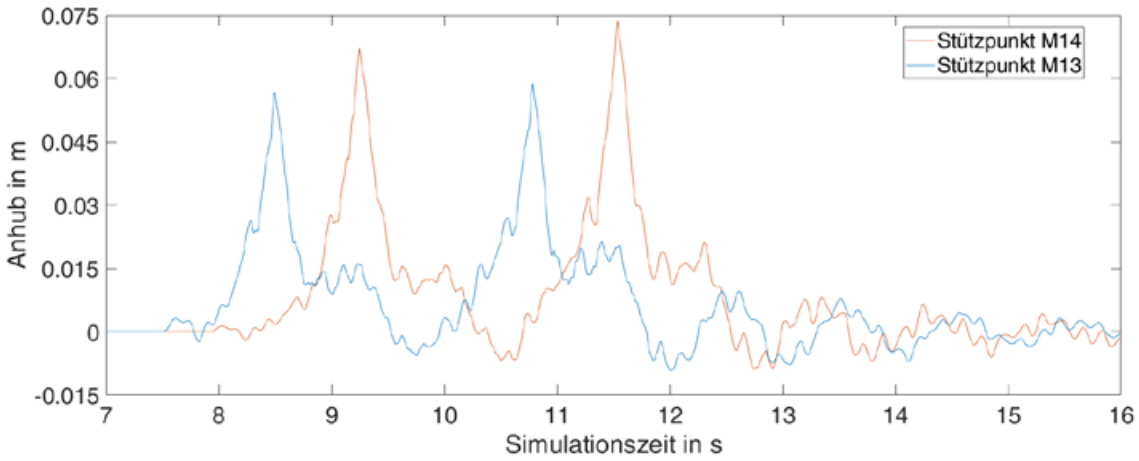


Abbildung 12: Anhub an den Stützpunkten

Tabelle 2: Normwerte Anhub EN 50318

Pantograph	vorlaufender	nachlaufender
	Anhub Stützpunkt M13	
Normwerte EN 50318	47 mm	49 mm
Simulationsergebnis	57.3 mm	59.5 mm
	Anhub Stützpunkt M14	
Normwerte EN 50318	52 mm	56 mm
Simulationsergebnis	67.7 mm	74.2 mm

4 Simulation Deckenstromschiene

Es wird der Simulationsansatz auf eine Deckenstromschiene übertragen und der besonders kritische Übergang von der AC-Oberleitung mit Y-Aufhängung und Stromschiene simuliert. Damit wird der im Beitrag skizzierte Anwendungsbereich auf Systeme erweitert, die in aktuellen Normen nicht abgedeckt sind.

Die Stromschiene wird analog zur Oberleitung ohne Zugkraft formuliert. Die mechanischen Kennwerte der Stromschiene wurden aus [10] entnommen. Dier Dämpfung der Lagerung ist geeignet zu wählen.

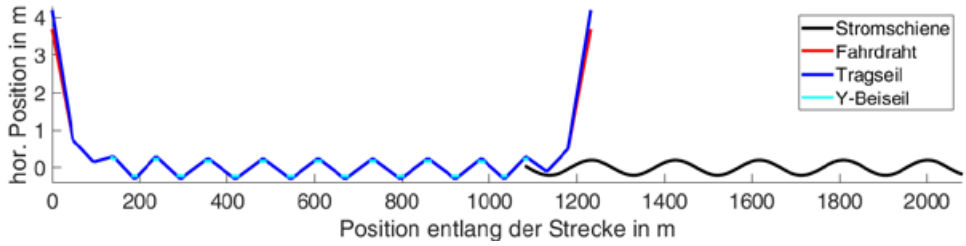


Abbildung 13: Draufsicht Oberleitung und Deckenstromschiene

Die Konvergenzuntersuchung zeigt für die Deckenstromschiene einem niedrigen relativen Fehler der Kontaktkraft bereits bei signifikant gröberer Diskretisierung als bei der elastischen Oberleitung. Eine Diskretisierung mit äquidistanten Elementen von 2.0 m und einer Abtastfrequenz von 1000 Hz führt zu einem Fehler von lediglich 0.219 % gegenüber der Referenzlösung (0.0625 m; 3200 Hz, vgl. Kapitel 3.2). Dies bestätigt die Robustheit des Verfahrens für steife Systeme.

Für die elastische Oberleitung ergibt sich, dass bei maximalen Elementlängen von 0.5 m (Fahrdraht), 2.0 m (Tragseil) und 1.0 m (Y-Beiseil) sowie einer Abtastfrequenz von 800 Hz der relative Fehler der Kontaktkraft 0.446 % beträgt, bezogen auf die Referenz mit einheitlich 0.1 m und 3200 Hz. Damit lässt sich die Oberleitung gegenüber der Referenz deutlich gröber und damit rechenzeiteffizient diskretisieren, ohne die Güte der Kontaktkraftergebnisse spürbar zu beeinträchtigen.

Der Übergangsbereich erweist sich, konsistent mit der Praxis, als kritischer Abschnitt. In den folgenden Abbildungen werden die Kontaktposition und die Kontaktkraft über der statischen Ruhelage der Oberleitung dargestellt.

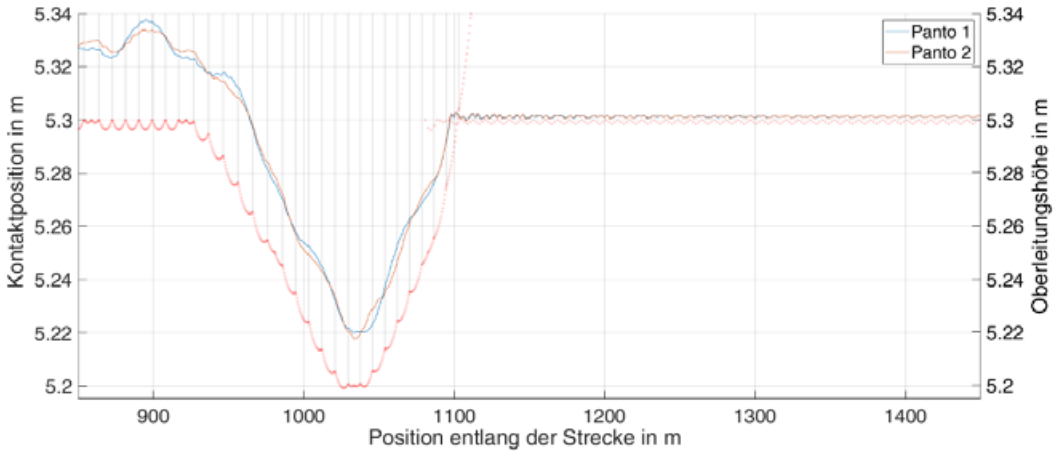


Abbildung 14: Kontaktverschiebung über Position

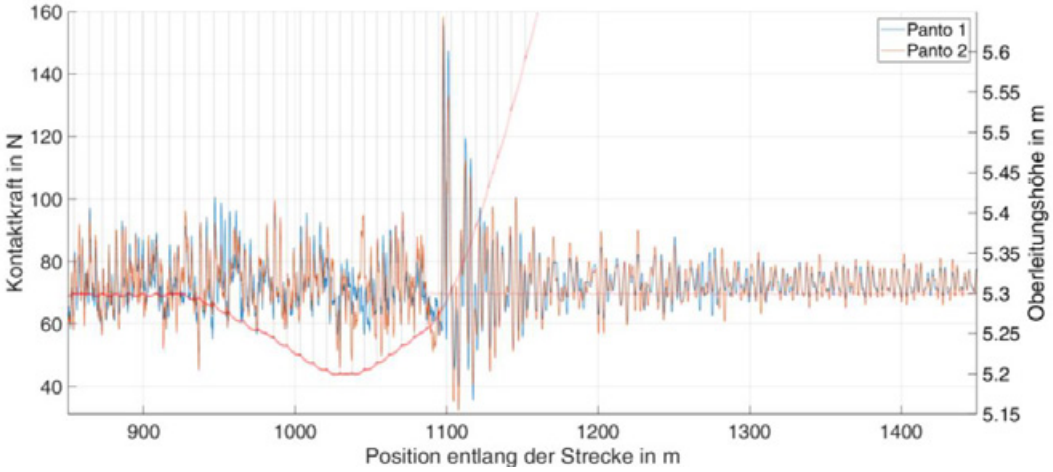


Abbildung 15: Kontaktkraft über Position

Die Übertragung des Simulationsmodells auf die Deckenstromschiene gelingt ohne strukturelle Änderungen am Lösungsverfahren. Die Resultate belegen die Machbarkeit und Funktionstüchtigkeit des Ansatzes für steife Stromschienen und schaffen die Basis für weiterführende Untersuchungen

5 Schlussteil

Im vorliegenden Beitrag wird ein Finite-Elemente-Modell zur Simulation der dynamischen Interaktion zwischen Pantographen und Oberleitung vorgestellt und umfassend validiert.

Das entwickelte Modell ist flexibel, robust und recheneffizient. Es ermöglicht die Identifikation kritischer Dachstromabnehmerabstände zur effizienten Planung und Durchführung von Messfahrten zur Fahrzeugzulassung. Zudem können Oberleitungsfehler analysiert werden und das Modell kann zur Optimierung von Oberleitungsinfrastrukturen verwendet werden.

Aktuelle Normen und Simulationsstandards bilden Deckenstromschienen sowie singuläre Elemente, wie Trennstellen, nicht ab. Dadurch bleibt das dynamische Verhalten dieser besonders kritischen Stellen unberücksichtigt. Der vorgestellte Simulationsansatz bildet sowohl Vorspann- als auch Steifigkeitsterme der vorgespannten Komponenten ab, wodurch auch steife Komponenten wie die Deckenstromschienen präzise, robust und besonders effiziente bewertet werden können.

Literatur

[1] OVE, „EN 50318: Bahnanwendungen – Stromabnahmesysteme – Validierung von Simulationssystemen für das dynamische Zusammenwirken zwischen Dachstromabnehmer und Oberleitung," Wien, 2023.

[2] OVE, „ÖVE/ÖNORM EN 50317: Bahnanwendungen – Stromabnahmesysteme – Anforderungen und Validierung von Messungen des dynamischen Zusammenwirkens zwischen Stromabnehmer und Oberleitung," Wien, 2012.

[3] F. Kießling, R. Puschmann und A. Schmieder, Fahrleitungen elektrischer Bahnen – Planung, Berechnung, Ausführung, Betrieb, 3. überarb. u. erw. Aufl. Hrsg., Erlangen: Publicis Publishing, 2014, pp. 555-561.

[4] A. Schirrer, E. Talic, G. Aschauer, M. Kozek und S. Jakubek, „Optimization based determination of highly absorbing boundary conditions for linear finite difference schemes," *Journal of Sound and Vibration,* Bd. 365, p. 45 – 69, 2016.

[5] D. Ritzberger, A. Schirrer und S. Jakubek, „Absorbing boundary layer control for linear one-dimensional wave propagation problems," *Journal of Vibration and Control,* Bd. 24, Nr. 6, p. 1019 – 1031, 2018.

[6] G. R. Liu und S. S. Quek, The Finite Element Method – A Practical Course, 2. Auflage Hrsg., Oxford: Butterworth-Heinemann (Elsevier), 2014.

[7] G. Aschauer, A. Schirrer und S. Jakubek, „Realtime-capable Finite Element Model of Railway Catenary Dynamics in Moving Coordinates," in *2016 IEEE Conference on Control Applications (CCA), Part of the 2016 IEEE Multi-Conference on Systems and Control*, Buenos Aires, Argentina, 2016.

[8] Q. Luo, G. Mei, G. Chen und W. Zhang, „Study of pantograph-catenary system dynamic in crosswind environments," Vehicle System Dynamics, 2023.

[9] M. Simarro, S. Postigo, P. Reyes, R. J. Luque, P. V. Baćić und J. J. Castillo, „A study on the dynamic response of the double pantograph interaction with rigid catenary," *Engineering Structures,* 2025.

[10] X. Feng, S. Gao, Y. Song, Z. Hu, L. Chen und T. Liang, „Static and Dynamic Analysis of Conductor Rail with Large Cross-Sectional Moment of Inertia in Rigid Catenary Systems," *Energies,* Bd. 16, Nr. 4, 2023.

Session 5

Matthias Zelinka, Martin Joch, Alexander Schirrer, Stefan Jakubek

Autore



Zelinka, Matthias

Matthias Zelinka arbeitet als Simulation Engineer bei der PJ Messtechnik GmbH in Graz und entwickelt und validiert das MATLAB Finite-Elemente-Modell zur Abbildung der dynamischen Wechselwirkung zwischen Dachstromabnehmern und Oberleitung. Er ist im Masterstudium Maschinenbau an der Technischen Universität Wien.



Joch, Martin

Martin Joch ist der Geschäftsführer von PJ Messtechnik GmbH und Leiter der Prüfstelle für Schienenfahrzeuge. Er war 8 Jahre am Institut für Maschinenelemente als Universitätsassistent tätig. 2006 gründete er mit Günter Petschnig PJ Messtechnik.



Schirrer, Alexander

Alexander Schirrer ist Privatdozent am Institut für Mechanik und Mechatronik der Technischen Universität Wien. Er lehrt und forscht in den Bereichen Modellierung, Simulation, Optimierung und Regelung linearer und nichtlinearer Systeme sowie die Modellierung und Regelung verteilter (ortsabhängiger) Systeme.



Jakubek, Stefan

Er ist Leiter des Instituts für Mechanik und Mechatronik der Technischen Universität Wien und Universitätsprofessor für Regelungstechnik und Anwendungen. Seine Forschungsinteressen umfassen Fehlerdiagnose, nichtlineare Systemidentifikation, Modellierung und Simulation sowie Regelungstheorie und deren Anwendungen.

Session 5
Philipp Leibner, Jannik Goersch, Benedikt Neubauer, Thomas Hempel, Christian Schindler, Raphael Pfaff

Entwicklung eines integrierten digitalen Zwillings auf Basis von Open Data für die Gleisinstandhaltung mittels Smartphone Sensorik

Leibner, Philipp[1]; Goersch, Jannik[1]; Neubauer, Benedikt[2];
Hempel, Thomas[2]; Schindler, Christian[1]; Pfaff, Raphael[1]

[1] Institut für Schienenfahrzeuge,
RWTH Aachen University
[2] Siemens Mobility GmbH

Zusammenfassung

Mit dem kontinuierlichen Anstieg der Fahrgastzahlen im Schienenverkehr und dem wachsenden Fokus auf nachhaltige Mobilität steigen die Belastungen der Bahninfrastruktur. Daraus ergibt sich die Notwendigkeit, Instandhaltungsprozesse effizient zu gestalten. Vor diesem Hintergrund wurde ein digitaler Zwilling auf Basis von Open-Source-Daten entwickelt, der als Werkzeug zur Optimierung der Instandhaltung dient. Der Zwilling verknüpft Sensordaten aus Fahrzeugen direkt mit der Gleisinfrastruktur und bildet so die Grundlage für die Entwicklung individueller Algorithmen für die zustandsorientierte Instandhaltung der Fahrwege. Zur praktischen Erprobung wurden über mehr als zwölf Monate Smartphones in Straßen- und Stadtbahnen in Frankfurt am Main eingesetzt, um Beschleunigungs- und Standortdaten zu erfassen. Die Auswertung zeigt, dass mit dem digitalen Zwilling und der kostengünstigen Messtechnik Veränderungen der Infrastruktur automatisiert erkannt werden können. Dies eröffnet neue Möglichkeiten für vorausschauende Wartungsplanung und gezielte Ressourcennutzung, wodurch Betreiber Instandhaltungsmaßnahmen effizient priorisieren und einen zuverlässigen Bahnbetrieb sicherstellen können.

Keywords: Data Analytics, Datafusion, Digital Twin, Open Data, Eisenbahninfrastruktur, Mobile Anwendung, Monitoring, Smartphone-Sensorik, Track Monitoring

Session 5
Philipp Leibner, Jannik Goersch, Benedikt Neubauer, Thomas Hempel, Christian Schindler, Raphael Pfaff

1 Einleitung

Angesichts der globalen Zielsetzung, Verkehre nachhaltiger zu gestalten, werden Bahnsysteme weltweit erweitert und neu aufgebaut. Gleichzeitig fließen erhebliche Investitionen in moderne Technologien wie den automatisierten Fahrbetrieb, um die Kapazität und Leistungsfähigkeit vorhandener Strecken weiter zu steigern. Der stetige Anstieg des Schienenverkehrs geht jedoch mit einer intensiveren Beanspruchung der Fahrwege einher, was wiederum einen wachsenden Bedarf an Instandhaltung nach sich zieht. Vor dem Hintergrund, dass der Infrastrukturbereich etwa 30 Prozent der gesamten Betriebskosten im Bahnwesen ausmacht [1], sind optimierte Instandhaltungskonzepte entscheidend, um sowohl Kostenstrukturen effizient zu halten als auch die Wettbewerbsfähigkeit gegenüber anderen Transportmitteln zu bewahren. Hierfür wird zunehmend der Wechsel von einer ausschließlich korrektiven oder präventiven Instandhaltung zu einer prädiktiven und zustandsbasierten Instandhaltung (engl.: Condition-based Maintenance (CBM)) favorisiert, um die Kosten weiter zu senken (vgl. Abb. 1). Wird ein Mangel frühzeitig erkannt oder prognostiziert, profitieren Betreiber zusätzlich durch einen längeren zeitlichen Horizont für die Planung und Personaldisposition sowie die Organisation von ggf. externer Dienstleistern und technischen Ressourcen.

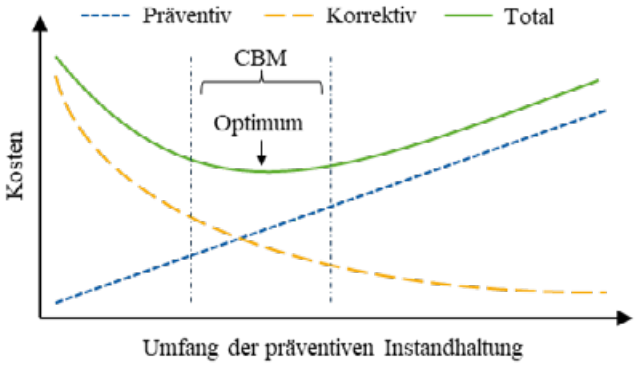


Abbildung 1: Kosten für die präventive und korrektive Instandhaltung

Um einen solchen Paradigmenwechsel in der Instandhaltung umsetzen zu können und die Entwicklung von Schienen- und Gleisdefekten möglichst präzise vorherzusagen, ist eine fortlaufende, automatisierte Datenerhebung und -analyse unabdingbar. Sensorik kann dabei sowohl an streckenseitigen Anlagen, beispielsweise Brücken, Weichen oder Bahnhofsbereichen, als auch direkt auf den Fahrzeugen im Regelbetrieb installiert werden. Durch die Kombination beider Ansätze lässt sich eine umfassende und kontinuierliche Infrastrukturüberwachung realisieren [2]. Besonders innovativ ist die neuere Herangehensweise, Sensorik nicht in speziellen Messzüge, sondern in Fahrzeugen im Fahrgastbetrieb zu installieren. Diese Züge erfassen im Regelbetrieb permanent Zustands- und

Umgebungsdaten, wodurch die Instandhaltungsplanung unmittelbar von aktuellen Betriebsdaten profitieren kann [3]. Gerade für städtische Verkehrsbetriebe mit einem kleinen Streckennetz ist der Einsatz dedizierter Messfahrzeuge häufig nicht wirtschaftlich. Daher werden gesetzlich geforderten Inspektionen oftmals an externe Anbieter ausgelagert.

Gleichwohl bestehen bei etablierten Monitoring-Verfahren Schwächen, insbesondere was die nahtlose Integration in Geoinformationssysteme (GIS) und die exakte Zuordnung von Analyseergebnissen zu spezifischen Infrastrukturteilen betrifft. Ergebnisse werden häufig lediglich als Punkte auf einer Karte visualisiert, wodurch für das Instandhaltungsmanagement nicht immer ein direkter Asset-Bezug ersichtlich ist. Ebenso hemmt die uneinheitliche Nutzung von standardisierten Geodatenformaten in den verschiedenen Betrieben sowie eine teils lückenhafte Digitalisierung der Infrastrukturdaten das Aufsetzen von digitalen Zwillingen. Um hierfür eine skalierbare und übertragbare Lösung zu schaffen, prüft der vorliegende Beitrag den Einsatz von OpenStreetMap (OSM) als zentraler Geodatenquelle. OSM enthält bereits ein breites Spektrum an, von der Nutzergemeinschaft bereitgestellten, bahnspezifischen Informationen. Dazu zählen unter anderem der Verlauf von Strecken, die Positionen von Weichen, Bahnübergängen, Betriebsstellen und weiteren Infrastrukturelementen. Hierbei variiert die Dichte der Daten allerdings je nach geographischer Region, wobei sie im europäischen Raum allgemein als sehr hoch gilt.

2 Schienen- und Gleisfehler

Innerhalb des Eisenbahnbetriebs lassen sich eine Vielzahl unterschiedlicher Fehlerbilder an Schiene und Gleis beobachten. Die folgende Auflistung beschränkt sich auf jene Defekte, die im Zusammenhang mit der Zustandsüberwachung der Fahrwege von Bedeutung sind. Zu den zentralen Fehlerbildern zählen sogenannte Gleislagefehler, also das Abweichen der tatsächlichen Gleislage von den vorgesehenen geometrischen Vorgaben. Längshöhenfehler, Überhöhungsfehler, Richtungsfehler sowie Spurweitenfehler stellen die vier Kategorien dieser Defekte dar (vgl. Abb. 2).

Darüber hinaus existieren schienenspezifische Schäden, die direkt den Schienenkörper betreffen. Typische Beispiele hierfür sind die Ausbildung von Wellenstrukturen auf der

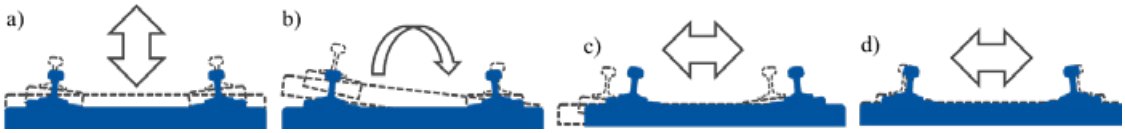


Abbildung 2: Gleislagefehler: a) Längshöhenfehler; b) Überhöhungsfehler; c) Richtungsfehler; d) Spurweitenfehler

Schienenoberfläche, auch Riffel genannt (vgl. Abb. 3), sowie Ausfahrungen an der inneren Seite der Außenschiene in Kurven. Kurzwellige Deformationen entstehen meistens durch das wiederholte Beschleunigen und Abbremsen im Bereich von Haltestellen oder engen Kurvenradien; sie führen einerseits zu erhöhten Schwingungen, andererseits auch zu einer verstärkten Geräuschbelastung und zu zusätzlichen Beanspruchungen des gesamten Oberbausystems.



Abbildung 3: Schienenriffel [4]

Speziell im Bereich von Straßen- und Stadtbahnen treten die, oben erwähnten, ausgeprägten Ausfahrungen an der Außenschiene auf. Verantwortlich hierfür ist der verstärkte Kontakt zwischen Radkranz und Schienenflanke, was infolge der dort üblichen engen Gleisbögen wesentlich häufiger auftritt als im klassischen Vollbahnbereich [4]. Sämtliche beschriebene Problemstellungen bringen eine Minderung des Fahrkomforts mit sich und können – je nach Ausprägung und Schweregrad – zu sicherheitsrelevanten Risiken im Bahnbetrieb führen. Daher ist eine regelmäßige und systematische Inspektion der Infrastruktur unverzichtbar.

2.1 Gleisüberwachung mit kostengünstigen Geräten

In den letzten Jahren ist das Interesse an kostengünstigen Messsystemen, die meist auf mikroelektromechanischen Sensoren (MEMS) basieren, im Bereich der Zustandsüberwachung im Bahnbereich deutlich gestiegen [5], [6]. Hervorzuheben sind hierbei Smartphones: Sie sind nahezu überall verfügbar, einfach zu transportieren und können rückwirkungsfrei, also ohne Eingriffe in die Fahrzeugsoftware oder mechanische Modifikationen an den Zügen, eingesetzt werden. Daraus ergeben sich keine regulatorischen Hemmnisse, wodurch sich diese Geräte problemlos für den mobilen Einsatz in Bestandsfahrzeugen eignen. Darüber hinaus eröffnet die integrierte Internetanbindung die Möglichkeit, erfasste Messdaten unmittelbar und automatisiert an eine Cloud-Umgebung zu übertragen. Typische Einschränkungen solcher Smartphone-Lösungen bestehen jedoch im Bereich der Sensorqualität, einer begrenzten Abtastrate sowie möglichen Abweichungen durch uneinheitliche Platzierung der Geräte im Fahrzeug [7].

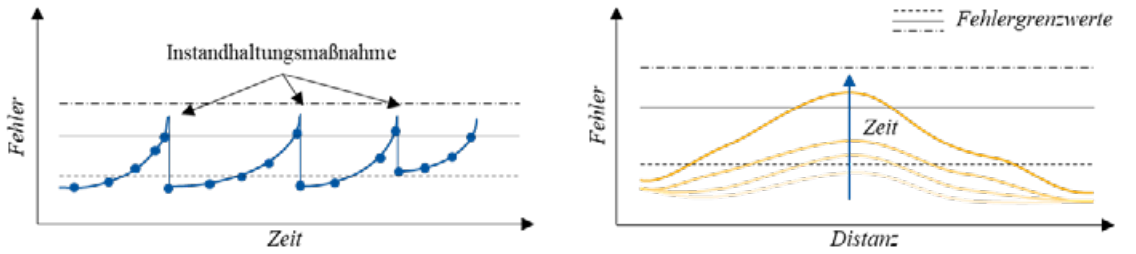


Abbildung 4: Fehlerentwicklung bei der zustandsbasierten Instandhaltung

Um Nachteile bzgl. der Sensorgenauigkeit sowie fahrzeugseitiger Einflüsse ausgleichen zu können, ist es daher vorteilhaft eine möglichst regelmäßige Datenaufnahme durchzuführen. Zudem ermöglicht diese kontinuierliche Datenerhebung eine fundierte Trendanalyse, da durch die erhöhte Erfassungsdichte der Abstand zwischen den einzelnen Messungen verringert wird. Letztlich lassen sich so Fehlerentwicklungen beobachten und frühzeitig prognostizieren (vgl. Abb. 4), was einen wichtigen Beitrag zur Umsetzung eines zustandsbasierten Instandhaltungskonzept leistet.

Bereits in früheren Untersuchungen wurde die Leistungsfähigkeit von Smartphone-Sensorik bestätigt, insbesondere im Hinblick auf die Ermittlung von Längshöhenfehlern [8]. Diese Vorgehensweise wurde auch in dieser Studie verwendet. Für die Erfassung von Defekten an Schiene und Gleis werden die gemessenen Beschleunigungsdaten mit Informationen aus dem Globalen Navigationssatellitensystem (GNSS) kombiniert. Die Beschleunigungsdaten werden mit 100 Hz aufgenommen, während die GNSS-Informationen zunächst mit 1 Hz aufgezeichnet und anschließend zur Vereinheitlichung ebenfalls auf 100 Hz interpoliert werden. Die genaue Zuordnung der Messwerte zu spezifischen Streckenpositionen ermöglicht eine präzise Lokalisierung und Zuordnung von Infrastrukturfehlern. Die Berechnung des Längshöhenfehlers basiert auf einer zweifachen Integration der vertikalen Beschleunigung, geteilt durch das Quadrat von der Fahrzeuggeschwindigkeit, wodurch eine ungefähre Abschätzung der vertikalen Fahrzeugbewegung und der damit korrelierenden Gleislagefehler erzielt wird.

Ein Vergleich der Aufnahmen von Smartphones, die im Fahrerraum von Straßenbahnfahrzeugen installiert wurden, mit den Messwerten eines Handmesswagens zeigt eine hohe Übereinstimmung (vgl. Abb. 5). Die Ergebnisse der Smartphone-Messungen werden allerdings durch die Fahrzeugfederung beeinflusst, so dass eine frequenztreue Vorhersage nicht ganz möglich ist. Dadurch erklären sich zum Teil die Unterschiede zwischen beiden Messmethoden. Während dieser Effekt bei Vollbahnfahrzeugen stärker ins Gewicht fällt, ist er im Bereich von Straßen- und Stadtbahnen aufgrund der in der Regel

steiferen Fahrwerksauslegung weniger relevant und führt zu geringeren Verfälschungen durch die Federung. Positiv anzusehen ist, dass die Smartphone-Messung aufgrund des Fahrzeuggewichts eine belastete Messung darstellt, während die Messung mittels Handmesswagen daher zum Beispiel keine Schwellenhohllagen erfassen kann.

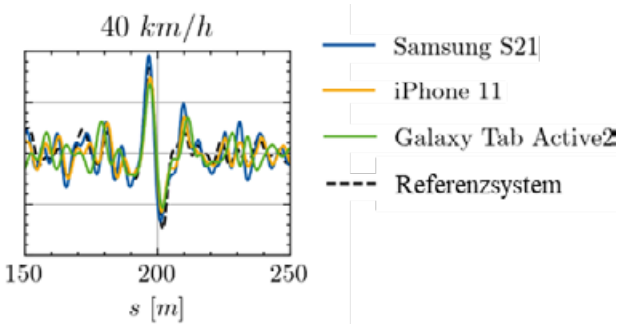


Abbildung 5: Längshöhenfehler gemessen mit Smartphones und einem konventionellen Messsystem [8]

3 Digitaler Zwilling

Ein digitaler Zwilling ist eine virtuelle Repräsentation, die sowohl den aktuellen Zustand als auch das zukünftige Verhalten eines physischen Systems abbildet. Zwischen der realen Anlage und ihrer digitalen Version erfolgt ein wechselseitiger Austausch von Daten [9]. Je nach Ausmaß des Datenflusses und der Komplexität in der Verhaltensabbildung lässt sich das Konzept weiter differenzieren, beispielsweise in Form eines digitalen Modells oder eines digitalen Schattens (vgl. Abb. 6; [10]). In dieser Arbeit wird aus Gründen der Übersichtlichkeit einheitlich die Bezeichnung „Digitaler Zwilling“ verwendet.

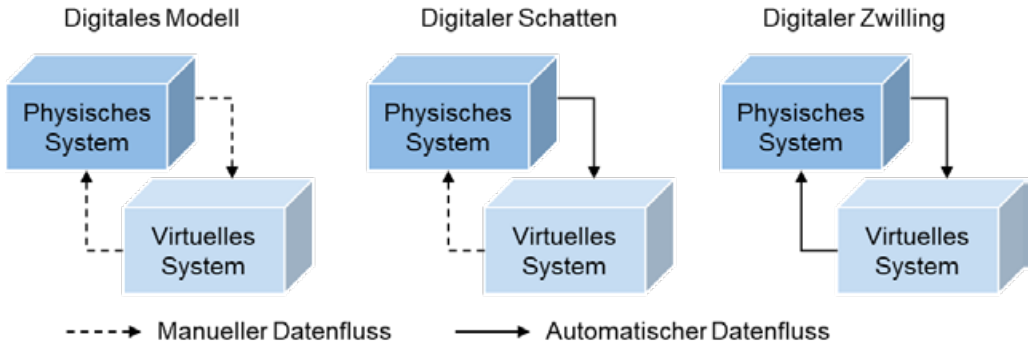


Abbildung 6: Digitales Modell, Schatten und Zwilling

3.1 Architektur des digitalen Zwillings

Der entwickelte digitale Zwilling basiert auf einer mehrschichtigen Softwarearchitektur, die verschiedene Datenquellen über Schnittstellen integriert (vgl. Abb. 7). Die Datenerfassung erfolgt über eine eigens entwickelte Smartphone Applikation: Nach Abschluss einer Messfahrt werden die erfassten Werte automatisiert in eine Cloud-Datenbank übertragen. Täglich prüft das System, ob neue Datensätze vorliegen, und startet für diese selbstständig den Auswerteprozess. Dabei wird jeder Datensatz auf Vollständigkeit untersucht, gefiltert und anschließend den passenden Streckenabschnitten zugeordnet. Das im Detail in Abschnitt 3.2 beschriebene Vorgehen bildet hierfür die Grundlage. In dem Aufbereitungs- und Analyseprozess fließen zum Beispiel Prüfungen ein, ob das Smartphone ordnungsgemäß ausgerichtet war, eine minimale Messdauer erreicht wurde und ob sämtliche erforderlichen Daten fehlerlos übermittelt wurden.

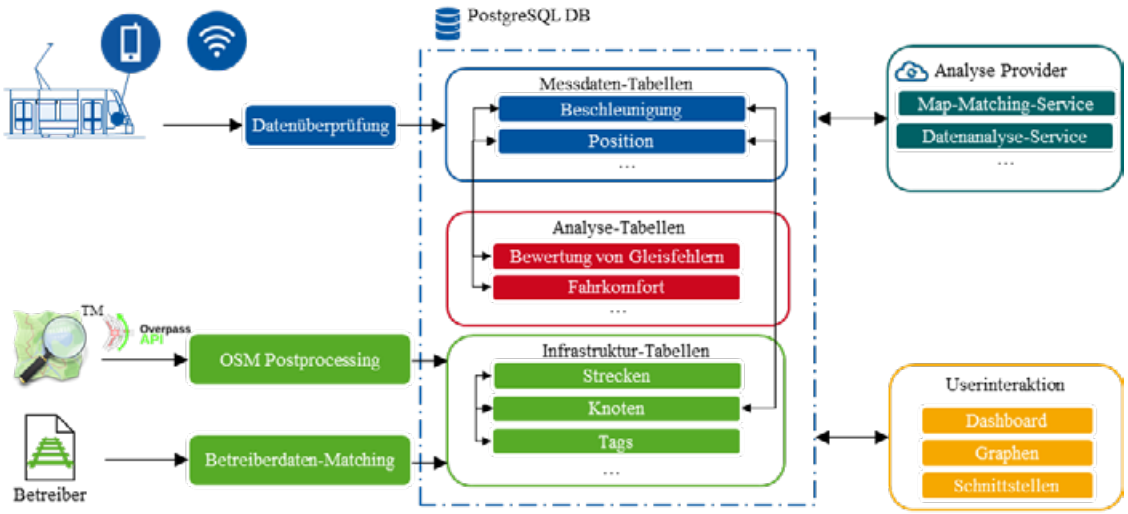


Abbildung 7: Architektur des digitalen Zwillings

Zur Abbildung der Infrastruktur wird auf einen kombinierten Ansatz zurückgegriffen, bei dem sowohl Open Source Geodaten aus OpenStreetMap (OSM) als auch betriebsspezifische Informationen einfließen. Das untersuchte Straßen- und Stadtbahnnetz in Frankfurt am Main ist in OSM angemessen detailliert erfasst und enthält Infrastrukturelemente wie Weichen, Kreuzungen, Haltepunkte, Signale sowie alle Linienverläufe. Zusätzlich werden die OSM-Daten mit betriebsinternen Informationen angereichert. Dazu zählen etwa die Zuordnung und Identifikation einzelner Streckenabschnitte oder Detaildaten zu Weichenanlagen. Punktobjekte wie Weichen und Kreuzungen werden routenweise mittels der „Closest Point“-Funktion von PostGIS bestimmten Linienverläufen zugeordnet. Streckenabschnitte werden darüber hinaus ebenfalls entlang der Linienführung gematcht. Zur Optimierung der Genauigkeit wurde ein Map-Matching Verfahren implementiert (vgl.

Abschnitt 3.2). Der Einbezug betriebsspezifischer Daten sichert nicht nur eine höhere Datenqualität, sondern bietet den Mitarbeitenden weiterhin die Möglichkeit, bekannte Benennungen und Notationen zu verwenden. Gleichzeitig wird durch die offene OSM-Basis sichergestellt, dass auch kleinere Verkehrsbetriebe ohne umfassende Digitalisierungsstände das System problemlos nutzen können.

Für die Nutzerinteraktion ist ein eigenes Frontend vorgesehen. Gegenwärtig befindet sich ein webbasiertes Dashboard in Entwicklung, das es erlaubt, Messdaten und Streckenabschnitte im Detail auszuwerten und über geeignete Performanceindikatoren zu visualisieren.

Das zugrunde liegende Datenmodell von OSM setzt sich aus georeferenzierten Punkten („Knoten“) zusammen, die zu linearen Elementen („Wegen“) kombiniert werden. Komplexere Relationen ergeben sich durch die Verknüpfung mehrerer Wege, wodurch Strecken und Linien eindeutig beschreibbar sind [11]. Metainformationen wie Liniennamen werden als Schlüssel-Wert-Paare angelegt. Für den Einsatz im digitalen Zwilling wird diese OSM-Struktur vereinfacht, indem nur relevanten Knoten und Relationen samt Metadaten verwendet werden. Die Speicherung sämtlicher Daten erfolgt in einer PostgreSQL-Datenbank mit aktivierter PostGIS-Erweiterung, um Infrastrukturdaten als räumliche Objekte (Punkte und Linien) im Datenbanksystem zu repräsentieren. Ein Import der OSM-Daten in die Datenbank erfolgt durch die Nutzung der Overpass API unter Berücksichtigung von geographischen Grenzen des jeweiligen Betreibergebiets. Zur genauen Verknüpfung (Mapping) von Messdaten wird die oft weitmaschige OSM-Punktstruktur durch Interpolation verdichtet, sodass etwa alle Meter ein neuer geographischer Punkt bzw. Knoten eingefügt wird. Die Wahl des Meterabstands wurde als Kompromiss aus angemessener Datengenauigkeit bei gleichzeitigem akzeptablen Datenvolumen gewählt. Grundsätzlich werden sämtliche Metadaten der OSM übernommen, wobei aktuell primär Attribute zu Weichen, Kreuzungen und Haltestellen aktiv ausgewertet werden. Die Vielschichtigkeit der in OSM bereitgestellten Zusatzinformationen lässt Raum für weitergehende Nutzungen in künftigen Entwicklungen.

Unterschiede und Unstimmigkeiten zwischen OSM- und Betreiberdaten machen weitergehende Datenverarbeitungen notwendig. So kann es vorkommen, dass etwa Kreuzungen oder Weichen in OSM fehlen, oder mit unpassenden Metadaten versehen sind, beispielsweise wenn ein Knoten gleichzeitig eine Weiche und einen Fußgängerübergang repräsentiert, aber nur der Übergang als Merkmal gespeichert ist. Zur Adressierung dieser Problematik wurde im Rahmen des Projekts ein Algorithmus entwickelt, der fehlende bzw. inkonsistent erfasste Weichen und Kreuzungen in OSM erkennt und die Metadaten lokal anpasst.

Aktuell liegt der Schwerpunkt auf der Verarbeitung von Beschleunigungsdaten. Jedoch lassen sowohl die zugrundeliegende Datenbankarchitektur als auch die Systemstruktur des digitalen Zwillings die spätere Einbindung von weiteren Datentypen wie Bild- oder Videomaterial zu. Die Verknüpfung eingehender Messdaten mit den Infrastrukturabschnitten erfolgt, wie im nächsten Abschnitt im Detail dargestellt, auf Basis von Geokoordinaten.

3.2 Map-Matching

Um eine präzise Zuordnung der erfassten Sensordaten zu den Infrastrukturelementen des Betreibers zu gewährleisten, wurde ein Map-Matching-Algorithmus in die Systemarchitektur des digitalen Zwillings eingebettet. Das Ziel dieses Verfahrens besteht darin, die mithilfe von GNSS erfassten Positionsdaten den Meter-genauen Knotenpunkten in der Datenbank zuzuordnen. Während zahlreiche etablierte Algorithmen auf Hidden-Markov-Modellen (HMM) basieren, die jedoch einen erheblichen Rechenaufwand verursachen [12], ergibt sich für den hiesigen Anwendungsfall eine Vereinfachung: Da die verwendeten Smartphones ausschließlich in planmäßig eingesetzten Straßen- und Stadtbahnfahrzeugen aufzeichnen und deren Fahrwege bereits im Voraus feststehen, ist es nicht erforderlich, jede beliebige GNSS-Spur mit dem gesamten Netz abzugleichen. Die eigentliche Aufgabe reduziert sich darauf, aus einer begrenzten Anzahl bekannter Trajektorien den wahrscheinlichsten Linienverlauf zu wählen. Nach erfolgreicher Identifikation der korrekten Route werden die Positionsmessungen den nächstgelegenen Knotenpunkten entlang dieses Streckenverlaufs zugeordnet. Ein Nachteil dieses Ansatzes ist, dass Depotgleise und weitere Abschnitte, die nicht zum Linienbetrieb gehören, nicht zuordbar sind. Diese stellen jedoch lediglich einen sehr geringen Anteil des Gesamtnetzes dar.

Zu Beginn des Prozesses werden potenzielle OSM-Relationen bestimmt, indem alle relevanten OSM-Punkte in ein metrisches Koordinatensystem, beispielsweise Gauß-Krüger, transformiert und in einem k-dimensionalen Baum (k-d-Baum) organisiert werden. Diese effiziente Datenstruktur wird dann genutzt, um die GNSS-Messpunkte aus der Smartphone-Aufzeichnung abzufragen. K-d-Bäume wurden aufgrund ihrer geringeren Berechnungskomplexität gegenüber klassischen paarweisen „Nearest-Neighbor“-Suchen gewählt, was angesichts der großen Datenmengen vorteilhaft ist. Damit unnötige Verknüpfungen vermieden werden, dient eine Distanzschwelle von zehn Metern als Filterkriterium. Danach erfolgt die Extraktion aller OSM-Relationen mit den identifizierten, zugehörigen Knoten. Um die Fahrtrichtung zu bestimmen, werden die drei häufigsten Relationen darauf untersucht, ob sie mit der Bewegungsrichtung der Messfahrt übereinstimmen. Dies geschieht, indem jeweils Start- und Endpunkte der gemessenen Trajektorie

bzw. der Relationen als Vektoren abgetragen und miteinander verglichen werden. Das Skalarprodukt dieser Vektoren fungiert als Kriterium: Liegt das Ergebnis im positiven Bereich, indiziert dies einen Winkel unter 90° und eine Übereinstimmung gilt als wahrscheinlich. Nachdem die entsprechende Relation eindeutig festgelegt wurde, wird für deren Verlauf erneut ein k-d-Baum für die gefundene Strecke erzeugt, bestehend aus den Knotenpunkten im Ein-Meter-Abstand. Im nächsten Schritt wird jede aufgezeichnete GNSS-Position mit ihrem nächstliegenden Punkt innerhalb eines maximalen Abstands von fünfzehn Metern verknüpft. Auf diese Weise werden Fehlzuordnungen bei abweichenden Fahrten vermieden. Diese Beziehungen werden für jede einzelne Messung in der Datenbank hinterlegt. Die Auswertung zeigt, dass der entwickelte Algorithmus sowohl effizient als auch äußerst genau arbeitet und alle Datensätze den korrekten Linien zuordnen konnte. Gelegentliche falsche Zuordnungen treten lediglich dann auf, wenn sich zwei Linien große Abschnitte teilen und erst an den Streckenenden auseinander gehen. In solchen Fällen kann es passieren, dass bestimmte Messpunkte nicht mehr im regulären Abstandsbereich liegen und daher keiner Strecke zugeordnet werden. Da die Auswertungen meist abschnittsbasiert erfolgen und relevante Knoten ohnehin mehreren Linien zugeordnet werden können, hat sich diese Limitation in der Praxis bislang kaum als problematisch erwiesen. Bisher konnten 88 % der Fahrten einer Linie zugeordnet werden. Werden nur Fahrten betrachtet, die länger als 10 Minuten waren, sind es sogar 99,2 %.

4 Ergebnisse

Wie bereits in Abschnitt 2.1 erörtert, können Smartphones bei der Erkennung von vertikalen Gleislagefehlern im Bereich von Straßen- und Stadtbahnfahrzeugen vergleichbare Resultate wie etablierte Referenzmesssysteme liefern. Durch ein Monitoring-Konzept, das auf kurz getakteten, regelmäßigen Messintervallen basiert, können Veränderungen im Zustand der Gleise zuverlässig über längere Zeitspannen nachverfolgt werden. Im Rahmen dieser Untersuchung wurden in Frankfurt auf nahezu allen Linien über mindestens zwölf Monate hinweg mehrere hundert Messfahrten mit Smartphones durchgeführt. Zusätzlich lagen ausgewählte Messergebnisse aus früheren Erhebungen im Jahr 2023 vor.

Zur Darstellung der Messresultate wurde eine Visualisierungsstrategie gewählt, bei der Abweichungen der Längshöhe durch verschiedene Farbcodierungen abgebildet werden. So lassen sich Veränderungen im Zustand der Infrastruktur übersichtlich nachvollziehen, und die Entwicklung der Gleisqualität kann eindeutig abgebildet werden, was wiederum eine frühzeitige Identifikation von möglichen Auffälligkeiten oder Degradationen ermöglicht. Ein Beispiel einer solchen grafischen Aufbereitung auf Basis von Smartphone-Daten ist in Abbildung 8 dargestellt. Hier werden die Veränderungen auf einem rund vier

Kilometer langen Streckenabschnitt über einen Zeitraum von etwa eineinhalb Jahren gezeigt. Besonders signifikante Änderungen traten zum Beispiel in den Streckenkilometern 0,7 und 0,8 auf, wo einerseits eine signifikante Verbesserung (Kilometer 0,7) als auch eine Verschlechterung (Kilometer 0,8) durch das System erkannt wurden. Die Entwicklung an dieser Stelle, einschließlich der Einzelbewertungen zur Gleisqualität, ist in Abbildung 9 zusammengefasst.

Abbildung 9 zeigt außerdem die Streubreite der Messwerte, bestätigt aber, dass sich Trends erkennen lassen, sofern regelmäßige Messungen durchgeführt werden.

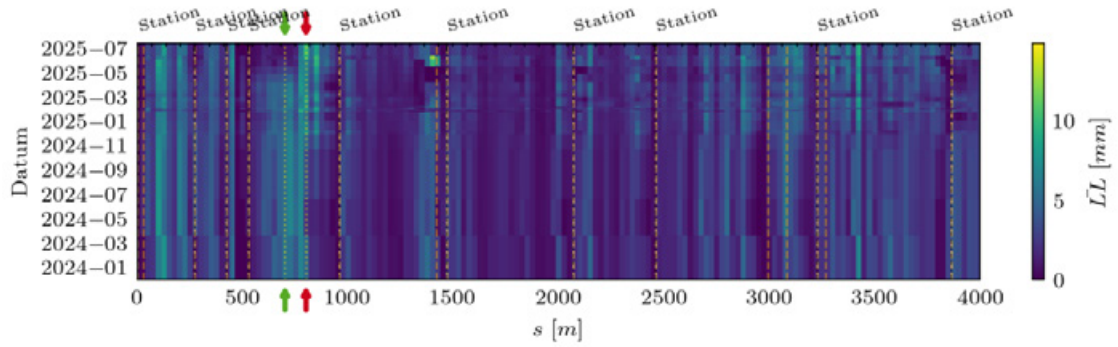


Abbildung 8: Entwicklung eines vier Kilometer langen Gleisabschnitts hinsichtlich der Abweichungen der Längshöhe über einen Zeitraum von mehr als einem Jahr

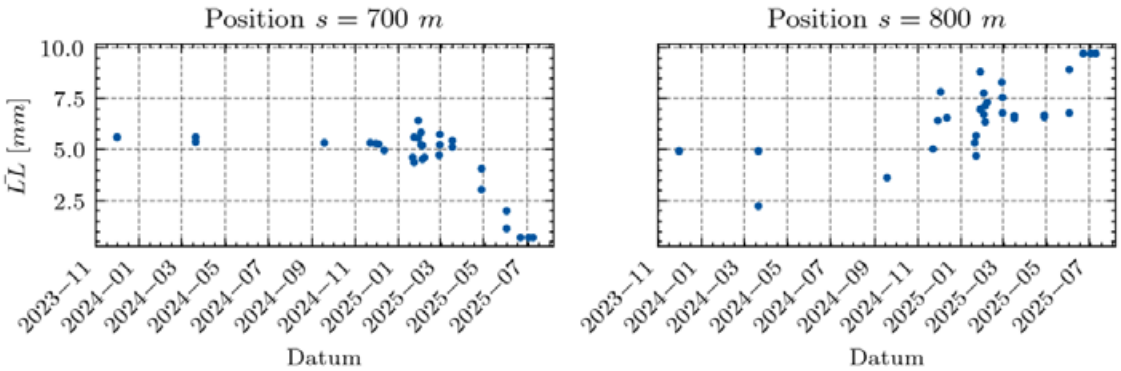


Abbildung 9: Veränderung der Infrastrukturqualität über Zeit

Die Verbesserungen im Streckenabschnitt bei Streckenkilometer 0,7 sind auf im Sommer 2025 durchgeführte Instandhaltungsarbeiten zurückzuführen. Für die zweite gezeigte Stelle im Bereich Streckenkilometer 0,8 ist kein exakter Grund für die Verschlechterung bekannt. Es davon auszugehen, dass es sich hierbei um normalen Verschleiß des Oberbaus handelt.

5 Prozessintegration

Die Instandhaltung von Schieneninfrastrukturen erfolgt bei vielen Verkehrsunternehmen aktuell vor allem durch jährliche Sichtprüfungen. Dabei wird das gesamte Gleisnetz durch Inspektoren zu Fuß begangen, um potenzielle Mängel visuell festzustellen und zu dokumentieren. Gegebenenfalls werden mobil einsetzbare (z.B. handgeschobene) Messsysteme verwendet. Parallel werden Weichen gesondert überprüft. Spezielle und dementsprechend kostenintensive Messzüge kommen hingegen meist nur im Abstand von fünf Jahren zum Einsatz, wie es die gesetzlichen Regelungen in Deutschland [13] für periodische Pflichtinspektionen vorsehen. Die anschließende Wartungsplanung richtet sich nach den gefundenen Mängeln.

Mit der Einführung des digitalen Zwillings wird auf die Schaffung einer Schnittstelle zwischen der klassischen visuellen Inspektion und der operativen Einsatzplanung gezielt (vgl. Abb. 10). Durch die fortlaufende, digitale Zustandserfassung der Gleisinfrastruktur können Algorithmen Unregelmäßigkeiten automatisiert detektieren. Dadurch basiert die Wartungsplanung nicht mehr ausschließlich auf jährlichen Begehungen, sondern es kann in Echtzeit auf Veränderungen reagiert werden und Fehlerstellen gezielt untersucht werden.

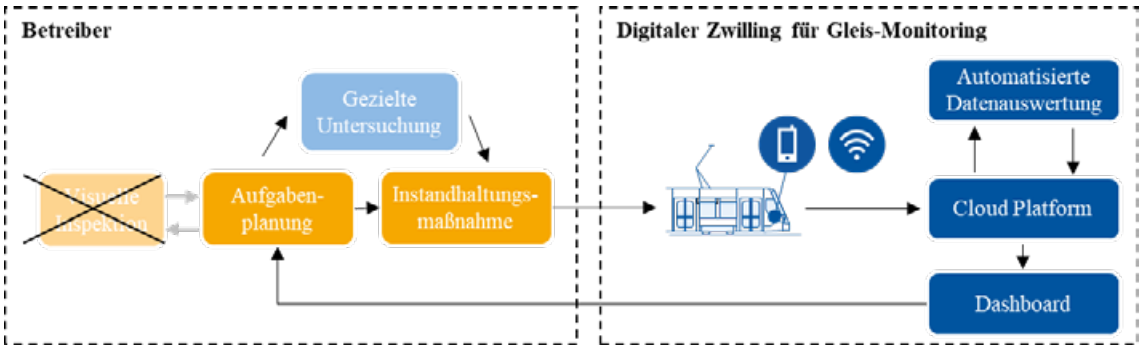


Abbildung 10: Integration in den Instandhaltungsprozess

Diese Umstellung ermöglicht es, stets über aktuelle und detaillierte Daten zum Zustand des Netzes zu verfügen, wodurch Ressourcen gezielter und effizienter eingesetzt werden können. Zudem verlieren klassische Aufgaben wie die manuelle Streckenbegehung angesichts des Strukturwandels zunehmend an Attraktivität, was die Personalsuche erschwert. Der Einsatz von digitalen Zwillingen in Kombination mit an Bord eingesetzter Messtechnik, beispielsweise Smartphones, besitzt daher das Potenzial, die Vorteile der zustandsbasierten Überwachung dauerhaft in die Wartungsprozesse von Stadt- und Straßenbahnnetzen zu integrieren. Durch diese digitale Unterstützung lassen sich nicht nur die limitierten Personalkapazitäten besser nutzen, sondern auch Betriebsstörungen durch vorausschauende und optimierte Planung minimieren.

6 Zusammenfassung

Im Mittelpunkt dieser Forschungsarbeit stand die Entwicklung eines digitalen Zwillings für städtische Schienenverkehrsnetze, der eine zustandsbasierte Überwachung der Infrastruktur erlaubt. Durch die Verknüpfung von mit Smartphones erhobenen Messdaten und einer cloudbasierten, auf Open Source beruhenden Plattform wird ein zeitabhängiges und zustandsorientiertes Infrastrukturmonitoring realisierbar. Besonders hervorzuheben ist das Potenzial, über die Zuordnung der Sensordaten zu einzelnen Infrastrukturobjekten spezifische Fehler an Gleis und Schiene lokal zu detektieren. Probleme, die sich aus der Heterogenität unterschiedlicher GIS-Datenquellen ergaben, konnten adressiert und mittels eines eigens entwickelten Map-Matching-Algorithmus gelöst werden, der eine präzise Verbindung von Messwerten mit konkreten geografischen Positionen ermöglicht.

Mit diesem Ansatz wird nicht nur eine kontinuierliche Überwachung des Anlagenzustands, sondern auch eine rückblickende Bewertung von Zustandsentwicklungen über längere Zeiträume hinweg möglich. Die Einbindung in bestehende Wartungsstrukturen verschafft Betreibern die Möglichkeit, die Vorteile einer zustandsorientierten Instandhaltung zu nutzen, darunter eine effizientere Ressourcen- und Personaleinsatzplanung. Gleichwohl ist zu empfehlen, die Methodik perspektivisch auf weitere Fehlerbilder auszuweiten.

Ein für die Zukunft relevantes Themenfeld stellt die Integration von Smartphone-Kameradaten dar, um in Kombination mit Beschleunigungssignalen zusätzliche Schadensbilder, wie Riffel oder Ausfahrungen, erkennen zu können. Angedacht sind hierzu praxisnahe Erprobungen verschiedener Kameraeinstellungen, Perspektiven, Belichtungszeiten und Beleuchtungsverhältnisse direkt im Betrieb mit den jeweiligen Verkehrsunternehmen. Für die technische Umsetzung muss der digitale Zwilling die Verarbeitung von Bilddaten unterstützen sowie auf aktuelle Algorithmen der Objekterkennung zurückgreifen. Die Auswertung von Bildmaterial könnte auch weitergehende Anwendungsfälle erschließen, etwa die automatische Erkennung von im Gleisbereich befindlicher Vegetation, Oberbaumängeln wie Schlaglöchern oder das automatisierte Erkennen außergewöhnlicher Ereignisse, zum Beispiel Überschwemmungen oder Unterspülungen.

Insbesondere bei Stadtbahnsystemen, in denen Tunnelstrecken einen wesentlichen Anteil bilden, gibt es derzeit Forschungsaktivitäten zur Entwicklung robuster Tunnel-Lokalisierungsverfahren. Diese Notwendigkeit resultiert aus der begrenzten Verfügbarkeit von GNSS-Signalen in unterirdischen Bereichen, welche die Präzision bisheriger Lokalisierungsmethoden mindert.

Darüber hinaus besteht weiterer Forschungsbedarf im Hinblick auf die Gestaltung der Benutzeroberfläche und die bedarfsgerechte Anzeige betriebsspezifischer Informationen, was in enger Abstimmung mit den Anwendern erfolgen sollte. Für die künftige Produktentwicklung ist außerdem eine Bewertung der wirtschaftlichen Vorteile gegenüber klassischen Instandhaltungsansätzen von entscheidender Bedeutung.

Danksagung

Diese Forschungsarbeit wurde im Rahmen des Forschungsprojekt „Smart.TRAM“ (Förderkennzeichen: 01FV2060A) durchgeführt, welches vom Bundesministerium für Verkehr gefördert wird. Für die Aufzeichnung der notwendigen Messungen im tagtäglichen Betrieb und der Bereitstellung von relevanten Daten und Informationen danken wir den Verkehrsbetrieben Frankfurt am Main mbH.

Literatur

[1] Steer Davies Gleave, „Study on the Cost and Contribution of the Rail Sector“, European Commission Directorate General for Mobility and Transport, 2015.
[2] W. Frenz, Hrsg., *Handbuch Industrie 4.0: Recht, Technik, Gesellschaft*. Berlin, Heidelberg: Springer, 2020. [Online]. Verfügbar unter: https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6144690
[3] T.-H. Yan, C. Hoelzl, F. Corman, V. Dertimanis, und E. Chatzi, „Integration of onboard monitoring data into infrastructure management for effective decision-making in railway maintenance“, *Railw. Eng. Sci.*, Bd. 33, Nr. 2, S. 151–168, Juni 2025, doi: 10.1007/s40534-024-00369-x.
[4] „Schienenfehler“, UIC, Paris, UIC report 712E, 2002.
[5] A. Paixão, E. Fortunato, und R. Calçada, „Smartphone’s Sensing Capabilities for On-Board Railway Track Monitoring: Structural Performance and Geometrical Degradation Assessment“, *Adv. Civ. Eng.*, S. 1–13, 2019, doi: 10.1155/2019/1729153.
[6] H. Tsunashima, R. Honda, und A. Matsumoto, „Track Condition Monitoring Based on In-Service Train Vibration Data Using Smartphones“, in *New Research on Railway Engineering and Transportation [Working Title]*, IntechOpen, 2023. doi: 10.5772/intechopen.111703.
[7] P. Leibner, A. Von Stillfried, und C. Schindler, „Potential of estimating track irregularities from in-service vehicles using smartphones“, gehalten auf der The Fifth International Conference on Railway Technology: Research, Development and Maintenance, Montpellier, France, 2022, S. 1–5. doi: 10.4203/ccc.1.10.21.
[8] P. Leibner, L. Stübinger, T. Hempel, L. Wolf, und C. Schindler, „Optimierung der Instandhaltung von Straßenbahngleisen durch eine intelligente Überwachung mittels Smartphone-Sensorik“, gehalten auf der IRSA, Aachen, 2023.
[9] M. Grieves und J. Vickers, „Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems“, in *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, F.-J. Kahlen, S. Flumerfelt, und A. Alves, Hrsg., Schweiz: Springer International Publishing, 2017, S. 85–113. doi: 10.1007/978-3-319-38756-7_4.
[10] W. Kritzinger, M. Karner, G. Traar, J. Henjes, und W. Sihn, „Digital Twin in manufacturing: A categorical literature review and classification“, *IFAC-Pap.*, Bd. 51, Nr. 11, S. 1016–1022, Jan. 2018, doi: 10.1016/j.ifacol.2018.08.474.
[11] N.N., „Elements“, *OpenStreetMap Wiki*. 2025. Zugegriffen: 11. Juni 2025. [Online]. Verfügbar unter: https://wiki.openstreetmap.org/wiki/Elements
[12] Paul Newson und John Krumm, „Hidden Markov Map Matching Through Noise and Sparseness“, *17th ACM SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst. ACM SIGSPATIAL GIS 2009 Novemb. 4-6 Seattle WA*, S. 336–343, Jan. 2009.
[13] N.N., *Verordnung über den Bau und Betrieb der Straßenbahnen*. 1987.

Session 5

Philipp Leibner, Jannik Goersch, Benedikt Neubauer, Thomas Hempel, Christian Schindler, Raphael Pfaff

Autoren



Leibner, Philipp

Studium des Maschinenbaus an der RWTH Aachen bis 2021. Seit 2021 wissenschaftlicher Mitarbeiter am Institut für Schienenfahrzeuge und Transportsysteme der RWTH Aachen im Bereich Zustandsüberwachung und Digitalisierung in der Schienenfahrzeugtechnik.



Goersch, Jannik

Jannik Goersch studierte bis 2023 Maschinenbau und Verkehrsingenieurwesen mit einem Fokus auf das System Bahn an der RWTH Aachen University. Im Anschluss begann er seine Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Schienenfahrzeuge der RWTH Aachen in den Bereichen Zustandsüberwachung und Produktentwicklung.



Neubauer, Benedikt

Benedikt Neubauer studierte Wirtschaftsinformatik und International Information Systems an der Friedrich-Alexander-Universität Erlangen-Nürnberg bis 2023. Er ist jetzt R&D-Projektmanager im Bereich Service Engineering für Bahninfrastruktur im Customer Service.



Hempel, Thomas

Thomas Hempel studierte Wirtschaftsingenieurwesen an der RWTH Aachen bis 2013. Anschließend arbeitete er bis 2017 als wissenschaftlicher Mitarbeiter am Werkzeugmaschinenlabor (WZL) der RWTH Aachen, wo er auch seine Promotion abschloss. Seitdem ist er bei der Siemens Mobility GmbH tätig. Heute leitet er den Bereich Service Engineering für Bahninfrastruktur im Customer Service.



Schindler, Christian

Christian Schindler studierte Maschinenbau an der RWTH und promovierte zum Thema Schienenfahrzeugdynamik. 2004 wurde er zum Professor für Konstruktion an die TU Kaiserslautern berufen. Seit 2016 leitet er das Institut für Schienenfahrzeuge und Transportsysteme an der RWTH Aachen.



Pfaff, Raphael

Raphael Pfaff studierte Mechatronik, Mathematik und Regelungstechnik in Bochum, Hagen und Coventry. Nach mehreren Stationen leitet er seit 2025 das Institut für Schienenfahrzeuge an der RWTH Aachen University und hat den gleichnamigen Lehrstuhl inne.

Session 5
Marcus Fischer, Felix Kröger, Thomas Rasel, Norman Kreisel, Sebastian Stumm

Wheel/rail adhesion-based support of automated train operation (ATO)

Subject area: Assisted, automated and autonomous driving (3A), automated transport

Dr. Fischer, Marcus[1], Kröger, Felix[2], Rasel, Thomas[1], Kreisel, Norman[1], Stumm, Sebastian[1]

[1]Knorr-Bremse Systeme für Schienenfahrzeuge GmbH, München
[2]DB Systemtechnik GmbH, Minden

Abstract

Expanding public transport and shifting freight transport from road to rail are essential for reaching ambitious climate goals. ATO over ETCS serves as one lever to increase capacity, in particular without cost-intensive expansion of the railway infrastructure. One starting point from brakes and wheel/rail adhesion perspective can be to improve the confidence level of the vehicle deceleration performance under different conditions at the same level of safety. Understanding the existing correlations in the wheel-rail contact forms the basis for potential solutions related to this. During yet another test campaign performed in 2024 by Knorr-Bremse and DB Systemtechnik, solutions for the development of this approach were successfully validated. Beyond, important basic data was generated.

Keywords: ATO, Adhesion, Adhesion Mapping, Adhesion Management, Homologation

1 Introduction/Context

The EU's legally binding goal is to achieve climate neutrality (net zero emissions) by 2050, enshrined in the European Climate Law and the European Green Deal. As a milestone on the way, the EU aims to reduce net greenhouse gas emissions by 2030 by at least 55% compared to 1990 levels. These targets are designed to align with the Paris

Agreement and ensure that all sectors, and especially transport, contribute to emission reduction [1].

It is widely recognized that shifting passenger and freight transport from road to rail is essential for reaching climate goals, as rail is significantly less carbon-intensive than road transport. Accordingly, the EU aims to expand public transport - especially rail- for both conventional and high-speed trains. By 2030, rail passenger transport is projected to reach 633 billion passenger-kilometers, accounting for nearly 10% of all transport activity [2]. To be able to manage this overall rising demand for transport, solutions need to be developed.

2 Means to expand railway capacity

Direct possibilities for the expansion of transport capacity in rail operation are limited, particularly due to infrastructure constraints [3]. The situation is additionally fueled by conflicting interests between increasing capacity versus maintaining stability and punctuality, both of which are further key performance indicators in rail operation. Since expanding rail infrastructure is both expensive and time-consuming, the potential to achieve higher transport capacities with this focus remains limited. Nevertheless, there are other ways such as operational topics which can be taken into account.

In [4] it can be seen that the dwell time of the trains, linked to passenger entry and exiting the trains, plays an important role in realizing more efficient rail operation. This is also confirmed in [5] with focus on transportation in the UK.

The topics of automation and trackside signaling offer additional measures that can be taken. One example is the Metropolitan Transportation Authority (MTA) in the US introducing CBTC on multiple lines in New York. This is also done with besides others the goal of more frequent train runs and higher transport capacity [6].

From [7] we can learn that the braking systems, their performance and responsiveness have a direct effect on the train headways. This in return influences the line's capacity. A way of how braking systems influence the considered deceleration and therewith the performance of trains being then relevant for operation can also be found in [8].

This brings us to the conclusion that improvements within the braking systems do have a positive effect on railway performance and line capacity. Nevertheless, if this approach is followed, the improvements need to be considered continuously within an integrated solution in combination with the signaling system. Otherwise, valuable potential in

braking performance as well as potential for improving transport capacity still remain unused.

3 Approach focusing on brake system and wheel/rail adhesion

The approach followed by the work described in the document on hand focuses on consideration of optimized and realistic braking performance, especially during low adhesion conditions. This is done using three pillars. Starting with Adhesion Mapping, giving an idea about the current wheel/rail adhesion condition up to the consideration of all available wheel/rail adhesion related measures such as adhesion enhancing measures and measures to make better use of the existing adhesion. By using these adhesion management systems, achievable decelerations of rail vehicles can significantly be impacted, especially under adverse environmental conditions, as present during autumn operation. This way, such systems can represent an important building block to improve the capacity, stability and punctuality of the overall railway system, especially when combined with future ATO operation or ETCS (European Train Control System) based train protection.

It is considered promising to even reduce the buffers required today using extended, situation-dependent system knowledge whilst maintaining the same or an even higher level of safety through higher accountable decelerations [9]. This task is also focused on by the R2DATO (Rail to Digital Automated up to Autonomous Train Operation) project, which is being worked on by Europe's Rail Joint Undertaking (Europe’s-Rail). Using technologies from its Reproducible Braking Distance (RBD) program, Knorr-Bremse aims to support the project by providing brake and adhesion management solutions, including systems to determine wheel/rail adhesion.

4 Interaction of functions to allow for best performance

Understanding the existing correlations in the wheel-rail contact forms the basis for potential solutions to reliably achieve the desired braking performance under different conditions, including degraded adhesion. Over the years, a plurality of basic data and knowledge was gathered during multiple tests on test rigs, e.g. the ATLAS (Advanced Test Laboratory for Adhesion-based Systems) roller rig, and test campaigns on test trains. Using this knowledge results in solutions for the determination of the adhesion level being present between wheel and rail and opportunities for its subsequent optimal utilization during the deceleration phase and where possible also its improvement. Nevertheless, to

ensure potential solutions can be taken into account, their use needs to be subject to certification and homologation.

This concept can be used for both ATO systems and train protection using ETCS. For this purpose, in both the ATO system as well as in the ETCS, the corresponding deceleration values can be stored and afterwards chosen based on external parameters such as wheel/rail adhesion. Whereas a similar approach can be used for ATO compared to ETCS, the procedure for ATO is less regulated compared to ETCS being responsible for train safety. This suggests that it is wise to start with the new solutions being considered in ATO first followed by ETCS.

An exemplary operating procedure is described in the following. A first vehicle applies the brakes on contaminated e.g., leaves-covered or wet, rails and detects an existing reduced adhesion situation. It then shares the relevant data with the infrastructure manager via a cloud environment to support the creation of an adhesion map. Using this information along with further data, e.g. current weather data, the infrastructure manager can continue processing this data and forward the update to the next approaching vehicle. This following vehicle, in turn, can then estimate its own deceleration to be expected while braking in the respective section. Using this information the infrastructure manager or the ATO system can adjust the traffic planning accordingly (Figure 1). Especially in an automated system, these additional insights could be leveraged to realise rail operation which is at the limit of what is currently possible and to reduce the previously very conservatively calculated time buffers between trains - without compromising safety - and thereby increase line capacity.

For ETCS (European Train Control System) the UNISIG Subset-026, Part 3 [10] describes how to handle low adhesion situations within the braking decelerations to be considered. For this, the correction factor Kwet is of central importance. Kwet is used to quantify how much the braking performance is reduced when braking in “wet condition” compared to dry conditions. This is currently based on an adhesion situation specified in the EN15595 standard [11] on wheel slide protection systems. Currently Kwet values are fixed for each train, and their use depends on the requirements of the infrastructure defining whether to use them or not without consideration of the true adhesion condition. As a fallback, based on driver reporting, other decelerations can be used (NV_MAXREDADH). More specifications on the determination of Kwet can be found in [8]. Higher benefits could be achieved if dynamic Kwet values, based on actual conditions, were used to reliably determine braking distances under low-adhesion rail conditions.

Whereas the Kwet approach is used for ETCS, similar approaches could be used for ATO operation, which would especially be useful to gather experience with such implementations as the ATO is less safety critical compared to the ETCS. First proposals showing the adaptation of the trains' deceleration parameters based on wheel/rail adhesion are being discussed already.

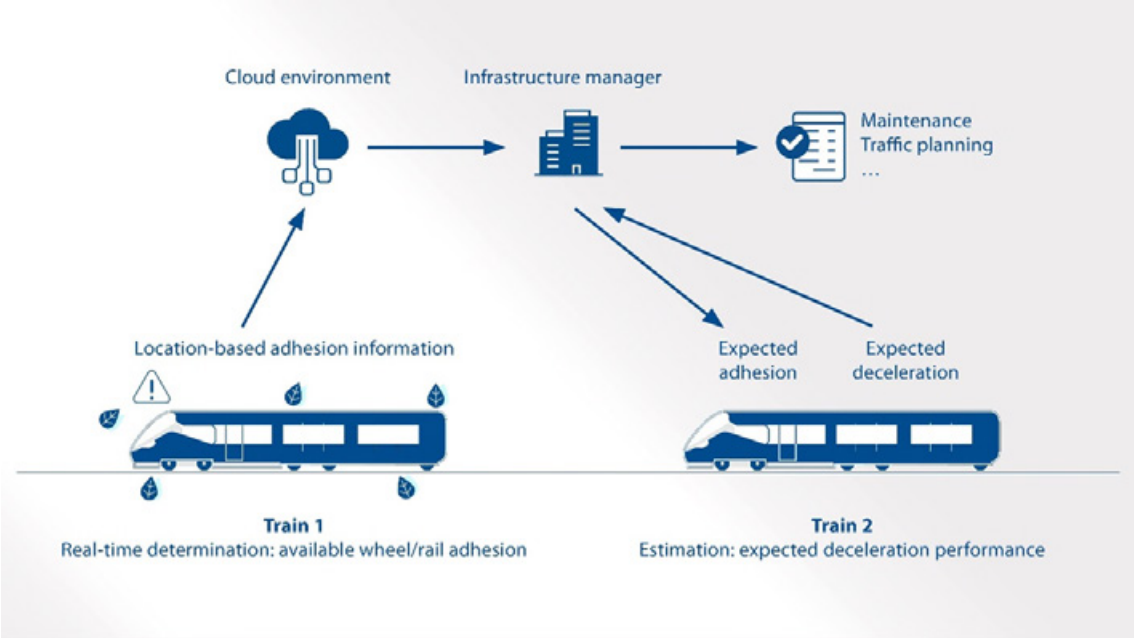


Fig. 1: Exemplary operational procedure starting with the Adhesion Mapping by one train followed by the performance estimation of the subsequent train to use this information for traffic planning. Beyond the information can be used for further purposes such as track maintenance planning.

As enabler for this approach, methods to determine or predict the current wheel/rail adhesion are known e.g., as implemented in the UK. Implementations of low adhesion prediction can be found from Met Office using weather models and environmental data [12].

Knorr-Bremse implemented first building bricks towards the realization of above ideas and put them under test using the DB aTL (advanced TrainLab) test train. Firstly, a functional prototype was developed allowing the available adhesion levels to be determined when passing over a defined position on the track whilst braking. Beyond, the prototype enables to transfer this information directly to a cloud environment linked to GNNS-generated position and time information. This way an "Adhesion Mapping" is realized. Related to improvement of adhesion/ improved adhesion usage, two other functionalities were put under test. A new wheel slide protection (WSP) algorithm, Wheelgrip Adapt, making better use of the available adhesion, was tested. The algorithm

uses various control strategies during the braking maneuver, depending on the adhesion levels prevailing at the time, to prevent wheel damage and to minimize the braking distance. Secondly, an Adhesion Management (ADM) system serves to improve adhesion levels. Its intelligent control system is used to define the application of sand between wheel and rail using distributed sanding systems. Unlike conventional systems applying predefined quantities of sand to a single wheelset per vehicle in a given direction of travel, the new function distributes the sand application in an intelligently controlled manner at multiple points along the entire train, dynamically adapting to the actual level of adhesion.

Using the systems and taking them into account, the deceleration relevant for rail operations (accounted service and emergency brake deceleration) can be increased.

5 Testing first building bricks of the future overall system

Knorr-Bremse and DB Systemtechnik carried out basic investigations regarding wheel-rail contact in the years of 2019 [13] and 2022 [14]. During yet another joint test campaign in 2024 performed by the project partners, a number of solutions were now successfully validated and additional basic data necessary for testing and development of such approaches was gathered. During the tests, DB's advanced TrainLab (aTL), a diesel-electric multiple unit train was used both on a public track close to Minden (North-Rhine Westphalia) and on a secluded branch line in Krakow am See (Mecklenburg-Western Pomerania). The aTL (VT605) is a "future laboratory" on rails used especially for testing and developing new technologies in real-world rail operations [15].

Before conducting the recent tests, additional sanding systems, adapted brake control units running the test wheel slide protection (WSP) algorithms (linked to the OE system) and extensive project-specific data acquisition systems (DAQs) were installed. The latter one was distributed across the train with over 1,000 channels for recording the performance of the systems under test. Next to sensors for the determination of axle speeds, brake- and pilot-pressures, the train deceleration, its position, sand control and sand flow signals, in addition all internal signals of the adapted brake control units were recorded. For the determination of the train-wide deceleration value, two Knorr-Bremse CubeControl brake control units were installed in addition to the actual system involved in the brake force generation.

Analogue signals from the existing braking system comprising the wheel speeds and signals for the control of the anti-skid valves were redirected to the test brake control system running the adaptive brake control under test. To ensure error-free vehicle operation, the inputs of the existing brake control units were stimulated via a secure

control path. This was possible using a special measurement and simulation environment developed by Knorr-Bremse. All measured data was consolidated synchronously via 12 CAN buses into one single data logger. In parallel, signals which were critical for the execution of the tests were evaluated and displayed live inside the train in a control center via a separate and independent DAQ.

The key parameters varied during the tests were the rail conditions and the parameter settings of the brake and adhesion management systems. Various challenging wheel-rail adhesion conditions were created using mixtures of water and soap, paper, oil, and leaves to simulate adverse test scenarios. In total, more than 350 braking maneuvers were performed.

6 Results of the dynamic testing

6.1 Proof of Concept: Adhesion Mapping

To perform a proof of concept (POC), a prototype of an Adhesion Mapping system was installed in the aTL. The system was active during all the test campaign. Its task was to determine data giving evidence of the experienced wheel/rail adhesion condition. It triggered, recorded and filtered the relevant data (identified adhesion, measured deceleration, speed, position, detected activity of magnetic track brake and sanding) and transferred it via GSM to a cloud server. An exemplary web-based user interface application was implemented which can be adapted to future user requirements.

Using this functionality, it could be shown that it is possible to determine the wheel/rail adhesion using systems that are already available on board of today's trains. The feasibility of representing adhesion data in real-time could be proven. Nonetheless, real-time requirements need to be specified for each specific application, e.g. depending on the distance spacings between the trains. There might be large differences between different modes of transport just as high-speed or suburban operation. Although the available network coverage must always be considered, the data transmission from the vehicle to the cloud environment for visualization worked smoothly. Once deposited in the cloud, the data can then be made available to different users (Figure 2). This way the information can, for example, be used by infrastructure managers to initiate maintenance measures such as cleaning the rail by means of pressured water. However, subsequent trains could also utilize the data to assess their expected deceleration performance on this

specific part of the track, which may be transferred into operational traffic planning and/or the ATO control sequences.

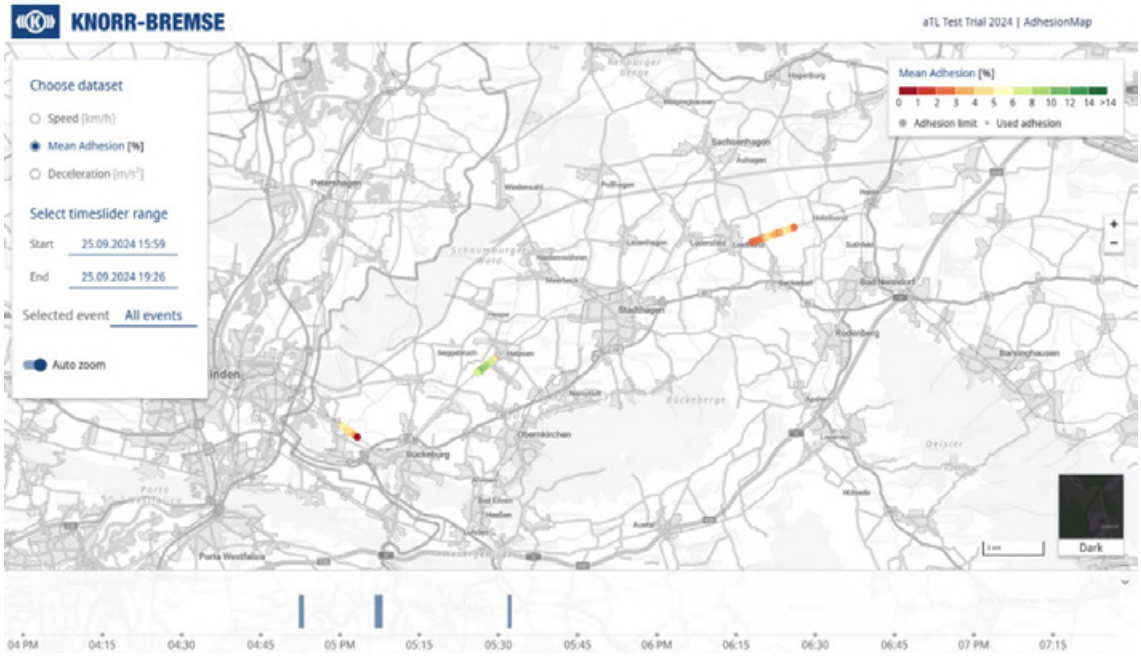


Fig. 2: Exemplary real-time representation of adhesion data (Adhesion Mapping) gathered during the test campaign. Transferred to a user specific format, it can be used by e.g., subsequent trains, infrastructure managers, or the ATO system.

6.2 Adaptive Wheel Slide Protection (WSP)

During the test campaign, the new WheelGrip Adapt system was tested and it replaced the OE wheel slide protection system originally implemented in the aTL. At its core, the employed adaptive algorithm dynamically switches whenever this enables higher brake forces, to an additional control range, specifically designed for extremely low adhesion conditions [16]. Main target of the tests carried out was to validate the WheelGrip Adapt wheel slide protection algorithm. Therefor it was implemented in the test brake control system and could be (de-)activated during the test alternating with a non-adaptive algorithm. Using this test setup, the new algorithm's impact on the utilization of available wheel-rail adhesion could be quantified in "autumnal conditions".

Each test series started with an initialization phase, where 500m to 600m of the track was covered with contamination, especially leaves. Hereafter, the prepared track section was overrun several times from both directions with brakes being applied. With each overrun the impact of the leaf contamination decreases, and therefore the available brake force increases. To be able to identify the benefit of the WheelGrip Adapt against the

WheelGrip Classic, after two runs in any of the modes, the configuration of the WSP-system was changed to the other variant.

During the data analysis, the mean deceleration of the train for the prepared section and achieved using the WheelGrip Adapt or the WheelGrip Classic algorithm was determined and plotted against the number of overruns (Figure 3).

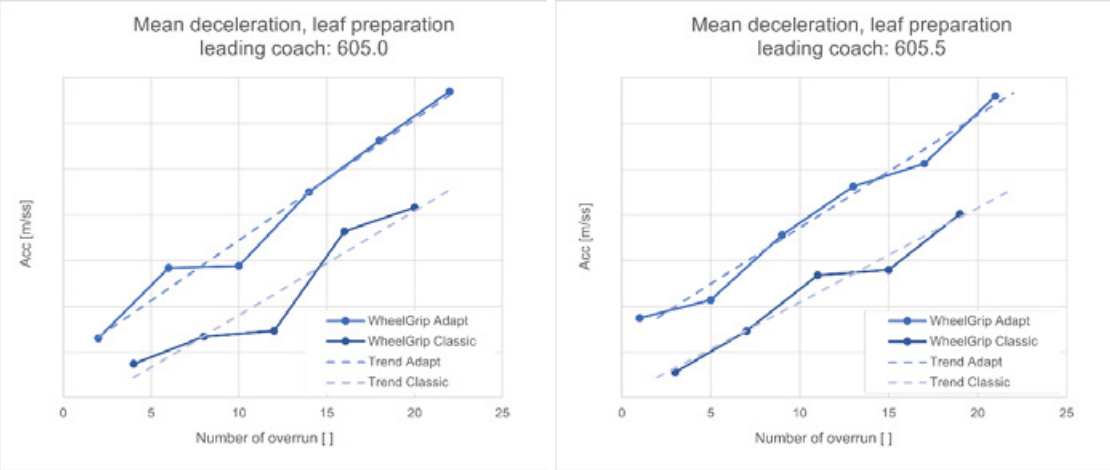


Fig. 3: Comparison of the mean decelerations achieved with the WheelGrip Adapt vs. WheelGrip Classic algorithm as a function of the number of overruns.

As a result, it could be shown when looking at consecutive overruns, that at all times the deceleration achieved using the WheelGrip Adapt was in a higher range compared to the one achieved using the WheelGrip Classic. To be able to compare the solutions using a numerical value, a linear trend line was determined for each of the resulting curves. Based on that, the ratio of the WheelGrip Adapt and WheelGrip Classic trendlines was calculated, split into the two driving directions.

EQU Ratio = trendline(WheelGrip Adapt)/ trendline(WheelGrip Classic) -1.

The performance comparison for the adhesion conditions considered with focus on leaf application shows a ratio between WheelGrip Adapt and WheelGrip Classic of approx. 0.25 up to 0.45. This indicates an increase of the brake force of 25% to 45% for the WheelGrip Adapt compared to the WheelGrip Classic algorithm (Figure 4).

The system ran stable and trouble-free during the entire test campaign. Based on the test results, the algorithm's suitability for approval in accordance with EN 15595 / UIC Leaflet 541-05 could also be confirmed for the first time using a multiple unit train at speeds up to 160 km/h.

In addition to the WSP specific tests, the system also collected relevant data for the Adhesion Mapping function.

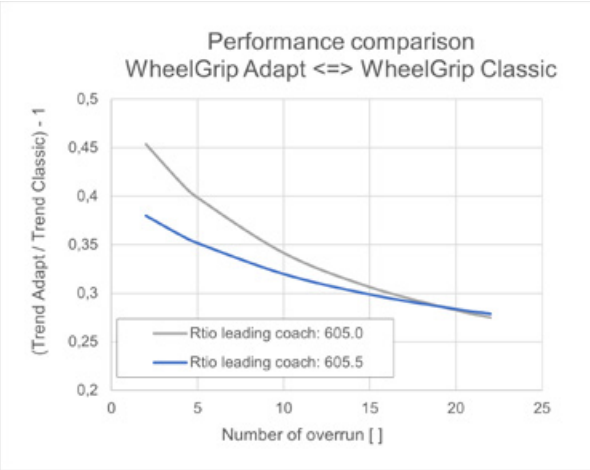


Fig. 4: Performance comparison: improvements achievable using WheelGrip Adapt compared to WheelGrip Classic.

6.3 Adhesion Management using braking sand

In addition to sensors and control components necessary for the Adhesion Management system, a additional sanding units were installed on two aTL bogies distributed along the train. Their activation was based on deceleration data provided by the wheel slide protection system.

After initial challenges related to bogie installation of the sanders, the ADM system worked as expected. The resulting test data sets were comparable with previous tests on the Knorr-Bremse ATLAS roller test rig in terms of control behavior and influence on wheel/rail contact. Particularly under low adhesion (nH) conditions, decelerations similar to dry conditions were achieved. Figure 5 shows a test data set for such nH conditions realized by a water/soap preparation. The ADM system was benchmarked against a 2 g/m and 4 g/m DVRS system and shows a lower sand amount usage. In addition, the demanded brake deceleration was almost reached. This is especially the case for brake stages with lower deceleration demands.

Under very low adhesion conditions (xnH), significantly increased decelerations were detected. To reach xnH condition, oil preparation was used. The ADM system improves adhesion in a way to realize a value which is over a minimal needed barrier (GBR). If possible, the ADM also tries to save sand in these scenarios. As a result, even under braking on oil-contaminated rails - considered to be the most adhesion-reducing condition - deceleration consistently remained above 0.55 m/s².

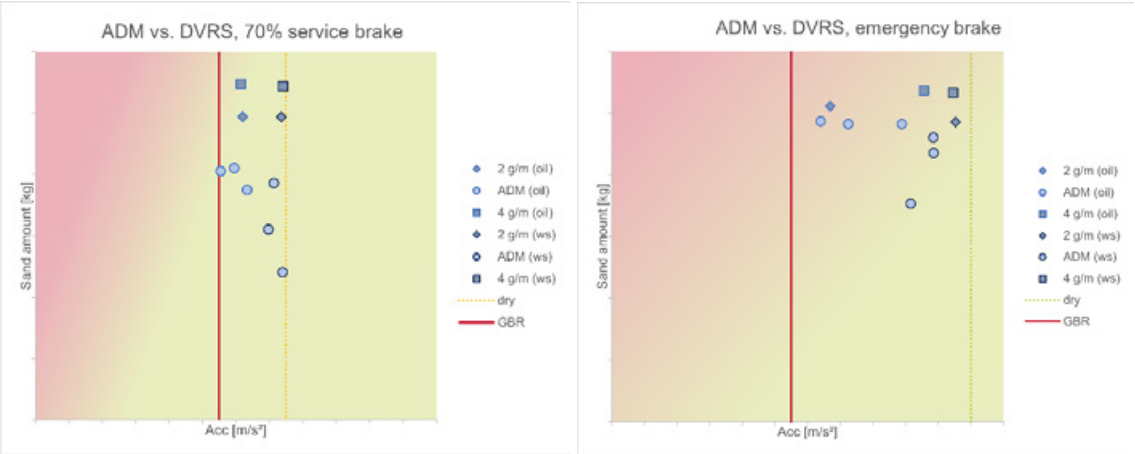


Fig. 5: The ADM system worked as expected on low (nH) and very low (xnH) adhesion conditions by automatically initiating sand extraction and adjusting the sand amount in a consumption optimized manner.

6.4 Impact of magnetic track brakes on leaves-contaminated track

The magnetic track brake (MTB) provides additional brake force for trains in case of emergency brake applications. Not less important is its cleaning effect in case of degraded adhesion conditions such as leaves in autumn.

Within the previous test campaign in 2022 [14] mentioned above, investigations on the adhesion improvement on track preparations with paper tape and oil were carried out. For this purpose, several consecutive braking tests were executed on the same track with just one initial preparation. This allowed us to evaluate the increase of adhesion between consecutive wheelsets of the train as well as for subsequent trains on the same track. The test sequences were performed with and without magnetic track brake (MTB), to be able to quantify the additional adhesion improvement caused by the MTB.

During the test campaign in 2019 [13], the comparability between contamination generated by paper tape and leaves related to the general adhesion conditions for single braking actions without MTB was already proven.

As a next step within the current test campaign one aim was to investigate the comparability between paper tape and leaves contamination regarding the long-term effects of brake applications with and without MTB as mentioned above.

Figures 6 and 7 each show two test series using leaves respectively paper tape as preparation for a 500 m section of the test track. Each of the test sequences starts with a new preparation of either leaves or paper tape and consists of 12 emergency brake

applications for this preparation. Throughout each single braking test, water was applied in front of the leading wheelset in the direction of travel. In the upper test sequence of both graphs, the MTBs installed in bogie no. 1 and 6 were activated in every second test to examine their impact on the adhesion conditions. The lower test sequences in the graphs consist exclusively of tests without MTB usage.

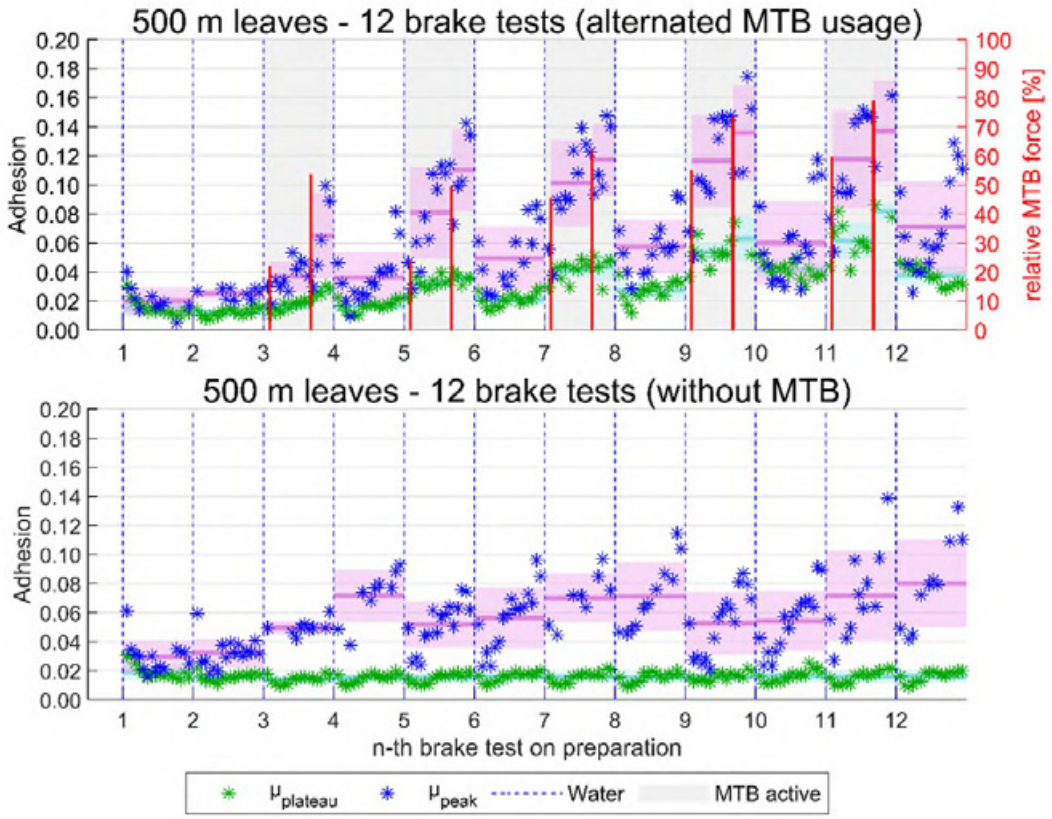


Fig. 6: Comparison of relative MTB force and adhesion behavior for 12 consecutive braking tests on a track prepared with leaves with and without MTB usage.

The plots show the relative MTB brake force of both MTBs (vertical red lines, scaled at the right y-axis) and two characteristic adhesion values for each of the 16 wheelsets (scaled at the left y-axis). The relative MTB force relates to the transmittable force on dry rail at the same speed. The adhesion of each wheelset was characterized by the two values μ_{peak} (blue asterisk) and $\mu_{plateau}$ (green asterisk), where μ_{peak} is the maximum adhesion value before the wheelset starts sliding. $\mu_{plateau}$ is the mean adhesion at sliding speeds above 10 kph, calculated as integral over the braking distance. The boxes and horizontal lines define the standard deviation resp. mean values of μ_{peak} (magenta) and $\mu_{plateau}$ (cyan) for a group of wheelsets. This group consists of all 16 wheelsets in case of inactive MTB. For tests with active MTB, the standard deviation and mean values were calculated for the wheelsets 2 to 11 between the first and the second MTB and for the wheelsets 12 to

16 behind the second MTB. This allows for showing the increase in adhesion that was caused by each active MTB.

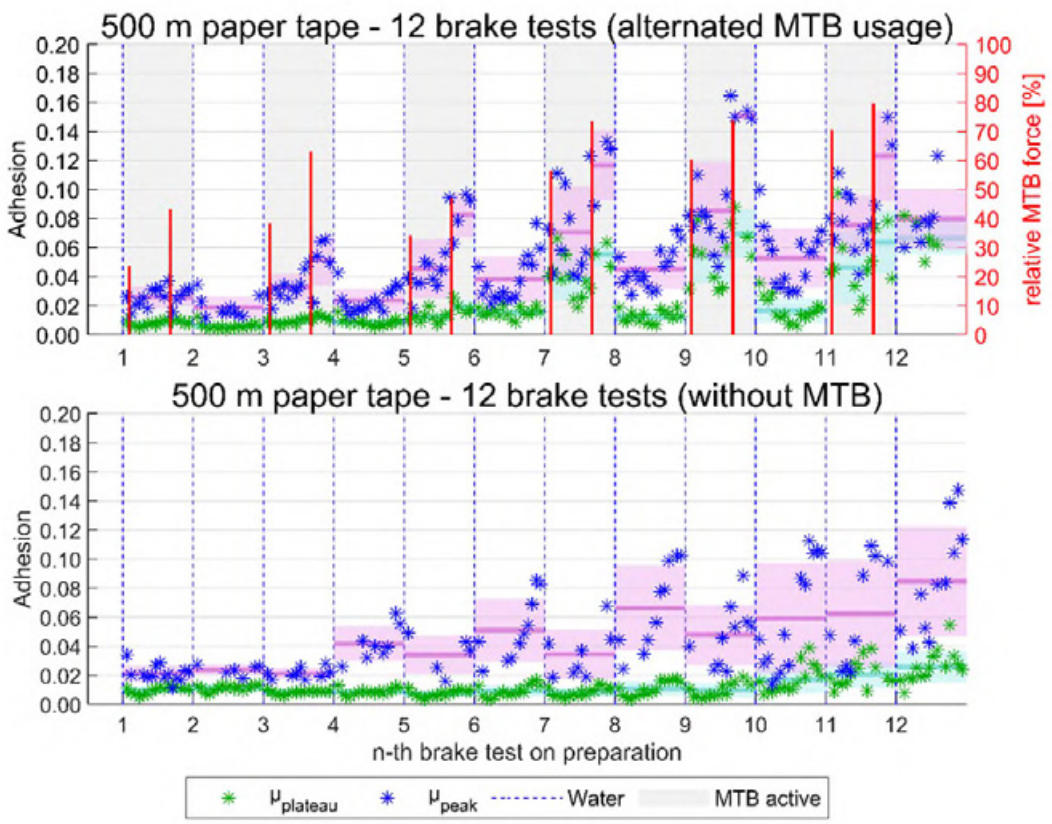


Fig. 7: Comparison of relative MTB force and adhesion behavior for 12 consecutive braking tests on a track prepared with paper tape with and without MTB usage.

Both the test series on leaves and on paper tape show quite comparable behavior regarding the cleaning effect with and without MTB as well as related to the transmittable brake force of the MTB itself.

In both test series (leaves/ paper tape) without MTB, the adhesion in macro slide ($\mu_{plateau}$) is neither increasing significantly from wheelset to wheelset within each single braking test, nor from one test to the next. However, there is a minor conditioning effect even in those extremely low adhesion conditions, that leads to increasing peak adhesion values from wheelset to wheelset within each test. Water being applied in front of the first wheelset in the direction of travel during the next braking test resets this effect almost ending up at the initial value of the previous braking test. Accordingly, this conditioning effect is mainly caused by displacing the water with each wheelset rolling over it. The reason for this is that the third body layer on the rail head (leaves as well as paper tape) only leads to extremely low adhesion conditions when combined with water. The wear of

the third body layer itself is a minor effect only as long as the MTB is inactive, so that the overall adhesion level does not increase significantly from test to test.

The test series with MTB on leaves and on paper tape shows that in both cases the wheelsets behind each MTB being active benefit from its cleaning effect, as the MTB displaces a significant amount of water from the rail head. This leads to a significant increase of peak adhesion as well as the transmittable braking force of the second MTB. By applying water during the next braking test, the adhesion level decreases, but compared to the test series without MTB, there is a major superimposed conditioning effect that increases the adhesion (especially μ_{peak}, but also $\mu_{plateau}$) from test to test. That this effect cannot completely be reset again by application of water, leads to the conclusion that it is caused by the wear of the third body layer due to the MTB being active.

Both track preparations, leaves and paper tape, lead to a third body layer on the rail head, the so called “black leaf layer”. This contaminant leads to extremely low adhesion conditions if water is applied to it. The analysis of the measured data leads to two conclusions, related to track preparation and MTB usage. First it showed that preparation with leaves and with paper tape both lead to comparable adhesion conditions that behave identically, especially regarding the conditioning effect at multiple overruns. Second, it can be derived that the use of MTBs immediately helps to increase the usable adhesion of the current train by drying the rail head but also has a long-term effect by wearing the third body layer.

Of strategic relevance is the consistency of the effects of the wheel slide protection system, the usage of sand and magnetic track brake experienced during the test runs with simulation results obtained previously. Being able to reproduce these effects in simulation models is a prerequisite for transferring aspects of real-world test runs to simulation test rigs. Given the (very) limited availability of suitable test tracks, this approach could be used to further develop DB Systemtechnik’s well-established procedure of combining the results of simulation test benches with real-world tests for the certification of wheel slide protection systems. Due to the limited availability of the necessary test tracks and in order to reduce the effort required for vehicle tests (costs, time), the partial implementation of the tests using simulations is an obvious choice.

7 Conclusions and next steps

The latest series of tests once again showed promising results, confirming the functionality and effectiveness of adhesion-based solutions to support later ATO (under

ETCS) operation and to support certification of wheel slide protection systems using test rig tests.

Related to the support of automated operation, as a first pillar, the capability of Adhesion Mapping with equipment already on board of today's trains was shown. Secondly, the effectiveness and readiness for approval of an improved adaptive WSP algorithm could be shown. Furthermore, the ability of an intelligent Adhesion Management system to efficiently increase prevailing wheel/rail adhesion conditions by using as little sand as possible was demonstrated. In the next step, especially the Adhesion Mapping functionality is aimed for being tested in an even more realistic environment under regular operation. The adhesion management will be enhanced by further functionalities.

Beyond, development efforts will increasingly focus on integrating the tested building blocks into a harmonized overall system. In parallel, it will be essential to advance the definition of certification/ homologation requirements for such adhesion-based solutions to enable their use for future ATO/ ETCS applications.

The basic data related to the effect of the MTB will be integrated into the simulation model of the WSP test rig environment. This enhances the possibility to migrate further tests from track testing to simulation test benches. Besides classical WSP testing, this simulation could also be used for testing and certification of other systems whose performance is affected by the available adhesion, like deceleration control, ATO or ETCS systems.

Acknowledgement

Co-funded by the European Union

Europe's Rail

Funded by the European Union. Views and opinion expressed are however those of the author(s) only and do not necessarily reflect those of the European union. Neither the European Union nor the granting authority can be held responsible for them. The project FP2-R2DATO is supported by the Europe's Rail Joint Undertaking and its members.

References

[1] European Commission: *Climate action and the Green Deal.* Unknown date. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/climate-action-and-green-deal_en; 14.07.2025.

[2] European Environment Agency: *Passenger transport activity.* 10.10.2024. https://www.eea.europa.eu/en/analysis/publications/sustainability-of-europes-mobility-systems/passenger-transport-activity; 14.07.2025.

[3] Umweltbundesamt. *Schienennetz 2025/2030 – Ausbaukonzeption für einen leistungsfähigen Schienengüterverkehr in Deutschland*. https://www.umweltbundesamt.de/sites/default/files/medien/461/publikationen/4005.pdf; 16.07.2025.

[4] TCRP Report 13: *Rail Transit Capacity*, 1996.

[5] Barron, Alex in: *Dwell time reductions: Good design plus people management*. Railway Technology Magazine. Dec/Jan 2016, https://www.railtechnologymagazine.com/Comment/dwell-time-reductions-good-design-plus-people-management; 14.07.2025.

[6] CBTC: *Upgrading signal technology.* https://www.mta.info/project/cbtc-signal-upgrades; 14.07.2025.

[7] Jähnichen, Dieter (Hrsg.): *Handbuch Bremstechnik von Eisenbahnfahrzeugen*; GRT, 2023.

[8] DIN EN 17997:2025: *Railway applications – Braking – Definition of ETCS brake curve parameters for Gamma trains.*

[9] Braeseke, J./ Fischer, M./ Friesen, U./ Gräber, J. in: *Fahrzeugseitige Optimierung der Infrastrukturauslastung: RBD-Testfahren & Feldversuche*. ZEVRail, Ausgabe 06/07 2023.

[10] ERTMS/ ETCS System Requirement Specification, *Chapter 3, Principles*, issue 3.6.0, SUBSET-026-3; 13.05.2016.

[11] DIN EN 15595:2018+A1:2023: *Railway applications – Braking – Wheel slide protection.*

[12] Met Office. *Rail Sector Services.* https://www.metoffice.gov.uk/services/transport/rail/index; 18.07.2015.

[13] Fischer, M./Haselsteiner, K./Szekely, F./ Heinz, S./Kröger, F. in: *Mehr Mobilität auf der Schiene: Erhöhung der Transportkapazität durch Optimierung des Kraftschlusses*. ZEVRail. Ausgabe 10 2020.

[14] Fischer, M./Rasel, T./Ebner, B./Kröger, F./Heinz, S. in: *Investigations of degraded adhesion conditions and interrelated methods for improving raking performance using the advanced TrainLab (aTL)*. EuroBrake 2023, Barcelona. 12. September 2023.

[15] Deutsche Bahn. *Our future lab runs on rails.* https://nachhaltigkeit.deutschebahn.com/en/measures/advanced-trainlab; 14.07.2025.

[16] Meyer, R./Rasel, T. in: *Höhere Zugtaktung: Neuartiger Gleitschutz für eine verbesserte Auslastung der Schieneninfrastruktur*. ZEVRail. Ausgabe 11/12 2020.

Authors



Dr. Fischer, Marcus

Degree in mechanical engineering at the Technical University of Munich. Since 2010 in various positions at Knorr-Bremse Rail Vehicle Systems (SfS GmbH), including systems engineer and project manager for innovation projects, leader of wheel/rail adhesion related work packages within Shift2Rail and Europe's Rail, recently focusing on the impact of brake systems on automated train operation (ATO) and ETCS considering testing and standardization.



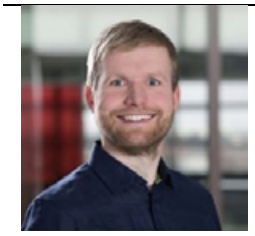
Kröger, Felix

Degree in mechanical engineering at the Leibniz University of Hanover. Since 2014, test manager for driving tests in the field of brake technology testing with a focus on wheel slide protection at DB Systemtechnik GmbH in Minden. Further development and support of the wheel slide protection test bench at DB Systemtechnik GmbH.



Rasel, Thomas

Graduated in electrical and communications engineering at Wiesbaden University of Applied Sciences. Joined Knorr-Bremse Rail Vehicle Systems (SfS) in 1995, involved in wheel slide protection development since 2005, including traction principles and utilisation, component design, system testing, test runs as well as specification, commissioning and basic tests on the Knorr-Bremse roller test bench (ATLAS).



Kreisel, Norman

Degree in mechatronics at the TU Dresden. Starting 2011 as field measurement engineer in the Center of Competence for bogie equipment at Knorr-Bremse SfS and project manager and system engineer for bogie diagnostics, Norman is now working in the Global Brake Systems department. Specifically supporting testing of innovative solutions from automated couplers via electromechanics brake up to adhesion management systems.



Stumm, Sebastian

Degree in computer science at the Hochschule Karlsruhe – University of Applied Sciences. Since 2011 at Knorr-Bremse Rail Vehicle Systems (SfS GmbH) as software engineer and project manager, focus on software architecture and functional standardization. Responsible for the topic Brake Force Distribution & Blending strategies. Evaluation, development and testing of a standardized functional approach to realize a train-wide adhesion management system.

IRSA 2025

SESSION 6

ASSISTIERTES UND AUTOMATISIERTES FAHREN

ASSISTED AND AUTOMATED DRIVING

Session 6
Tobias Hofmeier, Kilian Otto, Yan Cao, Martin Cichon

Vorentwicklung kamerabasierter on-sight-ATO-Funktionen – Schienenkopf- und Weichenlagendetektion sowie infrastrukturseitige Unterstützung der Lichtraumprofilüberwachung bei geschobenen Einheiten im skalierten Modell

Hofmeier, Tobias[1], Otto, Kilian[1], Cao, Yan[1], Cichon, Martin[1]

[1]Institut für Fahrzeugsystemtechnik, Karlsruher Institut für Technologie

Zusammenfassung

Die Automatisierung des Schienenverkehrs gilt als Schlüsselstrategie zur Erhöhung der Effizienz und Attraktivität des Systems im Zuge der Verkehrswende. Automatisierte Betriebskonzepte wurden bereits mit Erfolg in abgeschlossenen und kreuzungsfreien Umgebungen, wie etwa U-Bahnen, implementiert. Komplexe betriebliche Abläufe, wie das Fahren auf Sicht und Rangierprozesse, stellen hingegen neue Herausforderungen dar, insbesondere in Bezug auf Sicherheitsnachweise, Systemverhalten und Testbarkeit. Skalierte Testumgebungen bieten in diesem Zusammenhang eine kosteneffiziente Möglichkeit zur Vorentwicklung und Validierung.

Im Rahmen dieser Arbeit entsteht ein experimenteller Ansatz zur kamerabasierten Schienen- und Weichendetektion in einem skalierten ATO-System. Das Verfahren bestimmt die Fahrzeugposition im Gleisplan und liefert Informationen zur Ausrichtung. Im Fokus steht die Analyse sensitivitätsbestimmender Parameter auf die Detektionsleistung. Ergänzend wird die Selbstkalibrier- und Fehlererkennungsfähigkeit durch gezielte Störungen wie Versatz und Verdrehung der Kamera evaluiert.

Zudem wird eine infrastrukturseitige Überwachung mittels einer stationären Kamera

erprobt, um fahrzeugseitige ATO-Systeme im Rangierbetrieb zu unterstützen. Langfristiges Ziel der Methode ist es, Hindernisse im Gleisbereich auch ohne menschliche Sichtprüfung bei Fahrt mit geschobener Einheit zuverlässig zu erkennen. Hierfür werden Kommunikationsschnittstellen zwischen Infrastruktur und Fahrzeug entwickelt, sowie geeignete Bildverarbeitungsansätze analysiert und umgesetzt.
Die Ergebnisse zeigen, dass die eingesetzten kamerabasierten Verfahren in der skalierten Umgebung hinreichend zuverlässig arbeiten und wertvolle Erkenntnisse für reale Anwendungen versprechen. Der skalierte Modellansatz erweist sich als effektives Werkzeug zur Untersuchung automatisierter Fahrfunktionen und deren Übertragbarkeit auf reale Betriebsbedingungen.

Keywords: ATO, skalierte Modelle, Schienenkopfdetektion, Infrastrukturunterstützte ATO, way side monitoring

1 Einleitung

Die Klimapolitik der kommenden Jahrzehnte erfordert einen grundlegenden Wandel im Verkehrssektor. Eine nachhaltige Strategie ist die Verlagerung von Personen- und Güterverkehren von der Straße auf die Schiene [1]. Um einen erfolgreichen Modal Shift zu erreichen, müssen betriebliche Ineffizienzen im Personen- und Güterverkehr, insbesondere im Einzelwagenverkehr, reduziert werden [2], [3].
Automatisierung und Digitalisierung eröffnen neue Möglichkeiten, Effizienz und Planbarkeit des Schienenverkehrs zu verbessern [4]. Erste Automatic Train Operation (ATO) Projekte verdeutlichen das Potenzial technischer Innovationen, stoßen aber bei komplexen Betriebsformen wie dem Fahren auf Sicht an Grenzen. Hierfür sind erhöhte Systemperformance und neue Sicherheitsnachweise erforderlich. Ein szenariobasierter Testansatz mit virtuellen und realen Testumgebungen erscheint aufgrund struktureller Parallelen zur Automobiltechnik sinnvoll [5]. Die Umsetzung im Bahnbereich ist jedoch schwierig, da geeignete Simulationslösungen fehlen, experimentelle Plattformen nur begrenzt modelltreu sind und reale Tests hohe Kosten verursachen [6].
Zur Überwindung der Defizite wird ein skalierter Modellansatz entwickelt, der eine flexible und kosteneffiziente ATO-Entwicklung ermöglicht. In diesem Paper wird ein kamerabasiertes Verfahren vorgestellt, das Schienen- und Weichenlagen erkennt und Informationen zur Gleisposition und Fahrzeugausrichtung liefert. Sensitive Parameter werden per DoE-Analyse untersucht, und die Robustheit unter Störungen zur Bewertung von Kalibrier- und Selbstdiagnosefähigkeit getestet.
Zudem wird eine automatisierte Freiraumerkennung für geschobene Rangierfahrten entwickelt. Da bei fahrzeugseitigen ATO-Systemen die Sichtverbindung des Rangierers

an der Zugspitze entfällt [7], wird eine infrastrukturseitige Kameralösung mit Freiraumüberwachung, Bildverarbeitung und Kommunikationsschnittstellen integriert.

2 Automatisierter Bahnbetrieb

Die zunehmende Digitalisierung im Bahnsektor, sowie der politische Druck zur Verkehrsverlagerung auf die Schiene haben die Entwicklung automatisierter Zugsteuerungen beschleunigt. Der Automatisierungsgrad von Bahnsystemen wird über den „Grade of Automation“ (GoA) definiert. Dieser reicht vom manuellen Betrieb (GoA0) bis zum fahrerlosen Betrieb ohne Begleitpersonal (GoA4) [8]. Bereits 1981 nahm mit der Port Island Line (Japan) das erste GoA4-System den Betrieb auf [9]. Seitdem wurden weltweit über 40 GoA4-Strecken realisiert, z. B. in Paris, Kopenhagen und Vancouver [10]. In Deutschland ist vor allem das 2008 abgeschlossene Projekt „Rubin“ [11], der U2/U3 in Nürnberg und das Projekt „Digitale S-Bahn Hamburg“ [12], das ATO over ETCS erprobt, zu nennen.
Während ATO over ETCS vor allem für den automatisierten Betrieb auf signalisierten Hauptstrecken entwickelt wird, stellt der Sichtbetrieb im Rangier- und Nebenbahnbereich andere bzw. höhere Anforderungen an die Sensorik und Umfelderkennung. [13], [14].
Für die Entwicklung und Absicherung der Funktionen für das Fahren auf Sicht sind geeignete Testumgebungen notwendig. Skalierte Modelle ergänzen reale und virtuelle Erprobungen und bilden die Validierungsbasis für die in dieser Arbeit untersuchten Funktionen der Schienenkopf- und Umfelderkennung.

2.1 Reale, skalierte und virtuelle Testumgebungen für ATO auf Sicht

Die Entwicklung automatisierter Fahrfunktionen im Sichtbetrieb erfordert nachvollziehbare Testmethoden. Methodisch orientiert sich die Bahntechnik an erprobten Konzepten der Automobilindustrie, etwa Hardware- und Software-in-the-Loop-Verfahren. Unterschiede bestehen jedoch bei den verfügbaren Testumgebungen.
Feldtests bilden reale Bedingungen ab, sind im Bahnbereich jedoch aufwendig, da die Nutzung der Infrastruktur, die Verfügbarkeit der Fahrzeuge und der Personaleinsatz kostenintensiv und organisatorisch begrenzt sind [6]. Virtuelle Testumgebungen wie IPG CarMaker oder CARLA [15] sind im Automotive-Sektor etabliert. Im Bahnbereich wird zum aktuellen Forschungsstand primär mit noch nicht zulassungsfähigen Forschungsplattformen gearbeitet.
Skalierte Modelle schließen die Lücke zwischen Simulation und Realität. Sie ermöglichen realistische Sensorintegration und reproduzierbare Tests in kontrollierter

Umgebung. Die Forschungsgruppe nutzt diese Plattformen u. a. zur Untersuchung von Ortungsverfahren [16], zur Machbarkeitsanalyse skalierter Fahrzeuge mit Stellvertretertechnologien [6] sowie zur Synchronisation mit einem digitalen Zwilling in Unreal Engine 5 [17]. Die konkreten Skalierungsrandbedingungen, insbesondere hinsichtlich der Sensorik und der fahrzeugseitigen Signalverarbeitung, sind in [6] und [18] beschrieben. Die hier vorgestellte Plattform versteht sich als methodische Ergänzung zu MIL-, SIL- und HIL-Ansätzen, indem sie eine physisch skalierte, reproduzierbare Umgebung bereitstellt.

2.2 Schienenkopferkennung und Infrastrukturüberwachung

Schienendetektion und Freiraumüberwachung sind Schlüsselfunktionen von ATO-Systemen im Sichtbetrieb. In dicht genutzten Rangier- und Abstellbereichen müssen Umfelderkennungsalgorithmen unter engen Randbedingungen zuverlässig arbeiten.
Kamerasysteme ermöglichen die Erkennung von Schienen, Weichen und Hindernissen. In Kombination mit Routendaten lassen sich Gleisposition und Lichtraumprofil bestimmen. Visuelle Verfahren gelten dabei als Stand der Technik [19].
KI-basierte Verfahren gelten für komplexe Anwendungen wie die Weichenlagenerkennung als besonders geeignet [20]. Projekte wie ASaG [21] belegen das Potenzial infrastrukturbasierter Kamerasysteme zur Unterstützung von ATO bei geschobenen Einheiten.
Die kantengeführte Schienen- und Weichendetektion orientiert sich an etablierten Verfahren der bildbasierten Gleiserkennung [22], [23]. Dabei werden Kanten mittels Canny-Algorithmus extrahiert, durch Hough-basierte Verfahren segmentiert und anschließend über eine polynomische Modellierung zur Bestimmung des Schienenverlaufs zusammengeführt.
Die weiterführend in Kapitel 3 dargestellte Arbeit baut auf diesem Ansatz auf und beschreibt die experimentelle Umsetzung sowie Validierung im skalierten ATO-Modell.

3 Experimenteller Ansatz zur skalierten kamerabasierten Schienenkopfdetektion und Umfeldüberwachung

Kamerabasierte Sensordatenerfassung gewinnt im Bahnwesen an Bedeutung. Für sicherheitsrelevante Anwendungen sind Untersuchungen zu Robustheit und Sensitivität erforderlich. Das vorgestellte Verfahren detektiert den Schienenverlauf und ermöglicht eine präzise Positionsschätzung im Gleisplan. Ein systematischer DoE-Prozess

untersucht Einflussfaktoren und dient der methodischen Absicherung.
Ergänzend wird ein infrastrukturbasiertes Freiraumüberwachungssystem vorentwickelt, das bei geschobenen Zügen die fehlende Sichtverbindung kompensiert und automatisierte Fahrfunktionen unterstützt. Beide Ansätze werden im skalierten Modell validiert. Langfristig soll die Übertragbarkeit der Ergebnisse durch eine methodische Ableitung geeigneter Skalierungsparameter sichergestellt werden.

3.1 Versuchsaufbau und Testumgebung

Der „EduRail"-Versuchsaufbau [18] dient als flexible Testplattform für das Fahren auf Sicht. Der rund 12 m lange modulare Gleisplan mit Weichen, Bahnsteigen und Bahnübergang bildet typische Szenarien im Maßstab 1:32 (Normalspur) ab (Abb. 1).
Die Fahrzeuge sind mit Kamera-, LiDAR-, Ultraschall- und Ortungssystemen ausgestattet. Für die Datenverarbeitung und Steuerung kommen ein NVIDIA Jetson Orin Nano, Raspberry Pi und ESP32 zum Einsatz. Infrastrukturseitig ist eine Kamera mittig über dem Gleisplan installiert. Ein Raspberry Pi ermöglicht lokale Datenverarbeitung sowie die Weiterleitung der Bilddaten über eine Netzwerkschnittstelle.
Durch die Kombination aus realitätsnaher Infrastruktur und moderner Sensorik erlaubt der Aufbau umfassende Tests zur Umfelderkennung, Störverhalten und Verifikation virtueller Modelle. Weitere technische Details finden sich in [6], [16] und [17]. Perspektivisch ist ein deutlich erweitertes Testfeld nach den Kriterien aus [24] geplant.

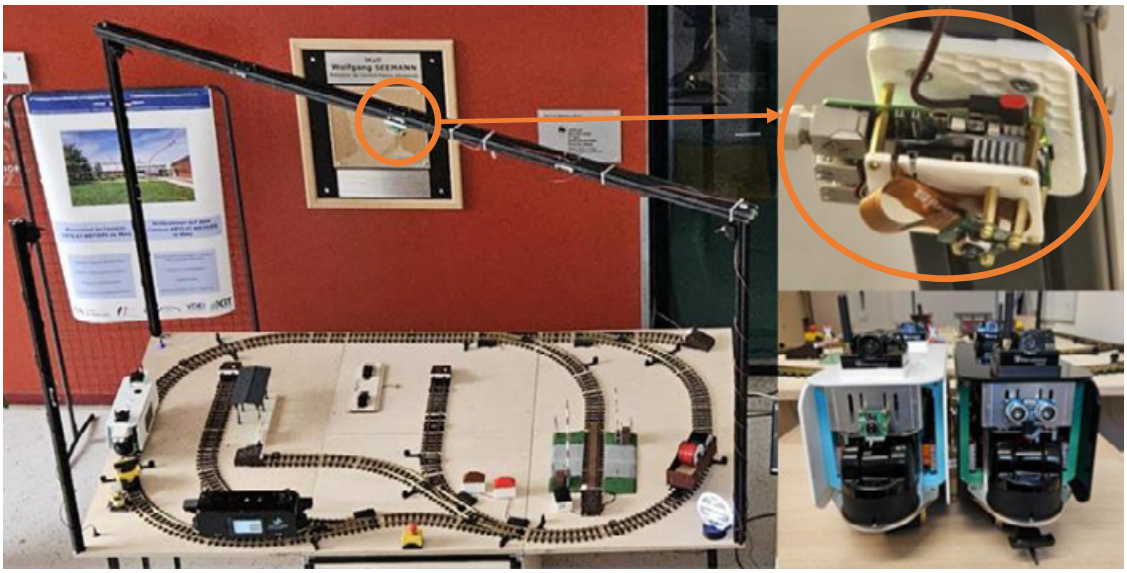


Abbildung 1: Gesamtaufbau EduRail (li.), Detailansicht Infrastrukturkamerasystem (o. re.), Detailansicht Multisensorplattform Demonstratorfahrzeuge (u. re)

3.2 Fahrzeugseitige kamerabasierte Schienen- und Weichendetektion

Im Rahmen der skalierten Testplattform wird ein kamerabasiertes System zur Schienenverlaufsdetektion entwickelt, das eine präzisierte Positionsbestimmung im Gleisplan ermöglicht und perspektivisch zur dynamischen Lichtraumprofilabschätzung beitragen soll.

Die Umsetzung folgt einem strukturierten Bildverarbeitungs-Workflow (siehe Abb. 2). Nach Laden der Systemparameter (Kamera- und Fahrzeuggeometrie, Kalibrierung, Streckendaten) wird ein entzerrtes Monochrombild erzeugt und die Kameraausrichtung automatisch validiert. Auf Basis einer initialen Ultraschall-Positionsmessung wird der Gleisplan in die Kameraperspektive überführt, die ROI bestimmt und ein Bird's-Eye-View generiert.

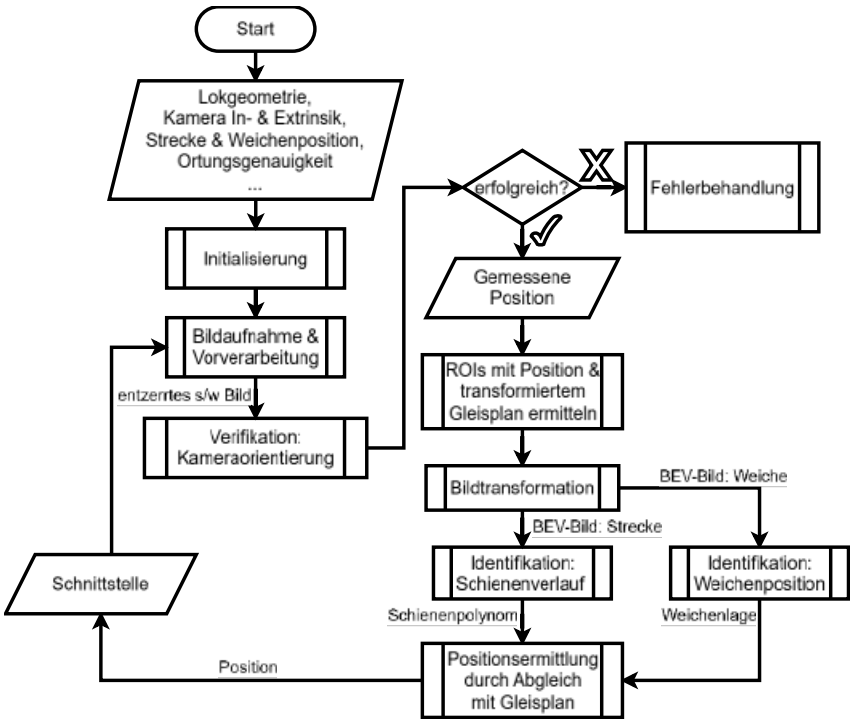


Abbildung 2: Flussdiagramm Schienenkopfdetektion

Die Detektion des Schienenverlaufs erfolgt auf Basis kantengeführter Algorithmen, die linienartige Merkmale extrahieren und zu einem Polynommodell aggregieren. Dieses wird mit dem erwarteten Streckenverlauf aus dem Gleisplan verglichen, um die Fahrzeugposition relativ zur bekannten Infrastruktur zu bestimmen. Dieser Prozess ist iterativ gestaltet und optimiert die gemessene Position je nach verfügbarer Rechenkapazität. Zur Optimierung der Kantenerkennung, werden dem Algorithmus Rahmenbedingungen zum Abstand der Drehgestelle, der Spurweite, sowie eine

Limitierung der minimal und maximal auftretenden Bogenradien vorgegeben.
Zur Gewährleistung der Sicherheit wird der gesamte Prozess kontinuierlich überwacht. Tritt ein Fehler auf, etwa durch Positionsunsicherheit wird eine Warnung ausgegeben.
Nachdem eine grundlegende Funktionalität des Systems zu beobachten ist, wird zur Untersuchung der Detektionsleistung und Systemrobustheit ein strukturierter DoE-Ansatz ausgearbeitet. Dieser erlaubt die gezielte Analyse des Algorithmusverhaltens bei systematischen Störungen und variierenden Bildverarbeitungsparametern. Dabei werden die Einflussgrößen Kameraneigung, Umgebungshelligkeit, Canny-Schwellenwerte sowie der Grad des Fit-Polynoms variiert. Bewertet wird anhand definierter Qualitätsmerkmale wie Detektionsrate, mittlerer quadratischer Fehler (MSE) sowie Rechenzeit. Basierend auf Voruntersuchungen werden für die Versuchsdurchführung zwei, zueinander möglichst inhomogene, Positionen im Gleisplan ausgewählt. Die DoE-Faktorstufen sind Tabelle 1 zu entnehmen. Für den Polynomgrad ist keine Faktorstufe 0 definiert, da nur kubische und quartische Polynome berücksichtigt werden.

Tabelle 1: Faktorstufen DoE

Faktorstufe	Neigungswinkel in °	Helligkeit in lx	Cannywert +	Cannywert -	Polynomgrad
–	24,14	10	0,05	0,1	3
0	33,88	50	0,1	0,2	-
+	44,01	600	0,15	0,25	4

Die physischen Einstellungen erfolgen durch den Austausch modularer Kamerahalterungen (siehe Abb.3) und gezielte Lichtsteuerung im Testraum. So wird sichergestellt, dass der Algorithmus unter kontrollierten, reproduzierbaren Bedingungen arbeitet. Die Kombination der Versuchsparameter ermöglicht eine differenzierte Bewertung der Kalibrierfähigkeit, Detektionsgüte und Fehlertoleranz. Die Ergebnisse werden in Kapitel 4 analysiert.



Abbildung 3: Standard Kamerahalter, DoE-Kamerahalter, AprilTag Befestigung

3.3 Infrastrukturseitige kamerabasierte Freiraumüberwachung

Zur Ergänzung der fahrzeugseitigen Sensorik wird ein stationäres, kamerabasiertes Überwachungssystem entwickelt, das Hindernisse im Gleisbereich zuverlässig erkennt.

Das stationäre System verarbeitet auf Anfrage des Fahrzeugs den relevanten Streckenabschnitt (Abb. 4). Vor der Erkennung wird die Kameraorientierung per ORB-Features und RANSAC gegen ein Referenzbild abgeglichen. Kleine Rotationsabweichungen können durch eine Rotation der Routeninformation kompensiert werden, sind jedoch beleuchtungs- und sichtfeldabhängig.
Die Bildauswertung erfolgt auf Grundlage eines speziell trainierten YOLOv11-Modells [25]. Dabei wird zwischen spezifisch bekannten Fahrzeugklassen der skalierten Umgebung differenziert, beispielsweise zwischen einem schwarzen und einem weißen Triebfahrzeug. Weitere im Bild erkannte Objekte, die keiner dieser Klassen zugeordnet werden können, werden automatisch als potenzielle Hindernisse klassifiziert.
Das Modell wird auf einem eigens erstellten Datensatz mit rund 2 000 annotierten Bildern des skalierten Gleisplans trainiert. Der Datensatz umfasst die Klassen „weiße Lokomotive“, „schwarze Lokomotive“ und „Wagen“ und wurde im Verhältnis 70 % Training, 15 % Validierung und 15 % Test aufgeteilt. Das Training erfolgt mit COCO-vortrainierten Gewichten (640 × 640 px, Batchgröße 16, 100 Epochen) unter Verwendung von PyTorch / Ultralytics.

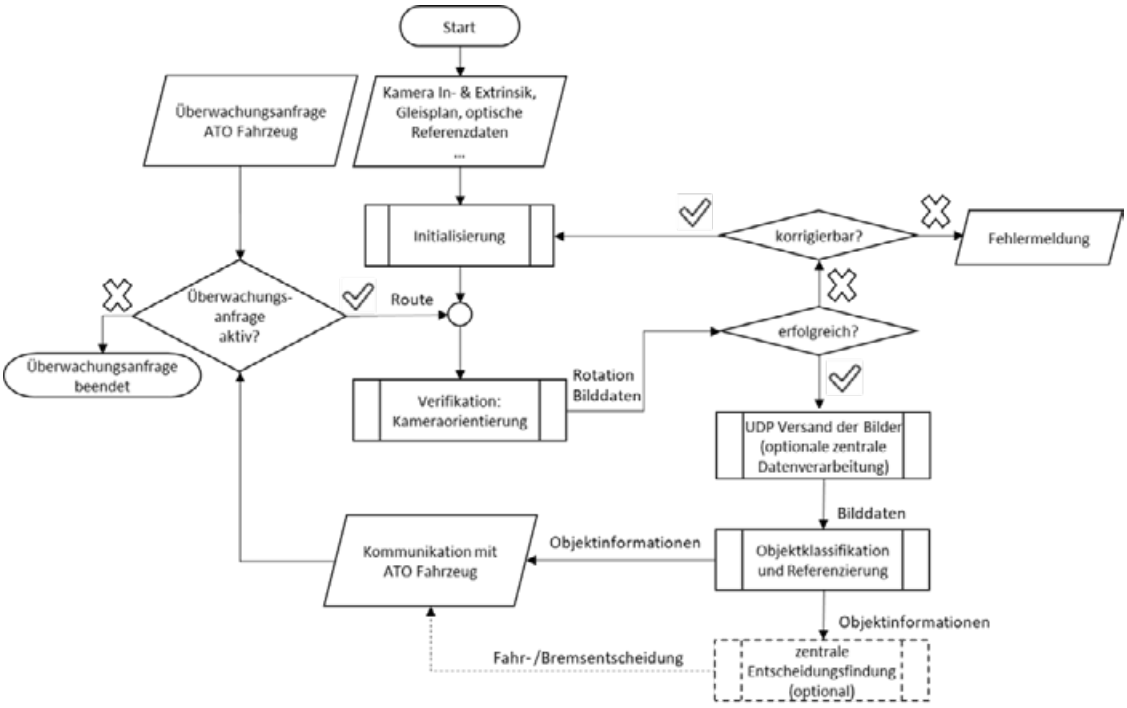


Abbildung 4: Flussdiagramm infrastrukturseitige Gleisüberwachung

Nach Empfang der Anfrage und Initialisierung wird ein aktuelles Kamerabild erfasst, analysiert und die Objektinformationen generiert. Diese umfassen sowohl die erkannte Klasse als auch die Position im Bild. Die ermittelten Ergebnisse werden entweder zur lokalen Entscheidung verwendet oder über definierte Schnittstellen an das ATO-Fahrzeug übermittelt.
Zur Validierung der Kameraposition erfolgt zu Beginn jeder Überwachungssequenz ein Abgleich mit einem zuvor aufgenommenen Referenzbild. Dieser Vergleich nutzt ORB-Feature-Extraktion sowie eine RANSAC-basierte Zuordnung charakteristischer Merkmale, um die Translation und Rotation der beiden Bilder zu bestimmen.
Der gesamte Ablauf ist als zyklischer Prozess konzipiert. Am Ende jeder Iteration wird überprüft, ob die Anfrage des Fahrzeugs weiterhin aktiv ist. Ist dies nicht der Fall, wird der Prozess abgebrochen und das System in den Ausgangszustand zurückversetzt.

4 Ergebnisse

Im Folgenden werden im ersten Schritt die Resultate der systematischen Parameterstudie der Schienenerkennung, sowie weiterführend die Testergebnisse der infrastruktur-basierten Freiraumüberwachung dargestellt und hinsichtlich ihrer Aussagekraft bewertet. Grundsätzlich liefernd die beiden durchgeführten DoE-Versuchsreihen unterschiedliche Ergebnisse. Insgesamt ergeben sich aus den Faktorstufen und Positionen 300 Iterationen. Versuchsreihe 1 wird vor einem Bogen durchgeführt, während in Versuchsreihe 2 das Fahrzeug aus dem Bogen heraus in eine Gerade positioniert ist. Abbildung 5 zeigt exemplarisch das Ergebnis der Schienendetektion an einer Weiche.
Die Bewertung der Schienen- und Weichendetektion erfolgt anhand der Kennwerte Detektionsrate und mittlerer quadratischer Fehler (MSE). Die Detektionsrate beschreibt den Anteil der Bildsequenzen, in denen der Algorithmus erfolgreich ein Polynom für den Schienenverlauf bestimmt, und dient als Maß für die Funktionszuverlässigkeit. Der MSE quantifiziert die geometrische Abweichung zwischen der detektierten und der idealisierten Schienenlinie und bildet damit ein Maß für die Genauigkeit der Schienenverlaufsschätzung. Für die Auswertung wird die Ground Truth manuell durch Referenzpunkte entlang des realen Schienenverlaufs definiert, wodurch geringfügige Abweichungen nicht ausgeschlossen sind.
Versuchsreihe 1 erreicht höhere Detektionsraten und deutlich geringere Standardabweichungen. Zudem liegen auch MSE-Werte und deren Streuung unter denen von Versuchsreihe 2. Die Rechendauer ist ähnlich. Häufige Fehlmuster sind fälschlich erkannte Querschwellenübergänge, während in Versuchsreihe 2 häufiger unvollständige Verläufe und seitliche Abweichungen im oberen Bildbereich auftreten.

Die besten Werte für eine hohe Detektionsrate und einen niedrigen MSE über beide Versuchsreihen und drei Helligkeitsstufen hinweg ergeben sich mit den in Tabelle 2 aufgeführten Kombinationen aus Neigungswinkel und Canny-Schwellwerten, bei Verwendung eines Polynommodells vierten Grades.

Tabelle 2: Ergebnisse Parameterstudie

Konfiguration		Ergebnisse	
Neigung in °	Cannyschwellwerte min/max	Detektionsrate	MSE in Quadratpixeln
44,0	0,15/0,25	1,00	217,46
24,1	0,10/0,20	0,97	172,68
24,1	0,15/0,25	0,92	147,39

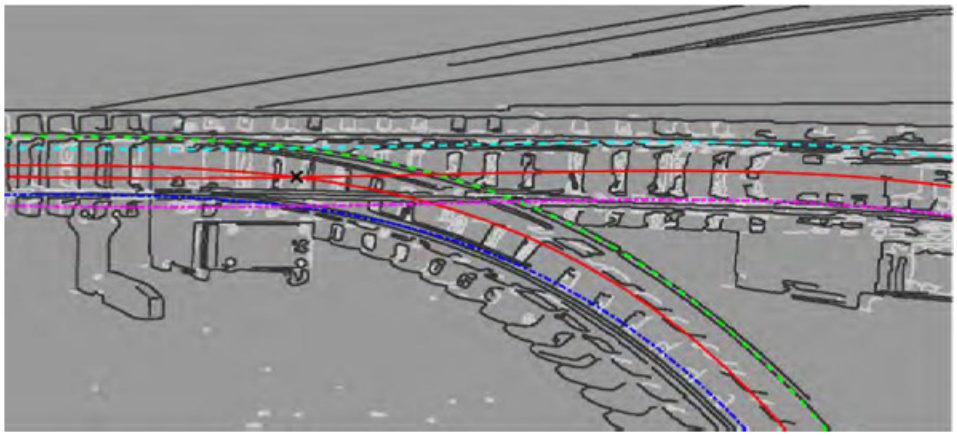


Abbildung 5: Beispiel der kantengeführten Schienendetektion

Die systematische Auswertung beider Versuchsreihen mittels DoE ergibt folgende zentrale Erkenntnisse. Höhere Canny-Schwellen verringern die Detektionsrate. Quartische Polynome senken den MSE in beiden Versuchsreihen deutlich und stabilisieren ihn gegenüber Änderungen von Helligkeit und Neigung, erhöhen jedoch die Rechenzeit um den Faktor 2,5–3. Helligkeit wirkt sich nur moderat auf die Rechenzeit aus (Faktor 1,1-2), während höhere Canny-Werte sie leicht reduzieren. Eine signifikante Wechselwirkung zwischen Helligkeit und Polynomgrad wird nicht beobachtet. Weitere Effekte zeigen kein konsistentes Muster.

Abschließend ist in Abb. 6 die Funktionalität des AprilTag-Detektors dargestellt. Durch die in hellblau gekennzeichneten ermittelten Abweichungen, ist die Verdrehung der Kamera bestimmbar, um Fehlermeldungen zu erzeugen oder die Kamerapose zu korrigieren.

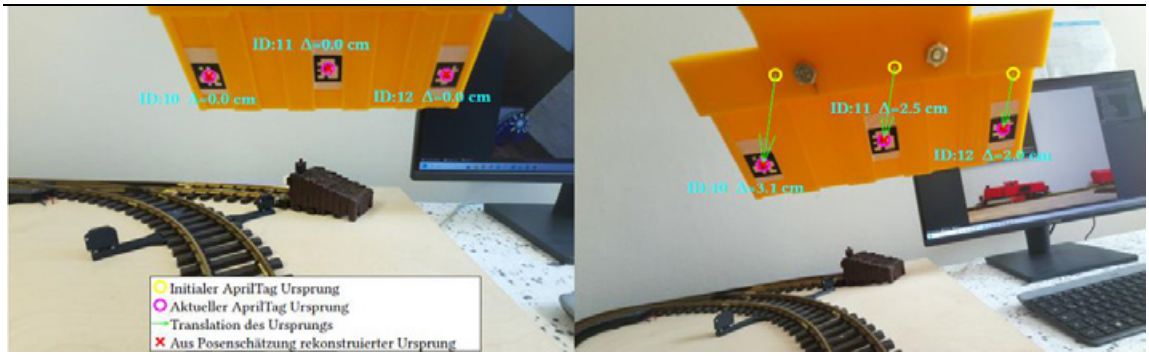


Abbildung 6: Vergleich AprilTag-Detektion in Referenzposition und Testposition

Weiterführend werden die Ergebnisse der infrastrukturbasierten Hinderniserkennung präsentiert. Zu Beginn der kamerabasierten Freiraumüberwachung ist es essenziell die Kameraorientierung zu verifizieren. Hierfür wird die in Abb. 7 dargestellte Feature-Extraktion durchgeführt. Daraus folgend wird die Rotation zwischen den beiden Bildern ermittelt. Gleichzeitig wird im Rahmen der Tests deutlich, dass die Helligkeit (bei gleicher Kamerapose) einen Einfluss auf die Präzision der Rotationsermittlung besitzt. Weiterführend ist es bei geringen Rotationsabweichungen möglich, die hinterlegten Routeninformationen ebenfalls zu rotieren, um weiterhin zuverlässige Daten zu liefern. Diese Möglichkeit hängt jedoch stark von der noch sichtbaren Gleisgeometrie und Belichtung ab und ist daher keine zuverlässige Lösung.

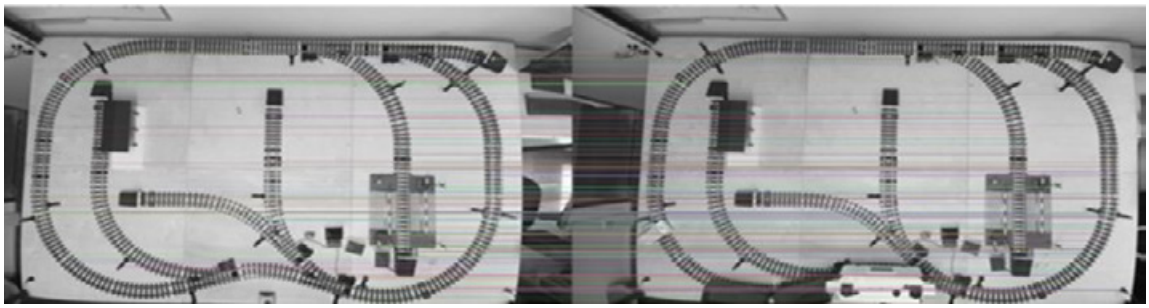


Abbildung 7: Feature-Extraktion zwischen Referenzbild und Testbild (bunte Linien)

Werden die Kameradaten als hinreichend präzise bewertet, wird im nächsten Schritt die Objekterkennung mit dem YOLOv11 Modell durchgeführt.

Für die Bewertung der kamerabasierten Freiraumüberwachung werden die in der Objekterkennung etablierten Kennzahlen *Precision*, *Recall* sowie *mean Average Precision (mAP)* nach dem COCO-Evaluationsstandard [26] verwendet. *Precision* beschreibt den Anteil korrekt klassifizierter Detektionen, *Recall* den Anteil erkannter Objekte an allen tatsächlich vorhandenen. *mAP* ergibt sich aus der Fläche unter der *Precision-Recall*-Kurve und stellt ein zusammenfassendes Maß für die

Erkennungsleistung des Modells dar. Diese Kennzahlen erlauben eine einheitliche Beurteilung der Detektionsgüte und der Generalisierungsfähigkeit des neuronalen Netzes über alle Szenarien hinweg.

Im Vergleich zu einem initial verwendeten YOLOv2-Modell, zeigt YOLOv11 durch verbesserte Feature-Extraktionsmodule eine deutlich höhere Robustheit. Die erzielten Leistungskennzahlen (Precision ≈ 99,5 %, Recall ≈ 1,0, mAP50 ≈ 0,995, mAP50–95 ≈ 0,84) belegen eine hinreichende Detektionsgüte und Generalisierungsfähigkeit innerhalb der skalierten Modellumgebung (siehe Abb. 8). .

Das Modell erkennt die Klassen Lokomotive (schwarz und weiß) sowie Wagen zuverlässig. Die Trainingsmetriken Precision und Recall zeigen übliche Größenordnungen. Für weitere Objekte wird ein ROI-Filter genutzt, um den Fokus auf den Gleisbereich bzw. die Route der geschobenen Einheit zu legen. Werden hier unbekannte Objekte detektiert, werden diese automatisch als Hindernisse klassifiziert.

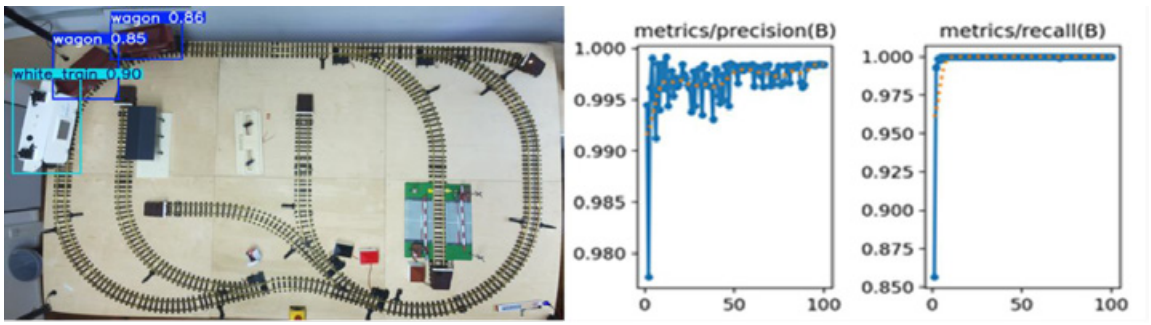


Abbildung 8: Detektierte Objekte mit Boundingboxen (links) die Performance Metriken precision und recall des YOLOv11 Modells (rechts)

Abschließend stellt Abb. 9 den vollen Funktionsumfang der infrastrukturbasierten Freiraumerkennung dar. Im Anwendungsfall wird das Fahrzeug, sowie die Wagen der geschobenen Einheit erkannt. Im Gleisbereich befindet sich ein Objekt, das als Hindernis klassifiziert wird. Für die weitere Datenverarbeitung werden die Objekte im Gleisplan hinterlegt und im nächsten Schritt zur weiteren Verarbeitung an das Fahrzeug gesendet.

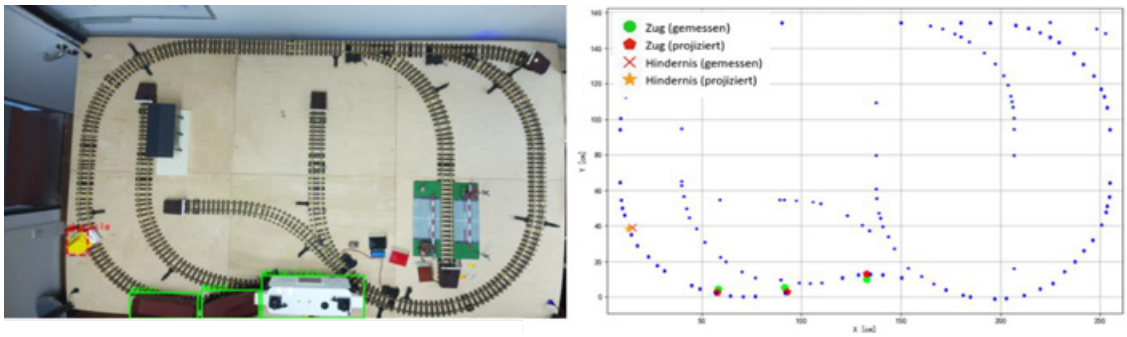


Abbildung 9: Objekterkennung, Hindernisklassifizierung und Lokalisierung

5 Fazit

Die Arbeit demonstriert, dass skalierte Modellumgebungen ein zeit- und kosteneffizientes Werkzeug zur Vorentwicklung kamerabasierter ATO-Funktionen im Sichtbetrieb darstellen, wie beispielsweise der fahrzeugseitigen Schienenkopferkennung und der infrastrukturseitigen Objektdetektion zur Absicherung geschobener Einheiten.

Die durchgeführte DoE-Analyse zeigt, dass die Parametrierung der Bildverarbeitung einen erheblichen Einfluss auf die Detektionsgüte besitzt. Insbesondere bei der Kantenerkennung konnten empfindliche Wechselwirkungen mit der Bildhelligkeit beobachtet werden. Zudem führt die Anwendung von quartischen Polynommodellen zwar zu einer signifikanten Verbesserung der Modellgüte, ist jedoch mit einer substanziellen Steigerung der erforderlichen Rechenzeit verbunden. Die Selbstdiagnose mittels AprilTags hat sich als praxistauglich erwiesen, jedoch ist die fest definierte Halterposition im Kamerabild nicht optimal für reale Anwendungen. Im skalierten Setup ist die Kamera geneigt, um auch im Bogen Schienen detektieren zu können. Diese Einschränkung muss bei der Übertragung der Erkenntnisse in die reale Anwendung berücksichtigt werden.

Das infrastrukturseitige Kamerasystem auf Basis eines YOLOv11-Netzes zeigt für die Vorentwicklung hinreichende Leistungen. Allerdings erlaubt die monokulare Perspektive nicht in allen Fällen eine eindeutige Beurteilung, ob detektierte Objekte tatsächlich in den Lichtraum hineinragen. Für die Übertragung in die reale Anwendung ist die Betrachtung der Situation aus mehreren Perspektiven oder mehreren Sensoren unabdingbar.

Insgesamt liefert die Arbeit eine methodische Grundlage für die Weiterentwicklung kamerabasierter Automatisierung im Sichtbetrieb. Die Kombination aus systematischer Parametervariation, gezielter Störungseinbringung und realitätsnaher Testumgebung erlaubt eine differenzierte Bewertung von Leistungsfähigkeit und Schwächen des Systems. Die erzielten Ergebnisse verdeutlichen das Potenzial der skalierten Entwicklungsumgebung für die Vorentwicklung kamerabasierter ATO-Funktionen, zeigen jedoch zugleich deren methodische Grenzen, insbesondere hinsichtlich der Maßstabsübertragbarkeit und Sensordynamik, auf.

Für künftige Arbeiten ist die Aufzeichnung realer Daten geplant, mit denen die Aussagen der Parameterstudie validiert und weiterführend Skalierungsmethoden entwickelt werden, um den Übertrag zwischen skalierter und realer Anwendung zu ermöglichen.

Danksagung

Die in dieser Arbeit eingesetzte Hardware wurde im Projekt EduRail durch das Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg im Rahmen des InnovationsCampus Mobilität der Zukunft (ICM) gefördert.

Literatur

[1] L. H. Kaack, P. Vasihnav, M. G. Morgan, I. L. Azevedo, und S. Rai, „Decarbonizing intraregional freight systems with a focus on modal shift“, *Environ. Res. Lett.*, Bd. 13, Nr. 8, Jan. 2018, doi: 10.1088/1748-9326/aad56c.

[2] M. Trepáčová, V. Kurečková, P. Zámečník, und P. Řezáč, „Advantages and disadvantages of rail transportation as perceived by passengers: A qualitative and quantitative study in the Czech Republic“, *Trans. Transp. Sci.*, Bd. 11, Nr. 3, S. 52–62, Dez. 2020, doi: 10.5507/tots.2020.014.

[3] P. Guglielminetti *u. a.*, „Study on Single Wagonload Traffic in Europe – challenges, prospects and policy options“. Zugegriffen: 31. Oktober 2025. [Online]. Verfügbar unter: https://transport.ec.europa.eu/system/files/2017-02/2015-07-swl-final-report .pdf

[4] European Union, „Digitalisation in railway transport“, 2019, [Online]. Verfügbar unter: https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/635528/ EPRS BRI(2019)635528_EN.pdf

[5] L. Greiner-Fuchs, S. Schäfer, T. Hofmeier, und M. Cichon, „Database-supported methodical approach for the development of a toolchain for the evaluation of ATO functions using a scenario-based test methodology“, in *Proceedings of the Fifth International Conference on Railway Technology: Research, Development and Maintenance*, in Paper 13.4. Montpellier: Civil-Comp Press, 2022. doi: 10.4203/ccc.1.13.4.

[6] T. Hofmeier und M. G. Cichon, „Introducing Scaled Model Development to on-Sight Automatic Train Operation“, in *Proceedings of the 10th International Conference on Vehicle Technology and Intelligent Transport Systems*, SciTePress, 2024, S. 378–385. doi: 10.5220/0012691500003702.

[7] DB InfraGO AG, „Konzernrichtlinie 408.4814 – Rangieren: Durchführen – Regelfall“, DB InfraGO AG, Frankfurt am Main, 2024.

[8] IEC, *IEC 62267:2009 Railway applications. Automated urban guided transport (AUGT). Safety requirements*, Definitive. 2010.

[9] J. P. Powell, A. Fraszczyk, C. N. Cheong, und H. K. Yeung, „Potential Benefits and Obstacles of Implementing Driverless Train Operation on the Tyne and Wear Metro: A Simulation Exercise“, *Urban Rail Transit*, Bd. 2, Nr. 3–4, S. 114–127, Dez. 2016, doi: 10.1007/s40864-016-0046-9.

[10] Y. Wang, M. Zhang, J. Ma, und X. Zhou, „Survey on Driverless Train Operation for Urban Rail Transit Systems“, *Urban Rail Transit*, Bd. 2, Nr. 3–4, S. 106–113, Dez. 2016, doi: 10.1007/s40864-016-0047-8.

[11] G. Trummer und J. Rappe, „RUBIN: Die erste fahrerlose U-Bahn im Mischbetrieb | ZEVrail.de“, *ZEVrail*, Bd. Jahrgang 132, Nr. 9, 2008, Zugegriffen: 31. Oktober

2025. [Online]. Verfügbar unter: https://www.zevrail.de/artikel/rubin-die-erste-fahrerlose-u-bahn-im-mischbetrieb

[12] DB Systemtechnik, „Digitale S-Bahn Hamburg - A safety concept to operate ATO over ETCS“. Zugegriffen: 31. Oktober 2025. [Online]. Verfügbar unter: https://www.era.europa.eu/system/files/2022-11/control_major_risks_-_marc.pdf

[13] R. Tagiew *u. a.*, „Onboard Sensor Systems for Automatic Train Operation“, in *Dependable Computing – EDCC 2022 Workshops*, S. Marrone, M. De Sanctis, I. Kocsis, R. Adler, R. Hawkins, P. Schleiß, S. Marrone, R. Nardone, F. Flammini, und V. Vittorini, Hrsg., Cham: Springer International Publishing, 2022, S. 139–150.

[14] R. Tagiew *u. a.*, „Sensor system for development of perception systems for ATO“, *Discov. Artif. Intell.*,Bd.3, Nr.1, S.22, Juni 2023, doi:10.1007/s44163-023-00066-4.

[15] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, und V. Koltun, „CARLA: An Open Urban Driving Simulator“, in *Proceedings of the 1st Annual Conference on Robot Learning*, S. Levine, V. Vanhoucke, und K. Goldberg, Hrsg., in Proceedings of Machine Learning Research, vol. 78. PMLR, Nov. 2017, S. 1–16. [Online]. Verfügbar unter: https://proceedings.mlr.press/v78/dosovitskiy17a.html

[16] T. Hofmeier, M. Kleinlein, und M. Cichon, „Development of Positioning Methods for ATO in a Scaled Model Environment“, in *Proceedings of the Sixth International Conference on Railway Technology Research, Development and Maintenance*, J. Pombo, Hrsg., in Online volume CCC 7. Edinburgh, UK: Civil-Comp Press, 2024, S. Paper 21.1. doi: 10.4203ccc.7.21.1.

[17] T. Hofmeier und M. Cichon, „Scaled Model Development and Testing for Automatic Train Operation: Creating a Digital Twin“, 31. Oktober 2025. [Online]. Verfügbar unter: https://www.researchgate.net/publication/390359078

[18] T. Hofmeier und M. Cichon, „Abschlussbericht EduRail - Educational Railway Demonstrator for Future Mobility“, Forschungsbericht/Preprint, Jan. 2025. doi: 10.5445/IR/1000180934.

[19] D. Ristić-Durrant, M. Franke, und K. Michels, „A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways“, *Sensors*, Bd. 21, Nr. 10, Art. Nr. 10, Jan. 2021, doi: 10.3390/s21103452.

[20] A.-R. Alexandrescu, A. Manole, und L. Dioșan, “Railway Switch Classification Using Deep Neural Networks,” in Proc. 18th Int. Joint Conf. Comput. Vis., Imaging Graph. Theory Appl., Lisbon, Portugal, 2023, pp. 769–776, doi: 10.5220/0011611600003417.

[21] Eisenbahn-Bundesamt, „EBA - ASaG - Automatische Schaderkennung an Güterwagen“. Zugegriffen: 31.Oktober 2025. [Online]. https://www.eba.bund.de/Z-SGV/Projekte/laufende_Projekte/ASaG/asag_node.html#doc3712042bodyText3

[22] R. Ross, „Vision-based track estimation and turnout detection using recursive estimation“, in *13th International IEEE Conference on Intelligent Transportation Systems*, 2010, S. 1330–1335. doi: 10.1109/ITSC.2010.5625015.

[23] M. Ziegler, V. Mhasawade, M. Köppel, P. Neumaier, und V. Eiselein, „A Comprehensive Framework for Evaluating Vision-Based on-Board Rail Track Detection“, in *2023 IEEE Intelligent Vehicles Symposium (IV)*, Anchorage, AK, USA: IEEE, Juni 2023, S. 1–8. doi: 10.1109/IV55152.2023.10186659.

[24] T. Hofmeier und M. Cichon, „A Scaled Lab Proposal for Development and Testing of On-Sight Automatic Train Operation“. Zugegriffen: 31. Oktober 2025. [Online]. Verfügbar unter: https://www.researchgate.net/publication/393267484

[25] J. Redmon, S. Divvala, R. Girshick, und A. Farhadi, „You Only Look Once: Unified, Real-Time Object Detection“, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, S. 779–788. doi: 10.1109/CVPR.2016.91.

[26] T.-Y. Lin *u. a.*, „Microsoft COCO: Common Objects in Context“, in *Computer Vision – ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, und T. Tuytelaars, Hrsg., Cham: Springer International Publishing, 2014, S. 740–755. doi: 10.1007/978-3-319-10602-1_48.

Autoren



Hofmeier Tobias, M.Sc.

Tobias Hofmeier ist wissenschaftlicher Mitarbeiter am Karlsruher Institut für Technologie (KIT). Im Rahmen seiner Forschung entwickelt er skalierte Testumgebungen zur Untersuchung und Erprobung automatisierter Zugfahrten auf Sicht. Darüber hinaus liegt ein besonderer Fokus auf der Umsetzung virtueller Simulationsumgebungen zur Absicherung und Validierung von ATO-Systemen. Ergänzend befasst er sich mit der Analyse und Berechnung von Energieverbräuchen im Bahnbetrieb.



Kilian Otto, B.Sc.

Kilian Otto absolvierte sein Bachelorstudium Maschinenbau im Schwerpunkt Informationstechnik am KIT und vertieft sich im Masterstudium in Bahnsystem- und Automatisierungstechnik. Seine Masterarbeit befasst sich mit optischer Schienendetektion und kamerabasierter Lokomotivortung.



Yan Cao, B.Sc.

Yan Cao studiert Elektrotechnik und Informationstechnik am KIT mit Schwerpunkt Elektromobilität. In ihrer Masterarbeit entwickelt sie ein kamerabasiertes Infrastruktursystem auf einem maßstabsgetreuen Zugmodell zur Unterstützung automatisierter Zugkoordination.



Cichon Martin, Prof. Dr.-Ing

Prof. Dr.-Ing. Martin Cichon ist Leiter des Institutsteils Bahnsystemtechnik am Institut für Fahrzeugsystemtechnik (FAST) des KIT. Sein Forschungsgebiet umfasst die Thematik des Maschinenbaus im Bereich der Schienenfahrzeugtechnik. Er leitet Kooperationsprojekte mit DB Cargo und weiteren Industrie- und Forschungspartnern und begleitet maßgeblich die Technologie-migration von ATO-Systemen.

Session 6
Klaus Mößner, Sam Münchow, Charbel Lahoud

Hybride 5G Echtzeitübertragung in der Anschlussbahn im Kontext von SAMIRA 2.0

Prof. Dr. Mößner, Klaus[1], Münchow, Sam[2], M.Sc. Lahoud, Charbel[3]

[1] Inhaber Professur Nachrichtentechnik
Technische Universität Chemnitz - Reichenhainer Str. 70 09107 Chemnitz
[2]Leiter Forschung und Entwicklung
IKADO GmbH - Auf der Hüls 198 52068 Aachen
[3] Wissenschaftlicher Mitarbeiter
Technische Universität Chemnitz - Reichenhainer Str. 70 09107 Chemnitz

Thematischer Schwerpunkt

Die drahtlose Echtzeitsignalübertragung von Videodaten im öffentlichen 5G Netz und in privaten 5G-Campusnetzen als Grundlage für die Digitalisierung und Automatisierung von Rangiervorgängen in der letzten Meile im Kontext des digitalen Rangierassistenten SAMIRA 2.0.

Keywords: 5G; Echtzeitsignalübertragung; digitaler Rangierassistent; letzte Meile; ATO

1 Einleitung

Für sicherheitskritische digitale Systeme ist die Echtzeitverarbeitung der Daten häufig von großer Bedeutung. Bei räumlich verteilten Systemen spielt neben der Verarbeitungszeit auch die Übertragungszeit eine wichtige Rolle. Der aktuelle Entwicklungstrend von Systemen für Automatic Train Operation (ATO) im Rangierbetrieb erhöht die Bedeutung das Eisenbahnumfeld. Bei autonomen geschobenen Fahrten von Rangierabteilungen ist es nötig Daten von der Zugspitze zum Triebfahrzeug zu übertragen. Die 5G Übertragungstechnologie ermöglicht in diesem Kontext eine drahtlose Übertragung und liefert so einen wichtigen Baustein für die weitere Digitalisierung im Bahnumfeld. Dabei werden, wegen Sicherheitsbetrachtungen, in erster Linie private Campus Netzwerke als Lösung gesehen [1]. Diese sind von der Leistungsfähigkeit und vor allem von der Verfügbarkeit den öffentlichen 5G-Netzen weit überlegen, verursachen aber auch erhöhte Kosten. Die Anzahl privater Campusnetze in Deutschland ist in den letzten Jahren deutlich gestiegen. Bis Oktober 2024 wurden von der Bundesnetzagentur 424 Frequenzzuteilungen im Bereich 3.700–3.800 MHz vergeben – das entspricht einem Zuwachs von 23 % im Vergleich zum Vorjahr. Die Wachstumsrate hat sich im Jahr 2024 etwas verlangsamt, was auf eine Verlagerung des Fokus von neuen Installationen hin zur qualitativen Weiterentwicklung bestehender Netze hindeutet.

Diese Zahlen zeigen, dass auch perspektivisch nicht mit einer flächendeckenden Abdeckung zu rechnen ist. Um die eingeschränkte Funknetzwerkabdeckung von privaten Campus Netzwerken zu erweitern, wird auch die hybride Nutzung von privaten und öffentlichen 5G Netzwerken untersucht (5GRacom, 2025). In dem vom Bundesministerium für Verkehr geförderten Forschungsprojekt SAMIRA 2.0 wird der Einsatz von solchen hybriden Verbindungen zur Vernetzung von Assistenzsystemen im Rangierbetrieb untersucht.

2 Projektkontext SAMIRA 2.0

SAMIRA ist ein Assistenzsystem für Rangierfahrten von Schienenfahrzeugen in der letzten Meile, bei der moderne Sensorik in einem tragbaren Sensormodul (SAMIRAmobil) an dem letzten Waggon einer Rangierabteilung befestigt wird und es dem Lokrangierführer damit ermöglicht, auch ohne einen Rangierbegleiter geschobene Fahrten sicher durchzuführen [2]. Abbildung 1 zeigt das Projektbild des Vorhabens, welches verschiedene technische Aspekte der Umfelderkennung und Visualisierung dargestellt.

Abbildung 2 zeigt den aktuellen Prototypen des Sensormoduls in einem kompakten Koffer, wodurch ein flexibler und mobiler Einsatz möglich ist. Auf einem Display im Führerhaus der Lokomotive wird der Echtzeit-Videostream vom letzten Waggon angezeigt, in dem zusätzlich mittels Augmented Reality wichtige Merkmale wie Personen in Gleisnähe, Abstände und Geschwindigkeiten angezeigt werden. Dazu werden die Videodaten des Sensormoduls sicher, mittels der Mobilfunk-Schlüsseltechnologie 5G, an das Display auf der Lok übertragen. Für diese Übertragung ist eine Mindestbandbreite von 14 Mbit/s pro Videosignal nötig.



Abbildung 1: SAMIRA 2.0 Rail Shunting System – Projektübersicht.

Es ist essenziell, dass dem Lokrangierführer aktuelle Videodaten zur Verfügung stehen und es zu keinen signifikanten Ausfällen in der Bildübertragung kommt. Die maximal tolerierbare Verzögerung vom Kamerasensor bis zur Anzeige auf dem Display beträgt 1000 Millisekunde.

Der Verlust einzelner Videopakete kann kompensiert werden. Ab einer Verlustrate von >1% kommt es zu deutlichen Artefakten, besonders bei Verlusten von I-Frames oder bei aufeinanderfolgenden Paketen [3]. Dies wiederum schränkt die Wahrnehmung des Lokrangierführers signifikant ein und ist im Anwendungsfall nicht tolerierbar.



Abbildung 2: SAMIRAmobil - die Sensoren sind in einem handlichen, tragbaren Koffer untergebracht.

Demgegenüber stellt die letzte Meile eine Anspruchsvolle Umgebung für die drahtlose Signalübertragung dar. In der industriellen Umgebung befinden häufig große Strukturen, die den Empfang stark einschränken. Hierdurch ist die generelle Abdeckung durch das öffentliche Netz teilweise stark eingeschränkt.

Im Rahmen des Projekts wurden eine Methode zur verschlüsselten Echtzeitübertragung von Videodaten entwickelt, welche eine hybride Nutzung von öffentlichen und privaten 5G Netzen ermöglicht und so besonders für den Einsatz in der letzten Meile geeignet ist.

3 Vergleich öffentliches 5G Netz und privates 5G Netz

Öffentliche Netzbetreiber nutzen in vielen Bereichen Frequenzbänder im Bereich 700 MHz und 1800 MHz für 5G-Netze, um eine wirtschaftliche und großflächige Abdeckung zu erreichen. Diese Frequenzen bieten eine gute Ausbreitung, aber geringere Datenraten und Flächenkapazität als höhere Frequenzen. Das n78 ist ein 5G Frequenzband, das im Bereich um 3.6 GHz liegt. Es ist ein beliebtes und leistungsstarkes Band für 5G, das vor allem in dicht besiedelten Gebieten genutzt wird, da es relativ kurze Reichweiten hat [4].

Höhere Frequenzen (3,5 GHz - 3,8 GHz) erfordern mehr Infrastruktur, bieten jedoch höhere Datenraten. Für private 5G-Campusnetze stellt die Bundesnetzagentur den Frequenzbereich von 3,7 bis 3,8 GHz bereit. Die Datenraten hängen von verschiedenen Faktoren ab, wie Bandbreite, Frequenzen, Netzwerktopologie und Netzlast. Campusnetze erreichen Spitzenwerte von über 1 Gbit/s, während öffentliche Netze typischerweise zwischen 100-400 Mbit/s liegen. Die Latenzzeiten sind bei Campusnetzen niedriger (<10 Millisekunde) als bei öffentlichen Netzen (bis zu 30 Millisekunde). [1] Hauptvorteil von Campusnetzten für die Echtzeitübertragung ist aber die Möglichkeit Kapazitäten im Netz für bestimmte Geräte oder Anwendungen zu reservieren.

Aufgrund der geringeren Ausbreitung in den Frequenzbändern der Campusnetze eigene diese sich besonders gut für den Schienengüterverkehr in der letzten Meile, da hier die räumliche Ausdehnung des Schienennetzes limitiert ist und so schon wenige 5G-Basisstationen einen großen Teil der Betriebsgelände abdecken.

Die Verbreitung privater 5G-Netze in Deutschland hat in den letzten Jahren erheblich zugenommen. Besonders in Branchen wie Fertigung, Gesundheitswesen, Transport und Logistik sowie Energie und Versorgung wird der Einsatz privater 5G-Netze immer wichtiger [5]. In vielen dieser Bereiche ist der Schienengüterverkehr von großer Bedeutung. Gleichzeitig wird eine Vollabdeckung der Betriebsgelände durch die privaten 5G-Netze aus wirtschaftlichen Gründen nur in wenigen Ausnahmen realisiert.

Netzwerk-Slicing und private 5G-Netze in Kombination mit Multi-RAT-Auswahl bieten ein robustes Rahmenwerk zur Erfüllung heterogener Anforderungen von Unternehmen. Durch den Einsatz von lokalen privaten Kernen (On-Premise) und dedizierten Spektrum-Slices können Organisationen strenge QoS-Vorgaben, extrem niedrige Latenzzeiten und erhöhte Sicherheit für geschäftskritische Anwendungen gewährleisten, während weniger priorisierter Datenverkehr auf alternative RATs (Radio Access Technology) wie LTE oder Wi-Fi ausgelagert wird [6] [7].

In industriellen oder Campus-Umgebungen bewertet der RBSE-Algorithmus („Reference Base Station Efficiency“) dynamisch die Sendeleistung der Basisstation, die Verkehrslast und die spektrale Effizienz der Nutzer, um die optimale RAT auszuwählen. Dadurch werden unnötige Übergaben reduziert, die Ende-zu-Ende-Verzögerung gesenkt und der Durchsatz über 5G-, LTE- und WLAN-Schichten hinweg maximiert [8].

In [9] werden fortschrittliche Nutzerassoziationsstrategien für heterogene Netzwerke untersucht. Einschließlich Massive MIMO, mmWave und energiegewinnender Szenarien. Ziel ist es, eine ausgewogene Lastverteilung und eine bessere Spektrumausnutzung zu erreichen, indem Endgeräte zur jeweils geeignetsten Zugangstechnologie und Slice-Instanz geleitet werden.

Feldversuche, die im Ericsson Mobility Report dokumentiert sind, zeigen, dass eine private 5G-Implementierung in der Logistik eine Produktivitätssteigerung von 20 % und CAPEX-Einsparungen von 15 % im Vergleich zu Wi-Fi ermöglichen kann, während gleichzeitig mehrere Anwendungsfälle auf einer einzigen Infrastruktur unterstützt werden [7].

Insgesamt bildet die Synergie aus privatem Netzwerk-Slicing und adaptiven Multi-RAT-Algorithmen die Grundlage für zuverlässige, leistungsstarke Konnektivität und energieeffiziente Abläufe in der Industrie 4.0 und neue Prozesse im Schienenverkehr.

4 Realisierung der Echtzeitübertragung im Hybriden 5G Netz

Durch mehrere Testmessungen im Rahmen von SAMIRA 2.0 konnte gezeigt werden, dass auch das öffentliche 5G-Netz prinzipiell für die Echtzeitübertragung von Videodaten geeignet ist. Hier steht sowohl die benötigte Bandbreite zur Verfügung und gleichzeitig ist auch die Latenz der Übertragung mit maximal 100 Millisekunde innerhalb des benötigten Bereiches.

Gleichzeitig kommt es in Bereichen mit schlechten Empfangsbedingungen oder sehr hoher Auslastung schnell zu Störungen im Videosignal und die benötigte Qualität und Verzögerung der Anzeige kann nicht sichergestellt werden. In diesen lokalen Bereichen kann durch den gezielten Einsatz von privaten 5G-Campusnetzen sichergestellt werden, dass das System zuverlässig funktioniert.

Durch eine simultane Verbindung zu privatem und öffentlichen 5G Netz ist es möglich, auf Anwendungsebene die Signalwege zu steuern und adaptiv an die aktuellen Empfangsbedingungen anzupassen und so stehts den optimalen Signalweg zu verwenden.

KPI	Abkürzung (Einheit)	Beschreibung	Normalisierungsbereich
Reference Signal Received Power	RSRP (dBm)	Misst die empfangene Signalstärke	−120 bis −70 dBm
Reference Signal Received Quality	RSRQ (dB)	Gibt die Signalqualität an (Verhältnis Signal zu Stör-/Rauschleistung)	−30 bis −5 dB
Signal-to-Interference-plus-Noise Ratio	SINR (dB)	Bewertet die Signalreinheit gegenüber Interferenz und Rauschen	−10 bis 30 dB
Subcarrier Spacing	SCS (kHz)	Abstand zwischen benachbarten Subcarriern (15 kHz oder 30 kHz)	0 bis 1 (kodiert: 0 = 15 kHz, 1 = 30 kHz)
Downlink-Bandbreite	Bandbreite (MHz)	Repräsentiert die verfügbare Übertragungsbreite / Kapazität	1,4 bis 100 MHz

Tabelle 1: Abgerufene physikalische KPIs und Normalisierungsbereiche.

Um das Verfahren zu erproben, wurde ein Multi-RAT Testaufbau implementiert, der mit zwei 5G-Quectel-Modulen ausgestattet ist und sich zeitgleich mit dem öffentlichen 5G-Netz und dem privaten 5G-Campusnetz verbindet. Der im Projekt entwickelte Algorithmus überwacht in Echtzeit Leistungskennzahlen (Key Performance Indicators, KPIs) der physikalischen Schicht, z. B. Signalstärke und -qualität. Basierend auf dem berechneten Leistungsscore wechselt der Controller zwischen den Verbindungen, um eine zuverlässige und effiziente Netzwerkleistung sicherzustellen.

Der Algorithmus fragt periodisch jedes Quectel-Modul nach Signalmetriken ab. Die abgerufenen KPIs der physikalischen Schicht werden wie folgt in Tabelle 1 beschrieben. Anschließend wird eine gewichtete Summe der normalisierten KPIs berechnet:

$$Score = w1 \cdot RSRP + w2 \cdot SINR + w3 \cdot RSRQ + w4 \cdot SCS + w5 \cdot Bandbreite$$

Die Gewichtungen, in Tabelle 2 dargestellt, wurden aus realen Messdaten urbaner und suburbaner Mobilitätsszenarien abgeleitet und anschließend mittels Grid-Search weiter verfeinert. Ziel der Optimierung war die Minimierung der mittleren quadratischen Abweichung (Mean Squared Error, MSE) zwischen vorhergesagter und gemessener Round-Trip-Time (RTT). Dieser Ansatz verbesserte die Anpassung an die Messdaten und führte zu einer allgemeinen Reduktion der RTT. Auf einzelnen Streckenabschnitten traten jedoch Perioden mit sehr schwacher oder fehlender Netzverfügbarkeit auf. Unter diesen Bedingungen führten eingeschränkte Funkverhältnisse zu zusätzlichen Verzögerungen bei der Videoübertragung und zu temporären Verbindungsabbrüchen.

Gewichte	$w1$	$w2$	$w3$	$w4$	$w5$
Gewichtswerte	0,4	0,4	0,1	0,05	0,05

Tabelle 2: Gewichtungen der KPIs für die Berechnung des Scores.

Der RAT-Auswahlprozess überprüft kontinuierlich beide RATs, vergleicht deren Leistungs-Scores und wählt die Verbindung mit dem höheren Wert. Um häufiges Umschalten zu vermeiden, wird eine Hysterese von 0,06 Punkten angewendet: Ein Wechsel erfolgt nur, wenn der Score der alternativen Verbindung den aktuellen um mindestens 0,06 übersteigt. Auf Netzwerkebene erlaubt das Multi-RAT-Verfahren eine hybride 5G-Nutzung. Beim Videostreaming über verbindungsorientierte Protokolle wie TCP treten jedoch zusätzliche Probleme auf: Da TCP-Verbindungen an IP-Adresse und Port gebunden sind, führt ein Interfacewechsel zum Abbruch der Verbindung. In den Messungen resultierte daraus ein Videoausfall von mehr als 5 Sekunden, selbst nach Optimierungen im System SAMIRA 2.0.

Das QUIC-Protokoll, ursprünglich von Google entwickelt und inzwischen von der IETF als RFC 9000 standardisiert, bietet erhebliche Vorteile bei der Migration von Netzwerkverbindungen. Ein zentrales Merkmal von QUIC ist die Fähigkeit zur nahtlosen Verbindungsmigration. Im Gegensatz zu TCP, das Verbindungen fest an IP-Adressen und Ports bindet und daher beim Wechsel des Netzwerkinterfaces (z. B. vom öffentlichen zum privaten 5G-Netz) typischerweise abbricht, verwendet QUIC verbindungsunabhängige Connection IDs. Diese ermöglichen es, eine bestehende Sitzung trotz Änderung der IP-Adresse fortzuführen, ohne dass ein vollständiger Neuaufbau der Verbindung erforderlich ist [10].

Auf Basis von QUIC wurde im Rahmen des Forschungsprojektes SAMIRA 2.0 ein Videoübertragungsprotokoll entwickelt, welches einen fließenden Wechsel zwischen den Übertragungswegen ermöglicht, ohne dass es zu erkennbaren Störungen beim Videosignal kommt. Gleichzeitig wurde eine präzise Zeitinformation in das Protokoll eingebaut, um neben der Videoübertragung auch die Latenz der Signale auf dem Display beim Triebfahrzeugführer auswerten zu können und eine automatische Überwachung der Videoübertragung zu ermöglichen.

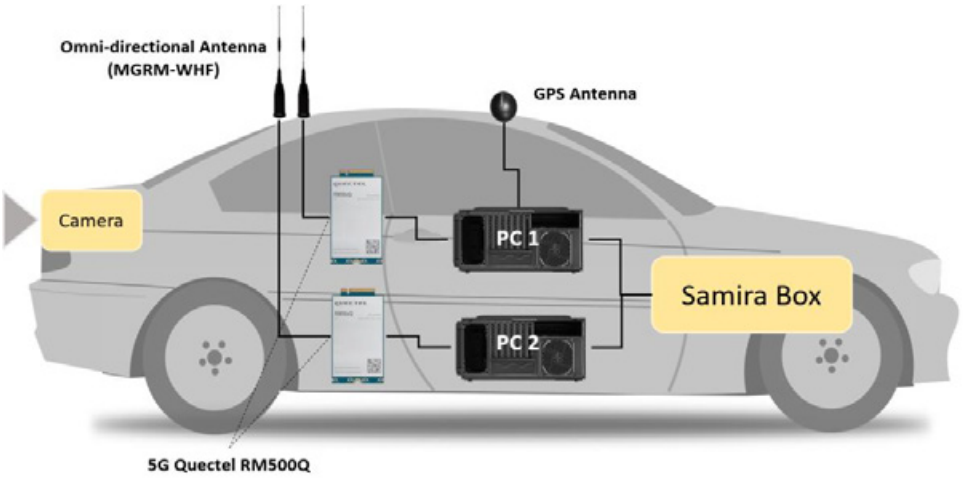


Abbildung 3: Versuchsfahrzeug mit Messaufbau zur Videoübertragung im 5G-Testfeld.

Die hybride 5G-Übertragung der Videosignale wurde im Rahmen mehrerer Fahrversuche auf dem Smart Rail Connectivity Campus (SRCC) in Annaberg-Buchholz untersucht. Der Campus verfügt über eine der modernsten 5G-Infrastrukturen entlang einer rund 25 km langen Strecke und kombiniert öffentliche sowie private Netze. Für die Versuchsfahrten wurde ein speziell ausgerüstetes Testfahrzeug eingesetzt, das den Multi-RAT-Aufbau als Netzwerkinterface für das SAMIRA 2.0-Sensormodul trug. Das Fahrzeug war mit einer omnidirektionalen Antenne, einer GPS-Antenne, zwei 5G-Quectel-Modulen und der zentralen SAMIRA-Box bestückt. Über diesen Aufbau wurden Videodaten mittels des projektintern entwickelten Protokolls an einen Webserver im Internet gesendet und dort empfangen. Dank vorheriger Zeitsynchronisation konnten sowohl Übertragungsverzögerungen als auch die Übertragungsqualität präzise gemessen werden. Der vollständige Messaufbau ist in Abbildung 3 dargestellt.

5 Ergebnisse

Wie in Tabelle 1 dargestellt, basiert der verwendete 5G-Netzwerkaufbau auf einem 5G-Standalone-Netz (5G-SA), das in den Bändern n28 (FDD), n3 (FDD) und n78 (TDD)

betrieben wird. In diesem Aufbau nutzen die Router N_p und N_u unterschiedliche Arten von Netzwerkslices, um sich mit ihren Zielservern zu verbinden. N_p nutzt einen öffentlichen 5G-Slice, der allen Nutzern zur Verfügung steht und eine direkte Konnektivität über das 5G-Netz bietet. N_u hingegen nutzt einen privaten 5G-Slice, der dafür reserviert ist, den Datenverkehr zunächst über die Server der Universität Chemnitz zu leiten, bevor der Zielserver erreicht wird. Die 5G-Signal- und Netzwerkkonfigurationen sind in Tabelle 3 zusammengefasst.

Modus	Band	Netz
5G-SA	N28 (FDD)/ N3 (FDD)/ N78 (TDD)	➔ Private slice N_p: Anbindung über Server der Universität Chemnitz ➔ Öffentlich slice N_u: Verbindung direkt zum Provider-Server

Tabelle 3: 5G-Signal- und Netzwerkkonfigurationen.

Die Abbildung 4 zeigt die vom Multi-RAT-Auswahlalgorithmus gewählte aktive Netzwerkverbindung. Die beiden Netzwerkpfade, N_p und N_u, wurden hinsichtlich ihrer Leistungsfähigkeit bewertet. Der Multi-RAT-Auswahlalgorithmus wechselte zwischen den Netzwerken N_p und N_u basierend auf den jeweiligen Bewertungen (Score). Das Ergebnis „Offline“ tritt auf, wenn das empfangene Signal zu schwach ist und keine erfolgreiche Verbindung zum Zielserver hergestellt werden kann. Auf der y-Achse sind die möglichen Zustände aufgelistet: die beiden Netzwerkslices N_p und N_u, der Zustand „Offline“ sowie Messpunkte, bei denen zwar ein Slice ausgewählt, aber keine Internetverbindung vorhanden war („kein Internet“). Jeder Punkt entspricht einer einzelnen Stichprobe/Entscheidung des Algorithmus.

Im dargestellten Messintervall wurde der private Slice N_p (blaue Markierungen) über einen großen Anteil der Zeit kontinuierlich gehalten. Der öffentliche Slice N_u (rote Markierungen) wurde phasenweise in Form diskreter Abschnitte genutzt. „Offline“-Zustände (magenta) traten vereinzelt auf, und punktuelle Messwerte ohne erreichbares Internet (schwarz) sind über den gesamten Zeitraum verteilt. Die Grafik veranschaulicht damit die zeitliche Abfolge der Slice-Wahl und die Häufigkeit von Verbindungsunterbrechungen während der Messung.

In Abbildung 5, das Testfeld wurde in einem mobilen Szenario von Schlettau nach Annaberg-Buchholz eingesetzt. Die Route ist mit einem Farbverlauf kartiert, der RTT darstellt und Werte von 22.8 bis 200 Millisekunde entlang der Strecke zeigt.

Das Zeitreihendiagramm in Abbildung 6 zeigt die zeitliche Entwicklung der Round-Trip-Time (RTT) für den privaten Slice N_p, den öffentlichen Slice N_u sowie für den jeweils gewählten aktiven Pfad. Die RTT des ausgewählten Pfads (schwarz Linie) folgt jeweils dem zur betreffenden Zeit leistungsstärkeren Pfad, N_u (rot) oder N_p (blau), und veranschaulicht damit die Reaktionsfähigkeit des Multi-RAT-Algorithmus auf sich ändernde Funkbedingungen.

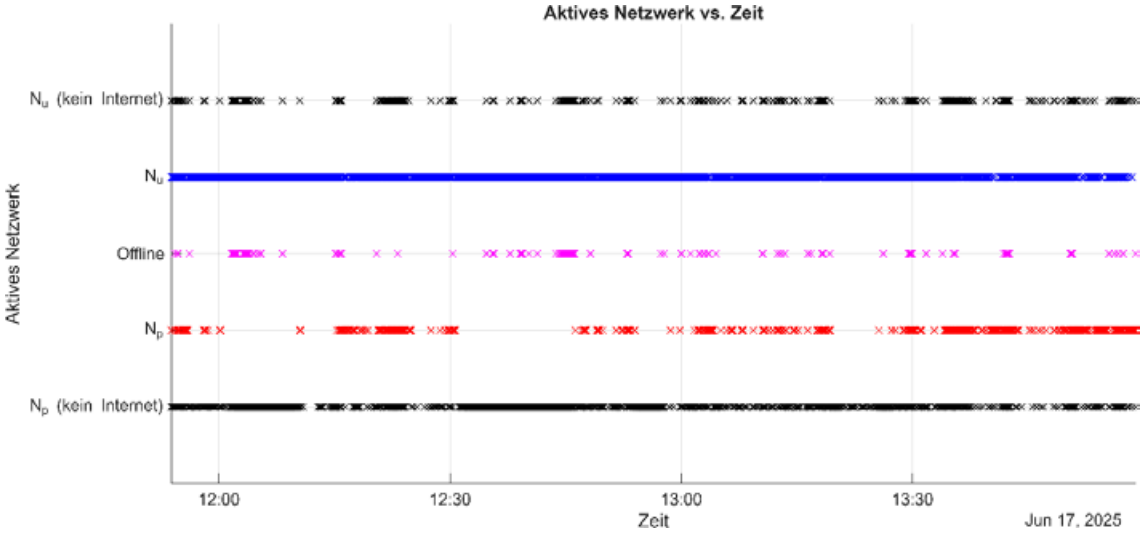


Abbildung 4: Zeitliche Entwicklung der vom Multi-RAT-Algorithmus ausgewählten aktiven Netzzustände.

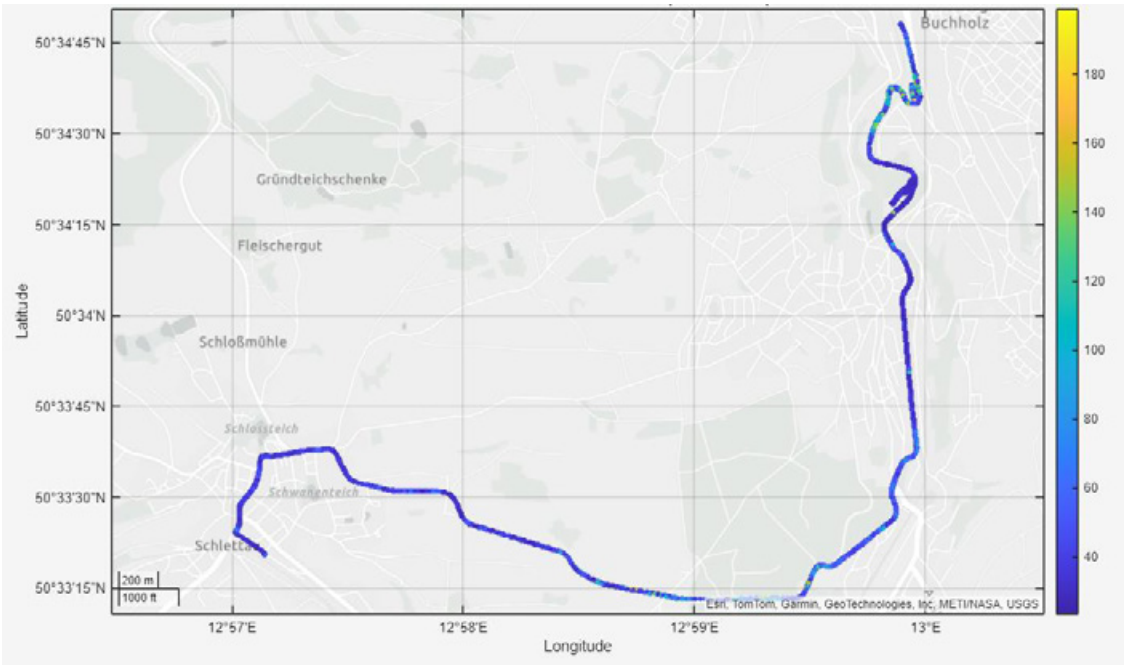


Abbildung 5: GPS-Pfad mit RTT des aktiven Netzes (Millisekunde).

Das Diagramm Abbildung 7 stellen die Scores von N_p und N_u im Bereich von 0 bis 1 dar; beide Werte zeigen häufige Schwankungen. N_p erreicht dabei oft wenigere Score, was auf die zusätzliche Latenz durch die Weiterleitung über die Server der TU Chemnitz zurückzuführen ist. Der Algorithmus wählt konsistent den Pfad mit dem höheren Score und wechselt entsprechend den Bedingungen. Dies lässt sich zeitlich korreliert mit Änderungen der RTT beobachten. Darüber hinaus hängt die Auswahl des aktiven Netzwerks nicht nur von der Punktzahl ab, sondern berücksichtigt auch, ob der jeweilige Pfad eine erreichbare Internetverbindung über das Kernnetzwerk hinaus zum Zielserver gewährleistet.

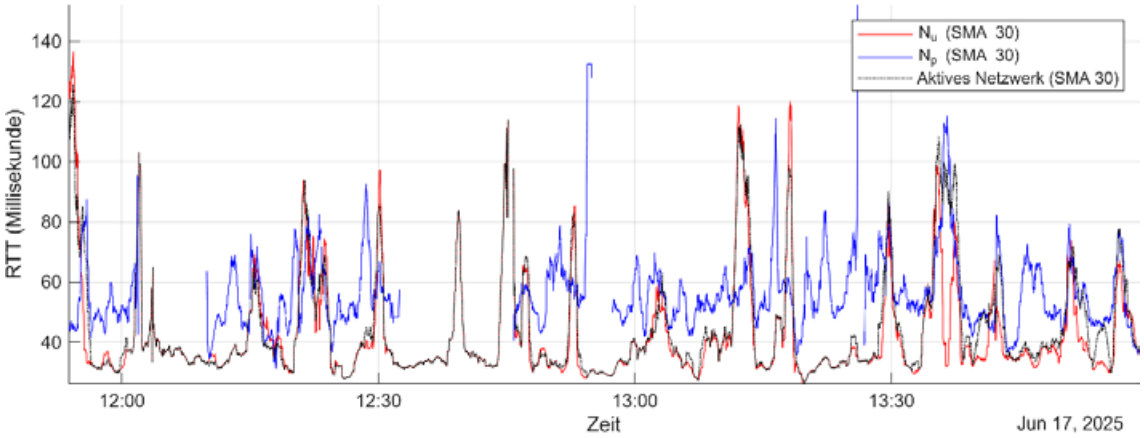


Abbildung 6: RTT über die Zeit für privaten Slice N_p, öffentlichen Slice N_u und den gewählten aktiven Pfad.

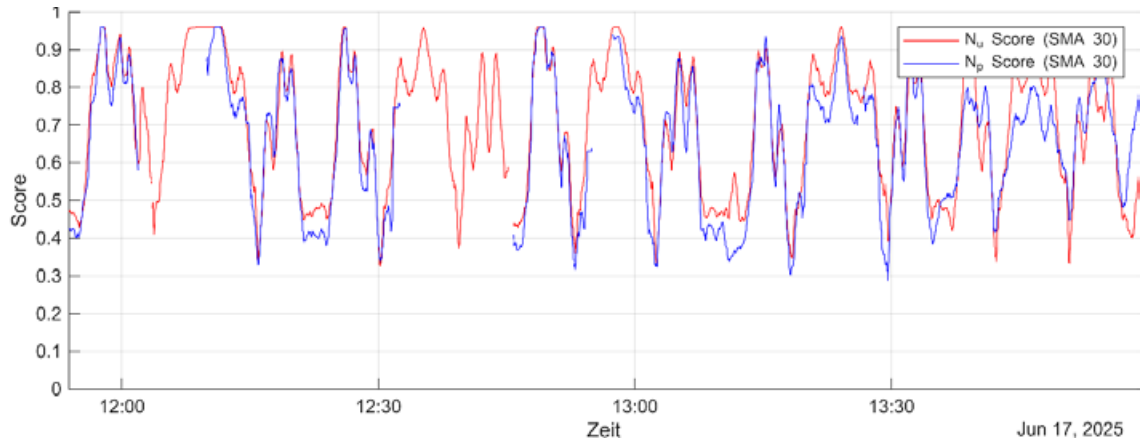


Abbildung 7: Netzwerke Leistungs-Score über die Zeit.

Während der Versuche kam es bei der Videoübertragung zu keinem sichtbaren negativen Effekt durch den Wechsel des Übertragungsweges. Mögliche Verzögerungen durch den Netzwechsel konnte durch den Video Buffer der Anzeige (80 Millisekunde) kompensiert

werden. Bei Referenzversuchen mit er typischen TCP basierten Verbindung, kam es beim Wechsel zwischen den Netzen zu einem Abbruch Videoübertragung. Erst nach manueller Wiederverbindung konnte die Verbindung nach mehr als 40 Sekunden wiederhergestellt werden konnte. Ursache hierfür waren vor allem Paketverluste beim Verbindungsaufbau.

Um die Effekte genauer untersuchen zu können wurde eine Reihe von Laborversuchen durchgeführt. Bei den Versuchen wurde ein ähnlicher Aufbau wie bei den Testfahrten verwendet. Hier wurde der Wechsel zwischen den Interfaces allerdings alle 2 Sekunden durch eine steuernde Software erzwungen. Es wurde ausgewertet wie sich die Dauer zwischen den Paketen bei Verwendung einer klassischen TLS basierten Übertragung und bei dem im Projekt entwickelten auf QUIC basierten Übertragungsprotokoll entwickelt. Abbildung 8 zeigt in Orange die Zeitpunkte des Interfacewechsels. In Schwarz/Grau ist die Verzögerung der Pakete bei TLS und in Blau die Verzögerung bei der Übertragung mit QUIC dargestellt. Es wird deutlich, dass es bei der TLS basierten Übertragung zu Verzögerungen >2 Sekunden kommt, da die Verbindung komplett neu aufgebaut werden muss.

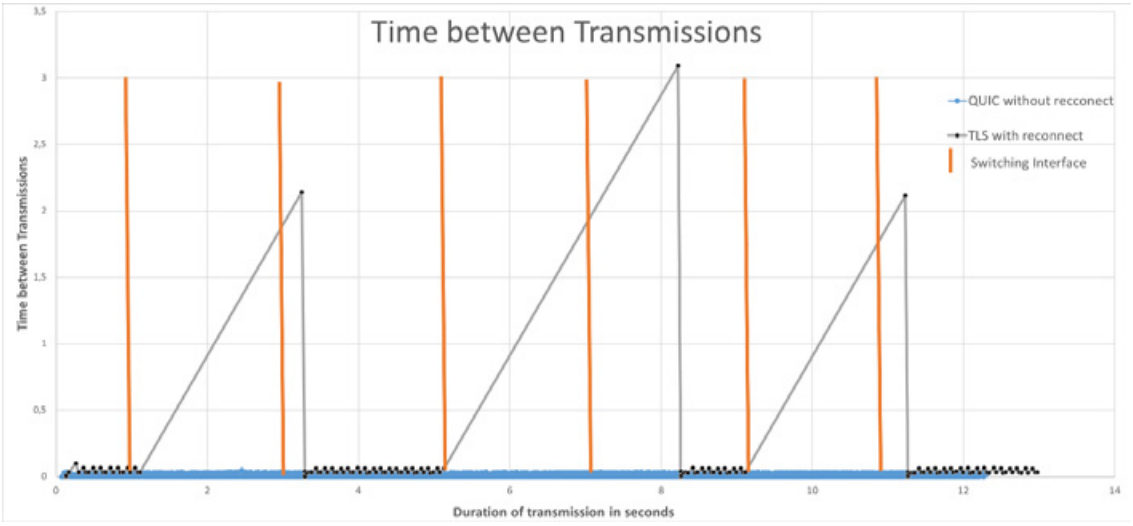


Abbildung 8: QUIC-Protokoll Dauer zwischen Übertragungen.

Bei der Übertragung mit dem QUIC basierten Protokoll kommt es zu keinen sichtbaren Signalverzögerungen durch den Wechsel des Interfaces. Dies ist in Abbildung 9 dargestellt.

Durch den Einsatz von QUIC als Transportprotokoll für die Videoübertragung wird es möglich, während einer Echtzeitübertragung den Übertragungsweg zu wechseln, ohne dass es zu sichtbaren Effekten kommt. Die hierdurch ermöglichte hybride Nutzung von 5G-Netzen bietet großes Potenzial in lokal begrenzte Anwendungsgebiete. Durch den

gezielten Ausbau der 5G Infrastruktur mit Hilfe von privaten 5G-Netzten in Bereichen mit schwacher Abdeckung kann so eine flächendeckende Abdeckung mit reduziertem Aufwand erreicht werden.

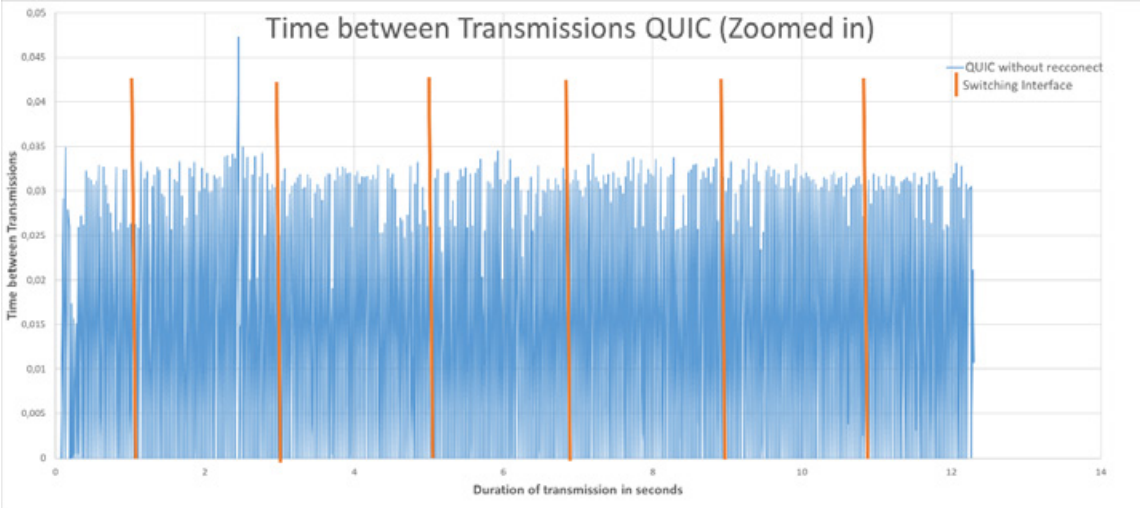


Abbildung 9: QUIC-Protokoll Dauer zwischen Übertragungen (herangezoomt).

Für die anspruchsvolle Umgebung des Schienengüterverkehrs in der letzten Meile, kann die hier vorgestellte Technologie genutzt werden, um drahtlose echtzeitfähige Signalübertragungen zu sich bewegenden Objekten sicher und gleichzeitig wirtschaftlich realisiert zu werden.

Im Rahmen der Untersuchungen wurde der Schwerpunkt auf die Nutzung hybrider 5G-Netze gelegt. Das Protokoll für die Videoübertragung lässt sich aber auch mit wenig Aufwand für weitere RATs optimieren. Die Verwendung von bereits vorhandener WiFi-Infrastruktur könnte zukünftig zusätzlich auch für die Videoübertragung genutzt werden, um so die Leistung zu verbessern und wichtige Redundanzen für einen sicheren und zuverlässigen Betrieb sicherzustellen.

Literatur

[1] K. C. P. Oberle, *Webinar: Accelerating rail digitalization with private wireless networks,* 2025.

[2] M. Blumenschein, „Das Fahrerassistenzsystem SAMIRA für rangierende Güterzüge," Ingenieurspiegel, 2020.

[3] S. Kanumuri, P. Cosman, A. Reibman und V. Vaishampayan, Modeling packet-loss visibility in MPEG-2 video, IEEE, 2006.

[4] N. M. D. K. J. Rangarao Orugu, Frequency Reconfigurable Antenna for 5G Applications at n77 and n78 Bands, Springer Nature Link, 2023.

[5] D. M. Steinhart, „Daten und Fakten," Lichtblick Kommunikation, 2021.

[6] J. L. A. C. T. Å. Henrik Basilier, „Applied network slicing scenarios in 5G," 2021.

[7] Extract from the Ericsson Mobility Report, „5G private networks enhance industry productivity," 2024.

[8] A. S. a. B. Kumar, „Effective RAT Selection Approach for 5G Dense Wireless Networks," 2015.

[9] J. L. e. al., „User Association in 5G Networks: A Survey and an Outlook," 2015.

[10] Adam Langley et al. (Google), „The QUIC Transport Protocol: Design and Internet-Scale Deployment," in *SIGCOMM*, 2017.

Session 6
Klaus Mößner, Sam Münchow, Charbel Lahoud

Autoren



Prof. Dr. Mößner, Klaus

Klaus Moessner (Senior Member, IEEE) ist derzeit Professor für Nachrichtentechnik an der Technischen Universität Chemnitz sowie Professor für kognitive Netzwerke am Institute for Communication Systems und am 5G Innovation Centre der University of Surrey. Er war an zahlreichen Projekten im Bereich kognitive Kommunikation, Dienstbereitstellung und Internet der Dinge (IoT) beteiligt. Im Rahmen des Cognitive-Radio-Projekts ORACLE war er verantwortlich für die Arbeiten zu kognitiven Entscheidungsmechanismen. Zudem leitete er die Arbeiten zur Funkbewusstheit im ICT-FP7-Projekt QoSMOS und war Projektleiter des H2020-Projekts Speed5G.



Münchow, Sam

Nach dem Elektrotechnik Studium mit Fachrichtung Nachrichtentechnik an der FH Aachen begann Sam Münchow 2017 mit der Arbeit als Projektingenieur bei Automotiv and Rail Innovation Center GmbH. Dort Entwickelte er Sensorsysteme für den Eisenbahnbereich und Drohnen mit Schwerpunkt auf Lokalisierung. Er ist seit 2024 Head of Research and Development bei der IKADO GmbH mit Sitz in Aachen und leitet unter anderem in dieser Funktion das Projekt SAMIRA 2.0.



M.Sc. Lahoud, Charbel

Wissenschaftlicher Mitarbeiter an der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität Chemnitz. Seit 2020 widmet er sich in seiner Forschung drahtlosen Kommunikationssystemen, darunter 5G, LTE, und WLAN. Seine Arbeit untersucht Anwendungen in Bereichen wie der Fahrzeug-zu-Infrastruktur-Kommunikation (V2X), softwaredefinierten Netzwerken und dem Internet der Dinge (IoT).

IRSA 2025

SESSION 7

RAD-SCHIENE INTERAKTION

WHEEL-RAIL INTERACTION

Session 7
Sören Barteldes, Andreas Bittner, Franz Murawa

Vorstellung einer neuen Radform zur Verminderung thermisch verursachter Zugeigenspannungen und deren möglichen Folgen im Radkranz klotzgebremster Räder

Barteldes, Sören[1], Bittner, Andreas[1] Murawa, Franz[1],

[1]Bochumer Verein Verkehrstechnik, Alleestr. 70, 44793 Bochum

Zusammenfassung

Bei Bremsungen mit Klotzbremsen entstehen hohe Aufheizraten an der Lauffläche von Schienenfahrzeugrädern. Während bei Stoppbremsungen an der Lauffläche oberflächennah und kurzzeitig hohe Temperaturen entstehen, liegen bei Beharrungsbremsen längere Einwirkzeiten vor, die einen Wärmefluss vom Radkranz in Richtung der kälteren Radscheibe entstehen lassen. Der Wärmefluss nimmt in Richtung Radscheibe und Radnabe kontinuierlich ab. Die Folgen sind hohe Thermospannungen und im Fall längerer Einwirkzeiten Plastifizierungen im Bereich der Radscheibe und Zugeigenspannungen im Radkranz, ggf. begleitet von Spurmaßveränderungen in Abhängigkeit von der Radform. Die entstehenden Zugeigenspannungen können in Verbindung mit Oberflächendefekten an der Radlauffläche zu Risswachstum führen, welches nach Möglichkeit auszuschließen ist. Durch die vorgestellte Radneuentwicklung wird das bereits bestehende hohe Sicherheitsniveau, der bei klotzgebremsten Schienenfahrzeugen eingesetzten spannungsarmen Radbauarten, weiter angehoben und damit werden gleichzeitig Reserven geschaffen im Hinblick auf im Betrieb möglicherweise zukünftig auftretende höhere Belastungen, z.B. bei Erhöhungen der Radsatzlast.

Keywords: Güterverkehr, Klotzbremsen, Zugeigenspannungen, Risse, Radentwicklung

1 Einleitung

Der Einsatz von Klotzbremsen ist im Schienengüterverkehr üblich. Bei Bremsungen wird im Kontakt der Bremssohlen mit der Lauffläche thermische Energie durch Reibung erzeugt und in den Radkranz eingetragen. Der Wärmeeintrag in die Lauffläche ist bei

Stoppbremsungen lokal begrenzt. Es entstehen an der Lauffläche oberflächennah hohe Temperaturen und damit verbundene Thermospannungen. Auch bei Beharrungsbremsungen entstehen in Abhängigkeit von Bremsleistung und –dauer hohe Temperaturen, die je nach Verschleißzustand des Radkranzes dort für starke Aufheizungen sorgen. Der Wärmefluss nimmt in Richtung Radscheibe und Radnabe dann kontinuierlich ab. Der Wärmeeintrag ist bei Kompositsohlen im Vergleich mit Graugußsohlen höher, da sie wärmeisolierend wirken: Sie werden wegen ihrer lärmmindernden Eigenschaften eingesetzt.

Zur vorgestellten Radentwicklung werden verschiedene Vorgänge aufgrund Temperatureinwirkung durch langes Bremsen, d.h. bei einer Bremsleistung von 50 kW pro Rad über 45 min entsprechend DIN EN 13979-1 [1] betrachtet.

Eine längere Temperatureinwirkung auf den Radkranz bedeutet eine Ausdehnung des Werkstoffs. Dies wird behindert durch die vergleichsweise kühlere Radscheibe. In Folge dessen und begünstigt durch den thermisch bedingten Abfall mechanischer Eigenschaften u.a. der Festigkeit der eingesetzten Stähle, können Überschreitungen der Dehngrenzen an Radscheibe und Radkranz entstehen. Während des anschließenden Abkühlungsprozesses nimmt, aufgrund der Plastifizierung der überdehnten Radscheibe, der Radkranz nicht mehr seine Ursprungsposition ein. Als Resultat entstehen Zugeigenspannungen im Radkranz und Spurmaßveränderungen in Abhängigkeit von der Radform. Abbildung 1 zeigt die neue Radform und die temperaturabhängige Verformung während und nach der Bremsung im abgenutztem Zustand.



Abbildung 1: Simulation des Spannungs- und Verformungsverhaltens der neu entwickelten Radform während der Bremsung (links) und nach der Abkühlung (rechts) für den abgenutzten Zustand

Weitere Folgen durch lange Bremsvorgänge können Anrissbildungen im Laufflächenbereich sein, die in Verbindung mit den entstandenen Zugeigenspannungen zu Risswachstum führen können und zu Laufflächenschäden führen bzw. im Extremfall zu Querrissen im Radkranz führen können.

Durch die neu entwickelte Radform wird die Ausbildung von Zugeigenspannungen im Radkranz auf ein Minimum reduziert. Auffälligstes Merkmal ist der Versatz zwischen dem Einlauf der Radscheibe in den Radkranz, d.h. unterhalb des Spurkranzes. Hier besteht die größte Steifigkeit des Radkranzes und das geringste Temperaturniveau. Damit sind hier die eigenspannungsfördernden Zugkräfte des Radkranzes während der Bremsung aufgrund der geringsten Aufweitung am niedrigsten. Unzulässige axiale Verformungen beim Bremsen und nach dem Abkühlen treten in Verbindung mit der speziell entwickelten Radscheibenform nicht auf.

Auch bei einer Erhöhung der Radsatzlast sind ausreichend Reserven vorhanden, dass mit der neuen Formgebung der Radscheibenquerschnitt erhöht werden kann, ohne das thermische Verhalten und die Ausprägung von Zugeigenspannungen im Radkranz in kritische Bereiche zu verlagern.

Ein weiterer Vorteil ist das schalltechnische Verhalten des Rades. Es verursacht im Vergleich zu anderen Rädern günstigere Geräuschemissionen, die nach STARDAMP und TWINS gemäß EN 13979-1 [1] berechnet wurden.

2 Auslegung von Vollrädern nach den Regeln der Technik

Das neu entwickelte Low Thermal Stress-Rad (LTS-Rad) entspricht selbstverständlich den Vorgaben nach DIN EN 13979-1 für den Einsatz im europäischen Verkehrsbereich. Diese dienen der schienenverkehrstechnischen Zulassungsbewertung. Die Norm gliedert sich in vier technische Teilaspekte.

- Geometrischer Aspekt (Austauschbarkeit verschiedener Lösungen für den gleichen Anwendungsbereich)
- Thermomechanischer Aspekt (Beherrschung der Verformung des Rades und Absicherung gegenüber Radbruch bei Bremsungen)
- Bruchmechanischer Aspekt (Verhinderung eines Ermüdungsrisses im Radsteg und Verhinderung einer dauerhaften Verformung im Falle einer außergewöhnlichen Belastung
- Akustischer Aspekt (Entwicklung erreicht einen Normschallpegel gleich eines Referenzrades)

Nachfolgend werden die normativen Anforderungen sowie die in der Entwicklungsphase durchgeführten Untersuchungen kurz vorgestellt.

3 Bewertung und Auslegung des thermomechanischen Verhaltens gemäß DIN EN 13979-1

Normativ umfasst der thermomechanische Nachweis für Güterwagenräder zehn Dauerbremsungen auf einem Bremsenprüfstand je für ein neues und ein abgenutztes Rad. Zwischen den Bremsungen wird das Rad auf eine Vergleichstemperatur von unter 50° Celsius abgekühlt und der Eigenspannungs- wie auch der Verformungszustand gemessen. Der Nenndurchmesser bestimmt die Höhe der Nennbremsleistung. Tabelle 1 gibt die durchmesserabhängige Nennleistung der Bremsung wieder [1].

Tabelle 1: Parameter für die Dauerbremsung für 100 % laufflächengebremste Güterwagen

Nenndurchmesser des neuen Rades [mm]	1000>ø> 920	920> ø>840	840> ø>760	760>ø>680
Nennleistung P_b [kW]	50	50	42,5	38
Anwendungsdauer t_b [min]	45			
Mittlere Fahrzeuggeschwindigkeit V_b [km/h]	60			

Der thermomechanische Nachweis gilt als erbracht, wenn

Für das Rad mit neuem Radkranz:

1. Die maximale seitliche Verformung des Radkranzes im Warmzustand während der Bremsung +3/-1mm beträgt
2. Die Eigenspannungen im Radkranz beim Werkstoff ER7 nach Abkühlung:
    a. 200 MPa für den Mittelwert von drei Messungen
    b. 250 MPa für jede Einzelmessung betragen
3. Die maximale seitliche Verformung des Radkranzes bei Abkühlung des Radkranzes von +1,5/-0,5 mm beträgt.

Für das Rad mit Durchmesser nach der letzten Reprofilierung:

1. Die maximale seitliche Verformung des Radkranzes im Warmzustand während der Bremsung +3/-1mm beträgt
2. Die Eigenspannungen im Radkranz beim Werkstoff ER7 nach Abkühlung:
    a. 275 MPa für den Mittelwert von drei Messungen
    b. 300 MPa für jede Einzelmessung betragen
3. Die maximale seitliche Verformung des Radkranzes bei Abkühlung des Radkranzes von +1,5/-0,5 mm beträgt.

Die Eigenspannungen werden mit Hilfe von Ultraschall und die Verformungen im Scanverfahren via Laser- oder Induktionssensoren gemessen.

4 Rechnerische Simulation zum Bremsverhalten von Vollrädern

Bei der Entwicklung neuer Radformen ist die rechnerische Simulation unverzichtbar. Die Verlässlichkeit der Simulation richtet sich nach der Qualität der Eingangsparameter. Für die nachfolgenden Ergebnisse wurden folgende Regelwerke betrachtet (siehe Literaturverzeichnis). Im ORE B169/RP1 sind physikalische Eigenschaften aufgeführt, um eine thermomechanische Simulation durchzuführen [2]. In der UIC B169/RP17 sind Angaben zum Tangentenmodul enthalten, das es erlaubt, die kinematische Verfestigung von duktilen Stählen zu berücksichtigen und nicht allein ideal plastisch (vergl. Spanungs-Dehnungs-Diagramm) zu rechnen [3]. Weitere Werte, auch zu Absicherung der benutzten Quellen, liefern die AAR Regelwerke S660/S669 [4]. Das Stahl-Eisen-Werkstoffblatt SEW310 zeigt temperaturabhängige physikalische Eigenschaften von Stählen auf [5].

4.1 Vergleichsberechnungen zur Beurteilung thermophysikalischer Eigenschaften nach UIC B69/RP17

Für die Vergleichsberechnungen zur Beurteilung thermophysikalischer Eigenschaften wurde ein ORE-Rad als Beispielrad wegen seiner freien Zugänglichkeit herangezogen. Die Berechnungen berücksichtigen eine 50 kW Bremsung über 45 Minuten mit einer anschließenden Abkühlungsphase auf Basis der DIN EN 13979-1 [1] auf unter 50° Celsius. Nach erfolgter Simulation werden die Verformungen am Spurkranz ausgewertet und nach der Abkühlung die mittlere Eigenspannung aus der abgebildeten Fläche im Radkranz berechnet.

4.2 Betrachtung der temperaturabhängigen physikalischen Eigenschaften

Die Berücksichtigung verschiedener temperaturabhängiger Eigenschaften, die für eine thermomechanische Simulation erforderlich sind, hat Auswirkungen auf die Berechnungsergebnisse. Daher wurden umfangreiche Vergleichsrechnungen der thermischen Stoffeigenschaften, wie Wärmeleitfähigkeit, Wärmeausdehnungskoeffizient, spezifische Wärmekapazität und Wärmeübergangskoeffizient durchgeführt. Dabei zeigte sich, dass insbesondere der Wärmeübergangskoeffizient einen erheblichen Einfluss auf die Berechnungsergebnisse hat. Dieser beschreibt die Kühlung des Rades durch den Fahrbetrieb. Hier sind zwischen dem amerikanischen und dem europäischen Regelwerk große Unterschiede festzustellen, die einen erheblichen Einfluss auf die Verformungswerte haben (siehe Abbildung 2). Aufgrund der geringeren Kühlung gemäß AAR erwärmt sich das Rad bis zu 40 % stärker als gemäß UIC.

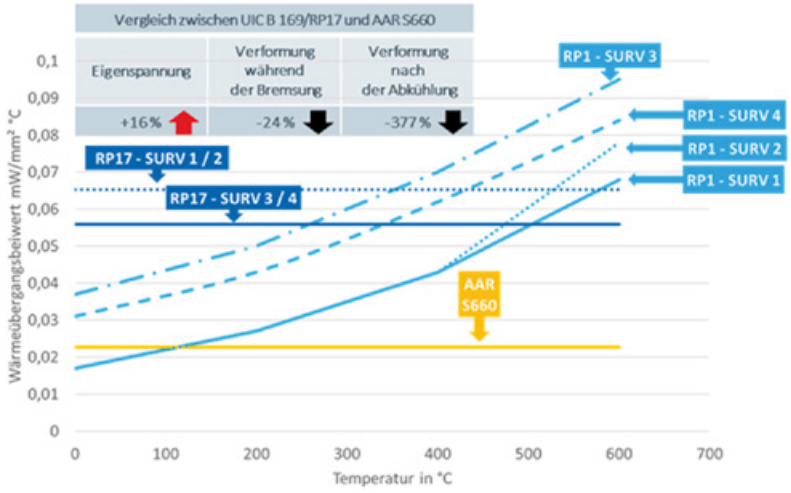


Abbildung 2: Resultierende Verformung während der Bremsung und nach der Abkühlung, sowie Eigenspannungsverhalten basierend auf dem Wärmeübergangsbeiwert verschiedener Literaturquellen und Zonen am Rad

Die weiteren betrachteten Stoffeigenschaften weisen nur eine geringe Beeinflussung des mechanischen Verhaltens auf.

Auf dieser Grundlage wurden die Materialdaten für die Untersuchungen gewählt, welche die besten Übereinstimmungen zu den Ergebnissen der durchgeführten Bremsversuchen liefern.

5 Strukturoptimierung

Die Grundidee für die Form eines idealen Rades für die thermomechanische Simulation wäre ein bereiftes Rad. Radkörper und Radreifen sind derart miteinander verbunden, das temperaturbedingte Geometrieveränderungen des Radreifens während der Aufheiz- und Abkühlphase durch äußere Einflüsse unberücksichtigt bleiben, d.h. auf ein Vollrad übertragen könnte sich der Radkranz in Folge ungehindert und spannungsarm ausdehnen und wieder zusammenziehen. Dieses ideale Rad wurde zur Maxime bei der hier vorgestellten Neuentwicklung.

5.1 Zusammenhang von Eigenspannung und radialer Ausdehnung

Zur Ermittlung des thermomechanischen Verhaltens wurde eine Vielzahl von Simulationen typischer eigenspannungsarmer Radausführungen durchgeführt. In Abbildung 3 ist in der Reihe 1 ein AAR-Rad dargestellt, es verfügt über eine leicht gewellte Radscheibe mit einem Sturz. Es zeigt sowohl während der Bremsung, als auch bei der Abkühlung moderate Eigenspannungen, aber hohe Verformungen, die eine TSI Zulassung verhindern würden. Die Variation des Steges zeigt die Reihe 2. Die geometrische Veränderung beeinflusst die Eigenspannungen und Verformungen. Der schräg gestellte Einlauf des Steges in den Radkranz in Nähe des Spurkranzes führt zu reduzierten Zugeigenspannungen. Dies führt jedoch zu einem starken Abknicken des Radkranzes während der Bremsung und entsprechenden Spurmaßänderungen. Reihe 3 zeigt, dass eine Verschiebung des Einlaufs des Steges direkt unter dem Radkranz bei gleichzeitig radialem anfänglichen Radscheibenverlauf dieses negative Verhalten ausgleichen kann.

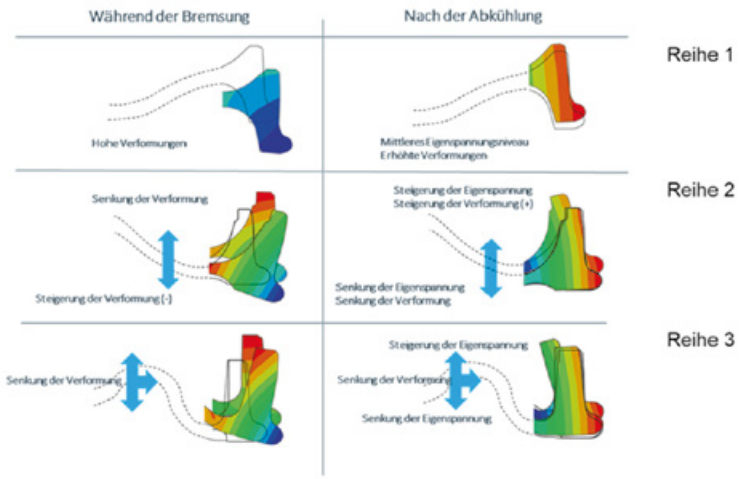


Abbildung 3: Strukturoptimierung an Merkmalen verschiedener Räder

5.2 Low Thermal Stress Rad (LTS-Rad)

Folgende Bauform hat sich daher aus der Strukturoptimierung ergeben. In dieser Radform, vergleiche Abbildung 4, findet sich der Einlauf des Steges direkt unterhalb des Spurkranzes in den Radkranz zur Vermeidung hoher Zugeigenspannungen, wie auch die gewellte Form zur Verhinderung unzulässiger seitlicher Verformungen wieder.

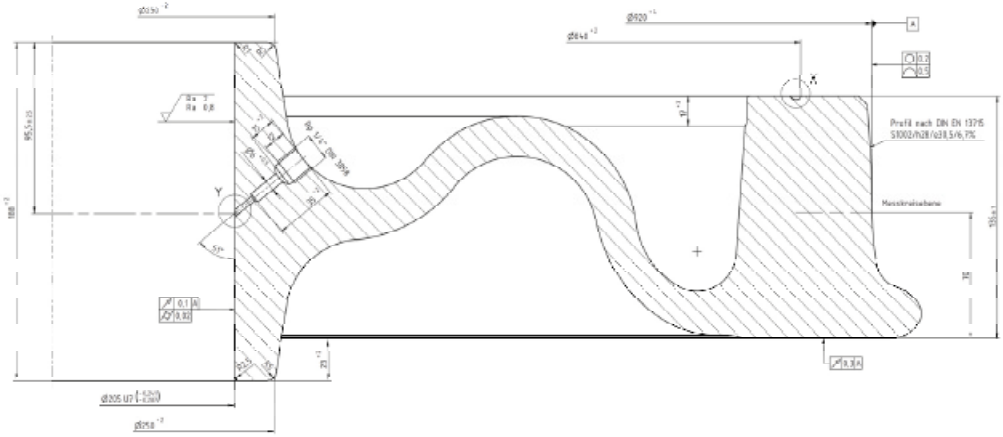


Abbildung 4: Zeichnung Low Thermal Stress Rad

6 Rechnerische Untersuchung des mechanischen Verhaltens

Der normativ geforderte rechnerische Ermüdungsfestigkeitsnachweis für eine Radsatzlast von 25 t stellt eine Selbstverständlichkeit für die Neuentwicklung eines Schienenfahrzeugrades dar. In Abbildung 5 ist der Spannungsverlauf während der Dauerschwingbelastung und der außergewöhnlichen Belastungssituation dargestellt. Beide Prüfungen wurden bestanden.

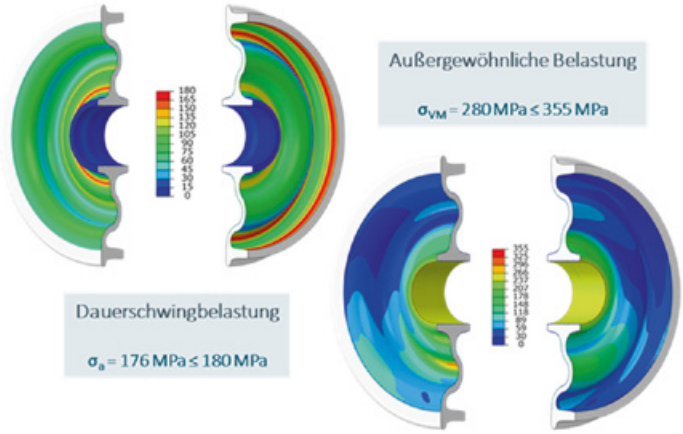


Abbildung 5: Darstellung der Spannungsverläufe für das LTS-Wheel während der Dauerschwingbelastung und der außergewöhnlichen Belastung

7 Nachweis durch Versuche

Im Weiteren werden die mechanischen und thermomechanischen Nachweise im Versuch geführt und die Versuchsergebnisse mit den Simulationswerten verglichen.

7.1 Bremsversuch und Eigenspannungen

Das LTS-Rad wurde im Bremsenprüfstand 10 Dauerbremsungen mit einer Bremsleistung von 50 kW innerhalb 45 Minuten unterzogen. Das verwendete Bremsklotzmaterial waren Kompositsohlen. Mit Hilfe der Ultraschallmethode wurde der Eigenspannungszustand für einen neuen Radzustand, als auch für einen abgenutzten Radzustand jeweils nach der Bremsung gemessen. Die Ultraschallmethode berechnet aus den Laufzeiten der Ultraschallwellen, der Polarisationsrichtung und der akusto-elastischen Materialkonstante den Eigenspannungszustand. Tabelle 2 gibt den Eigenspannungszustand im neuen und abgenutzten Zustand nach dem 10. Bremszyklus an. Im Mittel erreicht das LTS-Rad im Neuzustand 36 MPa < 250 MPa und im abgenutzten Zustand im Mittel 145 MPa < 275 MPa. Die Messstellen wurden alle 120° an der Radaußenseite gewählt.

Im Vergleich dazu zeigt Tabelle 3 die ermittelten Werte der Eigenspannungen im abgenutzten Zustand im Trennschnittverfahren nach DIN EN 13262:2020 als zerstörendes Verfahren. Messmittel in diesem Fall sind aufgeklebte Dehnungsmessstreifen. Beim Trennschnittverfahren wird der Gleichgewichtszustand des Materials gestört und das Material hat die Möglichkeit sich auszudehnen oder zu schrumpfen, je nach Höhe der Eigenspannungen, bis ein neuer Gleichgewichtszustand sich einstellt. Die Verformung wird mit Dehnungsmesstreifen erfasst und kann damit die Höhe der Eigenspannungen bestimmen.

Die ermittelte Höhe der Eigenspannung beträgt an der Lauffläche 122 MPa < 275 MPa, in einer Tiefe von 39 mm unterhalb der Lauffläche liegen noch 39 MPa vor.

Die ermittelten Werte aus den Ultraschallmessungen und im abgenutzten Zustand im Besonderen aus dem Trennschnittverfahren bestätigen die ermittelten Werten aus der Simulation.

Tabelle 2: Eigenspannungsmesswerte nach dem 10. Bremszyklus im Neuzustand und abgenutzten Zustand des LTS-Rad

Neuzustand		
Messstelle	Eigenspannung [MPa]	Zulässiger Wert [MPa]
1	37	250
2	35	
3	35	
Mittelwert σ_r	36	200
Abgenutzter Zustand		
Messstelle	Eigenspannung [MPa]	Zulässiger Wert [MPa]
1	145	300
2	145	
3	145	
Mittelwert σ_r	145	275

Tabelle 3: Eigenspannungsmesswerte nach dem 10. Bremszyklus im abgenutzten Zustand des LTS-Rad ermittelt nach dem Trennschnittverfahren DIN EN 13262:2020

Rad	Bezeichnung	Eigenspannung [MPa]	Tiefe unterhalb der Lauffläche [mm]
225642-002	σ_1	122	0
	σ_a	18	39

7.2 Nachweis der axialen Verformung

Mit Hilfe von Laser- und Induktionssensoren wurde die axiale Verformung nach 10 Dauerbremsungen mit 50 kW über 45 Minuten Bremsleistung mit Kompositsohlen bestimmt. Auch hier konnte der Nachweis erbracht werden. Die ermittelte Abweichung beträgt im Neuzustand 0,2 mm < 1,5 mm / -0,5 mm und 1,4 mm < 1,5 mm / -0,5 mm im abgenutzten Zustand.

7.3 Schwingfestigkeitsprüfung nach DIN EN 13262:2020

Der Nachweis gilt als erbracht, wenn bei der Prüfung zwei Räder eine Radialspannung von ± 240 MPa bei $1 \cdot 10^7$ Lastwechsel anrissfrei ertragen wird. Diese Prüfung wurde bestanden und im Nachgang wurde das zweite Rad ergänzend in der Prüfstufe 265 MPa getestet. Auch diese Prüfstufe wurde anrissfrei ertragen.

7.4 Akustischer Nachweis

Der akustische Nachweis wurde für das LTS-Rad im Vergleich mit einem BA004-Rad erbracht. Sowohl die Berechnungsmethode nach STARDAMP als auch TWINS zeigen leichte akustische Vorteile des LTS-Rades. Zugehörige Werte sind der Tabelle 4 und 5 zu entnehmen [6].

Tabelle 4: Mit STARDAMP berechnete Schallpegel für verschiede Geschwindigkeiten und Zustände

Geschwindig-keit [km/h]	Zustand	Schallpegel absolut dB(A)		
		LTS	BA004	Delta
80	Rad	90,9	92,1	1,2
100	Rad	94,1	95,5	1,4
80	Rad und Schiene	94,8	95,7	0,9
100	Rad und Schiene	97,9	98,9	0,9

Tabelle 5: Mit TWINS berechnete Schallpegel für verschiede Geschwindigkeiten und Zustände

Geschwindig-keit [km/h]	Rad	Schallpegel dB(A)			
		Rad Axial	Rad Radial	Rad Total	Rad und Schiene Total
80 km/h	BA 004	88.3	89.3	91.9	95.7
	LTS	86.3	88.8	90.8	95.2
100 km/h	BA 004	91.7	92.7	95.3	98.7
	LTS	89.6	92.1	94.1	98.2

8 Resümee und Ausblick

Das neu entwickelte LTS-Rad zählt zur Gruppe der spannungsarmen Vollräder bei Klotzbremsbeanspruchungen und zeigt vielfältige Vorteile für den Einsatz im Schienengüterverkehr. Durch seine spezielle Formgebung entsteht ein besonders hoher Widerstand gegenüber der Bildung thermisch induzierter Zugeigenspannungen im Radkranz – auch bei Erreichen des Betriebsgrenzmaßes – in dessen Konsequenz mögliches Risswachstum vermieden wird. Die Optimierung des Wärmeflusses zwischen Radkranz und Radscheibe verhindert zudem, zusammen mit der Geometrie der Radscheibe, zusätzlich unzulässige Verformungen und sorgt für die erforderliche Spurmaßstabilität während und nach den Bremsungen. Damit steht mit dem LTS-Rad ein klotzbremsbares Vollrad zur Verfügung welches den aktuellen und zukünftigen Anforderungen eines leistungsfähigen Schienengüterverkehrs entspricht.

Bei der Diskussion Klotzbremse versus Scheibenbremse im Schienengüterverkehr darf darüber hinaus nicht vergessen werden, dass die Verwendung von Klotzbremsen in der Regel zu einer signifikanten Gewichtseinsparung und Reduktion der ungefederten Massen führt, indem auf Bremsscheiben und Bremsanlagen verzichtet werden kann. Entsprechend kann so die Zuladung gesteigert werden.

Unter Kosten- und Gewichtsaspekten könnte es aus unserer Sicht für bestimmte Bereiche des Personenverkehrs zukünftig ebenfalls wieder interessant werden, über den Einsatz von Klotzbremsen in Verbindung mit dem LTS-Rad nachzudenken.

Literatur

[1] DIN EN 13979-1:2024-04 Bahnanwendungen - Radsätze und Drehgestelle – Vollräder - Technische Zulassungsverfahren - Teil 1: Geschmiedete und gewalzte Räder; Deutsche Fassung EN 13979-1:2023

[2] ORE Report B 169/RP 1. Selection of parameters for testing thermal limits of wheels and brake blocks. ORE, Utrecht, September 198

[3] D_B169/RP17 Standardisierung der Radsätze. Definition der Parameter zur thermomechanischen Berechnung der Vollräder. Korrelation der Berechnungen /Versuche. UIC

[4] Association of American Railroads MANUAL OF STANDARDS AND RECOMMENDED PRACTICES SECTION A, PART I SAFETY AND OPERATIONS The Association of American Railroads, 2014

[5] SEW 310: Physikalische Eigenschaften von Stählen, Verlag Stahleisen GmbH, Düsseldorf, 1992

[6] Fehndrich, M.: The Acoustics of the Low Thermal Stress Railway Freight Wheel, DAGA 2025, Kopenhagen

Session 7
Sören Barteldes, Andreas Bittner, Franz Murawa

Autoren

	**Barteldes** Seit Januar 2025 beim Bochumer Verein Verkehrstechnik als Innovationsmanager tätig, zuvor 10 Jahre Erfahrung im Bereich Messtechnologien für die Produktion im Automotive Bereich.
	**Bittner** Über 30 Jahre Berechnungsingenieur beim BVV für Radsatzkomponenten.
	**Murawa** Nach Abschluss der aktiven Berufslaufbahn, zuletzt als Leiter Engineering als Consultant beim BVV tätig

Session 7
Jakob Moder, Erik Gerstl

Kontaktlose Messung von Rad-Schiene Kräften auf Schienenfahrzeugen

Moder, Jakob[1], Gerstl, Erik[1]

[1] Zyglox GmbH

Zusammenfassung

Im vorliegenden Beitrag wird eine neue Messmethode vorgestellt, welche Kräfte in der Rad-Schiene-Schnittstelle kontaktlos messen kann. Die Sensorik wird dabei fix am Drehgestell verbaut, wobei sich die Messelemente des Sensors in einem Abstand von 2 mm zum Radsteg befinden. Durch die aktive Einbringung und Messung von Magnetfeldern kann basierend auf dem magnetostriktiven Effekt, die Aufstands-, Lateral-, und Traktionskraft gemessen werden.

Die Validierung des Prototyps erfolgte auf der elektrischen Lokomotive Siemens ES64U2 im Rahmen von Werksfahrten, wobei eine Plausibilisierung der Messresultate durch einen Vergleich mit aus der Literatur bekannten Werten durchgeführt wurde.

Keywords: Sensor; Messtechnik, Messradsatz, Rad-Schiene-Kontakt

1 Einleitung

Die Rad-Schiene-Schnittstelle stellt einen zentralen Bestandteil des Systems Eisenbahn dar. Maßgebliche Aspekte wie die **Entgleisungssicherheit** und die **Fahrzeugdynamik**, welche die Betriebsstabilität und den Fahrkomfort bestimmen, hängen direkt von den Kräften und Relativbewegungen im Rad-Schiene-Kontakt ab. Darüber hinaus sind diese Kräfte die primäre Ursache für **Verschleiß und materielle Schädigung** von Schienen und Rädern, was die Lebenszykluskosten und Instandhaltungsintervalle des Gesamtsystems entscheidend beeinflusst. Die etablierte Methode zur Messung dieser Kräfte, der Einsatz von **Messradsätzen mit Dehnmessstreifen (DMS)**, ist jedoch mit erheblichem **technischem und wirtschaftlichem Aufwand** verbunden. Dies umfasst nicht nur die aufwendige Applikation und Kalibrierung der Sensoren, sondern auch die

komplexe Datenübertragung von rotierenden Bauteilen und eine hohe Anfälligkeit für die rauen Umgebungsbedingungen im Bahnbetrieb. Um diese messtechnische Hürde zu überwinden, stellen wir in diesem Beitrag eine **neuartige Technologie** vor, welche die Kräfte im Rad-Schiene-Kontakt **durch ein vollständig kontaktloses Messsystem** ermitteln kann.

2 Technologie

Die vorgestellte Messtechnologie nutzt den **magnetostriktiven Effekt**, der den fundamentalen Zusammenhang zwischen der Magnetisierung und der mechanischen Dehnung in ferromagnetischen Werkstoffen beschreibt (Linnemann, 2007). Konkret wird der **inverse magnetostriktive Effekt** (auch Villari-Effekt genannt) herangezogen, bei dem eine mechanische Spannung eine Änderung des externen Magnetfelds eines Bauteils hervorruft. Das Bauteil selbst wird somit zum eigentlichen Sensorelement.

In seiner passiven Anwendung ist dieser Effekt jedoch für präzise Messungen ungeeignet. Die resultierenden Magnetfeldänderungen sind sehr schwach und, was noch gravierender ist, von einer starken **Hysterese** geprägt. Das bedeutet, dass das Messergebnis nicht nur von der aktuell wirkenden Kraft abhängt, sondern auch von der magnetischen Vorgeschichte des Materials, was exakte und reproduzierbare Messungen schwierig macht.

Die hier vorgestellte Technologie löst dieses grundlegende Problem durch ein **aktives Messverfahren**. Wie in Abbildung 1 (Mitte) schematisch dargestellt, wird simultan eine **kontrollierte magnetische Anregung** in das Material eingebracht und dessen Antwort gemessen. Diese Anregung überlagert die magnetische Vorgeschichte und zwingt die magnetischen Bezirke in einen definierten, ausgerichteten Zustand. Dadurch wird, wie das Diagramm in Abbildung 1 (Rechts) zeigt, ein **linearer und hysteresefreier Zusammenhang** zwischen der mechanischen Dehnung (und somit der Kraft) und dem gemessenen Magnetfeldsignal hergestellt (Moder & Gerstl, 2024).

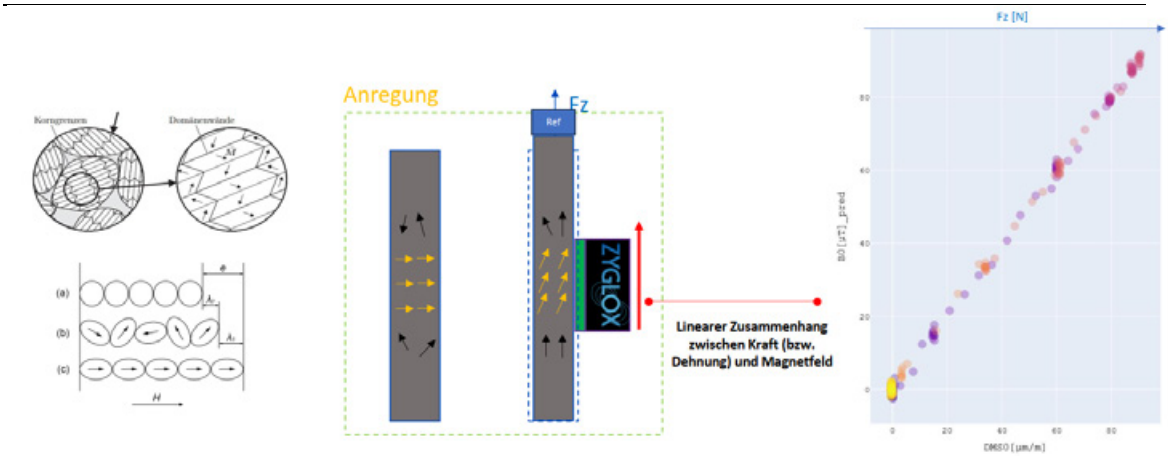


Abbildung 1: Technische Funktionsweise der Messmethode. Links: Zufällige Orientierung der magnetischen Bezirke in den Körnern von ferromagnetischen Werkstoffen (Linnemann, 2007) Mitte: Anregung der magnetischen Bezirke um lokal eine gleichförmige Ausrichtung zu erreichen. Rechts: Dadurch kann ein linearer, hysteresefreier Zusammenhang zwischen mechanischer Dehnung und gemessenen Magnetfeld erreicht werden (Moder & Gerstl, 2024).

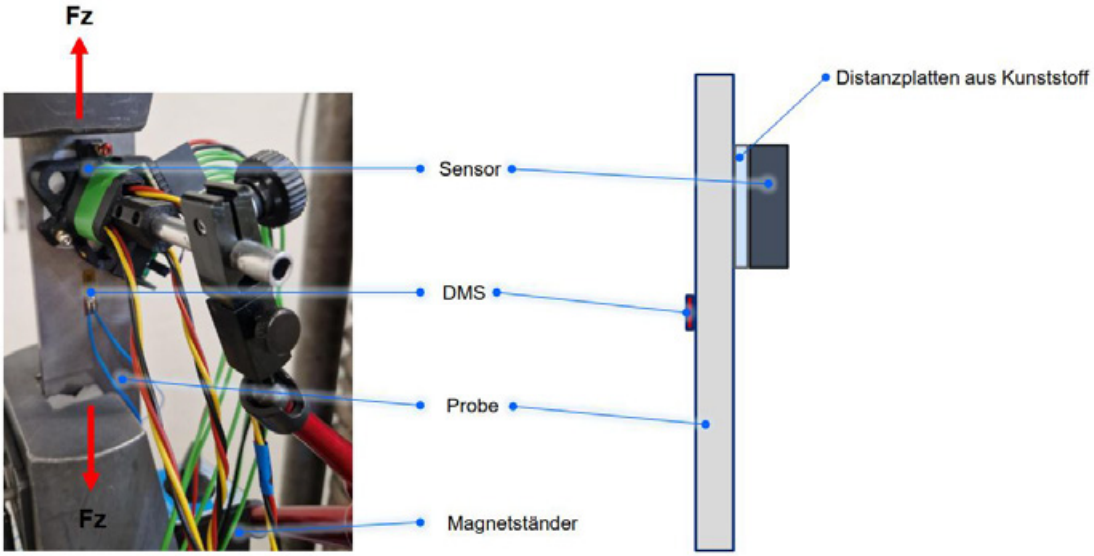


Abbildung 2: Versuchsaufbau für den Vergleich des Zyglox Sensor-Systems mit DMS an Flachstahlproben. Die mit DMS instrumentierte Probe wurde wurde auf Zug belastet, wobei der Zyglox Sensor mithilfe eine Magnetständers und einer Distanzplatte in einem Abstand von 1 mm zur Probenoberfläche positioniert wurde.

Um die grundlegende funktionsweise des Sensors zu testen, erfolgten wie in Abbildung 2 dargestellt, Versuche mit einer Zugprüfmaschine an Flachstahlproben mit Abmessungen 120x40x10 mm. Es wurden verschiedene, insbesondere für die

Eisenbahnindustrie, relevante Stähle getestet. Zum Vergleich wurde die Probendehnung mittels Dehnmessstreifen (DMS) erfasst.

Aus Abbildung 3 ist ersichtlich, dass eine ausgezeichnete Übereinstimmung zwischen beiden Messsystem besteht. Aus der Abweichungsverteilung geht hervor, dass diese in der überwiegenden Mehrheit der Messpunkte kleiner als 2 µm/m ist, was bezogen auf den Messbereich von 90 µm/m einer Genauigkeit von 2.22% entspricht.

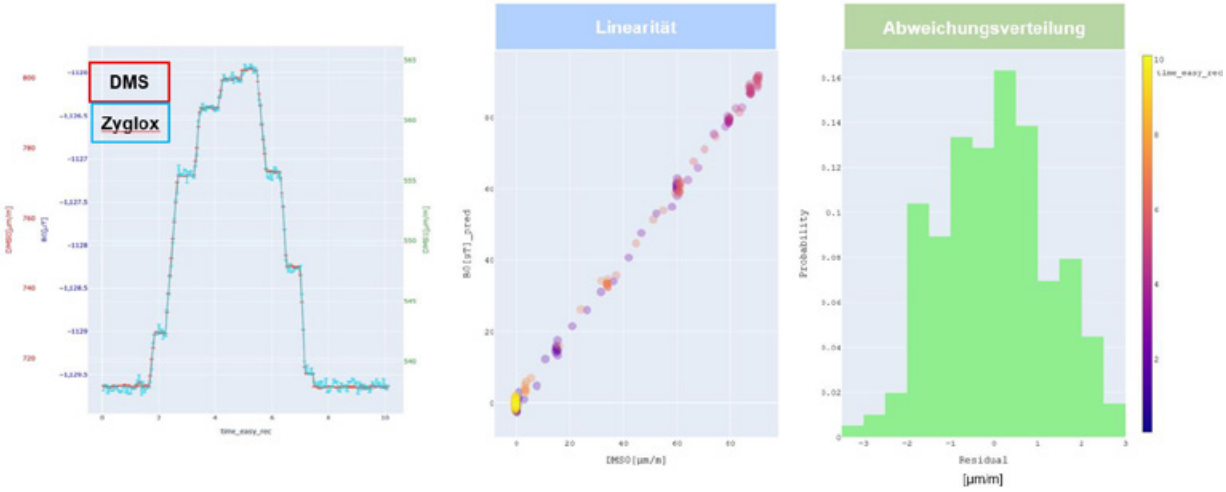


Abbildung 3: Vergleich unserer Sensorik mit herkömmlichen DMS: Links Vergleich der ermittelten Dehnung im Zeitbereich. Mitte: Darstellung der Linearität. Rechts Abweichungsverteilung in µm/m zwischen den Beiden Messsystemen. Probe aus 42CrMo4.

3 Aufbau Rad-Schiene Sensor

Der Aufbau der Sensorik ist in Abbildung 2 dargestellt. Das Gehäuse besteht aus Aluminium und die Verbindung mit dem Lagergehäuse ist aus Stahl ausgeführt, um eine möglichst hohe Steifigkeit zu gewährleisten. Neben Sensor- und Magnetisierungseinheiten befinden sich im Gehäuse auch ein Mikrocontroller, der

Kompensations- und Berechnungsaufgaben übernimmt, sowie Schnittstellen für eine Datenübertragung mittels USB oder CAN.

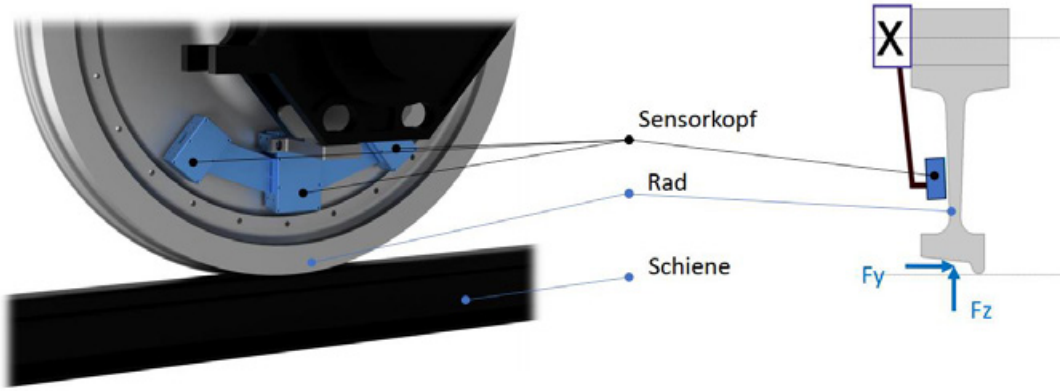


Abbildung 4: Einbausituation des Sensor-Systems. Links: Die Befestigung erfolgt direkt am Radsatzlagergehäuse. Rechts: Der minimale Abstand zwischen Radsteg und dem Sensor beträgt lokal 2 mm.

In Tabelle 1 sind die wichtigsten technischen Eigenschaften aufgelistet.

Tabelle 1: Eigenschaften des Messsystems

Eigenschaft	**Wert**
Messgrößen	Fz (Q), Fy (Y), Radaufstandspunkt (Y), Planlauf, Temperatur, Beschleunigung (x,y,z), Geschwindigkeit
Temperatur	-30°C bis 85°C
Genauigkeit	Kraft: <3.5% Accel. <2%
Bandbreite	>100 Hz (1 kHz Burstmode)
Energiebedarf	<5 W
Radgeometrie	Grundsätzlich keine Einschränkungen
Verwendungsdauer	Langzeit

Für die im vorliegenden Artikel dargestellten Resultate erfolgte eine Montage des Sensorsystems an einer Siemens ES64U2 Lokomotive. Hierfür wurde ein Klemmsystem entwickelt, das ohne weitere Bearbeitungsschritte eine Befestigung des Sensors am

Radsatzlagergehäuse ermöglicht. Die Abfolge der Montage ist in Abbildung 5 dargestellt. Das Gesamtgewicht von Sensorik und Halterungen beträgt 4,5 kg.

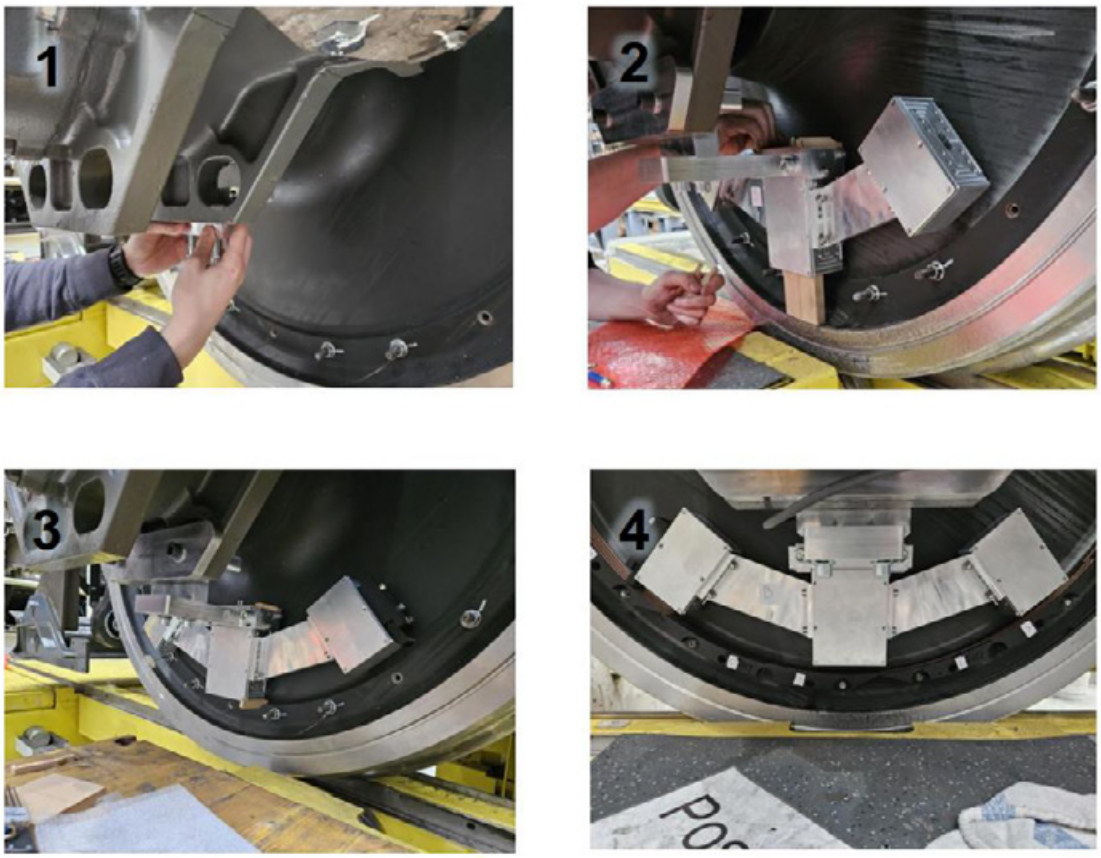


Abbildung 5: In Schritt 1 erfolgt die Ausrichtung einer Platte an den Aussparungen der Unterseite des Lagergehäuses. Diese wird dann in Schritt 2 mittels eines Gegenstücks mit dem Lagergehäuse verklemmt, wobei auch das Sensorgehäuse direkt verschraubt wird. In den Schritten 3 und 4 erfolgt dann eine geometrische Feineinrichtung des Sensors, wobei spezielle Distanzlehren verwendet werden, um eine korrekte Positionierung relativ zum Rad sicherzustellen.

4 Kalibrierung

Die Kalibrierung der Radaufstandskraft erfolgte auf einem Drehgestelldruckprüfstand des Herstellers Nencki Railway Technology. Dabei wurde durch einen Druckzylinder das Drehgestell mit Radlasten zwischen 45 kN (Eigenewicht des Drehgestells unter Annahme symmetrischer Lastverteilung) und 120 kN (dynamische Last) belastet. Abbildung 6 zeigt den Vergleich der Kraftmessung am Prüfstand mit dem Zyglox Sensorsystem.

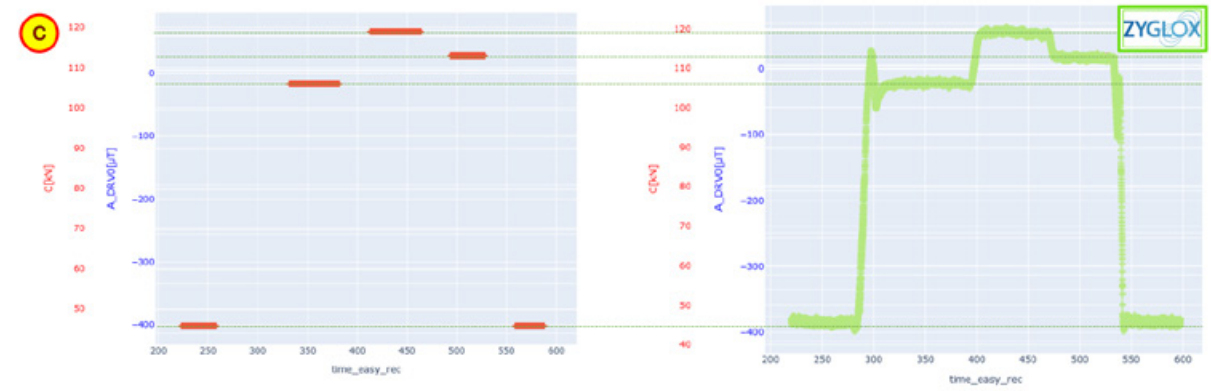


Abbildung 6: Vergleich Referenzkraft mit Zyglox. Links: Referenzkraftmessung des Prüfstands, wobei der Kraftwert am Prüfstand abgelesen und während der Haltedauer als konstant angenommen wurde. Rechts: Zyglox Sensorsystem Aufstandskraft. Hier konnte im Zeitbereich während der gesamten Messdauer gemessen werden. Die Genauigkeit bezogen auf die abgelesenen Referenzkraftwerte beträgt 2%.

Für die Kalibrierung der Querkräfte wurde die in Abbildung 7 darstellte Vorrichtung entwickelt, welche über eine Positionierung an der Radinnenseite mittels Gewindestange, Mutter und Druckstück die Einleitung von Querkräften in beide Räder ermöglicht. Durch eine Instrumentierung der Vorrichtung mit Dehnmessstreifen kann eine Referenzkraft im Zeitbereich ermittelt werden.

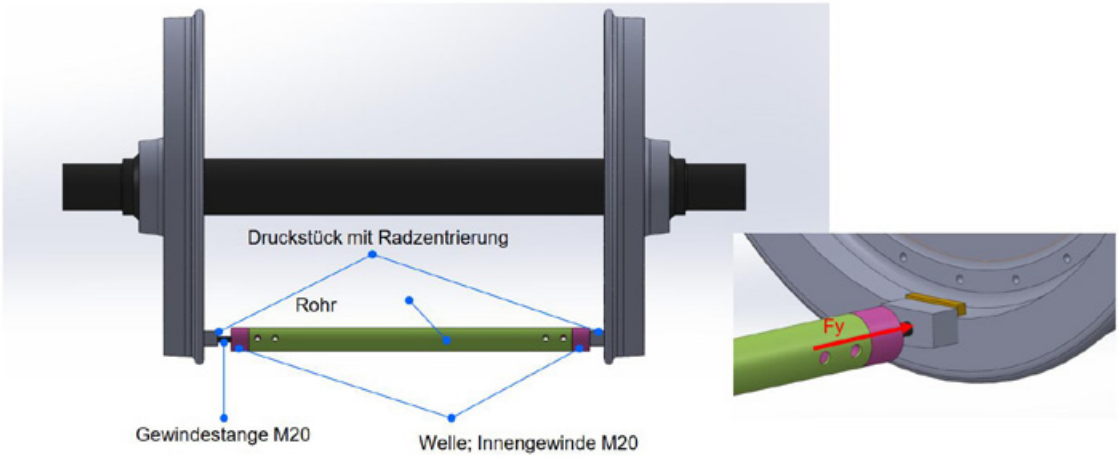


Abbildung 7: Vorrichtung für das symmetrische Einleiten von Querkräften in einen Radsatz. Über eine M20-Mutter wird mittels Gewindestange und Druckstück die Last eingeleitet. Dehnmesstreifen an der Vorrichtung messen die entstehende Dehnung und werden genutzt um die Querkräfte zurückzurechnen.

Ein Vergleich der DMS-Messung und des Zyglox Sensorsystems ist in Abbildung 8 dargestellt.



Abbildung 8: Vergleich der Querkräfte gemessen durch DMS und Zyglox-Sensor. Links: Gegenüberstellung im Zeitbereich DMS (rot) – Zyglox (grün). Rechts: Bilder während der Messung bzw. Lasteinleitung. Die Stufen in den Kurven entsprechen den Rastphasen während dem Anziehen der Mutter.

5 Messungen

Um die Funktionsweise des **Sensorsystems** im Fahrbetrieb zu erproben, erfolgten Testfahrten auf einem Werksgelände unter moderaten Geschwindigkeiten von **weniger** als 30 km/h. Insbesondere wurden die Auswirkungen realer Betriebsbedingungen, wie z.B. Erdungsströme, Überfahrten von Gleisschaltmitteln oder das Anheben und Absenken des Pantographen, untersucht. Beispielhaft ist in Abbildung 9 eine etwa 2 Minuten lange Fahrt dargestellt. Dabei wurden zwei Bögen sowie mehrere Weichen und Bahnübergänge durchfahren. Dargestellt ist die Querkraft (Y-Richtung) des bogenäußeren, vorlaufenden Rades. Die Maximalwerte im Bogen (etwa 250 m Radius) von 50 kN entsprechen aus der Literatur bekannten Werten (Schmid, et al., 2022).

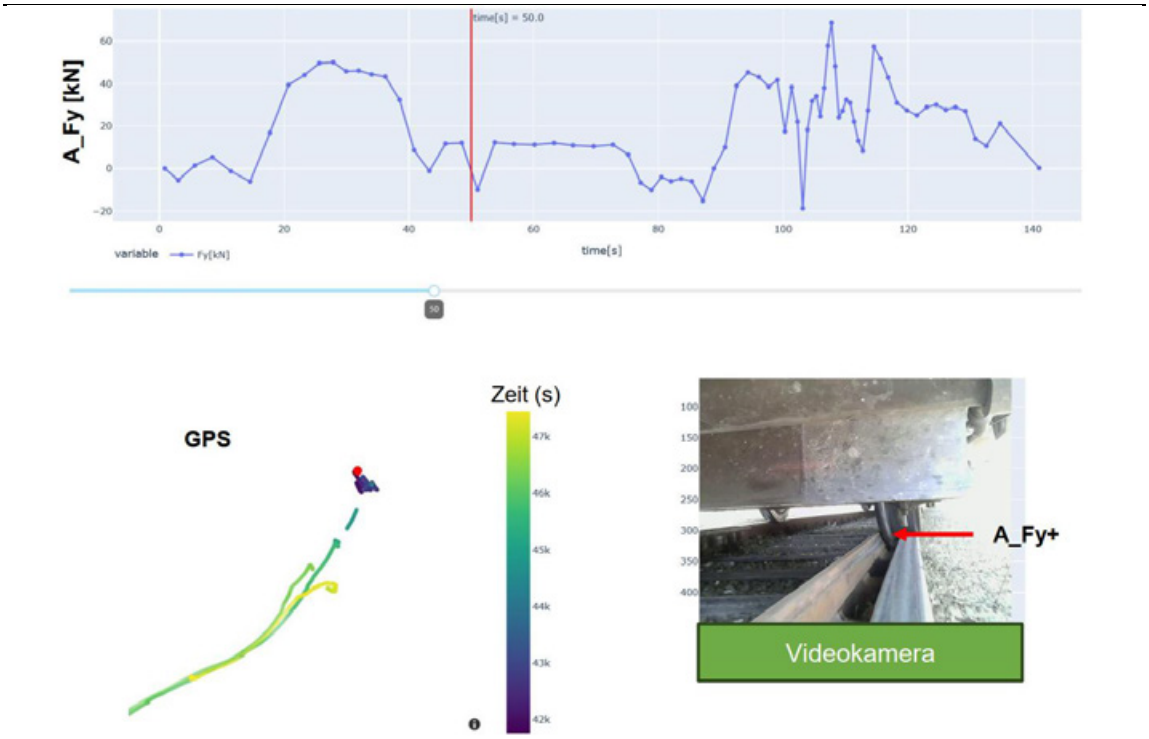


Abbildung 9: Darstellung der Y-Kraft des bogenäußeren Rades beim vorlaufenden Drehgestell. Nach 50 s (rote Linie) schlägt die Y-Kraft abrupt in die negative Richtung aus. Eine Videokamera, die ebenfalls Teil des Messaufbaues war, zeigt, dass gerade zu diesem Zeitpunkt ein Radlenker durchfahren wird.

6 Ausblick

Durch das vorgestellte Messsystem kann die Messung von Rad-Schiene-Kräften mit geringem Aufwand durchgeführt werden. Dadurch ergeben sich neue Möglichkeiten in den Bereichen vorausschauende Wartung, Streckenüberwachung oder Lastkollektivmessung. Die gewonnenen Daten ermöglichen eine präzisere Analyse des Betriebsverhaltens von Schienenfahrzeugen und der Infrastruktur, was zu einer erhöhten Sicherheit und Effizienz im Bahnbetrieb führen kann.

Um das Messsystem vom aktuellen Prototypenstatus in ein Produkt überzuführen sind umfassende Validierungen geplant. Diese werden einerseits einen systematischen Vergleich des Zyglox-Systems mit DMS-basierten Technologien (strecken- und fahrzeugseitig) beinhalten und andererseits aus umfangreichen Betriebsmessungen, unter Randbedingungen (Geschwindigkeit, Zugkraft, etc.), welche für den Regelbetrieb von Schienenfahrzeugen repräsentativ sind, bestehen. Zusätzlich wird aktuell ein weiteres

Sensorsystem, basierend auf derselben Technologie, entwickelt, welches die mechanische Belastung der Radsatzwelle (unter anderem Biege- und Torsionsspannungen) kontaktlos erfassen kann.

Schlussteil

Literatur

Linnemann, K., 2007. *Magnetostriktive und piezoelektrische Materialien - Konstitutive Modellierung und Finite-Elemente-Formulierungen.* s.l.:Universität Karlsruhe.

Moder, J. & Gerstl, E., 2024. *Vorrichtung zur Messung einer mechanischen Belastung sowie Verfahren hierzu.* WPT, Patentnr. 060371.

Schmid, R., Rosenberger, M., Obadic, D. & Breuer, W., 2022. Des einen Ferund des anderen Leid - die Spurkranzschmierung aus Sicht von Fahrzeug und Fahrweg. *Eisenbahntechnische Rundschau.*

Autoren



DI Dr. Moder, Jakob

Studierte Maschinenbau und verfasste anschließend eine Dissertation in den Bereichen Tribologie (Verschleiß & Reibung) und Messtechnik an der Montanuniversität Leoben. Danach Tätigkeit als Entwicklungsingenieur und Projektmanager bei österreichischen Industriebetrieben (AVL, Siemens Mobility). Im Oktober 2023 gründete er die Zyglox GmbH und ist dort für Anwendungsentwicklung, Software, Simulation und Business Development zuständig.



DI Gerstl, Erik

Er studierte an der TU Graz Verfahrenstechnik und schrieb dort seine Masterarbeit über Messtechnik und Produktentwicklung. Ab 2019 war er Projektmanager bei Piezocryst, mit dem Schwerpunkt auf Sensor-Komplettsysteme im Bereich der Bremsmoment-Messung. Nebenbei führt er eine erfolgreiche Gin-Destillerie als gewerberechtlicher Geschäftsführer. Er ist Mitbegründer der Zyglox GmbH (2023) und derzeit für Maschinenbau, Elektronikdesign, Fertigung und interne Operations verantwortlich

IRSA 2025

SESSION 8

FAHRPLANANALYSE & -VERBESSERUNG

TIMETABLE ANALYSIS AND IMPROVEMENT

Session 8
Thomas Graffagnino

Automated Delay Detection in Railway Networks

First Results from SBB's EFA System

Graffagnino, Thomas[1]

[1]SBB, Schweizerische Bundesbahnen

Summary

Operating a railway system with numerous trains on a constrained network results in high traffic density. To improve punctuality, railway planners strive to create timetables that are both stable and robust. Stability ensures trains can operate continuously on their scheduled paths without unforeseen stops, while robustness refers to the system's capacity to withstand delays and limit their propagation across the network. In this paper we will present the automatic system to identify delay propagation and show first results of data visualization while explaining related challenges.

Keywords: Timetable; Stability; Punctuality; Robustness; Data Science

1 Introduction

The Swiss Federal Railways (SBB) introduced the EFA system (from the German "Erfassung Fahrplanabweichung") in May 2021 [1]. This system employs algorithms to automatically identify chains of delays among trains, producing approximately 150 Mb of data, encompassing 3000 points and 1500 links per day. The system involves all trains on the network and the connections between them. Delays are categorized into two types: primary delays, which affect the first train due to issues like a customer holding a door, weather conditions, or technical problems; and secondary delays, also known as knock-on delays [2], which occur when another system element causes a blockage, such as waiting for a connection, uncleared platforms, or single-track occupation in the opposite direction. While primary delays can be mitigated in various ways, they cannot be entirely eliminated. However, adjusting timetables to reduce dependencies can prevent a single delay from affecting the entire system. SBB has been collecting data on delay chains since

2021. Time deviations are detected from operational data and analyzed to classify delays into the two categories. For each knock-on delay, responsible trains are identified and listed. This results in a graph where vertices represent individual train delay events at measurement points, and edges represent causal links between delays. Hence, a key question arises:

- how can the most "critical" points of the system be identified based on this data?



Figure 1: Example of data visualization

Knowing the most "critical" delay propagation points on the Swiss railway system will aid in their elimination. Providing this information to timetable planners will assist them in revising track usage and adjusting departure and arrival times, thereby reducing delay propagation across the network. The concept of a "critical" point is further discussed in point 5.

2 Related Work

Our research on identifying delay dependencies within trains began in 2009 [3]. The initial goal was to detect pairs of trains at stations with dependencies, using actual delays as the sole source of information. To achieve this, a substantial amount of data was required to identify possible interactions between trains at stations. This research was a preliminary step toward understanding all delay chain reactions within our network.

Various studies have been conducted to identify knock-on delays based on traffic data; recent examples include: [4], [5] or [6]. Researches conducted in Japan [7], [8] and [9] have explored metro-type systems on double-track lines, yielding inspiring results and visualizations. However, these findings cannot be directly applied to the complex infrastructure and traffic flows characteristic of European systems. Since SBB has implemented a system that automatically identifies delay chain reactions as detailed here [1], we now possess extensive data on delay chaining, prompting us to explore a new topic: how to define "critical" delay propagations and identify significant delay chains to make them accessible to internal operations and timetable specialists. This specific issue of filtering, sorting, and visualizing useful data for timetable improvement based on collected knock-on delays for large network is, to our knowledge, uncommon in the literature.

3 Data Source and Preprocessing

3.1 Delay detection

On the Swiss railway network, trains runs are planned on the network's almost 1,500 operating points. Operating points are mostly stations where For any train on a given day, its route is defined by series of operating points $p_1, p_2, \dots, p_n$. Its planned path is defined by the scheduled times $s_1, s_2, \dots, s_n$ at each operating point along its route. From operational data we then get real passage times $r_1, r_2, \dots, r_n$. For any point on the route, delay d_i is then defined by $d_i = r_i - s_i$. The difference between two successive delay is called delta-delay, $\delta_i = d_i - d_{i-1}$. EFA system then works with two distinct thresholds based on these measures. One is absolute: given a train path, the first operating point p_i with a delay d_i greater than 179 seconds is reported. The second is relative, any increase of delta-delay δ_i greater than 179 seconds is also referenced. For example, train X leaves 19 seconds ahead of schedule from operating point “BN” and then arrives 162 seconds late at next operating point “BNWD”, the system identifies a δ of 182 seconds and therefore references train X at “BNWD” as delay point and will look for the reason.

3.2 Chaining

In a second phase, the EFA system searches for potential causes for each individual recorded delay. It introduces the concepts of *source train* and *impact train*, and classifies delay causes into five categories:

- **Initial cause:** The delay is a primary delay not caused by prior events.
- **Shared resource usage:** A preceding train uses the same physical infrastructure within a close time window.
- **Connection:** The delayed train may be waiting for a late arriving connecting train.
- **Same rolling stock:** The rolling stock is scheduled to be reused in the opposite direction, and its availability is affected.
- **Imported delay:** After an initial event, a delayed train may trigger a cascade of subsequent delays.

The EFA system compiles a list of possible causes based on its knowledge of the physical infrastructure and information from ALEA, a system where train dispatchers record technical disruptions and timetable changes. EFA uses this data to assign probabilities to each potential cause, weighted by spatial and temporal proximity. Connections exceeding a defined threshold are considered relevant, allowing multiple source trains to be linked to a single impacted train.

3.3 Raw data for analysis

The data used for this analysis, generated by the EFA system, consists of two tables: a Chaining file and a Deviation file. The Deviation file contains all delays detected by EFA. The Chaining file captures the source-target relationships, representing the chaining between two delays from the Deviation file.

In practice, we select whether to download weekdays or weekends within the specified time window. This distinction is particularly important as the timetable and operational situations differs significantly on weekends.

The processed raw data then appears as follows:

OP	LATITUDE	LONGITUDE	TRAIN NUMBER	TRAIN FAMILY	EVENT_TYPE	OPERATIONAL_DAY	TRAIN TYPE	OP_FAMILY_EVENT	OP_ZN_EVENT	DEVIATION_ID
OL	47.35193157.90770536		707	701	departure	19.03.2025	IC	OL~701~departure	OL~707~departure	cAmnjx4TgGnHltGgONGgONF1-F19F7z7-zza7z7-zzamzm...
OL	47.35193157.90770536		608	602	arrival	19.03.2025	IC	OL~602~arrival	OL~608~arrival	cAmnkbOkxGnHltGgONGgON5R5-F1EzA-zzmEzA-zzmmzmu...

Figure 2: Deviation file with two example rows

IMPACT_DEVIATION_ID	REASON_DEVIATION_ID	CHAINING_TYPE	WINDOW_ID	TARGET	TARGET_TN	SOURCE	SOURCE_TN	OPERATIONAL_DAY
cAmnkbOkxGnHltGgONGgON5R5-F1EzA-zzmEzA-zzmmzmu...	cAmnjx4TgGnHltGgONGgONF1-F19F7z7-zza7z7-zzamzm...	Shared occupancy	20250322_0200	OL~602~arrival	OL~608~arrival	OL~701~departure	OL~707~departure	19.03.2025

Figure 3: Chaining file with example row

The columns represent the following information:

- **OP**: Operating Point
- **LATITUDE/LONGITUDE**: Coordinates of the Operating Point
- **EVENT_TYPE**: Indicates whether the delay was measured at arrival or departure at the operating point
- **OPERATIONAL_DAY**: Date of the event
- **BP_FAMILY/TN_EVENT**: A triplet describing the event, where *FAMILY* refers to the train family and *TN* to the train number
- **DEVIATION_ID**: A unique identifier for the specific delay event
- **IMPACT_DEVIATION_ID / REASON_DEVIATION_ID**: The *DEVIATION_ID* of the target (IMPACT) and source (REASON) train, respectively
- **CHAINING_TYPE**: The cause of the delay propagation, corresponding to one of the five defined chaining categories
- **WINDOW_ID**: Specifies the time window used for calculating the delay propagation
- **TARGET / TARGET_TN**: The triplet *BP_FAMILY_EVENT* or *BP_TN_EVENT* for the target train
- **SOURCE / SOURCE_TN**: The triplet *BP_FAMILY_EVENT* or *BP_TN_EVENT* for the source train

4 Graph Modeling and Weighting Algorithm

The processed data can be interpreted as a directed graph, where individual delay events (entries from the Deviation file) correspond to vertices, and causal relationships between these events (entries from the Chaining file) correspond to directed edges. In graph theory, graphs may possess various properties – for instance, edges can be assigned weights to indicate the strength or influence of a connection.

In this section, we first introduce key concepts from graph theory that are relevant to our analysis. We then describe how edge weights are assigned and explain the weighting algorithm used to quantify the influence between connected events.

4.1 Graph Definition

We follow the classical terminology from graph theory as introduced by Harary in 1969 [10]. For detailed explanations of the following concepts, we refer the reader to his foundational text.

Definition. A *Graph* G is a pair of two sets $G = (V, E)$ where $V = \{v_1, v_2, \dots, v_n\}$ are called the vertices and $E = \{e = (x, y) \,|x, y \in V \; and \; x \neq y\}$ the edges. We usually write $e = xy$ for a lighter notation. In this paper we consider *directed graphs* where the edges are ordered pairs (i.e., $(x, y) \neq (y, x)$), in opposition with undirected graphs.

Example. Let $G = (V, E)$, where $V = \{A, B, C, D, E, F, G\}$ and $E = \{AB, BC, BE, BD, CE, DE, EF, GD\}$. The graph then looks like:

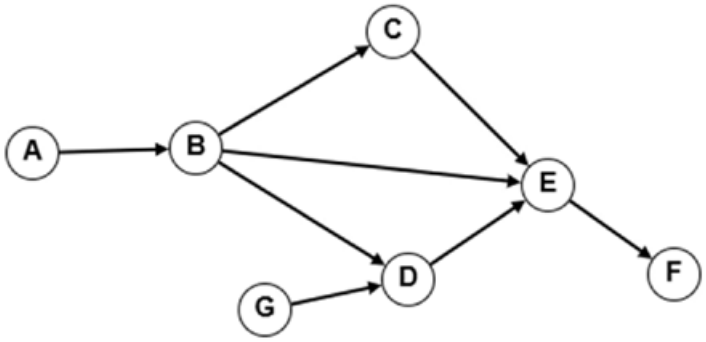


Figure 4: Example of graph

Each vertex $v \in V$ corresponds to a single delay event (e.g., arrival or departure) for a specific train at a defined operational point, with a recorded delay or delta-delay that exceeds the predefined thresholds. The name of the vertices is the triple OP_TN_EVENT. The directed edges represent the causal links between these events as determined by the EFA system. An example is shown below:

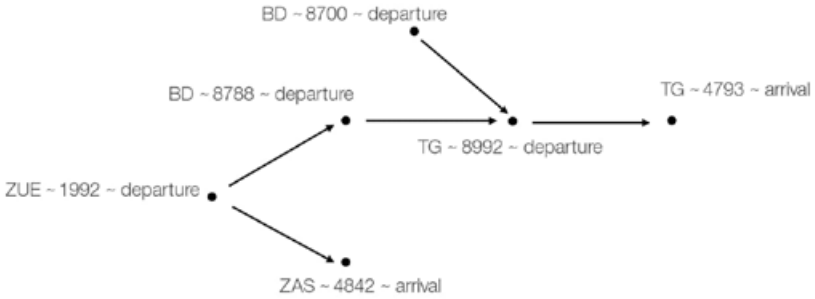


Figure 5: example of a graph with knock on delay

4.2 Weighting Function

Definition. Let $G = (V, E)$ be the graph defined above. A *weight* is a function $w: E \to \mathbb{R}$. A pair (G, w) is called a *weighted graph*.

Our aim is to define a weight on this graph to measure the severity of each link of a delay chain. The idea is to count the number of disturbances caused by a link in a clever way. To implement this, we first define the following sets for any vertex $x \in V$:

$$Par(x) = \{y \in V \mid yx \in E\}$$
$$Ch(x) = \{y \in V \mid xy \in E\}$$

Where $Par(x)$ can be understood as parents of x and $Ch(x)$ as children of x. A vertex x is *final* if $Ch(x) = \emptyset$.

Some edges in the graph are classified as „self-imported“. These represent delay propagation within the same train – for example, a train that departs late from operational point A and therefore arrives late at operational point B. Although these delays are relevant (since the train may further delay other trains at operational point B), they are not considered *new* disruptions. Thus, we treat self-imported edges differently when calculating the weight.

Let $xy \in E$, we define $w: E \to \mathbb{R}$ by

$$w(xy) = \begin{cases} \dfrac{1}{|Par(y)|} & \text{if } y \text{ is final} \\ \dfrac{\sum_{z \in Ch(y)} w(yz)}{|Par(y)|} & \text{if } xy \text{ has type "self-imported"} \\ \dfrac{1 + \sum_{z \in Ch(y)} w(yz)}{|Par(y)|} & \text{else} \end{cases}$$

This weight simply sums the number of impacted trains by an edge, without over numbering the edge type “self-imported”.

Example. Assuming that the green edge has type "self-imported", the above defined weighting function assigns weights to the edges as follows:

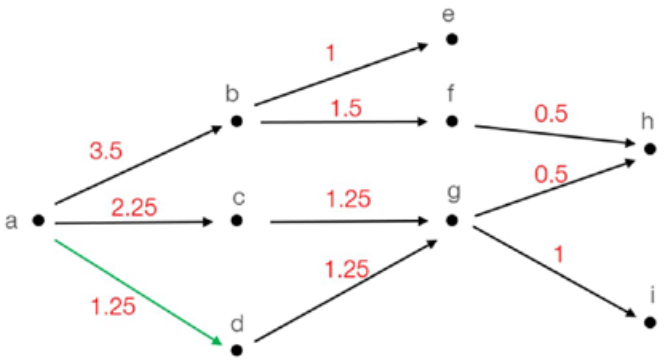


Figure 6: weighted graph

4.3 Data Aggregation

So far, our analysis has focused on daily graphs, i.e., chains of delays that occurred on the same day. However, the goal of this study is to uncover systematic dependencies rather than isolated daily incidents. To achieve this, the *chaining file* – now containing weighted edges – is aggregated by edges in this second step.

This aggregation can be performed either by train number or by train family. In this paper, we focus exclusively on aggregation by train number.

The aggregation is based on the columns *Source_TN* and *Target_TN*, representing the train numbers of the source and target events. During this process, two additional columns are added to the Chaining file:

- *Sum_of_Weights*: the total sum of the edge weights across all days where the same causal link was observed.
- *edgeRealisedNumber_TN*: the number of distinct days on which this connection between train numbers occurred.

5 Edge Filtering

Note that up to this point, we have not yet constructed complete graphs; we still have only individual edges. That is why we now turn to the final step: filtering the edges for

constructing the graph. However, we are not interested in visualizing all graphs, but only those that contain critical points – that is, highly influential delays within the network. To identify them, we first need to determine the most significant edges.

In our context, we filter the important edges which are those that

(i) have a large weight (i.e., they lead to many subsequent delayed trains),
(ii) occur frequently across different days.

Based on this reasoning, we define the following metric:

$$\text{PFS} = \text{Frequency} \times \text{Sum of Weights}$$

We found this combination especially suitable, as it implicitly emphasizes frequency by incorporating it quadratically. Without this adjustment, the frequency component could easily be overshadowed by the sum of weights – especially when analysing shorter time periods.

To now identify the critical points, we define a threshold on the PFS values. For this, we examined the cumulative distribution function (CDF) of the PFS values. The threshold was chosen at the point where the spacing between successive PFS values increases noticeably, indicating a natural break in the distribution and the emergence of a smaller subset of dominant edges.

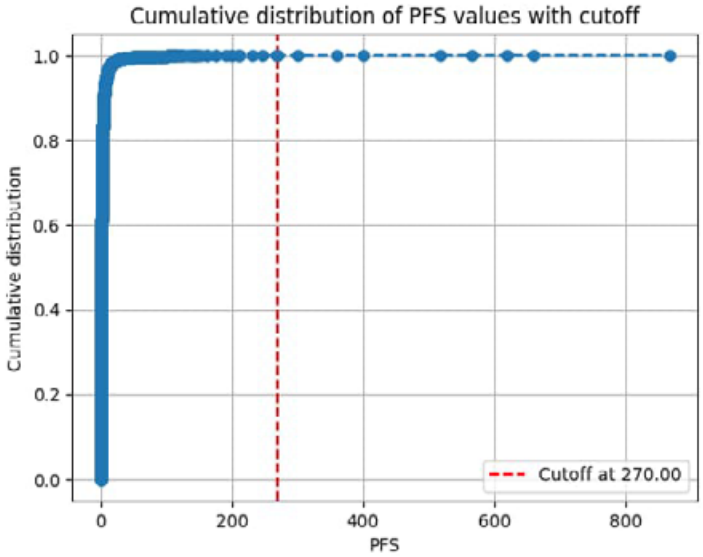


Figure 7: CDF of PFS values

Edges with a PFS value above the threshold are our top edges – those representing the critical points. The challenge now is to not only show these critical edges, but also all other edges also involved in their respective chains of delay.

We now aim to construct the graphs that contain such top edges. To do this, we start by selecting a top edge. For demonstration, we assume that the edge with the 2-uple OL~602~arrival__OL~701~departure (which corresponds to Target_TN__Source_TN) is a top edge. This top edge appears on multiple days, although the structure of the corresponding daily graphs can vary significantly. In the Figure 8 below we can see the daily graphs integrating the top edge in red for the 13.03.2025, 23.03.2025 and 07.02.2025. For illustration purposes, the 23.03.2025 also includes another disjunct graph:

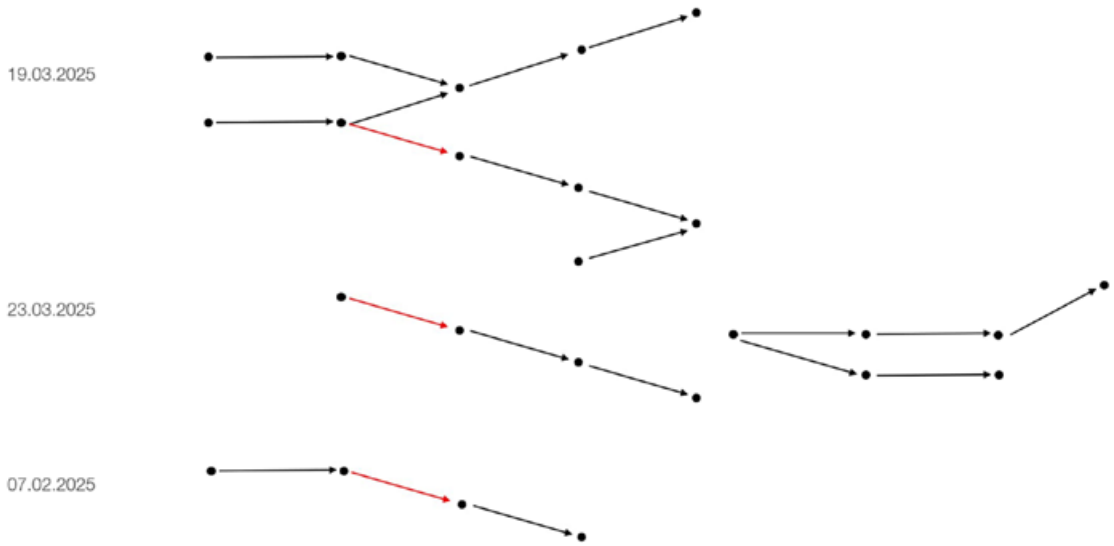


Figure 8: Example of similar and disjunct graphs

In our visualization, we want to display an aggregation of all daily graphs that contain this top edge in red – meaning, every edge that occurs within at least one of these daily graphs containing the red edge should be included in the final aggregated graph.

We also identify all edges which repeats on different days and compute their frequency for each.

6 Visualization

Now that we have compiled all daily graphs for the top edges, we can proceed with visualization. These graphs are mapped onto an OpenStreetMap background, with each graph assigned a unique color. The thickness of each edge indicates its frequency of occurrence.

Additional useful information includes the source and target details, the sum of weights, the average weight, and the number of days the edge was realized (*edgeRealisedNumber*) for edges, as well as the event identifier (*operating point_train number_event*) and the scheduled time for nodes.

6.1 Results

Below are examples of visualizations of top edges based on data from January 6 to March 31, 2025:

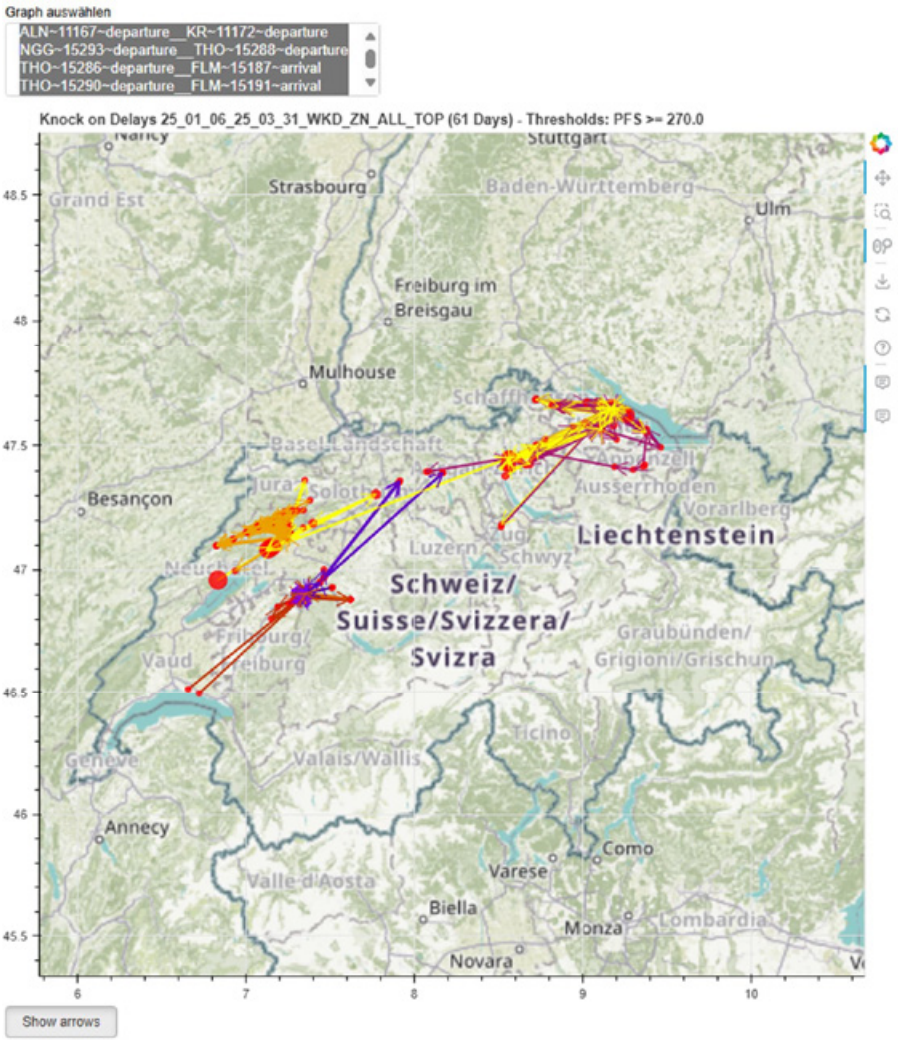


Figure 9: Example of diferent disjunct graphs

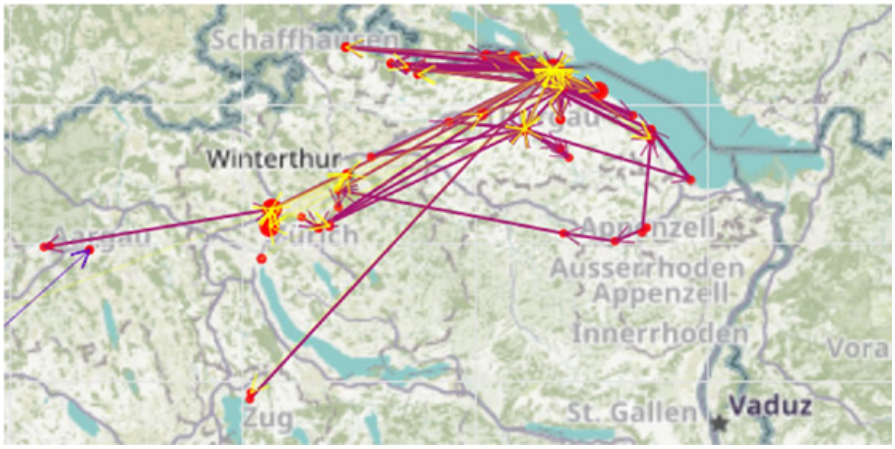


Figure 10: Graph of top edge ALN~11163~departure__KR~11168~departure with arrows

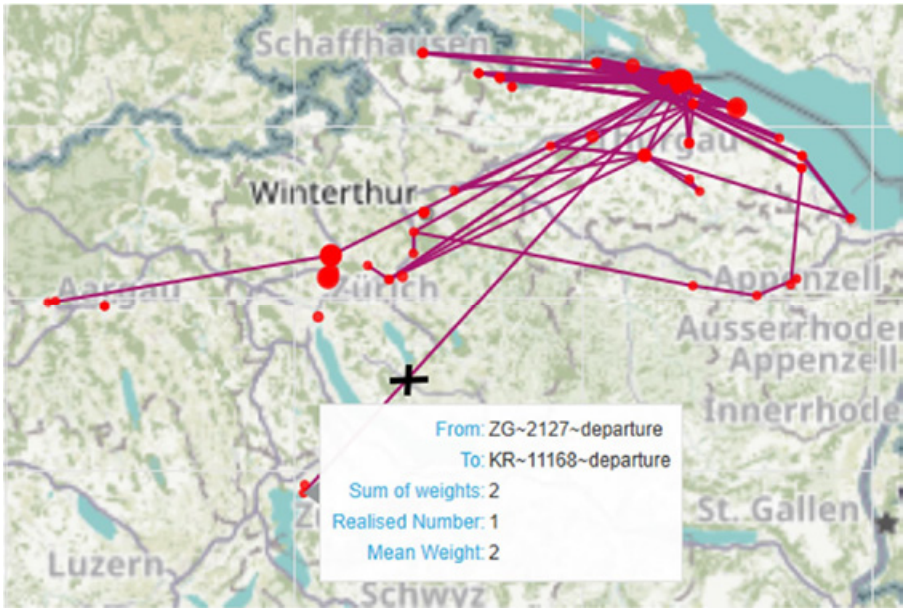


Figure 11: Example of edge information

7 Discussion of results and Future work

The initial visualizations are promising, and we plan to create more to enhance our understanding and refine our approach. One limitation is our delay propagation recognition system, which only identifies delays of three minutes or longer. This constraint may result in missing data where shorter delays could be significant. Additionally, the visualization aspect needs improvement. Further research is required to

determine the most effective way to present the data, enabling us to efficiently interpret the information and adjust train routes or timetable schedules to mitigate the risk of knock-on delays.

8 Conclusion

The approach presented in the paper involves the use of the EFA system by Swiss Federal Railways (SBB) to automatically identify and analyze delay propagation across the railway network. By modeling delays as directed graphs with weighted edges, the system identifies critical points where delays are most influential. Key findings include the effective visualization of these delay chains, allowing for targeted timetable adjustments to minimize knock-on delays.

This work would not have been possible without the contributions of many individuals. I would like to thank the entire EFA team, along with Lionel Brütsch, Enea Peter, and Amanda Zeltner, for their unique and valuable contributions to the development of this method. I would like to thank Prof. Norio Tomii for many insightful discussions on topics related to this study, which helped inform its development.

Literature

[1] G. Auer, "SBB – Data Analytics für Verspätungsursachen," Corporate Communications Detecon International GmbH, 18 February 2022. [Online]. Available: https://www.bigdata-insider.de/sbb-data-analytics-fuer-verspaetungsursachen-a-75a98462773b3f9bc46ed60237696c44/. [Accessed 11 July 2025].

[2] I. Hansen and J. Pachl, Railway timetabling and operations : Analysis, modelling, optimisation, simulation, performance, evaluation, Eurail press, 2014.

[3] H. Flier, R. Gelashvili, T. Graffagnino and M. Nunkesser, "Mining railway delay dependencies in large-scale real-world delay data," *R.K. Ahuja et al. (Eds.): Robust and Online Large-Scale Optimization,,* p. 354–368, 2009.

[4] B. F. N. O. A. N. S. S. H. Fabrizio Cerreto, "Application of Data Clustering to Railway Delay Pattern Recognition," *Journal of Advanced Transportation,* vol. 1, 2018.

[5] J. G. S. L. F. C. Ping Huang, "Explainable train delay propagation: A graph attention network approach," *Transportation Research Part E: Logistics and Transportation Review,* vol. 184, 2024.

[6] B. G. P. D. C. R. e. a. Monechi, "Complex delay dynamics on railway networks from universal laws to realistic modelling," *EPJ Data Sci. 7,* vol. 35, 2018.

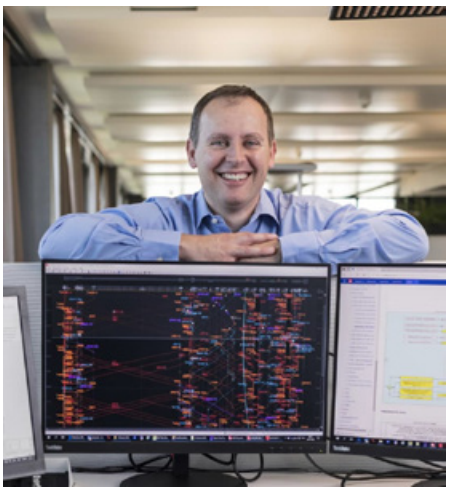
[7] S. Y. T. N. Ochiai Yasufumi, "An algorithm to identify delay propagation routes based on visualization of association rules," in *Proceedings of the 2nd International Railway Symposium Aachen*, Aachen, 2019.

[8] K. M. S. T. N. Yamamura A., "Identification of causes of delays in urban railways," in *Computers in Railways XIII*, Southampton, 2012.

[9] K. A. Tomii Norio, "Identifying the Cause and the Propagation Route of Delays of Trains using Association Rules," in *Proceedings: 1st International Railway Symposium Aachen*, Aachen, 2017.

[10] F. Harary, Graph Theory, Westview Press, 1969.

Autoren



Graffagnino, Thomas

Thomas Graffagnino is the Business Owner for Analytics Tools in SBB Infrastructure's Timetable and Operation Department. He oversees tools and methods that improve timetable quality, dispatching, punctuality, and network stability. As part of the vPro project, he served as an expert in planning and traffic control systems and developed algorithms for a train advisory system to improve energy efficiency and punctuality. Following its deployment on the Swiss standard-gauge network, punctuality increased by one percentage point. Thomas holds a Master's degree in Mathematics and is an external lecturer at the University of Aachen in Germany.

Session 8
Takuya Kaneko, Toshihiro Hiura, Kazushige Yonemoto,
Kiyoshi Narita, Hideyuki Yabuki, Shuho Watanabe, Norio Tomii

Estimation of Train Dwell Times using a Data-Driven Approach

Kaneko, Takuya[1], Hiura, Toshihiro[1], Yonemoto, Kazushige[1],
Narita, Kiyoshi[1], Yabuki, Hideyuki[1], Watanabe, Shuho[1], Tomii, Norio[2]

[1] Tokyo Metro Co., Ltd.

[2] Nihon University & Institute of Science Tokyo

Summary

In urban railways, planned dwell times of trains are set to the minimum necessary to enable high-density operations. Consequently, when actual dwell times are extended for any reason, delays occur and propagate to many other trains. Such extensions often result from increased alighting and boarding times at specific train cars. One potential countermeasure is to equalize congestion levels among cars, and railway operators are interested in understanding how such measures might affect dwell times. To support this need, we developed a dwell time estimation based on a random forest using historical congestion data as training data. Validation using real-world data confirmed that the proposed algorithm achieves high estimation accuracy.

Keywords: Dwell time, congestion rate, urban rail, random forest, machine learning

1 Introduction

In urban railway systems, where many trains are operated at short intervals, the scheduled dwell times are kept to a minimum. Hence, when the number of boarding and alighting passengers is large, the dwell time tends to exceed the scheduled value, resulting in delays. Since the trains are operating at high density, such delays propagate to subsequent trains and can ultimately impact the entire network.

It has been generally observed that passengers tend to choose train cars that are close to stairs or escalators at their destination station. As a result, there are more alighting

passengers from those specific cars (and, because of their convenience, more boarding passengers as well). This leads to longer dwell times. To improve this situation, it is effective to balance the passenger load across the train cars. For example, adjusting the stopping position of the train can change the positional relationship between the cars and the stairs/escalators, potentially helping to equalize the congestion rate. Railway operators are interested in estimating the effects of such measures. Specifically, they wish to estimate the dwell times that would result when such measures are implemented—before taking action. In order to fulfil their requirement, it is necessary to analyze the data at the individual car level.

Previous studies have used the number of boarding and alighting passengers to estimate dwell time. While passenger count is a key determinant of dwell time, other factors—such as crowding inside the train or on the platform—may also influence it. It is also known that the ratio of alighting to boarding passengers affects dwell time. Moreover, additional yet-unidentified factors may be involved. However, identifying all relevant factors and accurately quantifying them is extremely challenging (e.g., estimating platform congestion and using it as an explanatory variable is difficult).

To address this issue, we developed a data-driven algorithm that estimates dwell time based on historical data. The algorithm uses car-specific congestion data collected every second by the load-compensation device as input. It is based on the random forest, a widely used supervised machine learning algorithm. We constructed the dwell time estimation model using historical car-specific congestion data as training data. By inputting congestion data in an improved situation to the model, we can estimate train dwell times under those improved situations.

This paper describes the data utilized, the structure of the proposed algorithm and the results of the application.

2 Related Works

Dwell time is one of the most critical factors affecting punctuality on high-frequency railway lines. Consequently, a substantial body of research has been conducted, as summarized in the review by Kuipers [1]. Daamen [2] explores the possibility of estimating dwell times at short stops without passenger demand data through statistical analysis of track occupation. With recent advances in data collection technologies, it has become feasible to obtain passenger-related data. Kuipers [3] investigates the relationship between passenger volumes and both the frequency and duration of dwell times for commuter trains, using data collected from automatic passenger counting systems.

Tortainchai [4] applies Data Envelopment Analysis (DEA) to evaluate dwell time efficiency at each station relative to passenger volume. Cornet [5] proposes a data-driven approach to assess the influence of boarding, alighting, and onboard passenger numbers on dwell time. Their model decomposes dwell time into a deterministic component—termed the *Minimum Dwell Time*, which depends on passenger flow—and a stochastic component. A method for estimating the minimum dwell time is also presented. Coulaud [6] attempts to develop a unified statistical model to estimate actual dwell times across all stations. Kuipers [7] proposes the use of quantile regression to examine the conditional distribution of dwell times across different percentiles, based on one year of detailed train operation and passenger count data. Pritchard [8] utilizes On-Train Monitoring Recorder (OTMR) data collected over 14 months, which includes wheel stop and start timings at 1-second resolution, to investigate small fluctuations in station dwell time. Ge [9] introduces a model to determine the optimal dwell time scale using boarding and alighting data, focusing on the most congested car.

Despite these efforts, few studies have examined alighting time, boarding time, and confirmation time as separate components, nor have they accounted for variations between individual train cars.

3 Data obtained by Load Compensation Device

A load-compensation device is a system that maintains constant acceleration and deceleration regardless of the number of the passengers on board. This system can output the congestion rate of each car by converting the pressure values inside air springs with a resolution of one second. In addition, we can get the times when the doors opened and closed. By extracting the data during a train is stopping at a station, we can get the data concerning the change of the congestion rate as shown in Fig.1 which shows the result for one car, but we can get the data for each car in reality.

More in detail, we can get the data shown in Table 1. Nos 1 through 12 represent values corresponding to each individual car, whereas Nos. 13 and 14 correspond to the entire train. Nos. 1 to 13 are the (candidates of) explanatory variables and No. 14 is the objective variable.

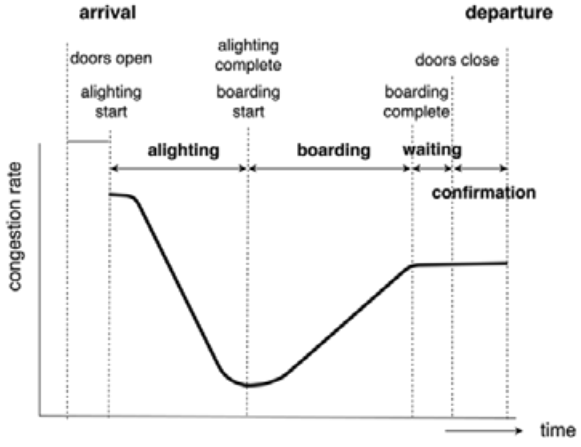


Fig. 1 Load compensation device data

Table 1: Data obtained from the load compensation device.

No.	Data	Note
1	Congestion rate when doors opened	
2	Number of alighting passengers	Calculated as No.6 - No.1
3	Congestion rate when alighting completed	
4	Average alighting speed	Calculated as No.2 ÷ No.5
5	Alighting time	
6	Congestion rate when boarding started	We assume this is the same as No.3
7	Number of boarding passengers	Calculated as No.8 - No.6
8	Congestion rate when boarding completed	
9	Average boarding speed	Calculated as No.6 ÷ No.10
10	Boarding time	
11	Waiting time	
12	Congestion rate when doors closed	
13	Confirmation time	
14	Dwell time	

4 Detailed Analysis of Train Dwell Times

4.1 Components of Train Dwell Times

Components of a train's dwell time are shown in Fig. 2. After the train arrives at a station and stops, the doors open (**open doors**). Passengers then alight. This time is called **the alighting time**. When alighting is complete, boarding begins. In this paper, it is assumed that the time at which alighting is completed is the same as the time at which boarding starts (i.e. boarding starts immediately after alighting is completed). The time taken to complete boarding is **the boarding time**. The doors are closed after all cars have been boarded and safety has been confirmed. In this paper, the time between the completion of boarding and the closing of the doors is called **the waiting time** (of the car concerned). Because the time of the completion of boarding may be different among cars, waiting times may be different from one car to another. After the doors have been closed and safety has been re-confirmed (the time required for this process is called **the confirmation time**), the train departs (**departure**).

The time between the opening of the doors and the departure of the train is the target of the estimation of this paper (originally a dwell time refers to the time between the arrival and the departure of the train, but in this paper, we call the time between the opening of the doors and the departure of the train a dwell time).

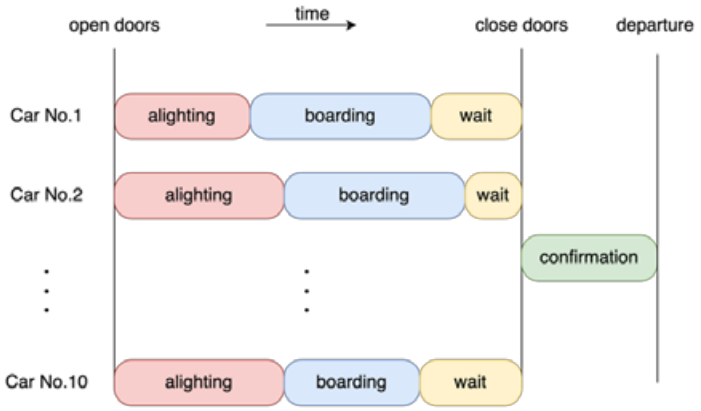


Fig. 2 Components of a dwell time.

4.2 Train Dwell Times

The actual dwell times of trains at a certain station of Tokyo Metro (hereafter referred to as Station A) between 7:00 and 9:00 are shown in Fig. 3. These data are from weekdays over the course of one month. Station A is located in the center of the city and has a high volume of both boarding and alighting passengers. The horizontal axis represents the actual arrival times, the vertical axis represents the dwell times in seconds. A polynomial regression curve is shown as blue line which is added to identify trends. The same data in the form of a histogram is shown in Fig. 4 . In this figure, dwell times of less than 60 seconds are shown in blue, and those of 60 seconds or more are shown in orange. From these results, it can be identified that dwell time varies over time, reaching its peak between around 8:15 and 8:30. Additionally, 4.8% of trains have dwell times of 90 seconds or longer, and 1.0% of trains have dwell times of 120 seconds or longer.

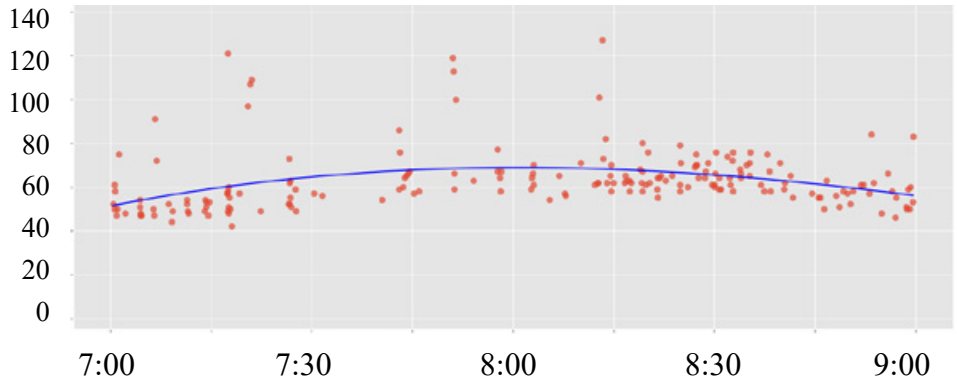


Fig. 3 Actual dwell times from 7 o'clock to 9 o'clock.

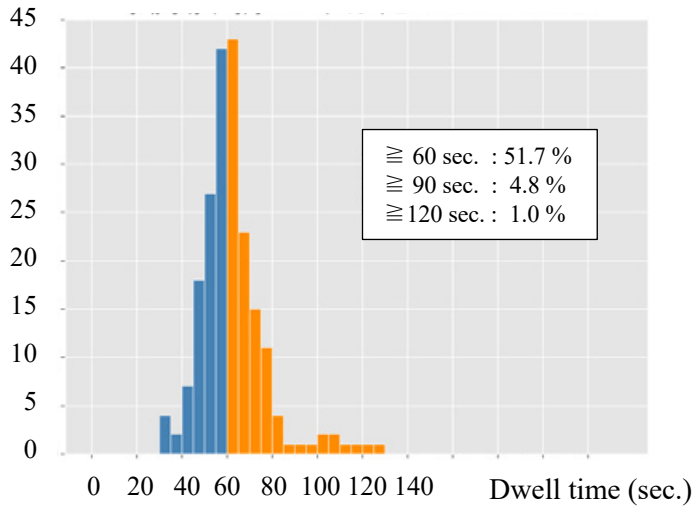


Fig. 4 Distribution of dwell times.

4.3 Analysis of Train Congestion

Trains consist of ten cars, and the congestion rate of the rearmost car (Car No. 1) and that of Car No. 7 of the same train when the train arrived at the station are shown in Fig. 5 and Fig.6, respectively. The station and trains are the same as those in Fig. 3. From Fig. 5, it can be identified that the congestion rate of Car No. 1 reached its peak between around 8:15 and 8:30, reaching approximately 210%, whereas from Fig. 6, it is observed that the congestion rate of Car No. 7 also reached its peak during the same time period, around 8:15 to 8:30, reaching approximately 150%.

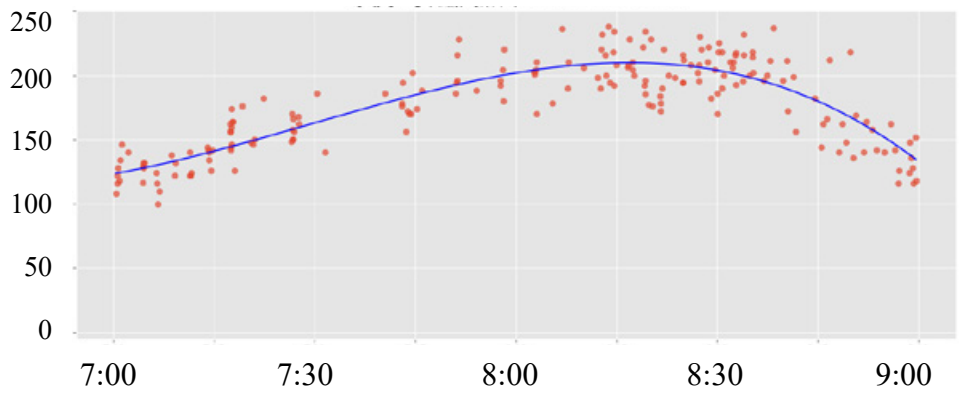


Fig. 5 Congestion rates of Car No.1.

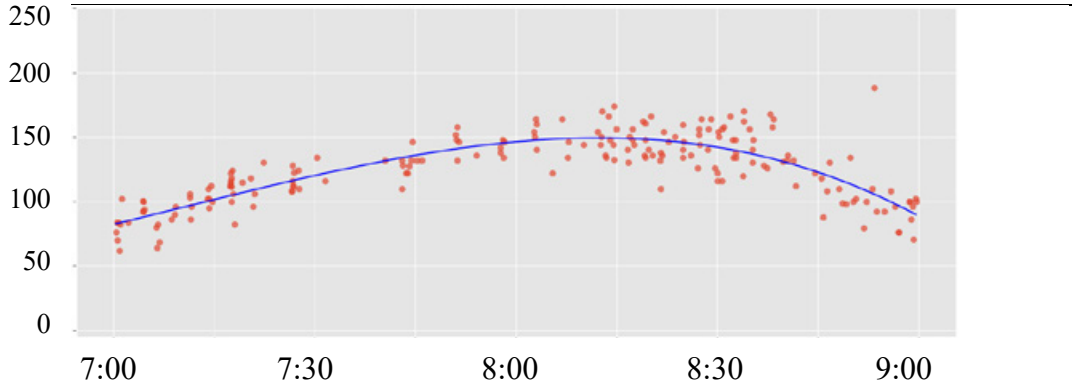


Fig. 6 Congestion rates of Car No.7.

4.4 Analysis of Components of Dwell Time

In this section, an analysis is conducted on each component of the dwell time. The data are the same as before but for the period between 8:00 and 9:00, when congestion is particularly heavy.

(1) Alighting time and boarding time

As reported in previous research, it is hypothesized that there is a relationship between the number of alighting passengers and alighting time, as well as between the number of boarding passengers and boarding time. Figures 7 and 8 present scatter plots illustrating the relationships between the number of alighting and boarding passengers and their respective times for Car No. 1. As shown in the figures, both alighting time and boarding time tend to increase with the number of alighting and boarding passengers, respectively.

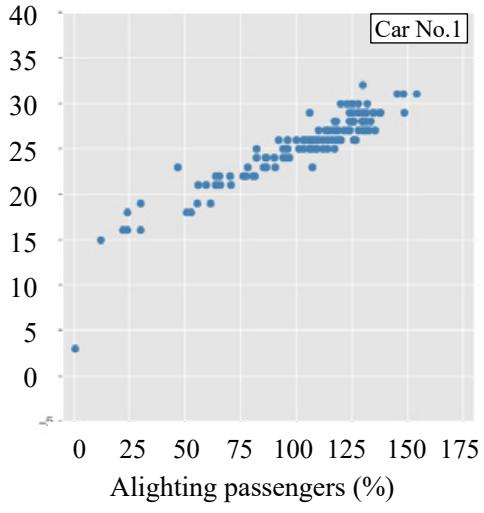


Fig. 7 Alighting time

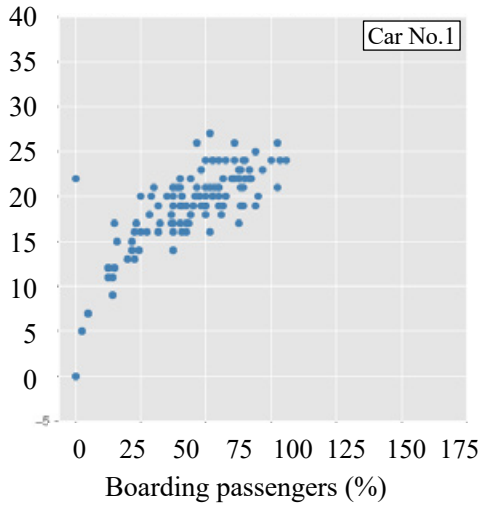


Fig. 8 Boarding time

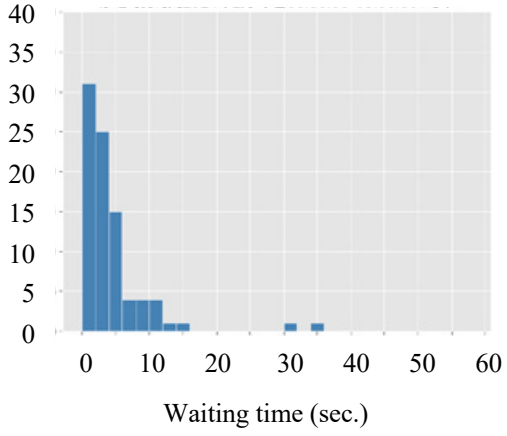
These figures indicate that the relationship between boarding time and the number of boarding passengers is non-linear and exhibits a weaker correlation compared to that between alighting time and the number of alighting passengers.

(2) Waiting Time

If the sum of the boarding and the alighting times of a given car is short, the time from the completion of boarding to door closure—i.e., the waiting time—is expected to be longer. Furthermore, when the congestion level of a given car is high, it is suspected that it takes longer to close the doors, resulting in a longer waiting time. Figure 9 shows a histogram of the waiting time for Car No. 1, while Figure 10 presents a scatter plot of the boarding rate of Car No. 1 at the time of door closure versus the waiting time. Corresponding graphs for Car No. 5 are shown in Figures 11 and 12, respectively.

For Car No. 1, the waiting time is less than 6 seconds for many trains. In contrast, for Car No. 5, there are very few trains that waiting time is less than 6 seconds. This indicates that for Car No. 1, the doors are closed in a shorter time after boarding is completed compared to Car No. 5. In other words, in Car No. 5, boarding is completed earlier than in Car No. 1, resulting in a longer time until door closure.

From these results, it can be concluded that Car No. 1 must be a very influential car in determining the dwell time at Station A. It can be also observed that there are cases where the waiting time for Car No. 1 is prolonged, indicating that some form of disruption occurred before door closure, requiring additional time. In addition, we cannot observe a correlation between the congestion rate and the waiting time.



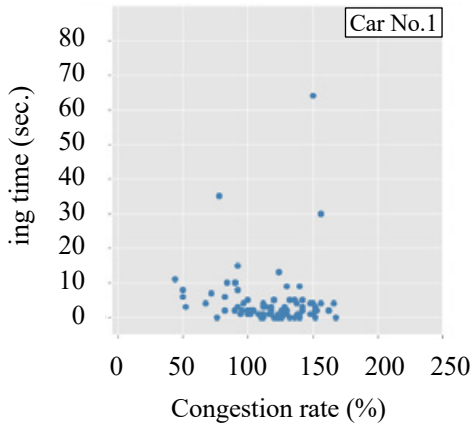


Fig. 9 Waiting time – Car No.1. *Fig. 10 Congestion rate vs waiting time.*

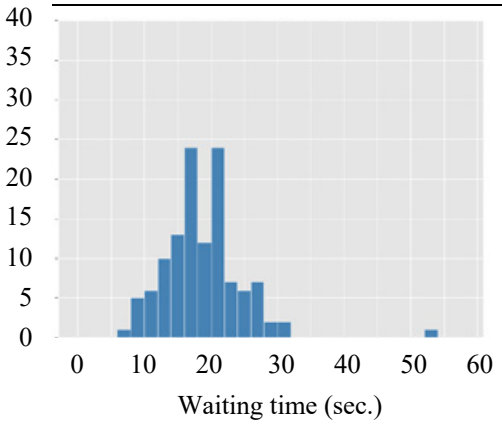


Fig. 11 Waiting time – Car No.5.

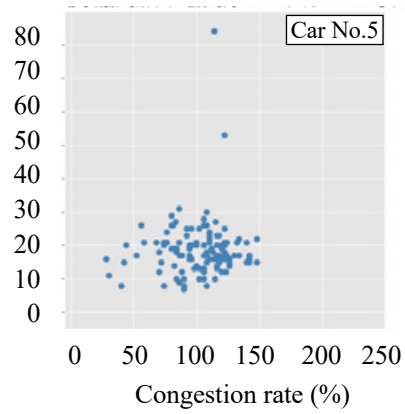


Fig. 12 Congestion rate vs waiting time.

(3) Confirmation Time

When the congestion level of certain cars is high, it is suspected that some trouble may occur, resulting in a longer time until the train departs. Figure 13 shows a histogram of the confirmation times, and Figure 14 presents a scatter plot of the congestion rate of the most crowded car at door closing versus the confirmation time. From Figure 13, it can be observed that the confirmation time is occasionally long; however, Figure 14 indicates that there is no correlation between the congestion rate and the confirmation time.

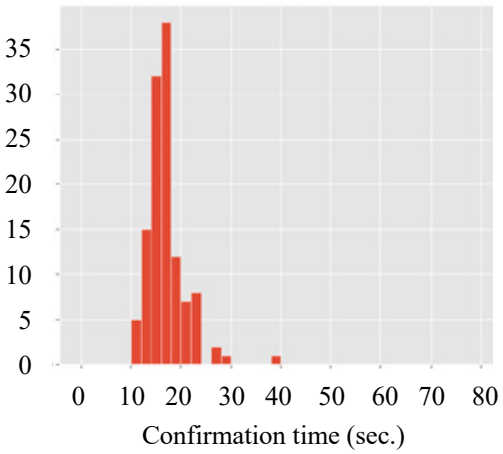


Fig. 13 Confirmation time.

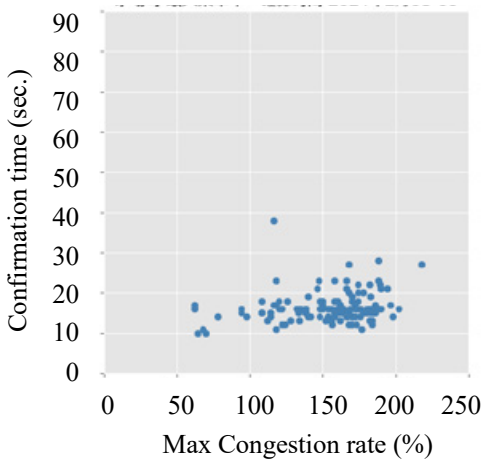


Fig. 14 Congestion vs Confirmation time.

4.5 Findings from the analysis of dwell times

The authors conducted similar analyses over extended time periods and other stations. As a result, the following findings were revealed.

- Dwell times vary between trains, and in some cases, certain trains exhibit exceptionally long dwell times.
- Congestion rates vary not only between trains but also between cars within the same train.
- A positive correlation exists between the number of passengers alighting and the alighting time.
- A positive correlation exists between the number of boarding passengers and the boarding time; however, this relationship is non-linear. Moreover, the correlation is weaker than that between the number of alighting passengers and the alighting time.
- No significant correlation is observed between waiting time and congestion rate.
- Cars with longer alighting and boarding times tend to have shorter waiting times.
- Waiting time may still increase due to unforeseen disruption. Since the causes cannot be identified in advance, such increases should be considered as stochastic phenomena.
- No significant correlation is observed between confirmation time and the maximum congestion rate.
- Confirmation time may increase due to certain disruption. As the causes are not known beforehand, the increase in confirmation time should also be treated stochastically.

5 DWELL TIME ESTIMATION MODEL

5.1 Overall structure of the proposed algorithm

The overall structure of the algorithm proposed in this paper is shown in Fig.15. The algorithm is based on the supervised learning framework. We prepare the historical time series congestion data as the training data set. This data set consists of instances of the explanatory variables and the target variable (details are explained in 5.2). A dwell time estimation model is built by using the random forest algorithm (details are given in

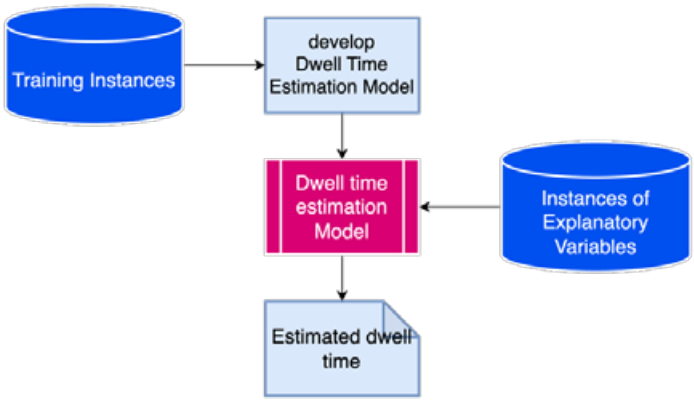


Fig. 15 Overall structure of the algorithm.

5.3). After finishing the development of the model, we can estimate dwell times of trains in a situation we are interested in (such as when the stopping position of trains is changed) by giving the instances of explanatory variables to this model.

5.2 Explanatory variables

The target variable of our algorithm is the dwell time. Each item shown in Table 1 except the dwell time could be an explanatory variable. It is known in the textbooks, however, that we should not use variables which are mutually dependent. It is also known that machine learning is generally prone to overfitting when there exist too many explanatory variables. The explanatory variables in Table 1 contain data that have mutual relationship. For example, the number of alighting passengers can be calculated as the difference between the congestion rate at door opening and the congestion rate at the end of alighting. The speed of alighting and boarding is calculated from the time of alighting and boarding. From the viewpoint of application phase of the algorithm, we should consider the workload to prepare the instances of the explanatory variables. As shown in the previous section, the waiting times and the confirmation times behave stochastically. Hence, we may well think it is almost impossible to estimate them from other explanatory variables. Thus, we include them in the explanatory variables.

To sum up, the explanatory variables we use are as follows:

1. Congestion rate when doors opened (for each car)
2. Congestion rate when alighting completed (for each car)
3. Congestion rate when doors closed (for each car)
4. Waiting time (for each car)
5. Confirmation time (for the train)

5.3 Dwell time estimation model

In this study, the random forest is used as the learning algorithm; Random forest is a machine learning algorithm that constructs many decision trees and adopts the results of majority decisions, which is known to provide higher accuracy than the decision tree algorithm [10].

When we apply the dwell time estimation model, we need to prepare the instances of the explanatory variables. As for the waiting time, if we want to estimate the dwell time in an ideal situation (i.e., doors are closed just after the completion of boarding), we can set the waiting time to 0 seconds (or very small value obtained by an analysis of actual data).

If we want to estimate the dwell time in more realistic situation, we can prepare the instances of the waiting time using some stochastic distribution based on the actual data (such as the one shown in Fig. 9). The situation is similar for the confirmation time. We should use a fixed value, or a value obtained by some stochastic distribution reflecting our purpose.

5.4 Basic Algorithm and improvement

We have implemented an algorithm (**Algorithm 1**) using the explanatory variables No.1 to No. 4 of all the 10 cars for each train and No.5 for each train.

When we estimate dwell times using Algorithm 1, however, we need to prepare the data on waiting times for each car of each train, which not only requires a lot of work but is also very difficult. Therefore, we should consider a method that does not use the waiting time of each car as an explanatory variable.

The dwell time is obtained by the following equation, where "max" denotes the maximum among all cars, and "waiting" refers to the waiting time of the car with the largest sum of alighting and boarding time.

Dwell time = max (alighting + boarding) + waiting + confirmation ..... (1)

On the other hand, for the car for which the sum of the alighting time and the boarding time is the largest, the waiting time must be the smallest among all cars in the train. This means that the following equation applies (min is the minimum among all cars in the train):

Dwell time = max (alighting + boarding) + min(waiting) + confirmation ..... (2)

From Equation (2), it can be seen that it is not necessary to prepare the waiting time for each car as an explanatory variable, instead, it suffices to prepare only the smallest waiting time among all cars. Based on this idea, we have developed Algorithm 2 which uses the smallest waiting time instead of the waiting time of each car.

When we apply **Algorithm 2** for estimation, we need to prepare explanatory variables No.1 to No.3 for all the cars. On the other hand, our preliminary analysis revealed that there exist cars that have significant influence on the dwell times (we call these cars "**influential cars.**" We should note that there may exist more than one influential cars). Thus, we consider the idea that only the data for the influential cars should be used when constructing the dwell time estimation model. **Algorithm 3** uses explanatory variables for only the influential cars instead of all the cars. Algorithm 3 has an advantage that it is

not necessary to prepare the instances of the explanatory variables for all the cars; it is sufficient to prepare the instances only for the influential cars.

5.5 Validation of the algorithms

We have conducted the validation for Station A. We have made a dwell time estimation model using the data of February 2024 (7 o'clock to 11 o'clock, weekdays). Seventy percent of the data was used as the training data, and the remaining 30 percent of the data were used to validate the accuracy of the model. As for the influential cars of Algorithm 3, we chose Car No.1 and Car No.2. The validation results for Algorithms 1 to 3 are presented in Table 2, which reports the mean absolute errors (in seconds).

Table 2 Results of Validation : Algorithms 1 – 3.

Algorithm	Explanatory Variables concerning waiting time	Cars	MAE(sec.)
Algorithm 1	Waiting times for each car	All Cars	3.92
Algorithm 2	Smallest waiting time	All Cars	5.81
Algorithm 3	Smallest waiting time	Car 1, Car 2	4.54

In Table 2, Algorithm 1 showed the highest accuracy. But considering both the estimation accuracy (the difference is less than one second) and the workload to prepare the instances of explanatory variables (the workload of Algorithm 3 is 20% of that of Algorithm 1), we decided to use Algorithm 3.

6 APPLYING DWELL TIME ESTIMATION MODEL

6.1 Estimate dwell times in the new situation

To evaluate the generalizability of the model under modified operational conditions, we applied it to the post-change data and compared the estimated and actual dwell times. Namely, the stopping position of the train at station A was changed on March 23rd 2024, and the congestion rates of each car have changed significantly. The dwell time estimation model, constructed using the data before the stopping position was changed (i.e., February 2024) was applied to the post-change situation (i.e., April 2024) to estimate the dwell time under the new stopping position.

As already explained, the algorithm is originally intended to estimate the dwell time when certain measures will be implemented in the future. However, in this paper, to verify the

accuracy by comparing the algorithm's estimation results with the actual outcomes, we have used the actual data of April 2024 as the instances of the explanatory variables. These include the congestion rates at the time when doors opened for Car No.1 and Car No. 2, the congestion rates when alight was completed for Car No.1 and Car No.2, the congestion rates when doors closed for Car No.1 and Car No.2, the smallest waiting time and the confirmation time of the train.

6.2 Results and discussions

Differences between the estimated values and the actual values are shown in Fig. 16 and Fig. 17. Fig. 16 says the mean absolute error was 4.36 seconds and Fig. 17 shows that approximately 64% and 95% of the estimated values fall within an error of ±5 seconds, and ±10 seconds, respectively. We conducted similar analyses for other months after the stopping position was changed and confirmed the algorithm is sufficiently accurate for practical use. It was found, however, the algorithm sometimes underestimates the dwell time when it is particularly long. We aim to clarify the cause of this issue and continue improving the algorithm.

The proposed algorithm utilizes the congestion rates, the smallest waiting time, and the confirmation time. A primary challenge lies in developing a methodology to prepare these data when the situation will be changed. While the random forest offers strong predictive performance, a known limitation is its lack of interpretability. Therefore, future work will also consider the development of interpretable models that can provide insights into the underlying decision-making process.

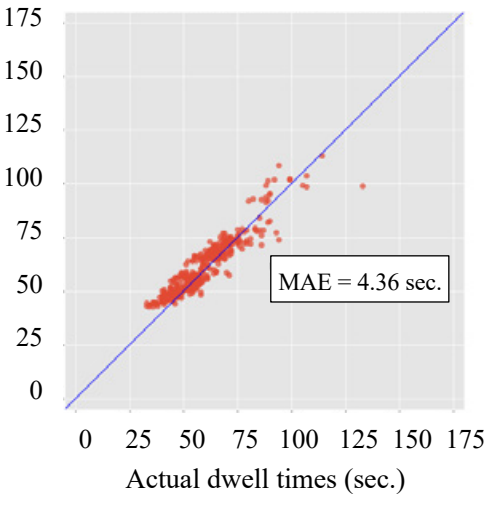


Fig. 16 Results of estimation.

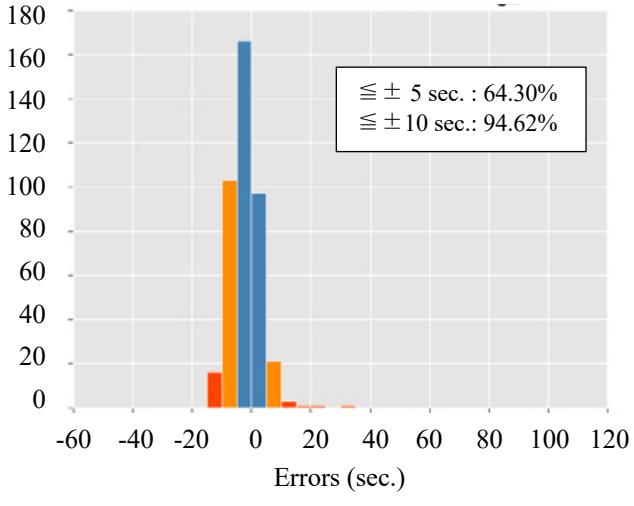


Fig. 17 Histogram of errors (sec.)

7 CONCLUSION

To estimate train dwell times under scenarios where certain measures—such as equalizing congestion levels across train cars—are implemented, we developed an algorithm based on the random forest. The model utilizes data obtained from load compensation devices as training instances. It was designed with the objective of reducing the number of explanatory variables, considering both the estimation accuracy and the workload involved in preparing the input data.

Based on actual operational data, we verified that the proposed algorithm, despite relying on a limited set of input variables, can yield reasonable estimation results. Specifically, over 64% of the estimated values fall within ±5 seconds of the actual dwell times, and more than 95% fall within ±10 seconds.

These results demonstrate the feasibility of estimating dwell times under hypothetical congestion patterns using a minimal set of explanatory variables. Nonetheless, further refinement is required to address underestimation in cases with particularly long dwell times, and to incorporate external operational factors such as the day of the week, weather conditions, delay propagation and so on.

Future work will include consideration of alternative ensemble learning methods, such as XGBoost and LightGBM, in pursuit of improved accuracy. On the other hand, simpler algorithms such as the decision tree may still be valuable for their interpretability, despite their comparatively lower predictive performance.

Literature

[1] Ruben A. Kuipers, Carl-William Palmqvist, Nils O. E. Olsson, Lena Winslott Hiselius, The passenger's influence on dwell times at station platforms: a literature review, *Transport Reviews*, 41:6, 721-741, 2021.

[2] Li, D., Daamen, W. and Goverde, R. M. P., Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station. *Journal of Advanced Transportation*, 50: 877–896. 2016, doi: 10.1002/atr.1380.

[3] Ruben A. Kuipers, Carl-William Palmqvist, Passenger Volumes and Dwell Times for Commuter Trains: A Case Study Using Automatic Passenger Count Data in Stockholm, *Applied Science*, 2022, 12(12), 5983; https://doi.org/10.3390/app12125983

[4] N. Tortainchai, H. Wong, D. Winslett, T. Fujiyama, Train Dwell Time Efficiency Evaluation with Data Envelopment Analysis: Case Study of London Underground Victoria Line, *Transportation Research Record*, Vol. 2676(3), 2022.

[5] Sélim Cornet, Christine Buissonb, François Ramond, Paul Bouvarela, Joaquin Rodriguez, Methods for quantitative assessment of passenger flow influence on train dwell time in dense traffic areas, *Transportation. Research Part C*, 106, 2019.

[6] Rémi Coulaud, Christine Keribin, Gilles Stoltz, Modeling dwell time in a data-rich railway environment: With operations and passenger flows data, *Transportation Research Part C*, 146, 2023.

[7] Ruben A. Kuipers, Understanding dwell times using automatic passenger count data: A quantile regression approach, *Journal of Rail Transport Planning & Management*, Volume 29, 2024. https://doi.org/10.1016/j.jrtpm.2024.100431

[8] James Pritchard, Jason Sadler, Simon Blainey, Ian Waldock, Jeremy Austin, Predicting and mitigating small fluctuations in station dwell times, *Journal of Rail Transport Planning & Management*, Volume 18, 2021, https://doi.org/10.1016/j.jrtpm.2021.100249.

[9] Jianhao Ge, Pengling Wang, Xiaofang Xiao, Yutao Ye, Yunqi Chen, Dwell Time Scale Optimization for Urban Rail Transit Based on Passenger Boarding and Alighting Data Using Improved Fisher Clustering Method, *RailDresden 2025 -11th International Conference on Railway Operations Modelling and Analysis*, Dresden, Germany, 2025.

[10] T. Hastie et al., *The Elements of Statistical Learning: Data Mining, Inference and Predictions*, Springer, 2009.

ACKNOWLEDGMENT

Part of this research was supported by JSPS Grant-in-Aid JP21K04554.

Session 8
Takuya Kaneko, Toshihiro Hiura, Kazushige Yonemoto,
Kiyoshi Narita, Hideyuki Yabuki, Shuho Watanabe, Norio Tomii

Authors



Kaneko, Takuya

Kaneko Takuya has worked as a conductor and a train driver after he joined Tokyo Metro. He is currently a member of Transport section and in charge of making timetables.



Hiura, Toshihiro

Hiura Toshihiro has worked as a train driver, a controller at operation control center, a manager of Transport Section in Tokyo Metro.



Yonemoto, Kazushige

Yonemoto Kazushige was working for Tokyo metro for 40 years. He devoted most of the time for improving train timetables. He is currently Director of Subway Museum operated by Tokyo Metro.



Narita, Kiyoshi

Narita Kiyoshi has worked as a train driver after he joined Tokyo Metro. He is now a section head of Transport Section in Tokyo Metro.



Yabuki, Hideyuki

Yabuki Hideyuki has worked as a conductor, a train driver, a controller at operation control center. He is now manager of Transport Section in Tokyo Metro.



Watanabe, Shuho

Watanabe Shuho has worked as a conductor and a train driver after he joined Tokyo Metro. He is currently a member of Transport section and in charge of making timetables.



Tomii, Norio

Professor Tomii worked for Japan National Railways and Railway Technical Research Institute (RTRI). He is currently a professor at Institute of Science Tokyo and senior researcher at Nihon University. His current interests are applying optimization technique to railway scheduling, applying machine learning to realize robust timetables etc. He is ex-President of International Association of Railway Operations Research.

Session 8
Frédéric Weymann, David Janecek, Katrien Pauwels, Anneleen Van de Mergel, Quentin Van Leeckwyck

Improving operational quality through global optimization of reserve and buffer times

Weymann, Frédéric[1], Janecek, David[1], Pauwels, Katrien[2], Van de Mergel, Anneleen[2], Van Leeckwyck, Quentin[2]

[1]Quattron GmbH
[2]Infrabel nv

Summary

One approach to improving operational quality is to plan more robust timetables that contain sufficient reserves to deal with disruptions and delays. We propose a novel approach of achieving this goal by globally optimizing buffer and reserve times using the OptDis method in a subsequent step after the actual timetable compilation process. The approach was implemented and evaluated in cooperation with the Belgian infrastructure manager Infrabel. Operational simulations in the railway operation research application LUKS showed the positive effects on the achieved operational quality.

Keywords: Timetabling, Timetable optimization, Robustness, Mathematical optimization, Operational simulation, OptDis, LUKS

1 Introduction

Reliable and punctual operation is a central requirement that customers expect from the rail system and a crucial point in keeping traffic on the rails and shifting more traffic to them. Unfortunately, the system often fails to meet this requirement.

One approach to improving operational quality is to plan more robust timetables that contain sufficient reserves to deal with disruptions and delays. The key features here are sufficiently dimensioned running and stopping time reserves, which can be used to reduce delays, as well as sufficient buffer times between consecutive trips to dampen the transmission of delays.

Both measures reduce the usable capacity and thus contradict the goal of maximizing the utilization of the available infrastructure. Therefore, general specifications in the form of minimum reserves or minimum buffer times are not very effective: instead, they should be distributed skillfully through a global, targeted optimization.

In recent years, different approaches to improve the robustness of timetables using optimization methods have been presented, among others by Burggraeve et al [1], Högdahl et al [2] and Van Hoeck et al [3]. However, these methods either work on a macroscopic level, do not consider route changes or do not correctly model the effect of route or speed changes on the blocking time staircase [4]. Particularly with modern safety systems such as ETCS, this can lead to a major error in determining capacity utilization [5]. This error is particularly critical in this area of application, as the aim is not the rough planning of a timetable structure, but a timetable fine construction that is accurate to the second. An optimization of travel time reserves and buffer times is of little value if the optimized timetable has blocking time conflicts and therefore cannot be executed without obstructions in reality. In some cases, other properties such as cyclical timetables are also required. We use the established OptDis method [6][7], which has been used successfully in practice and precisely models running and blocking times on a microscopic level, supports route changes and works with any input timetables.

In addition, we investigate the effectiveness of our method using established simulation methods that have been used in practice for many years to determine the robustness of timetables [8]. This allows a well-founded and realistic assessment of whether and to what extent our method would have an impact on delays in operation.

2 Motivation and Goal

Infrabel is the Belgian railway infrastructure manager, established in 2005 as part of the restructuring of the national railway system. It is an autonomous public company responsible for the design, construction, maintenance and modernization of the Belgian railway network. Infrabel also manages capacity allocation and real-time train traffic control across the entire network.

Infrabel plays a key role in ensuring a safe, reliable and efficient railway system that is accessible to both domestic and international railway operators. The company contributes significantly to sustainable mobility and the economic connectivity of Belgium within the broader European transport network.

The robustness team is part of the "Customer & Business Excellence" direction at Infrabel which manages all operational and costumer-related tasks within the company. The robustness team plays a crucial role in ensuring reliable and resilient rail traffic, today and in the future. It is a multidisciplinary group of experts dedicated to analyzing and improving the robustness of rail traffic in Belgium. Their aim is to make the train traffic as resilient as possible to planned works and unexpected minor disruptions. To achieve this, the team conducts both microscopic and macroscopic robustness studies.

Microscopic studies focus on detailed infrastructure-level simulations using the software tool LUKS [8]. These simulations model individual train movements and infrastructure elements to assess the impact of delays and evaluate timetable robustness at a local level. Macroscopic studies, on the other hand, are performed at the network-wide level using the software tool OnTime. These simulations provide a broader view of robustness and help identify structural weaknesses or bottlenecks across the entire Belgian railway network. By combining both levels of analysis, Infrabel is able to provide data-driven recommendations for improving timetable resilience and operational performance.

Continuous improvement of the operational quality of timetables requires more than analysis afterwards (retrospective). There is a need for a structural feedback loop whereby insights from operation are systematically being fed back into the planning process. Optimizing robustness and buffer times early in the process allows for better anticipation of disruptions and leads to increased timetable reliability.

In this contribution, we develop a process that can be easily and efficiently integrated into the existing planning process. By using global optimization of reserve and buffer times in a post-planning step, we want to provide a practical method which clearly improves the operational quality of the timetable.

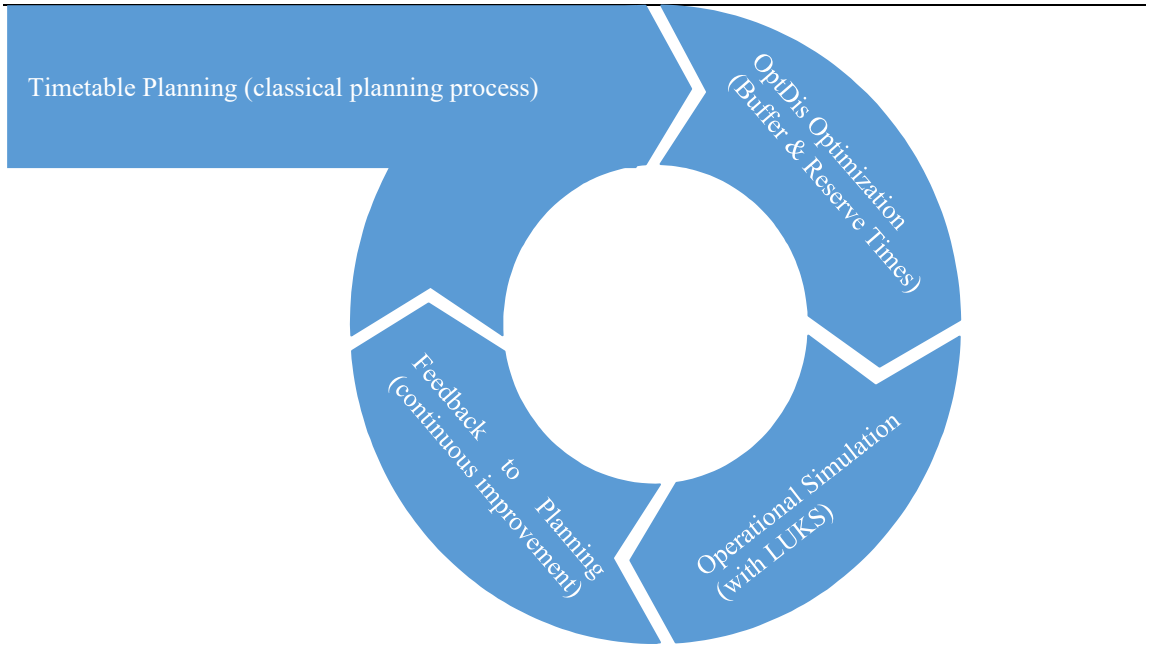


Figure 1: Overview of process

3 Methodology

3.1 Robustness and simulations

Robustness in railway timetables refers to the ability of a timetable to resorb disruptions without causing significant (secondary) delays or domino effects on other trains. By strategically incorporating buffer and reserve times and optimizing the timetable aiming to improve its resilience to delays, the impact of incidents can be minimized while maintaining operational quality. A robust timetable contributes to a more stable network, increases reliability for passengers and reduces the need for real-time interventions during operations.

To evaluate the robustness of timetables, our approach systematically compares two timetable variants: a reference timetable and a scenario timetable incorporating the proposed changes. The analysis begins with a conflict detection step to identify potential operational inconsistencies within the timetable. Subsequently, historical data is calibrated to generate realistic primary delays, which serve as input for the simulation phase. Using these inputs, operational simulations are conducted to assess the performance of both timetable variants. Finally, the resulting data is analyzed and visualized to quantify the effects on operational quality.

The simulation results are subjected to a number of statistical analyses to evaluate the robustness of the timetable. Both parametric (t-test) and non-parametric (Wilcoxon test) methods are used to determine the statistical significance of differences between the reference and scenario timetables. In addition, the median delay values are used as an indicator to identify the most affected trains and stations. This allows for a targeted interpretation of where and why trains are delayed and to have an insight in the underlying causes of delay propagation.

3.2 Timetabling

The construction of the annual timetable follows a demand based approach, whereby railway undertakings submit their path requests during predefined phases: the New Path Request phase (NPR) (starting on the second Monday of April for the following year's transport plan), the Late Path Request (LPR) phase and the Ad Hoc phase.

The starting point for the new timetable is the current operation timetable, which is incrementally adapted based on incoming demands. In case of conflicting path requests between different railway undertakings, a coordination procedure is initiated to reach a mutual agreement. If no consensus is achieved, the concerned infrastructure section is declared over-saturated, triggering formal over-saturation procedures.

Currently there are no strict guidelines or standardized methods for allocating buffer and reserve times within the timetable. The times are typically added where necessary based on operational experience, which does not always correspond to the most optimal placement from a robustness perspective.

3.3 Timetable optimization

To improve timetable robustness, we use the OptDis method integrated in LUKS, which uses the established approaches of blocking time theory and mixed-integer linear programming (MIP) to analyze timetables at the microscopic level and optimize them according to selected objectives.

As input, OptDis receives the pre-planned timetable agreed with all parties involved, which contains the planned route through the network with arrival and departure times at all stations for all planned trains. Additional information includes vehicle rotations between trains, connections and intervals.

From this, OptDis creates a mathematical MIP model that describes the set of all permissible and operationally sensible timetables. There is no rigid specification of rule sets, but rather a list of requirements for the result. These are in detail:

- The selected route for each train through the network must be able to be implemented by the interlocking system.
- The selected arrival and departure times and thus the running times must be realizable by the train. This is where the LUKS running time calculation comes into play, which takes into account the dynamic characteristics of the train, infrastructure characteristics such as gradients and speed restrictions as well as restrictions of the control and interlocking technology.
- During stops, the pre-planned tracks, departure times and minimum stopping times must not be changed. The running sequence between two conflict-free planned train runs must not be changed. This ensures that the coordinated timetable structure is maintained, and no further coordination is necessary.
- The blocking times of different trains resulting from the selected time-distance lines may only overlap if this is unavoidable due to the unchangeable constraints listed above.
- Linkages between trains such as decoupling, coupling, interval and connection must be observed.

In addition to constraints, a MIP model consists of an objective function that evaluates the quality of the solution with the help of an aggregated key figure. OptDis offers many setting options here. The following weighted combination of several objectives was selected for our approach:

- Primary: Minimizing the duration of conflicts that still exist in the input timetable. In other words: Minimization of “negative” buffer times; usually, this goal should be fulfilled by the timetable to optimize.
- Secondary: Maximization of buffer times and travel time reserves.
- Tertiary: Minimize deviations from planned arrival and departure times and routes. The constraints already contain hard conditions for maintaining the timetable structure. This objective further ensures that no unnecessary changes are made to the timetable if they do not lead to improved robustness.

A more detailed description of the OptDis approach can be found in [6] and [7].

The generated model is transferred to an external MIP solver. In our prototypical implementation, Gurobi is used [9]. The MIP solver searches the solution space given by

the model for an optimal solution, i.e. a valid adapted timetable which fulfills all the above constraints and whose objective function value is at least as good as the evaluation of any other valid timetable.

4 Improving robustness by timetable optimization

In theory, the best timetable quality and robustness can be achieved by integrating an optimization process directly into the timetabling process or even leaving the compilation completely to automatic optimization. In reality, such an approach is almost impossible to implement because timetabling relies heavily on coordination between the timetabler and the ordering transport companies. It would also require a complete reorganization of the existing timetabling process with correspondingly high risks and challenges.



Figure 2: Integration of OptDis into the current timetabling process at Infrabel

We are therefore pursuing a step-by-step introduction instead, which leaves the actual timetabling process unchanged in the first step and instead optimizes the finished timetable in a subsequent step as is shown in Figure 2. The existing timetable compilation process continues to focus on creating a market-driven, conflict-free timetable according to customer requirements.

In the pre-constructed timetable, the mathematical optimization procedure OptDis optimizes speed profiles, stopping times and routes in station heads without changing the existing timetable structure, including departure times, tracks, connections, vehicle rotations and intervals (cf chapter 3.3).

This approach takes advantage of the fact that all planned routes are known in their desired position and therefore a global optimization of all buffer times and travel time reserves in one fast step is possible.

The times and routes optimized by OptDis are fed back into the timetable creation process as changes and processed from there as usual. Since the actual timetable structure is unchanged, time-consuming coordination with the transport companies can be dispensed with.

5 Evaluation

5.1 Approach

Optimizing the finished timetable in a subsequent step limits the optimization possibilities through OptDis. To evaluate the actual effect on the operating quality, we examined different timetables with the help of operational simulations in the railway operation research application LUKS [8]. The simulations were carried out for selected zones of the Belgian infrastructure:

- Timetable “A”: Initial timetable without OptDis optimization
- Timetable “B”: Timetable optimized by OptDis with the aim of maximizing reserve and buffer times
- Timetable “C”: After the evaluation of delay data of simulations results of timetable “B”, a subsequent optimization with the aim of prioritizing reserve and buffer times takes place. This optimization focuses on relevant trains and locations based on historical knowledge.

The comparative evaluation of delays enabled the potential of this approach to be analyzed and demonstrated in a technically sound manner.

5.2 Considered zones of evaluation

For the microscopic robustness analysis of new transport plans, the Belgian rail network is divided into twelve regions. This division into different zones is to keep it feasible in terms of computational capacity and data processing. In this study, two regions were selected: Hasselt in the north of Belgium and Mons in the south of Belgium. These zones were chosen based on their high global delays, making them representative and relevant case studies for evaluating the impact of robustness optimization.

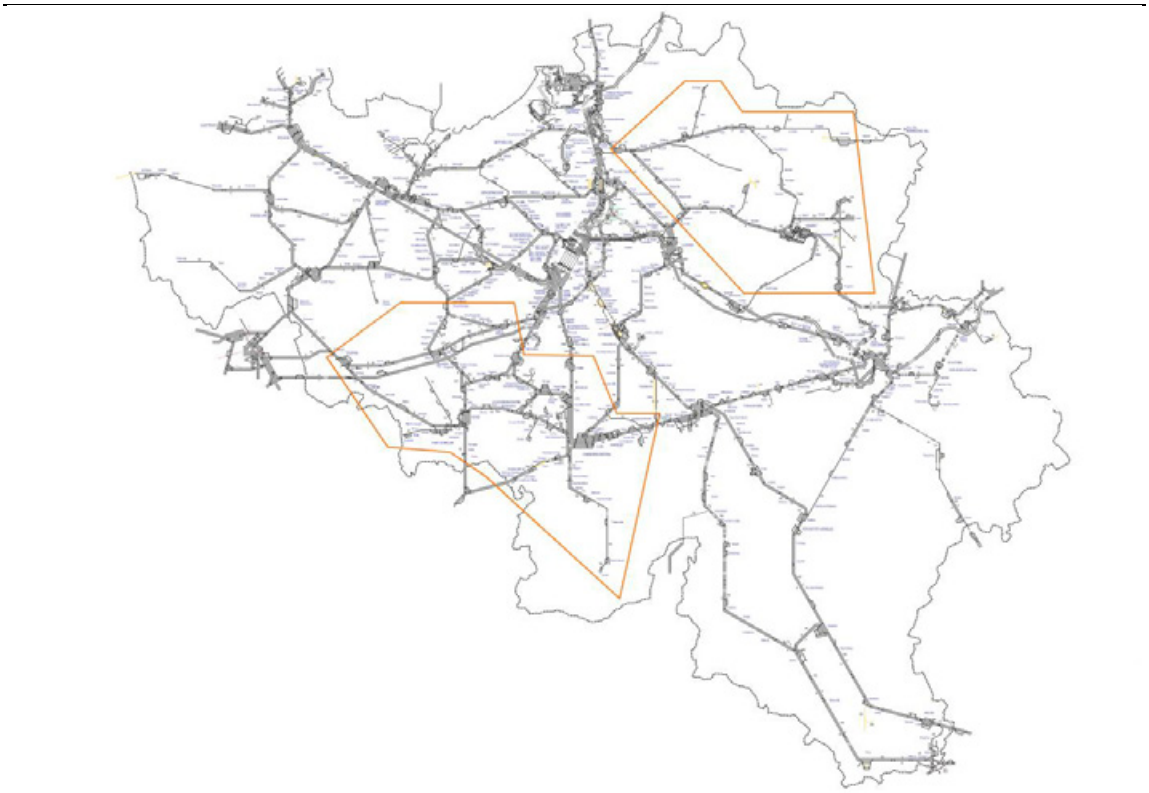


Figure 3: Considered zones in the Belgium network

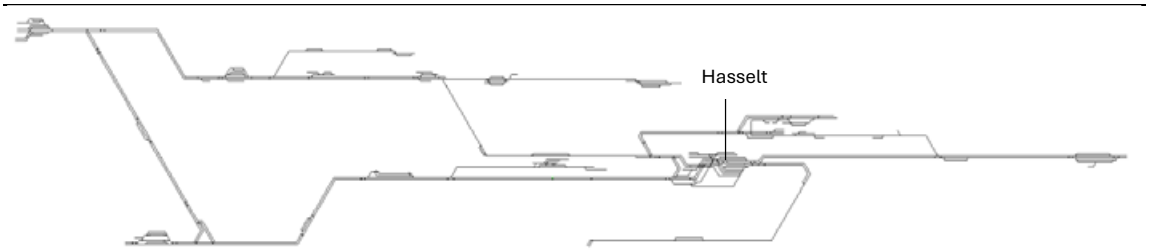


Figure 4: Visualization of LUKS infrastructure model of the zone Hasselt

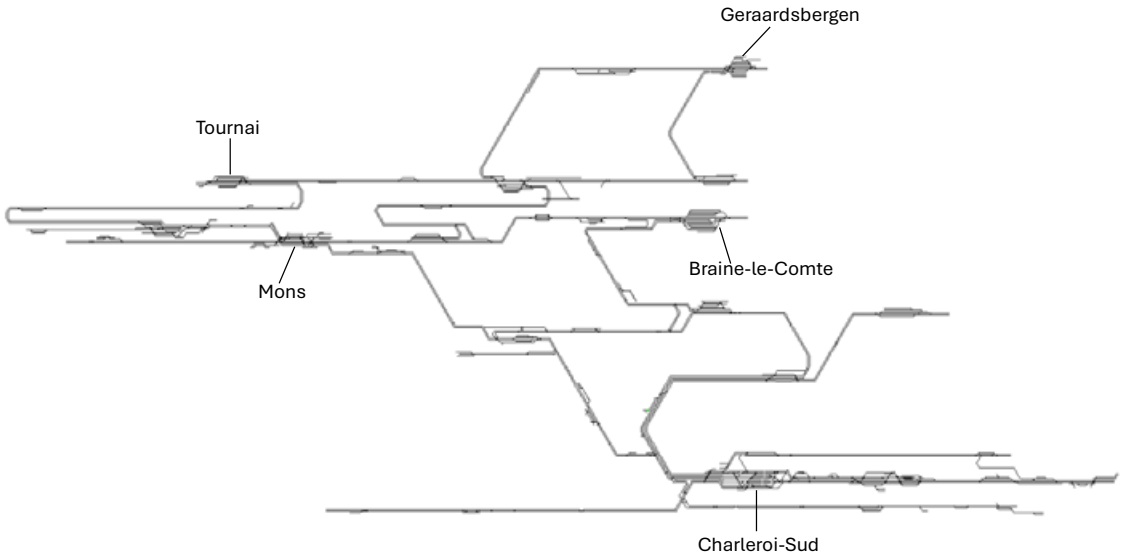


Figure 5: Visualization of LUKS infrastructure model of the zone Mons

The infrastructure used in the LUKS simulation environment is modelled at a microscopic level, incorporating detailed elements, such as signals, gradients and track configurations. This high level of detail ensures realistic and accurate simulation outcomes. Two distinct regions were analyzed: the region of Hasselt, including 89 stations and the region of Mons, with 168 stations. In the Hasselt area, the main lines are:

- L15 (Lier – Hasselt),
- L16 (Lier – Aarschot) and
- L35 (Aarschot – Hasselt).

For the Mons area, the main lines are

- L96 (Braine-le-Comte – Quévy),

- L124 (Nivelles – Charleroi-Central) and
- "la dorsale wallone" formed by the lines L78-L118-L130 (Tournai – Jemeppe-sur-Sambre).

The analysis is based on the first timetable of 2025, which started on 15 December 2024. The focus is on the peak of the morning (from 6h to 9h), when traffic density is at its highest. The selected timetable is not conflict free, which allows for a realistic optimization and robustness analysis under typical planning conditions. To maximize the number of trains, simulations were conducted for Wednesday, 22 January 2025, a representative weekday.

5.3 Results and comparison

The optimization procedure described in Chapter 4 was applied to the two zones, comparative simulations were carried out and the change in mean delay was evaluated.

The complexity of the input data and the required optimization times are shown in Table 1. It can be seen that even with complex zones, short calculation times can be achieved on a standard notebook (Intel Corel Ultra 7, 64 GB RAM), making a practical approach possible.

Table 1: Complexity of the input data and the required optimization times

Zone	Stations	Network size	Trains	Duration Optimization
Hasselt	89	532 km	155	2,9 min
Mons	168	1067 km	254	4,6 min

The mean delay is calculated across all train-station pairs within the selected zone. The analysis is restricted to the morning peak period from 06:00 until 09:00. For each case study, the same set of delays is used to ensure comparability, although the delay sets differ between the two zones. These delays are derived from historical delay data (via calibration). Each delay set consists of approximately 100 disruption lists, each representing a simulated day. For the zone Hasselt, 130 disruption lists were used, while 160 lists were used for zone Mons. The higher number for Mons reflects the greater complexity of the network in that region and ensures a sufficient number of valid simulations.

The delay is systematically compared between the baseline case (timetable "A") and each OptDis scenario (timetable "B" and "C"). Simulations that result in invalid outcomes – such as deadlocks or unresolved conflicts – are excluded from the analysis. If a disruption list leads to an invalid simulation in either the baseline case or the OptDis case, the corresponding results are discarded for both cases to maintain consistency.

Table 2: Mean delay and improvement

Area	A [s]	B [s]	C [s]	Improvement B vs A	Improvement C vs B
Zone Mons	121	115	---	4,96%	---
Zone Hasselt	153	147	---	3,92%	---
Charleroi (Mons)	92	86	86	6,52%	0,00%
Enghien (Mons)	181	179	179	1,10%	0,00%
Mons (Mons)	102	96	94	5,88%	2,08%
Nivelles (Mons)	188	190	189	-1,06%	0,53%
Aarschot (Hasselt)	145	141	141	2,76%	0,00%
Hasselt (Hasselt)	153	152	153	0,65%	-0,66%
Lier (Hasselt)	170	158	158	7,06%	0,00%
Mol (Hasselt)	128	128	132	0,00%	-3,13%

The results of the simulation are summarized in Table 2. For Mons, the mean delay was reduced by almost 5% across the entire zone. In view of the small amount of time required and the limited optimization options, this is an impressive improvement in timetable robustness. A noticeable reduction of almost 4% can also be achieved in the Hasselt zone.

For the focused "C" timetables, the four most relevant stations were identified based on the simulation results and given priority for optimization. In the Mons zone, the delay in the central Mons station can be reduced by a further 2% thanks to this procedure. The "A" timetable provides high buffer times in the area of Nivelles, which are redistributed as part of the optimization in the interest of a global improvement. This slightly worsens the delays in the "B" timetable by 1.06%. This effect can be partially reduced by focusing.

The Hasselt zone shows the limits of focusing. Here, focusing does not lead to any improvement and in some cases even worsens the delay values. This is due to network effects caused by the topology of the zone: disadvantaging some areas at the expense of others leads to higher delays in the disadvantaged areas, which propagate from there

through the entire network and therefore also have a negative impact on the preferred areas.

The usefulness of "C" timetables is therefore heavily dependent on the infrastructure under consideration and cannot be recommended without reservation. Global optimization appears to be more useful due to possible network effects.

6 Conclusions and outlook

Optimization processes can make a valuable contribution to improving operational quality by creating more robust timetables with better distributed buffer times and travel time reserves. However, their introduction into existing timetable processes is challenging because these require close coordination between the timetable creator and the customer.

As a first step, we have therefore developed a procedure that improves the robustness of a fully constructed timetable through subsequent optimization without changing the coordinated timetable structure. Realistic operational simulations with LUKS show that this approach can reduce the average delays by a significant value of 4% to 5%. The optimization only takes a few minutes and can therefore be integrated into an existing process without noticeably slowing it down.

In a further step, the optimization was focused on areas that are particularly relevant for delays. The simulations show that the approach only partially offers further potential for improvement due to network effects in delay propagation. Its use is therefore only recommended to a limited extent, preferably in combination with an evaluation using operational simulations.

The procedure is currently limited to two zones in the Belgian network. An extension to the entire network will therefore be investigated next. As the timetable structure is not changed by the procedure, we are confident that each zone can be optimized individually, and the results subsequently linked to form a network-wide timetable. In addition, the potential of other focusing approaches and a slightly greater freedom of OptDis in optimization will be investigated.

Our method is only a first step towards improving the robustness of timetables using optimization techniques. In subsequent steps, the impact can be improved by using the procedure during the actual timetable creation process, e.g. by providing the timetabler with suggestions for optimizing travel time reserves and buffer times during the editing process.

Literature

[1] Burggraeve, S.; Vansteenwegen, P.: (2017). Robust routing and timetabling in complex railway stations. In: Transportation Research Part B 101 (2017), S. 228–244.

[2] Högdahl, J.; Bohlin, M.; Fröidh, O.: A combined simulation-optimization approach for minimizing travel time and delays in railway timetables. In: Transportation Research Part B 126 (2019), S. 192–212.

[3] Van Hoeck, I.; Vansteenwegen, P. (2024): A MILP model to improve the robustness of a railway timetable by retiming and rerouting in a complex bottleneck area. In: Journal of Rail Transport Planning & Management, 32 (2024).

[4] Hansen, IA; Pachl, J (eds.) 2008. Railway timetable & traffic, Eurailpress, Hamburg.

[5] Büker, T.; Grafagnino, T.; Hennig, E.,; Kuckelberg, A.: Enhancement of blocking-time theory to represent future interlocking architectures. In: Proceedings 8th international conference on railway operations modelling and analysis–Rail Norrköping (2019), S. 219-240.

[6] Weymann, F.; Nießen, N.: Unterstützung der Fahrplanfeinkonstruktion mit Optimierungsverfahren. In: Eisenbahntechnische Rundschau 64 (2015) 3, S. 16 –19

[7] Weymann, F.; Nießen, N.: Verbesserung der Disposition des Eisenbahnbetriebs durch innovative Optimierungsverfahren. In: ZEVRail Glasers Annalen 139 (2015), S. 1-2

[8] Janecek, D.; Weymann, F.; Schaer, T.: LUKS – Integriertes Werkzeug zur Leistungsuntersuchung von Eisenbahnknoten und -strecken. In: Eisenbahntechnische Rundschau 59 (2010) 1+2, S. 25–32.

[9] Gurobi Optimization, LLC: Gurobi Optimizer Reference Manual. https://www.gurobi.com, 2025

Session 8

Frédéric Weymann, David Janecek, Katrien Pauwels, Anneleen Van de Mergel, Quentin Van Leeckwyck

Authors



Weymann, Frédéric

Dr rer. nat Frédéric Weymann has been working as a software architect at quattron GmbH since 2015. After studying computer science at RWTH Aachen University, he worked as a research assistant at the Institute of Transport Science at RWTH Aachen University, where he gained his PhD in 2011.



Janecek, David

David Janecek is a partner at quattron GmbH. After studying computer science at RWTH Aachen, he worked as a research assistant at the Institute of Transport Science at RWTH Aachen before joining quattron.



Pauwels, Katrien

Katrien Pauwels studied mathematics at University Ghent. She started working as a business analyst at Infrabel's IT department in 2000. In 2019 she made a switch to the department "Long term (train and works) planning" where she's now responsible for the communication and robustness analysis on this long term planning, as well as the supporting IT-development



Van de Mergel, Anneleen

After studying commercial engineering at KU Leuven campus Brussels, Anneleen Van de Mergel started her career at Infrabel in 2009. Initially focusing on capacity studies, her expertise later shifted to robustness studies. Since 2020, she has led the robustness team as team lead, where she is committed to the optimisation and resilience of the rail network



Van Leeckwyck, Quentin

Quentin Van Leeckwyck studied chemical engineering at ULB (Brussels). After some years in private companies and at SNCB, he joined the robustness team of Infrabel in 2022.

IRSA 2025

SESSION 9

FAHRKOMFORT
RIDE COMFORT

Session 9
Guozhen Jing, Hauke Schmidt, Xin Ding, Chutan Xue, Gang Chen

The sensitivity of simulated ride comfort of the tail coach in a high-speed EMU to unsteady aerodynamic loading and variation in suspension parameters

Jing, Guozhen[1], Schmidt, Hauke[1], Ding, Xin[2], Xue, Chutan[1] and Chen, Gang[1]

[1]CRRC CHANGCHUN GERMANY RailTech GmbH, Pascalstraße 19, 52076 Aachen, Germany
[2]CRRC Changchun Railway Vehicles Co., Ltd, Changchun, 130000 Jilin, China

Abstract

In recent years, unsteady aerodynamic loading has been identified as a key contributor to the deterioration of ride comfort due to low-frequency swaying in the tail car of certain Chinese high-speed EMUs. Prior studies applied precomputed aerodynamic loads as known loads to the multi-body dynamics model, neglecting variations in aerodynamic loading induced by the vehicle's own unsteady body motion. To capture and assess the effect of this neglected aspect of unsteadiness, a co-simulation approach utilizing panel aerodynamics was adopted in this study. The resulting impact on ride comfort in the tail car of a Chinese high-speed EMU is analyzed and compared against that of variations in suspension and environmental parameters.

Keywords: High-speed train; Co-simulation; Unsteady aerodynamics; Panel aerodynamics; Ride comfort

1 Introduction

With the increasing speed of modern high-speed trains, aerodynamic disturbances generated by trains are expected to have a growing impact on their running behavior, leading to a stronger and more complex coupling between the aerodynamics and running dynamics of high-speed trains. In recent years, researchers have identified unsteady aerodynamic loads as a key contributor to the low-frequency swaying observed in the tail car of certain Chinese high-speed EMUs, which in turn resulted in the deterioration of ride comfort. Li et al. [1] found that the application of unsteady aerodynamic loads, characterized by significant oscillations due to vortex shedding, on a 3-car train traveling at 400 km/h resulted in low-frequency oscillations in the lateral response of the tail car. Ding et al. [2] applied precomputed unsteady aerodynamic loads to a 3-car MBD model of a Chinese high-speed EMU, known to experience low-frequency swaying in its tail car when passing stations turnouts at 300 km/h. They found that the presence of unsteady aerodynamic loads substantially increased the oscillations in the lateral accelerations of the tail car, with the main frequency peak lying close to the experimentally measured value of 1.45 Hz.

Both studies adopted a unidirectional coupling approach, where aerodynamic loads are precomputed using prescribed vehicle motions and applied as known loads to the train's multi-body dynamics (MBD) model. While such approaches capture flow unsteadiness due to vortex shedding, they neglect the influence of unsteady body motion on the airflow in the near-field, which can in turn modify the aerodynamic loading on and thus dynamic response of the vehicle. To capture such interactions, a bidirectional coupling, or co-simulation, between aerodynamics and running dynamics is required. Existing research on co-simulations focuses on scenarios with prominent aerodynamic forces, such as crosswind [3, 4] or tunnel operations [5, 6, 7], but not the open-air design conditions, under which [1, 2] conducted their research on car body sway. Additionally, due to the high computation cost of typical high-fidelity computational fluid dynamics (CFD) solvers, the physical time simulated by such studies are often limited to about 10 s, which may not be sufficient for coupled interactions to fully evolve or to characterize low frequency oscillations with sufficient resolution.

This study addresses the above issue by replacing high-fidelity CFD by a lower fidelity aerodynamic solver, allowing for longer co-simulations with reduced computational cost. Under open-air design conditions, where the incoming flow is nearly axial, flow separation

typically only occur at sharp corners and on the nose of the tail car [8]. If vortex shedding effects, having been previously investigated by [1, 2], are omitted, the flow outside the boundary layer can be modeled by an inviscid, irrotational, and incompressible flow. These assumptions allow the flow velocity to be written as the gradient of a scalar potential φ, reducing the continuity equation to the Laplace's equation:

$$\nabla \cdot \mathbf{v}(\mathbf{x},t) = \nabla^2 \varphi(\mathbf{x},t) = 0 \tag{1}$$

which is the governing equation for *potential flow*. Due to linearity, solutions can be formed by superimposing elementary solutions such as sources and irrotational vortices, the so-called *singularity functions*. A numerical solution of the flow past 3D bodies can be obtained via *panel methods*, where solid boundaries are discretized by *panels* containing such singularity functions. Obtaining the coefficients of the singularity functions that result in zero normal velocity on solid boundaries yields the velocity field, from which pressure can be derived using Bernoulli's theorem[1].

Though viscous effects are neglected in panel methods, they provide fast solutions that accurately capture the pressure-based aerodynamic forces on streamlined bodies (cf. section 3). For this reason, they have been previously used for the evaluation of unsteady pressure variations during open-air passing in the TRANSAERO project [10, 11]. In the aerospace industry, panel methods remain an important early-stage design tool, where the computational costs associated with high-fidelity CFD methods are prohibitively high for design iteration [12]. In this study, a commercial panel method solver, USAERO, is adopted as the lower fidelity aerodynamic solver for the partitioned co-simulation of a Chinese high-speed EMU train. Through this approach, the effect of coupling such unsteady aerodynamic forces on the simulated ride comfort of the tail coach in a Chinese high-speed EMU will be analyzed and its impact will be characterized by comparison to the impact of variations in selected suspension parameters.

2 Multi-body dynamics model

In accordance with design specifications, the 3-car MBD model of a certain Chinese high-speed EMU were established in SIMPACK based on existing validated single-car models

[1] For a more thorough theoretical treatment on potential flow and panel methods, please refer to *Low Speed Aerodynamics* by Katz and Plotkin [9]

developed by Ding et al. [13] for wear prediction. New wheel and rail profiles are used, resulting in a low equivalent conicity of 0.1 at the rail-wheel interface. Unless otherwise configured, the friction coefficient μ was taken to be 0.4. The rigid-body modes of the standalone tail car for a speed of 300 km/h are summarized in Table 1 per the nomenclature used in the Handbook of Railway Dynamics by Iwnicki et al. [14]. Here, the two hunting modes can be distinguished by the motion of the bogies: in hunting A, the bogies move in-phase; in hunting B, the bogies move out of phase.

Table 1: Rigid body modes of the tail car

Mode	Frequency [Hz]	Damping [-]
Lower sway	0.37	7.5%
Yaw	0.60	65.8%
Bounce	0.70	12.9%
Pitch	0.84	14.8%
Upper sway	1.41	44.6%
Hunting A	1.79	17.1%
Hunting B	1.93	31.8%

The rigid car body modes of the 3-car train at 300 km/h is summarized in Figure 1, where the lateral modes of the car have been plotted with blue crosses and other modes with orange circles. Here, a greater degree of coupling is observed in the lateral dynamics of the 3-car train, with mixed modes of yaw and lower sway emerging at lower frequencies and mixed modes of upper sway and hunting at higher frequencies. Additionally, the presence of inter-car connections couple the motion of neighboring cars, leading to many of the modes involving multiple cars. Consequently, two coupled modes between longitudinal and pitching motions emerge at 1.47 and 1.48 Hz, one of which is illustrated in the bottom right of the figure. Apart from this, minimal coupling is observed in the vertical dynamics of the vehicle, with bounce and pitch modes remaining distinct and decoupled from other modes. Overall, the general range of frequencies and damping is consistent with those of the single car tabulated in Table 2.

For the purposes of the present study, a straight track with track irregularities generated from the Chinese high-speed rail ballastless spectrum reported by Kang et al. [15] was used. The full range of spatial frequencies defined in the spectrum, i.e. between 0.01 and 0.5 m^{-1}, was used for the generation of distance domain irregularities. To allow sufficient

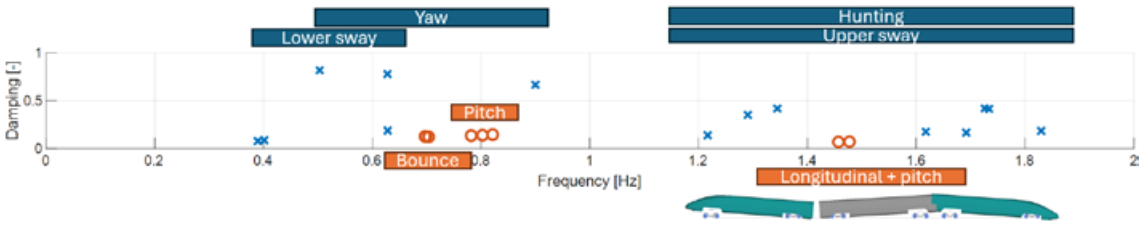


Figure 1: Distribution of car body modes in the 3-car train

time for the co-simulation (section 4) to reach equilibrium, the initial 420m of track contain no irregularities.

3 Aerodynamic modeling

The commercial unsteady panel solver USAERO is employed in this study as the aerodynamic solver. For this solver, the external geometry of the EMU has been simplified by omitting the pantograph, running gear, and inter-car gaps, resulting in the discretization shown in Figure 2. As only a relatively coarse surface mesh is required in panel solvers, the resulting mesh of the 3-car train contains only 2816 panels. Moreover, since the surface of the solid boundaries constitute the entirety of the computational domain in panel methods, general body motions of the train can be directly prescribed with no additional treatment. In contrast, typical high-fidelity computational fluid dynamics (CFD) approaches such as Reynold-Averaged Navier Stokes (RANS) require a relatively fine volume discretization of the entire fluid domain, resulting in typical cell counts in the tens of millions for trains. Additionally, as motion of the train significantly changes the fluid domain, techniques such as overset meshing are required to deal with general body motion of the train.

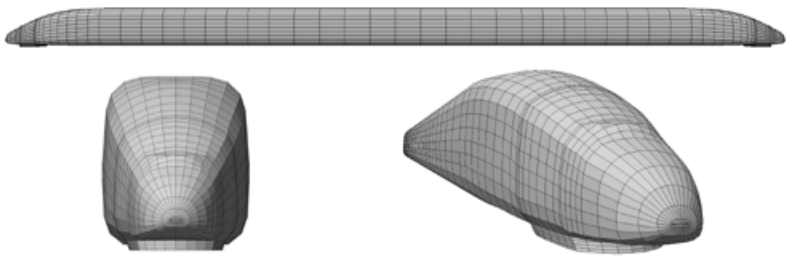


Figure 2: Paneling of simplified surface geometry used in USAERO

To assure the validity of using panel aerodynamics to obtain the surface pressure on the high-speed EMU, a RANS simulation was carried out in StarCCM+ using the k-ω-SST

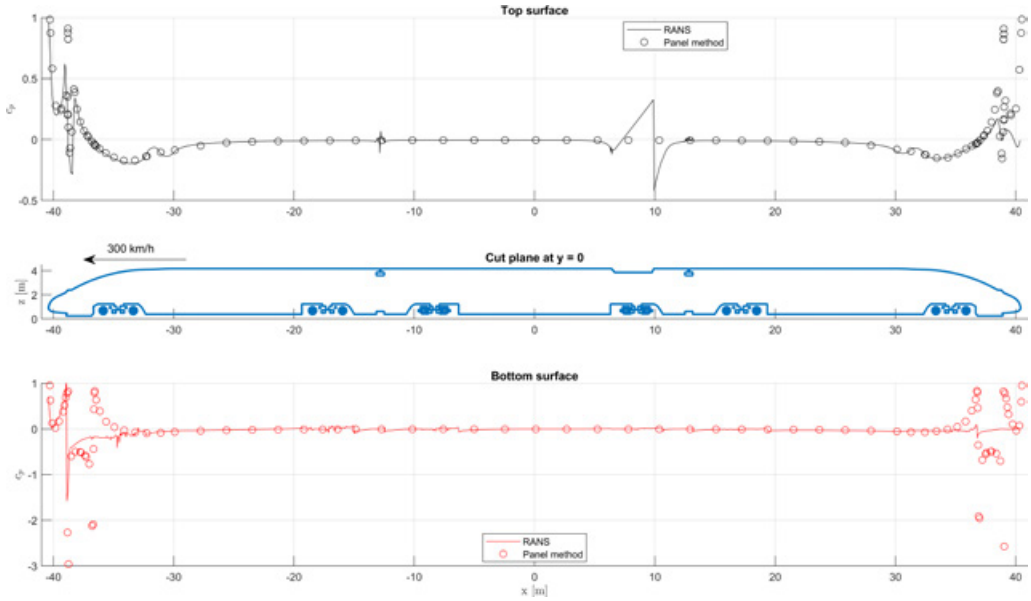


Figure 3: Comparison of c_p predicted by RANS and panel method along EMU centerline

turbulence model for the full-scaled 3-car train traveling at 300 km/h for comparison against USAERO. For the purposes of this comparative study, only the pantograph of the 3-car train was omitted in the RANS model, while inter-car gaps, pantograph bay and simplified running gear geometries were preserved, resulting in a total of 50 million cells. To simulate the steady-state motion of the train, an incoming airflow is prescribed in conjunction with a moving ground.

A comparison of the predicted pressure distribution by the two approaches are given in Figure 3. Here, the top and bottom plots show respectively the pressure coefficient along the centerline of the EMU on its top and bottom surfaces, with RANS results plotted as solid lines and panel method results as hollow circles. The middle plot shows the cross-section of the 3-car train model in the RANS simulation along its centerline.

Over most of the EMU's length, the pressure distribution along its centerline predicted by panel method and RANS are in good agreement, particularly on the upper surface. Discrepancies arise in two situations:

1. where effects of separated flow are prominent and hence the assumptions of potential flow are invalid, e.g. flow behind the leading edge of the cowcatcher or flow near the tip of the tail car;
2. where changes in the surface geometry have not been adequately captured by the

paneling used, most prominently the pantograph bay at x ∈ [7, 10] m, which has been omitted in the simplified USAERO model (Figure 2).

Considering that these regions of poorer agreement account only for a small part of the EMU's surface area and the significantly lower computation cost of panel methods, the discrepancies were deemed to be acceptable: while the panel solver USAERO took 20 s on 4 cpu cores, the RANS solution took StarCCM+ 63000 s to compute on 60 cores. Hence, under normal operating conditions with minimal crosswind, panel codes can capture the unsteady pressure-based aerodynamic loading acting on high-speed EMUs with reasonable accuracy.

4 Coupling framework and simulations setup

The MBD solver SIMPACK and panel-based unsteady aerodynamics solver USAERO are coupled via a non-iterative Gauss-Seidel co-simulation approach [16], as shown schematically in Figure 4. In such an approach, data is exchanged between the MBD and aerodynamic solvers at discrete communication intervals to model the physical interactions between the two subsystems, with SIMPACK providing USAERO with the motion of the car body and USAERO supplying SIMPACK with the aerodynamic forces. To facilitate this data exchange and to control the overall co-simulation process, an in-house coupling interface was developed in Python.

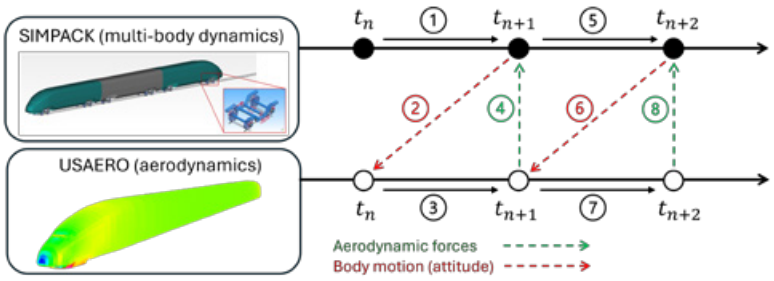


Figure 4: Co-simulation approach used in this paper

Using the aforementioned framework, time domain co-simulations were conducted at a constant speed of 300 km/h for a duration of 100 s on a straight track section. To better assess the effect of coupling with unsteady aerodynamic loads, the same simulations were repeated using SIMPACK only (i.e. in the absence of aerodynamic loads) and the results analyzed and compared. In the SIMPACK-only simulations, additional cases with modified

suspension characteristics are included for comparison. A summary of the cases considered is provided in Table 2.

Table 2: Cases considered for comparison

	Unsteady aerodynamic loads	Modifications
Case 1	NO	Nominal state
Case 2	NO	Low adhesion ($\mu = 0.2$)
Case 3	NO	Anti-yaw damper (AYD) stiffness +20%
Case 4	NO	Vertical damper (VD) damping -20%
Case 5	YES	Nominal state

5 Results and discussion

The rolling and pitching accelerations of the tail car about its center of gravity (CG) are plotted for each of the tested cases against time for the first 20 s in Figure 5. Apart from initial oscillations in the pitching acceleration of the co-simulation (purple curve, right plot), similar acceleration profiles with differing amplitudes are observed across all tested configurations. The initial oscillations observed are transient numerical artifacts resulting from the abrupt introduction of aerodynamic loads in the initialization phase of the co-simulation and will be precluded from subsequent analysis by discarding the first 10 s of the simulation.

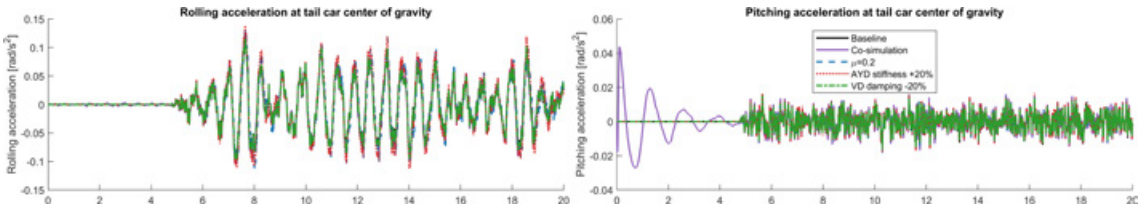


Figure 5: Rolling (left) and pitching (right) accelerations of the tail car

The accelerations at the middle of the car body floor were filtered and weighed per EN12299 to obtain 5-second weighted root-mean-square values, referred to as continuous comfort indices. These indices, plotted in Figure 6 for all simulated configurations, serve as quantitative metrics for ride comfort, with higher values indicating poorer ride comfort.

In the longitudinal direction, along which the relative wind mainly acts, the comfort index resulting from the co-simulation (purple curve) differs most significantly from that in

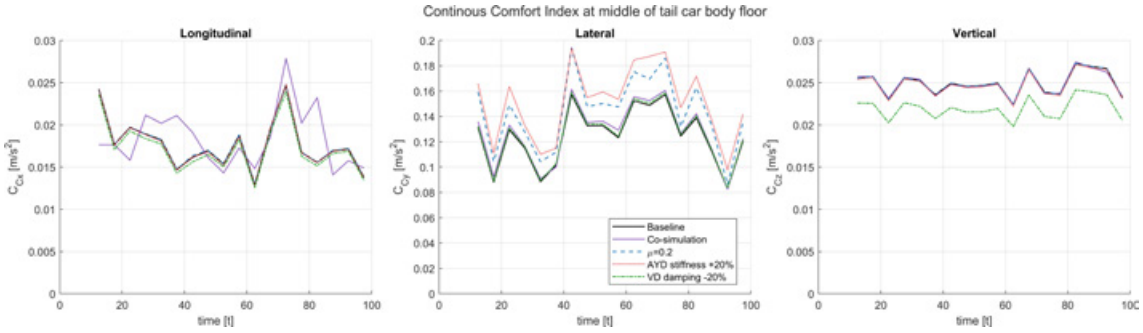


Figure 6: Continuous comfort indices at middle of tail car's floor

the reference configuration (black curve) across all simulated configurations. In contrast, variations in suspension parameters have stronger influence on the comfort indices in the lateral and vertical directions. While coupling with unsteady aerodynamics (purple curve) resulted in a 2-5% increase in the lateral comfort index over the reference configuration (black curve), a 10-20% and a 15-25% increase may be attributed to the cases of low adhesion (blue curve) and increased anti-yaw damper stiffness (red curve), respectively. In the longitudinal direction, while the decreased vertical damper damping (green curve) resulted in a 12% decrease in the lateral component index, no significant difference was observed in any other simulated configuration.

The insensitivity of vertical comfort index to the anti-yaw damper stiffness and that of the lateral comfort index to the vertical damper damping is consistent with the lack of coupling between the vehicle's lateral and vertical dynamics (cf. section 2). Overall, as the comfort indices in the longitudinal and vertical directions are one order of magnitude lower than that in the lateral direction, the ride comfort in the tail car is expected to be determined primarily by its lateral comfort index.

To examine the differences in the tail car's dynamic response across simulated configurations in greater detail, power spectral densities (PSD) for the tail car's acceleration were estimated for all six degrees of freedom (DoF) over the time interval from 10 to 100 s (Figure 7). The estimates were obtained by averaging overlapping periodograms using Welch's method [17], with 50s segments, 80% overlap, and flat-top windows. For consistency, the same averaging is used in this paper for all frequency domain estimates.

The amplifications of the main frequency peaks in the PSD of lateral, rolling and yawing accelerations (right column, Figure 7) relative to the reference configuration mirrors the previously observed increases in the lateral comfort index, with the stiffening of the anti-

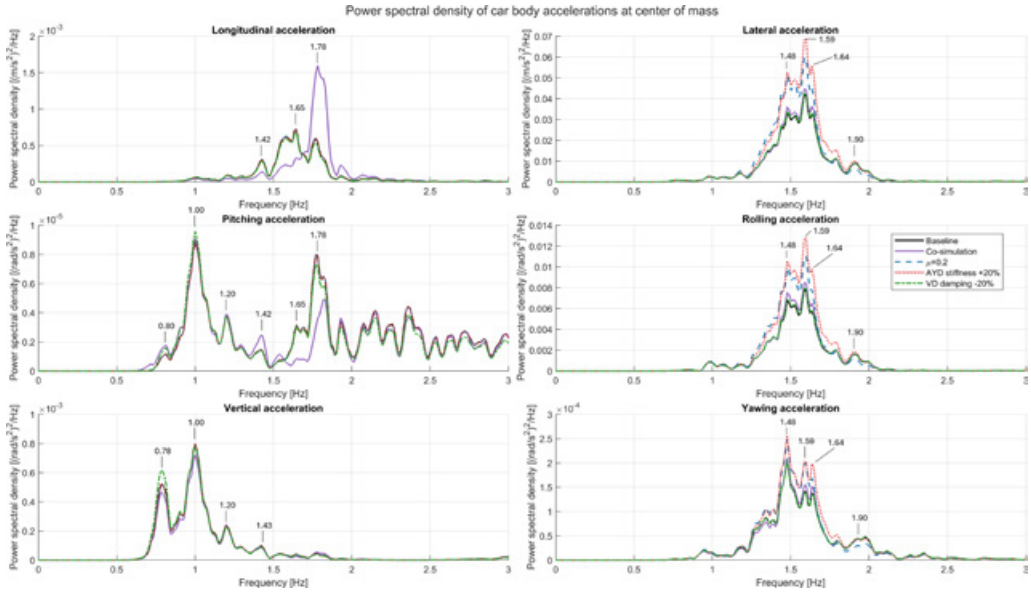


Figure 7: Power spectral densities of accelerations at the tail car CG

yaw damper (red curve) and low adhesion (blue curve) resulting in larger increases than aerodynamic coupling in the co-simulation (purple curve). For the longitudinal, pitching, and vertical accelerations (left column), however, frequency dependent effects are observed for both the co-simulation and decreased vertical damper damping: co-simulation generally amplifies the higher frequency components and attenuates lower ones, while the opposite is true for decreased vertical damper damping.

Despite these amplitude variations, the location of frequency peaks for a given DoF is largely consistent across all tested configurations. Additionally, matching frequency peaks are observed in the PSD of lateral, rolling, and yawing accelerations, with the highest responses occurring over the same frequency band between 1.4 and 1.7 Hz in all cases. Similarly, the PSD of pitching acceleration (middle left) shares peaks with those of longitudinal (top left) and vertical acceleration (bottom left) in the 1.4–2 Hz and 0.7–1.5 Hz bands, respectively.

In combination with the previously observed similarity in acceleration time histories (Figure 5), the consistency of frequency peak locations across configurations with different suspension and environmental conditions suggest that the tail car accelerations are primarily driven by a factor common to all cases — namely, excitation from the wheel-rail contact.

To confirm this, the magnitude-squared coherence (MSC) between the tail car's accelera-

tions and the track irregularities encountered by its leading wheelset was computed using Welch's estimate (Figure 8). For two signals $x(t)$ and $y(t)$ with cross power spectral density $P_{xy}(f)$ and respective PSDs $P_{xx}(f)$ and $P_{yy}(f)$, the MSC is given by

$$|\gamma_{xy}(f)|^2 = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)}$$

a scalar function of frequency bounded between 0 and 1 that provides an indication of the extent of linear relation between the two signals at each frequency [18]. A value close to 1 in a given frequency range indicates that the relationship between the two signals is well-described by a linear time-invariant system for that frequency range, while a value close to zero indicates the absence of a linear relationship.

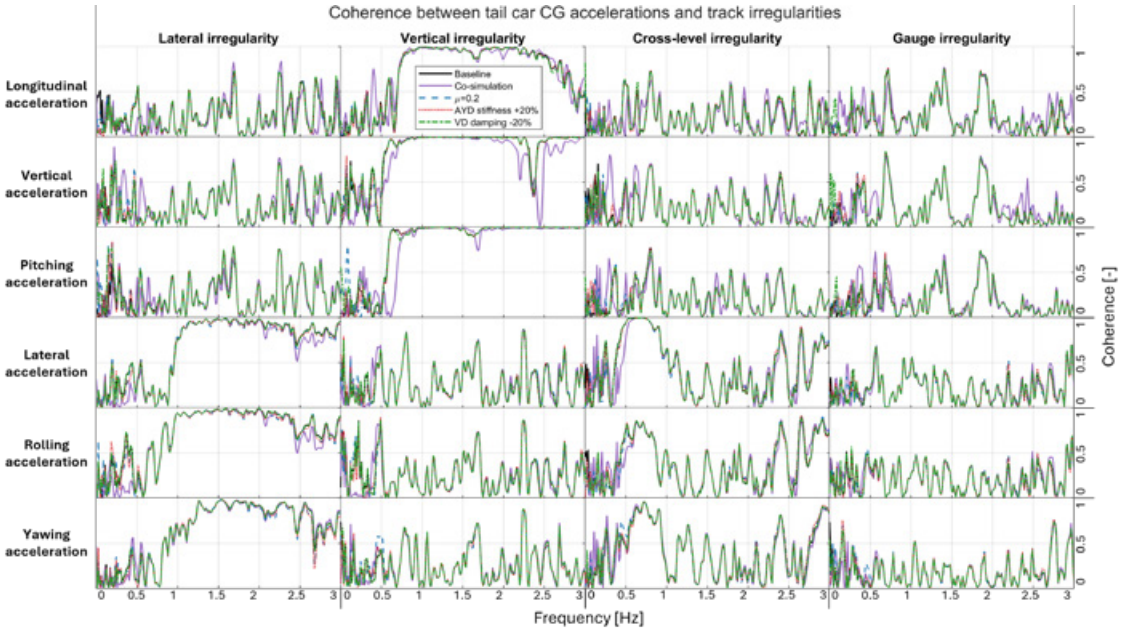


Figure 8: MSC between tail car CG accelerations and track irregularities

At low frequencies below 0.5 Hz, the MSC (Figure 8) between car body accelerations and track irregularities are low and inconsistent across configurations. This is expected: as the smallest spatial frequency contained in the generated track irregularity is 0.01 m^{-1}, the lowest time frequency domain excitation is 0.833 Hz at a speed of 300 km/h. At higher frequencies, the MSC values are largely consistent across configurations, differing only slightly in the case of the co-simulation (purple curves) due to the additional excitation from unsteady aerodynamic loads. MSC estimates approach unity in the following cases,

indicating strong linear correlation in these regions:

1. Between the lateral, rolling, and yawing accelerations of the tail car and:
   a) the lateral irregularities (first column) above 1 Hz;
   b) the cross-level irregularities (third column) from 0.5 to 1 Hz;
2. Between the longitudinal, vertical and pitching acceleration and the vertical track irregularity (second column) above approximately 0.6 Hz.

As these regions encompass all significant peaks in the PSD of tail car accelerations (Figure 7), track irregularities are likely the primary source of excitation in all simulated configurations. Considering the car body modes discussed in section 2, the following inferences can be drawn:

1. The cross-level irregularities excite the yaw and lower sway modes at the lower frequency ranges below 1 Hz. Due to the absence of substantial track irregularity excitations below 0.8 Hz and the high damping in the yaw mode, however, no significant car body accelerations occur.
2. Lateral track irregularities excite the mixed modes consisting upper sway and hunting, resulting in the pronounced peaks between 1.4 and 1.7 Hz observed in the PSD of lateral, roll, and yaw accelerations.
3. Vertical track irregularities excite the bounce, pitching, and coupled longitudinal-pitching modes, leading to the observed longitudinal, pitching, and vertical accelerations. This is supported by the lower frequency range of prominent vertical accelerations compared to longitudinal accelerations, reflecting the lower eigenfrequency of the bounce mode compared to the coupled longitudinal-pitching mode.

To characterize the influence of unsteady aerodynamic forces, the phase shift between aerodynamic loads and the velocity along the same DoF was determined in addition to the MSC for each DoF at the tail car's CG. This is presented in Figure 9, where blue lines show the MSC and orange circles indicate the phase shift. Note that the phase shift was only plotted where the MSC exceeds 0.6, as values below this threshold indicate weak correlation and hence unreliable phase information.

For the lateral, roll, and yaw DoFs, the aerodynamic loads on the tail car are well-correlated with the corresponding velocities between 1.3 to 1.8 Hz, which includes the frequencies

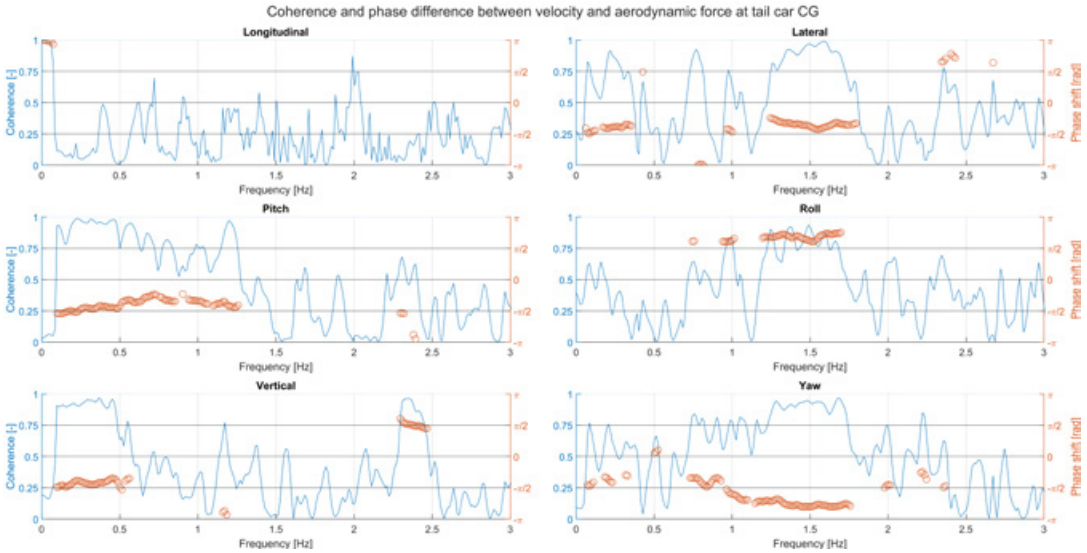


Figure 9: MSC and phase difference between tail car velocity and aerodynamic loads

associated with the highest acceleration responses (cf. Figure 7). Within this range, the phase shift for the lateral DoF lies closer to 0 rad, suggesting that the lateral aerodynamic forces act largely in phase with the lateral velocity and thus tend to amplify the motion. In contrast, the phase shifts for roll and yaw DoFs are closer to $\pm\pi$ rad, indicating that the aerodynamic moments act more out of phase with the corresponding rotational velocities and thus tends to dampen the motion. These opposing influences result in a net destabilization of the lateral dynamic of the tail car, as previously observed.

In contrast, the correlation between aerodynamic loads and body velocities is notably weaker for the longitudinal, pitch, and vertical DoFs at frequencies where the largest deviations in the corresponding acceleration PSDs exist between the co-simulation and reference cases (e.g. between 1.4 and 1.8 Hz for pitch). The poor linear correlation could be attributed to nonlinear aerodynamic effects or to time-varying system characteristics, with the latter being supported by the observation that the longitudinal comfort index in the co-simulation is neither consistently above nor below that of the reference case (Figure 6). Consequently, the previously noted frequency-dependent features in the PSD of longitudinal and pitching accelerations (Figure 7) may not accurately reflect the underlying vehicle dynamics.

Additionally, the simultaneous excitation of multiple modes across different frequencies made it difficult to identify specific mechanisms behind the bi-directional coupling of

unsteady aerodynamics and running dynamics. For instance, as the amplitude of the tail car's response and active modal content also vary with frequency, the influence of frequency, response amplitude, and active modal contents cannot be clearly separated.

Overall, the existence of unsteady aerodynamic effects due to body motion of the EMU has been clearly established in the present study. This was shown to amplify the dynamic response of the tail car, leading to a notable but slight deterioration of the overall ride comfort.

6 Conclusion

In the current study, the sensitivity of the ride comfort of the tail coach of a high-speed EMU to unsteady aerodynamics and variations in suspension parameters have been investigated via co-simulation of MBD and panel-based aerodynamics solvers. Overall, the effect of unsteady aerodynamic forces on the simulated comfort indices was relatively minor compared to that caused by variations in suspension parameters.

Nevertheless, the bidirectional coupling approach presented in this paper may be useful as a preliminary assessment for the presence of coupling effects when designing of future high-speed trains, which are expected to have higher speeds and lower masses. Additionally, as modern panel methods continue to be extended with additional capabilities [12], future developments may allow partial modeling of the presently ignored vorticial effects.

Furthermore, neither the unidirectional coupling with high fidelity CFD solvers nor the co-simulation with panel-based aerodynamics can capture possible changes in vortex shedding locations induced by the body motion of the train. Here, co-simulation with high-fidelity CFD solvers are required. Though Song et al. [6] performed this investigation for a 160 km/h EMU in single-track tunnels, no similar investigation has been done for high-speed EMUs to the authors' knowledge. This thus remains a point for future investigation regarding the swaying issue in high-speed EMUs.

Lastly, in the current study, multiple frequencies and modes are excited simultaneously by track irregularities, making an in-depth analysis difficult. Future studies may consider the use of artificial sinusoidal track irregularities to better isolate the coupling characteristics associated with individual eigenmodes.

Literature

[1] T. Li, Y. Li, L. Wei, and J. Zhang, "Study on lateral vibration of tail coach for high-speed train under unsteady aerodynamic loads," *Vibration*, vol. 6, no. 4, pp. 1048–1059, 2023.

[2] X. Ding, C. Chang, L. Ling, Z. Sun, Y. Yu, L. Zhang, and W. Zhai, "Mechanism analysis of low-frequency swaying motion of high-speed trains induced by aerodynamic loads," *Journal of Vibration and Control*, vol. 30, no. 13-14, pp. 3141–3153, 2024.

[3] W. Zhai, J. Yang, Z. Li, and H. Han, "Dynamics of high-speed train in crosswinds based on an air-train-track interaction model," *Wind and Structures*, vol. 20, pp. 143–168, 02 2015.

[4] T. Li, M. Yu, J. Zhang, and W. Zhang, "A fast equilibrium state approach to determine interaction between stochastic crosswinds and high-speed trains," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 143, pp. 91–104, 2015.

[5] Z. Ji, W. Liu, D. Guo, G. Yang, and J. Mao, "Analysis of the fluid–structure coupling characteristics of a high-speed train passing through a tunnel," *International Journal of Structural Stability and Dynamics*, vol. 22, no. 16, p. 2250185, 2022.

[6] Y. Song, T. Qin, Y. Yao, and F. Dai, "Investigation on aerodynamic fluid–structure coupling vibration of 160 km/h emu tail in single-track tunnels," *International Journal of Structural Stability and Dynamics*, vol. 24, no. 18, p. 2450197, 2024.

[7] Y. Hu, C. Chang, Q. Chen, X. Ge, L. Ling, and K. Wang, "Investigation on the dynamic performance of high-speed trains under tunnel and crosswind environments considering car-body flexibility based on a CFD-MBD co-simulation method," *International Journal of Structural Stability and Dynamics*, vol. 25, no. 07, p. 2550070, 2025.

[8] B. Schulte-Werning, "The transaero project," in *TRANSAERO — A European Initiative on Transient Aerodynamics for Railway System Optimisation*, pp. 11–24, Springer Berlin Heidelberg, 2002.

[9] J. Katz and A. Plotkin, *Low-Speed Aerodynamics*. Cambridge Aerospace Series, Cambridge University Press, 2 ed., 2001.

[10] T. Bérenger, A. Kessler, and R. Grégoire, "Part 1 : Panel method applied to the prediction of unsteady effects caused by high-speed trains passing in the open air," in *TRANSAERO — A European Initiative on Transient Aerodynamics for Railway System Optimisation*, pp. 136–147, Springer Berlin Heidelberg, 2002.

[11] T. Bérenger and R. Grégoire, "Part 2 : Panel method applied to problems of european

high-speed train interoperability," in *TRANSAERO — A European Initiative on Transient Aerodynamics for Railway System Optimisation*, pp. 148–159, Springer Berlin Heidelberg, 2002.

[12] V. Ahuja, "Panel methods: Relevant in the 21st century?." NASA Advanced Modeling & Simulation (AMS) Seminar, Oct. 2023.

[13] X. Ding, E. Khoramzad, R. Giossi, S. H. Nia, H. Netter, G. Chen, Z. Liu, and S. Stichel, "Prediction and control of wheel wear of a high-speed train based on measured data and simulation," in *Advances in Dynamics of Vehicles on Roads and Tracks III*, pp. 589–596, Springer Nature Switzerland, 2025.

[14] S. Iwnicki, M. Spiryagin, C. Cole, and T. McSweeny, eds., *Handbook of Railway Vehicle Dynamics*. Boca Raton: CRC Press, 2 ed., Dec. 2019.

[15] X. Kang, X. B. Liu, H. Y. Li, F. Yang, J. M. Gao, and W. M. Zhai, "Psd of ballastless track irregularities of high-speed railway [in chinese]," *Scientia Sinica Technologica*, vol. 44, no. 7, pp. 687–696, 2014.

[16] C. Gomes and H. Vangheluwe, "Co-simulation of continuous systems: A hands-on approach," in *2019 Winter Simulation Conference*, pp. 1469–1481, 2019.

[17] P. D. WELCH, "The use of the fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacous*, vol. 15, pp. 70–74, 1967.

[18] G. Carter, C. Knapp, and A. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast fourier transform processing," *IEEE Transactions on Audio and Electroacoustics*, vol. 21, no. 4, pp. 337–344, 1973.

Authors



Jing, Guozhen
As a Research and Development Engineer at CRRC CHANGCHUN GERMANY RailTech GmbH, his work involves the interdisciplinary modeling and simulation of high-speed trains. In particular, he focuses on the coupled interactions between aerodynamics, running dynamics, and structural dynamics.
Email address: g.jing@crrc-railtech.cc

Session 9
Guozhen Jing, Hauke Schmidt, Xin Ding, Chutan Xue, Gang Chen



Schmidt, Hauke
As a Senior Technical Expert at CRRC CHANGCHUN GERMANY RailTech GmbH, he specializes in the structural dynamics of bogie structures for high-speed trains. The main focus of his work is on vibration analysis, structural condition monitoring and the integration of intelligent systems to improve reliability.
Email address: h.schmidt@crrc-railtech.cc



Ding, Xin
As a Senior Engineer in CRRC Changchun Railway Vehicles Co., Ltd, he specializes in the running dynamics of railway vehicles, including both experimental testing and numerical simulation. He obtained his PhD in engineering from Dalian University of Technology, China.
Email address: wwwdingxin@163.com



Xue, Chutan
As an Aerodynamics Engineer at CRRC CHANGCHUN GERMANY RailTech GmbH, he works on computational fluid dynamics (CFD) studies of high-speed and maglev trains. He holds a Diplom-Ingenieur degree in Mechanical Engineering from Technische Universität Dresden, and his research interests include train aerodynamics, pressure wave phenomena, and aerodynamic optimization.
Email address: c.xue@crrc-railtech.cc



Chen, Gang
As the Managing Director of CRRC CHANGCHUN GERMANY RailTech GmbH, he oversees the strategic directions and day-to-day operations of the company.
Email address: g.chen@crrc-railtech.cc

Session 9
Yoshiki Sugahara

Active vertical secondary suspension using low power actuator and variable damper for railway vehicles

Sugahara, Yoshiki[1]

[1] Manager, Vehicle Noise and Vibration Labratory, Railway Technical Research Institute, Tokyo, Japan

Summary

This paper investigates a method for installing variable vertical dampers in parallel with actuators designed to suppress actuator temperature rise in active vertical secondary suspension systems for railway vehicles. This is achieved by controlling the dampers and actuators simultaneously to suppress the electrical current for driving the actuators without significantly degrading the vibration reduction performance of the active suspension. To verify the effectiveness of this method experimentally, the vertical vibration control system using the actuators and variable dampers to a Shinkansen-equivalent actual vehicle was applied and vehicle excitation tests simulating actual vehicle running on a railway vehicle roller rig testing facility were carried out. By limiting the maximum force of the actuators and controlling both the actuators and the vertical dampers, the test results show that the almost same vibration reduction effect could be achieved with approximately 15% less current than when controlling only the actuator without using the aforementioned force limiter. This corresponds to a reduction of approximately 27% in energy consumption by the actuator.

Keywords: Vertical vibration; Vibration control; Active suspension; Ride comfort; Secondary suspension

1 Introduction

Lateral vibration control systems have become standard equipment on Japan's Shinkansen trains to improve ride comfort. Lateral active suspension systems using electro-

mechanical or electro-hydraulic actuators, in particular, provide excellent vibration reduction performance thanks to the actuators' excellent response characteristics [1,2]. Various studies have also been conducted on vertical vibration control systems [3]. We have also studied vertical semi-active suspensions in primary and secondary suspensions, which are cost-effective and easy to implement [4-6]. Furthermore, similar to the evolution of lateral vibration control, we have also researched vertical active suspensions to achieve even higher vibration reduction performance.

Previous studies have reported that active vertical secondary suspension systems using hydraulic actuators reduces vertical carbody vibrations, as demonstrated in vehicle running tests [7,8]. However, these papers also reported delayed actuator response (time lag) causing a slight degradation in vibration suppression performance. Therefore, we proposed a method of applying vertical actuators with significantly lower force but shorter time-delay than those previously applied to secondary active suspension, and evaluated its vibration reduction performance through vibration excitation tests using an actual vehicle [9]. Although this test confirmed the excellent vibration-reduction effects of the control using the actuators, it also indicated that suppressing temperature rise of the actuator would be desirable for practical applications.

Installing a cooling mechanism is an effective approach to reduce temperature rise in actuators. However, adding a cooling mechanism without compromising the dustproof and waterproof functions of the actuator makes the structure more complex and increases the weight and size, which makes installation in a vehicle difficult. In particular, recent Japanese Shinkansen vehicles generally require bogie covers to reduce running noise [10]. Consequently, only limited space is available around the secondary suspension of the vehicle.

Reducing the energy consumed by motors is another effective approach to reduce temperature rise in actuators. Since the energy consumed by motors is proportional to the electrical resistance of the coil and the square of the electrical current driving the actuator motor, suppressing the current is more effective in reducing energy consumption. However, suppressing the current reduces the force of the actuator, resulting in a decrease in the vibration control performance of the active suspension.

Therefore, we propose a method of installing variable vertical dampers in parallel with the actuators of the active secondary suspension and controlling them simultaneously to suppress the current of the actuators without significantly degrading the vibration reduction performance of the active suspension. The variable vertical damper can be made slimmer, facilitating installation in secondary suspension. Additionally, our system

already used a variable vertical damper to provide the basic damping of the secondary suspension required during vehicle running when no control is applied [9]. Therefore, no additional space is required.

To verify the effectiveness of this method experimentally, we applied the vibration control system using a prototype actuator and variable vertical dampers to a Shinkansen-equivalent actual vehicle and carried out vehicle excitation tests simulating actual vehicle running on a railway vehicle roller rig testing facility. This paper reports an overview of the system configuration and the results of vibration excitation tests.

Although it is desirable to reduce motor power consumption and heat generation through electromagnetic optimization of the motor and motor driver, this is beyond the scope of this study and is therefore not discussed in this paper.

2 System configuration

2.1 Overall system configuration

Figure 1 shows the system configuration of the active vertical secondary suspension equipped with variable dampers. The system consists of four vertical accelerometers on the carbody, four variable vertical hydraulic dampers, four vertical actuators, and a control unit.

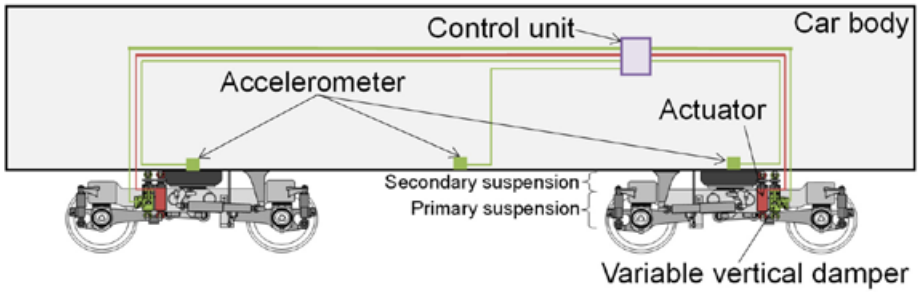


Figure 1: System configuration of active vertical secondary suspension with variable vertical damper

2.2 Linear actuator

In this test, we used the same actuator as in our previous test [9]. Figure 2(a) shows the actuator which has a mass of approximately 31 kg and an installation length of 570 mm. The actuator is powered by a linear motor. The actuator generates force from a supply of three-phase synchronous currents supplied from the motor driver (Fig. 2(b)) to the

actuator. This supply is adjusted to generate the desired force, according to the extension or compression of the actuator, which is detected by an in-built position sensor. The maximum RMS current is 12 A, which generates a force of approximately 2.5 kN. This force is only about 10 to 25% of the 10 to 20 kN force generated by actuators previously applied to actual vehicles [7,8]. However, this actuator is not affected by the compressibility of the working fluid, as is the case with hydraulic or pneumatic actuators, and therefore has excellent response to desired force commands. The characteristics of the generated force (output) in relation to the desired force (input) are a gain reduction of less than 3 dB and a phase delay of approximately 40 deg at 20 Hz, and the actuator has sufficient responsiveness to control vibrations of approximately 10 Hz. Additionally, in contrast to electric mechanical actuators, there is no backlash in the thrust-force generation mechanism of the actuator in principle, resulting in excellent responsiveness to small vibrations. However, care must be taken to prevent overheating of the actuator. Spherical bearings were used to connect the actuator to the vehicle carbody and bogie frame. This enables relative displacement between the bogie frame and the carbody in both lateral and longitudinal directions, providing a smooth movement even when the bogie rotates.

The motor drivers of the actuators are custom-made based on commercially available products, and the electrical safety functions of the actuator and driver are the functions commonly found in general-purpose drivers, and temperature monitoring functions for the actuator.



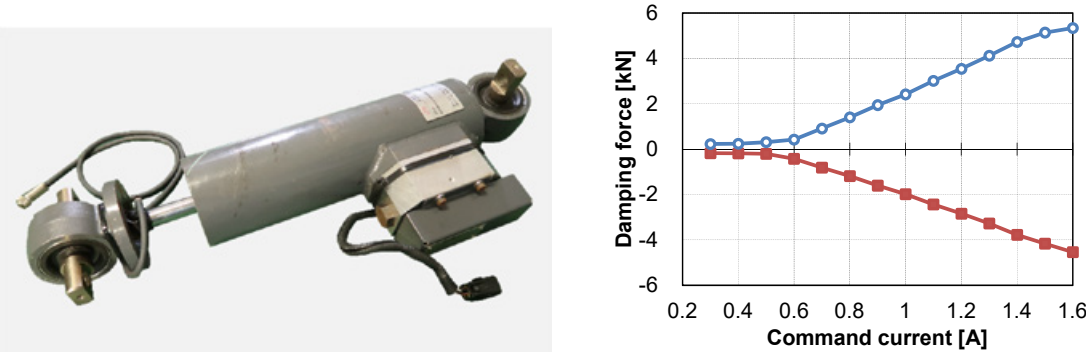
(a) Vertical actuator *(b)Motor driver*

Figure 2: First prototype vertical actuator and motor driver for a Shinkansen vehicle

2.3 Variable secondary vertical damper

Figure 3(a) shows a variable hydraulic vertical damper that is used to control the damping of the secondary suspension. This damper, which has been used in previous tests [6], has a mass of approximately 23 kg, an installation length of 510 mm and a maximum damping

force of 6-8 kN in the normal piston speed range. This damper has two damping force control valves in total. One controls the damping force in the extension direction and the other controls the damping force in the compression direction independently. Figure 3(b) shows the damping force characteristics with respect to the command current. The damping force is minimum at a command current of approximately 0.3 A, increasing to a maximum at approximately 1.5 A. Similar to the above-described actuator, spherical bearings were used to connect the dampers to the vehicle carbody and bogie frame.



(a) Appearance *(b) Damping force characteristics[6]*

Figure 3: First prototype variable vertical damper for a Shinkansen vehicle

2.4 Accelerometer and control law

In previous studies, Qazizadeh et al. used four accelerometers to control only the rigid body mode vibration of the carbody [8], whereas Kambayashi et al. used six accelerometers to control the vertical bounce, pitch, roll, and first-bending mode vibrations of the carbody [7]. In this study, as in our previous tests [9], a total of four accelerometers were attached in the positions shown in Fig. 4 to control the bounce, pitch, roll, and first-bending modes of the carbody.

A modal skyhook control law was applied to control the force of the actuators and the variable dampers as shown in Fig. 4. The vertical vibration acceleration of the carbody was separated into bouncing, pitching, rolling, and first-bending modes, and each component was integrated through a filter to obtain the vibration velocity for each mode. The vibration velocity is multiplied by the skyhook gain to calculate the skyhook force (desired force of the actuators or variable vertical damper) for each vibration mode, which is then commanded to the motor driver or solenoid driver. Bending vibration modes, however, are controlled using only actuators without variable vertical dampers. This control law does not directly take into account the interaction between the actuator and damper controls.

When controlling vibration using only actuators, the damping force of the damper should be minimized so that the basic damping of the secondary suspension should be kept to a minimum to minimize vibration transmission from the bogie to the carbody. This corresponds to setting the sky-hook gain of the variable vertical damper to zero.

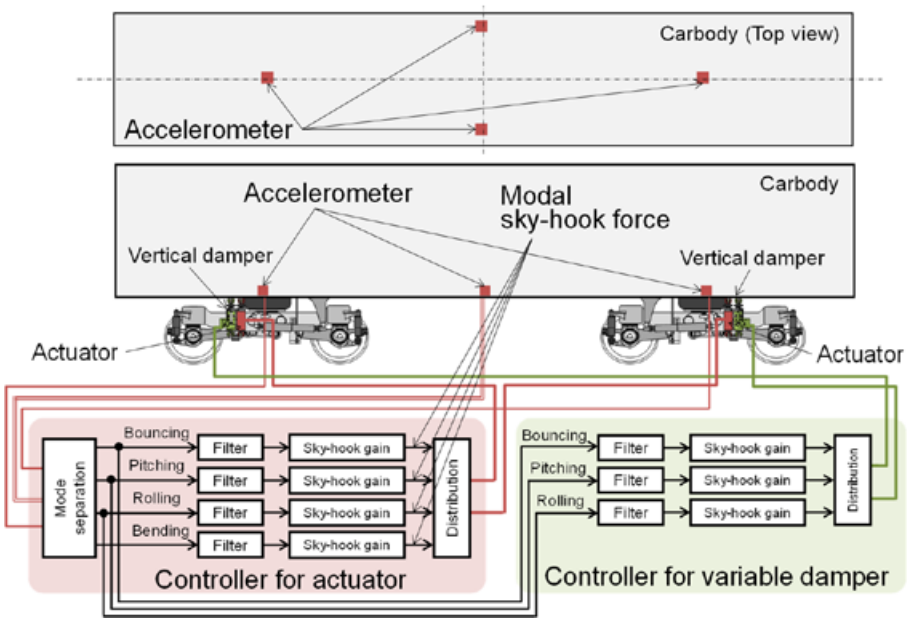


Figure 4: Accelerometers placement and control law for controlling active secondary suspension with variable vertical damper

3 Demonstration of system performance at the roller rig testing facility

3.1 Test conditions

Vehicle excitation tests using a test car were carried out at the RTRI railway vehicle roller rig testing facility to evaluate the vibration suppression performance of the active suspension with the actuator and variable damper. Figure 5 shows a photograph of the test vehicle, which is equivalent to a Shinkansen vehicle. The carbody is of single-skin aluminum alloy construction, with a car body length of 24.5 m, a distance between bogie centers of 17.5 m, and a sprung mass of approximately 32 ton, including steel deadweight.

However, the car body was unusual in its construction: two half-carbodies made of different structures were welded together to form a single carbody. As a result, this carbody has vibration characteristics which differ from standard Shinkansen carbodies. Bogies with a wheelbase of 2.5 m for the Shinkansen vehicle were used.

The vertical and rolling vibration disturbances that simulate actual running were calculated based on the Shinkansen line longitudinal and cross-level track irregularity data, and applied to each wheel with a phase difference corresponding to a running speed of 300 km/h. The maximum amplitude of the excitation displacement was limited to 90 % of the actual track irregularity due to the limitations of the rolling stock testing facility. The rotational speed of the roller rig was set to that of the vehicle running at 50 km/h to reduce the load on the axle bearings and the roller rig testing facility.

Figure 6 shows the installation of this system in vehicles. Adapters for mounting the actuator and variable vertical damper were fitted to the carbody and bogies, and the actuator and variable vertical damper were mounted in parallel with the air springs. As shown in Fig. 6, the actuator and the variable vertical damper extend beyond the rolling stock gauge because no on-track test on an actual commercial line was carried out. The adaptor is sufficiently strong, so no resonance frequencies exist, at least within the excitation frequency range. The control unit was set up on the test bench side, not inside the vehicle. Figure.7 shows the configuration of the control unit for this test. The control algorithm was implemented on a digital signal processor (DSP), and the sampling period was set to 1 msec. This control system set was configured for this vehicle excitation bench test and is not intended for use in actual railway line running tests.



Figure 5: Test vehicle on railway vehicle roller rig testing facility at RTRI

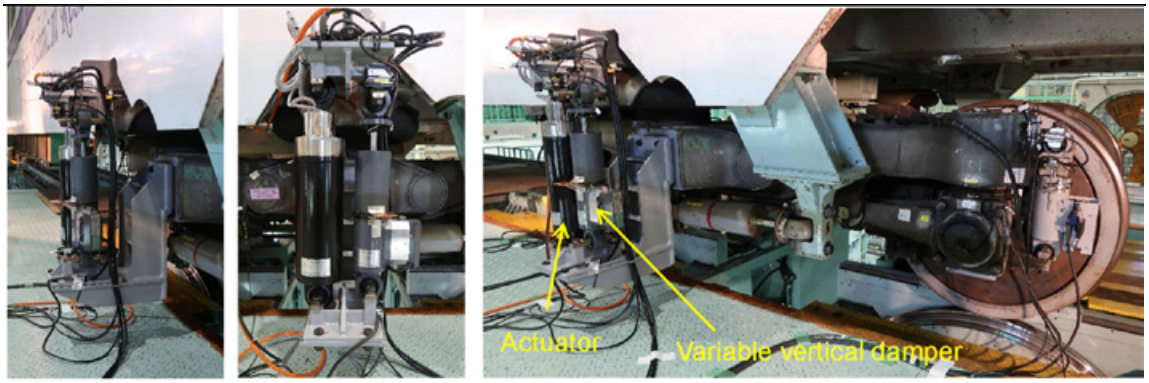


Figure 6: Installation of actuator and variable damper on test vehicle

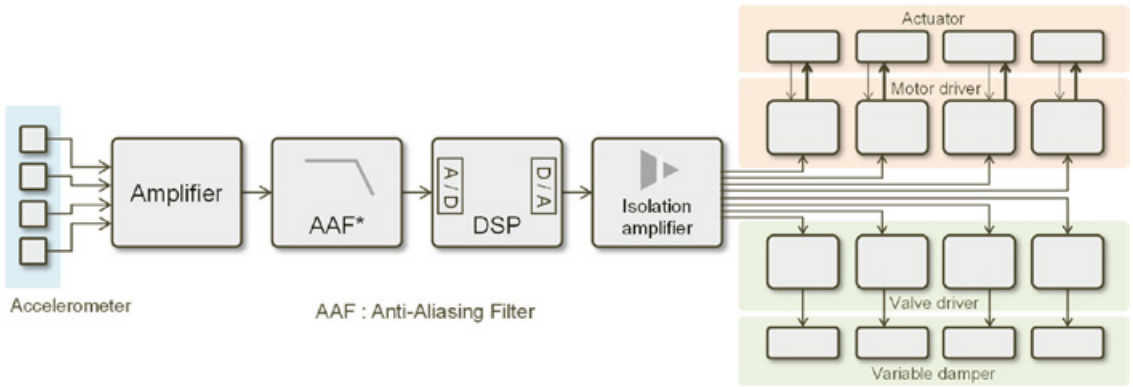


Figure 7: Configuration of the control unit for the vehicle excitation test

3.2 Basic performance verification test result

Firstly, we confirmed the basic vibration reduction performance of this system by controlling only the actuators and the variable vertical dampers. The control parameters (Fig. 4) for the active and semi-active suspension were tuned independently. When controlling only the actuators, the damping of the vertical dampers was set to a minimum. To prevent damage or deterioration of the actuator due to heat, the limiter for the desired force of the actuator was set to 1.8 kN.

Figure 8(a) and (b) show the vertical vibration acceleration power spectral density (PSD) at the carbody center and at the carbody above rear bogie center, respectively. Figure 8(c) shows the roll angular velocity PSD of the carbody.

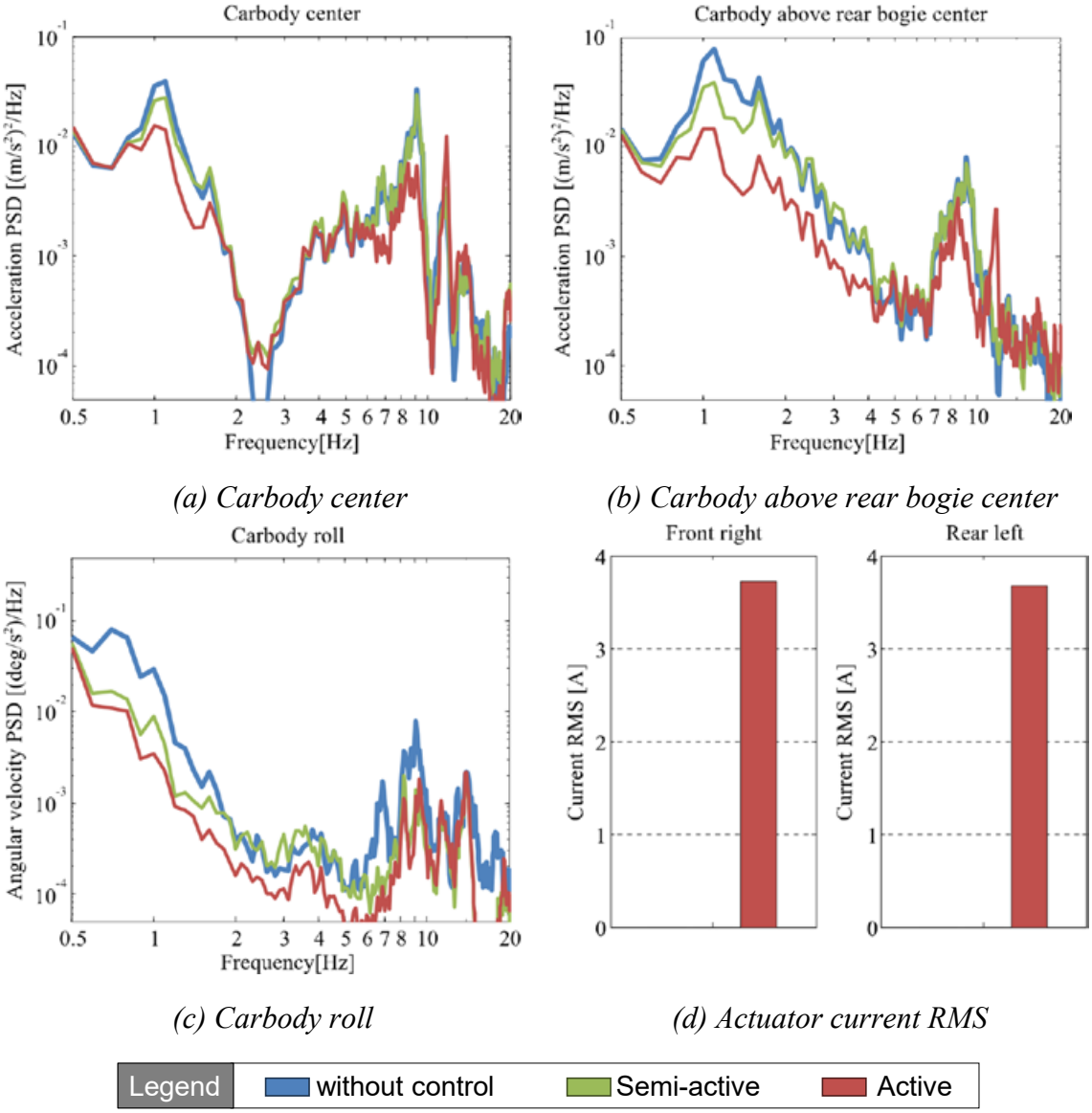
Without control (blue line in Fig. 8), PSD peaks were found at 1.1, 1.6 and 9.1Hz, both in the center of the carbody and just above the bogie. Figure 9 shows the carbody and bogie vibration shapes calculated from the vibration accelerations. Vibration accelerations were obtained from two types of accelerometer: accelerometers used for

vibration analysis and accelerometers used for controlling. In total vibration acceleration data were collected from 13 carbody accelerometers (of which 9 only collected data for analysis, while 4 collected data for both control and analysis), and 16 bogie accelerometers. Figures 8 and 9 show that the PSD peak around 1.1 Hz is due to vibrations mainly in the bounce mode of the carbody (1.1 Hz) in Fig. 9(b) and the pitch mode of the carbody (1.6 Hz) in Fig. 9(c). The PSD peak around 9.1 Hz is an elastic vibration with an anti-node at the center of the carbody, as shown in Fig. 9(d), similar to the first bending mode of the carbody but with a slightly more complex vibration shape. The PSD peak on roll angular velocity around 0.7 Hz in Fig. 8(c) is due to vibrations mainly in the roll mode and bounce mode of the carbody shown in Fig. 9(a). In addition, regardless of whether control was applied or not, a peak of 11.7 Hz was observed in the center of the carbody. As shown in Fig. 9(e), this was due to the elastic vibration of the carbody in the longitudinal direction and also in the transverse direction.

Next, when vibration was controlled using the actuators with the damping force of the variable damper fixed at a minimum (red lines in Fig. 8(a), (b) and (c), referred to as Case A), the PSD peaks at 1.1 Hz were reduced to approximately 40% at the center of the carbody and approximately 20% directly above the rear bogie. Additionally, the PSD peak at 9.1 Hz caused by elastic vibration of the carbody, was reduced to approximately 20% at the center of the carbody. Vibration reduction was achieved in the 0.6–10 Hz frequency ranges, with the most significant effects observed directly above the rear bogie center. However, when controlling vibration using actuators, vibrations around 11.7 Hz increased. Figure 9(f) shows the results of recalculating the vibration shape at 11.7 Hz with control applied. Compared to Fig. 9(e), which shows the results without control, Fig. 9(f) indicates that the control excites elastic vibrations in the carbody's transverse direction. In the case of recent Shinkansen carbodies, there are few examples of such vibrations occurring significantly at approximately 12 Hz, which can be attributed to the unusual structure of the carbody used in this study, as mentioned in section 3.1. Roll vibration was effectively reduced in the frequency range of 0.5 to 10 Hz by control with actuators, as demonstrated in Fig. 8(c). Specifically, the PSD peak of roll angular velocity at 0.7 Hz was reduced to 15% compared to that without control. The RMS value of the electrical current for driving actuator (hereinafter ‘actuator current’) when this control was performed was approximately 3.7 A, as shown in Fig. 8(d).

Finally, the green line on each graph in Fig. 8 shows the acceleration and angular acceleration PSD for the case where only the variable vertical damper is controlled. As described in section 2.4, active vibration control using actuators is primarily employed to mitigate vibrations in both the rigid-body and elastic modes of the carbody, and semi-

active vibration control using dampers is only employed to mitigate vibrations in the rigid-body mode. As demonstrated in each graph in Fig. 8, the damping control reduces the natural vibrations of the rigid body modes around 1 Hz. In particular, the reduction effect on roll vibrations was significant. In contrast, the vibrations magnitude in other frequency bands remained at almost the same values as that without control.



(a) Carbody center *(b) Carbody above rear bogie center*

(c) Carbody roll *(d) Actuator current RMS*

Figure 8: Measured results on carbody floor during vibration excitation tests for vertical acceleration PSD, roll angular velocity, and RMS of actuator current (Simulated running excitation corresponding to 300 km/h)

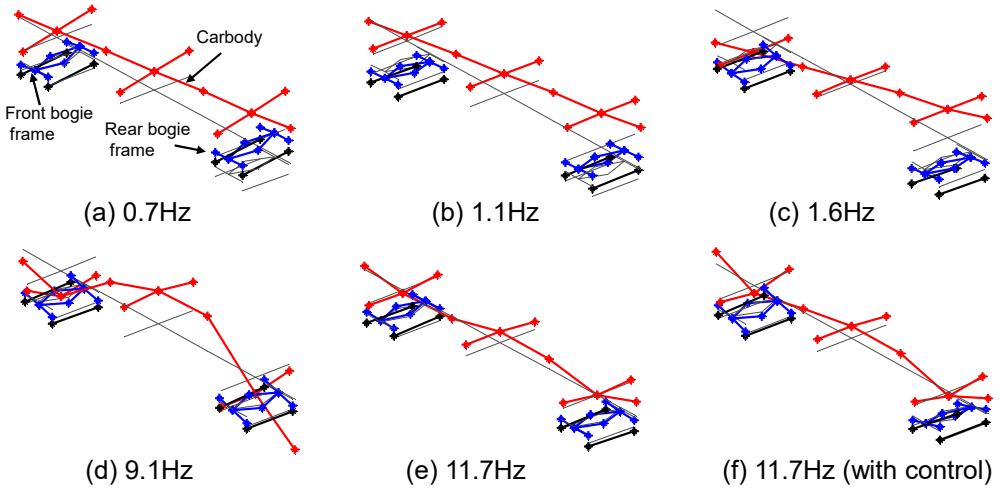


Figure 9: Mode shapes of test vehicle carbody observed during excitation tests

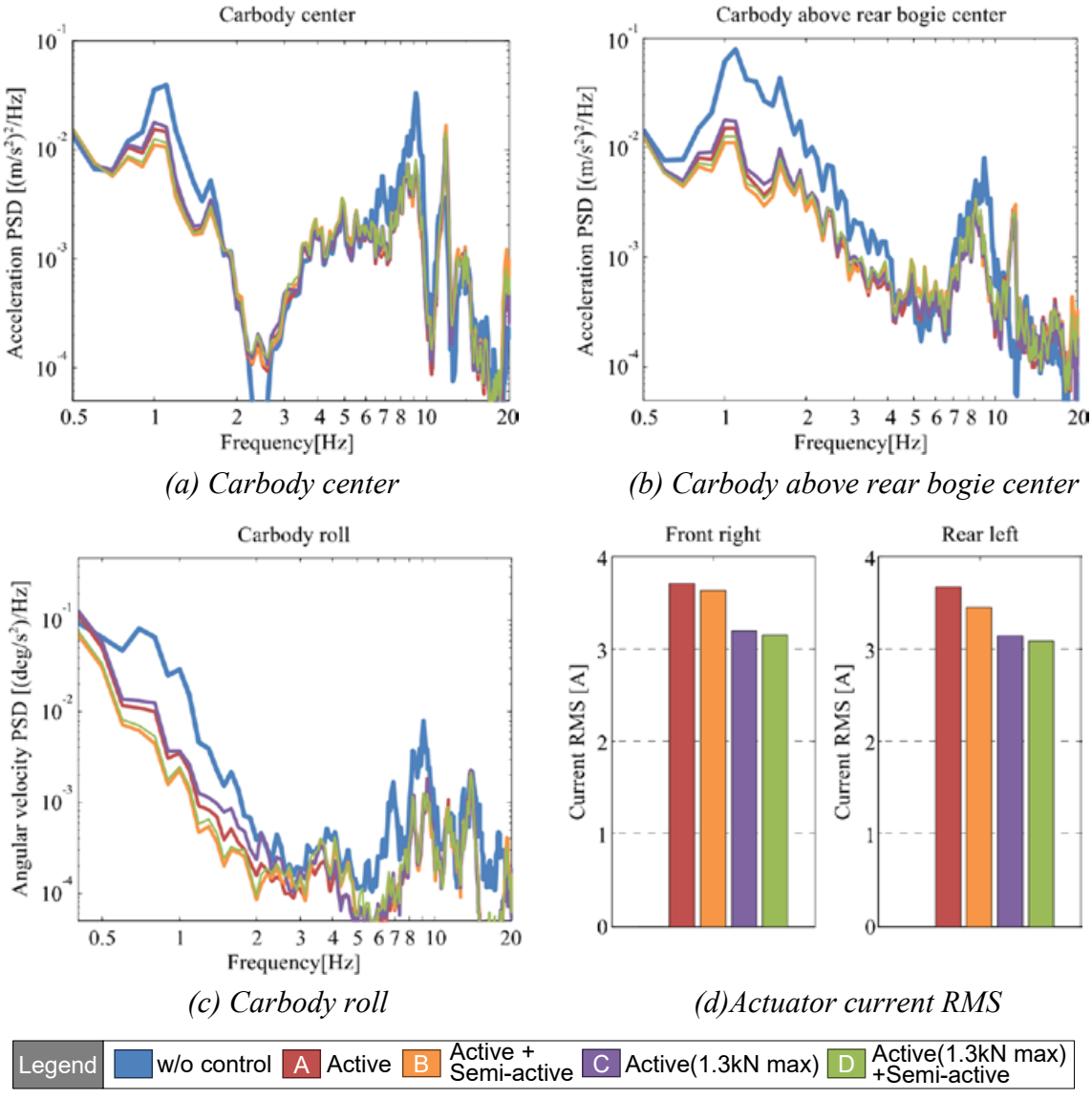
3.3 Performance verification test result of vibration control using actuators and variable dampers

Based on the results obtained in section 3.2, we evaluated the vibration reduction effect and actuator current when a semi-active suspension using variable vertical dampers was combined with an active suspension using low power actuators.

The orange line in Fig. 10 indicates the carbody vertical acceleration PSD and roll angular velocity PSD during excitation tests when the actuator control and variable damper control described in section 3.2 are used in combination (referred to as Case B). By controlling the actuator and damper together, the vertical vibration acceleration around 1 Hz was reduced, and the roll vibration angular velocity was further reduced at 0.5 to 3 Hz compared to when the actuator was controlled alone (referred to as Case A). However, as shown in Fig. 10(d), the RMS value of the actuator current was suppressed to only 2 to 6%.

Based on these results, the limiter for the desired force of the actuator was changed from 1.8 kN to 1.3 kN, and vibration tests were conducted to compare the carbody vibrations and actuator's current RMS values with and without damper control. First, the results of only controlling the actuator with 1.3 kN force limiter and with the damping of the variable damper fixed at a minimum are shown in Fig. 10 in purple (referred to as Case C). As a result of limiting the desired force of the actuator to 1.3 kN, the actuator's current RMS value was reduced by 14%, as shown in Fig. 10(d). However, on the other hand, the vertical and roll vibrations around 1 Hz caused by the rigid mode vibrations of the carbody increased slightly due to the decrease in control force. Next, the results of controlling both

the actuator with 1.3kN force limiter and damper are shown in green in Fig. 10 (referred to as Case D). The RMS value of the actuator's current decreased further from Case C to 16% less than Case A, it was the smallest among cases A to D. This corresponds to a reduction of approximately 29% in energy consumption by the actuator. Additionally, no reduction in the vibration reduction effect of the rigid body mode vibration around 1 Hz, as observed in Case C, was observed. In particular, the vibration reduction effect for roll vibration was better than that of Case A. However, as shown in Fig. 10(a)(b), there was a tendency for the vertical vibration acceleration PSD in the 3–10 Hz range to increase slightly compared to Case A and Case C, which did not include damper control.



(a) Carbody center

(b) Carbody above rear bogie center

(c) Carbody roll

(d)Actuator current RMS

Figure 10: Measured results on carbody floor during vibration excitation tests for vertical acceleration PSD, roll angular velocity, and RMS of actuator current (Simulated running excitation corresponding to 300 km/h)

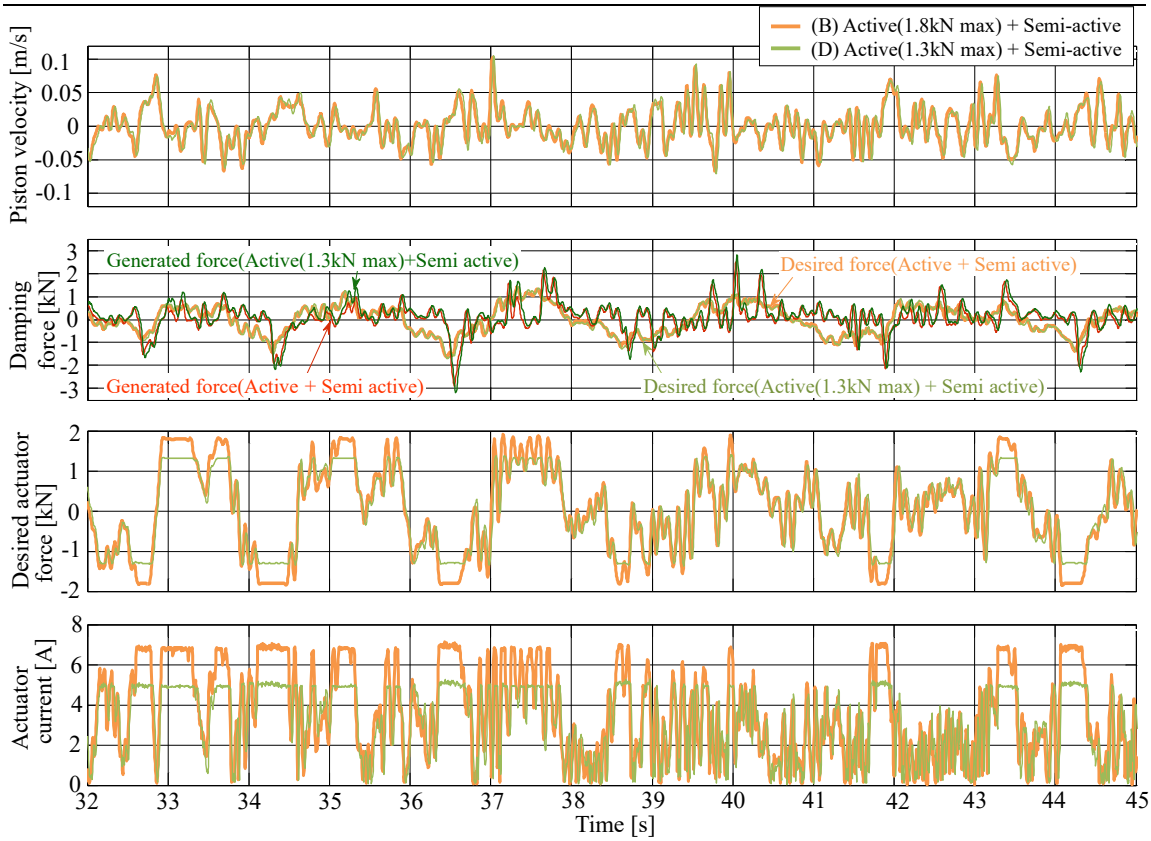


Figure 11: Piston speed, damping force (desired force and generated force), desired actuator force (command value), and RMS value of actuator current on the left side of the rear bogie during the vibration excitation tests (Simulated running excitation corresponding to 300 km/h)

The reason why simply using an actuator and a damper in combination does not significantly suppress the actuator current can be explained by the graph of the desired force and actual force of the damper as shown in Fig. 11. For example, in Fig. 11, a pulling force is commanded to the damper and actuator at approximately 36 seconds, however, the damper is stroking in the compression direction, which prevents the generation of a pulling force. Consequently, only the actuator is responsible for vibration control. At approximately 36.5 seconds, the damper strokes in the extension direction, resulting in the generation of a pulling force. As described above, the damper is unable to generate the desired force unless the phase conditions between the stroke direction and the desired force are met. Therefore, simply using dampers and actuators together does not necessarily lead to reduced actuator current.

On the other hand, limiting the force of the actuator directly limits the current, which leads to a reduction in current. This results in insufficient control force, which can be compensated for by controlling variable dampers to some extent, thereby mitigating the

deterioration of vibration reduction effects, in rigid body modes of the carbody. Of course, theoretically, the most efficient method would be to directly measure the damping force of the variable dampers and subtract the force generated by the variable dampers from the actuator's desired force. However, this approach requires sensors, such as strain gauges, to measure damping force, and it is necessary to evaluate the influence of delays and errors in the damping force measurement system (or estimation system) on control performance. There are many challenges to overcome for practical application in railway vehicles.

4 Conclusion

In this paper, as an example of a measure to suppress temperature rise in actuators for active vertical secondary suspension systems for railway vehicles, we propose a method of installing variable vertical dampers in parallel with the actuators and controlling them simultaneously to suppress the current of the actuators without significantly degrading the vibration reduction performance of the active suspension. To verify the effectiveness of this method experimentally, we applied the vibration control system using prototype actuators and variable vertical dampers to a Shinkansen-equivalent actual vehicle and carried out vehicle excitation tests simulating an actual vehicle running at 300km/h on a railway vehicle roller rig testing facility. A modal skyhook control law was applied to control the actuators and the variable dampers. The vibration modes of the vehicle carbody to be controlled are only rigid-body modes for dampers and, in addition to rigid body modes, the first-bending modes for actuators. This control law does not directly take into account the interaction between the actuator and damper controls.

By limiting the maximum force of the actuators to 1.3 kN and controlling both the actuators and the vertical dampers, we confirmed that the almost same vibration reduction effect could be achieved with approximately 15% less current than when controlling only the actuator using its maximum force of 1.8 kN. This corresponds to a reduction of approximately 27% in energy consumption by the actuator. In addition, combining damper control improved the reduction of vertical vibrations of approximately 1 Hz and carbody roll vibrations.

Note that the vibration reduction effect and the actuator temperature rise suppression effect due to current suppression depend on the excitation disturbance, vehicle vibration characteristics, actuator and damper characteristics, ambient temperature, and other factors. When applying this system to vehicles operating on commercial lines, it is necessary to refer to the track conditions of the running section, running conditions (train

performance curve), vehicle specifications, climate conditions, and other factors, and to specifically examine the abovementioned effects.

Literature

[1] K. Tanifuji, S. Koizumi and R. Shimamune, "Mechatronics in Japanese rail vehicles:active and semi-active suspensions," *Control Engineering Practice,* Vol.10, No.9, pp.999 – 1004, 2002, DOI:10.1016/S0967-0661(01)00165-4.

[2] K. Asano, "JR East High-speed Rolling Stock Development," *JR EAST Technical Review*, Vol.36, pp.1 – 6, 2017.

[3] B. Fu, R. L. Giossi, R. Persson, S. Stichel, S. Bruni and R. Goodall, "Active suspension in railway vehicles: a literature survey," *Railway Engineering Science*, Vol.28, No.1, pp.3 – 35, 2020, DOI:10.1007/s40534-020-00207-w.

[4] Y. Sugahara, A. Kazato, R. Koganei, M. Sampei and S.Nakaura, "Suppression of vertical bending and rigid-body-mode vibration in railway vehicle car body by primary and secondary suspension control (Results of simulations and running tests using Shinkansen vehicle)," *Proceedings of the Institution of Mechanical Engineers, Part F, Journal of Rail and Rapid Transit*, Vol.223, No.6, pp.517 – 531, 2009, DOI:10.1243/09544097JRRT265.

[5] Y. Sugahara, T. Kojima, Y. Akami and Y. Igarashi, "Development of a vertical semi-active suspension system using variable hydraulic dampers," *Proceedings of the 15th International Conference on Railway Engineering Design and Operation (COMPRAIL 2016)*, pp.469 – 477, 2016, DOI:10.2495/CR160421.

[6] Y. Sugahara, T. Kojima, D. Sannomiya and K. Miyahara, "Suppression of vertical vibrations in railway vehicles using variable primary and secondary hydraulic dampers," *Proceedings of the 1st International Railway Symposium Aachen 2017(IRSA2017),* pp.647 – 659, 2017, DOI:10.18154/RWTH-2018-01320.

[7] K. Kambayashi, S. Usui, T. Ohtsuka, H. Matsushima and K. Dambata, "Development of an active vibration control systems for high speed railway vehicles," *Proceedings of the 5th Jointed Railway Technology Symposium (J-Rail'98)*, pp.499 – 502, 1998 (in Japanese).

[8] A. Qazizadeh, R. Persson and S. Stichel, "On-track tests of active vertical suspension on a passenger train," *Vehicle System Dynamics*, Vol.53, No.6, pp.798 – 811, 2015, DOI:10.1080/00423114.2015.1015429.

[9] Y. Sugahara, "Preliminary performance test of active secondary vertical suspension using low power actuator on railway roller rig," *Proceedings of the 6th International Conference on Railway Technology: Research, Development and Maintenance (Railways 2024)*, Paper 5.1, 2024, DOI: 10.4203/ccc.7.5.1.

[10] T. Oku, "Development of Technologies for Weight Reducing and More Efficient High-speed Trains," *Hitachi Review*, Vol.70, No.6, pp.648 – 653, 2021.

Authors



Sugahara, Yoshiki
Ph.D. Manager, Vehicle Noise and Vibration Laboratory, Vehicle Technology Division, Railway Technical Research Institute.

IRSA 2025

SESSION 10

ZUSTANDSÜBERWACHUNG II

CONDITION MONITORING II

Session 10
Yan Niu, Dietmar Kraft, Kai Heimann

Landing data science and AI in condition monitoring for bogie components and track

Niu, Yan[1], Kraft, Dietmar[2], Heimann, Kai[1]

[1] Alstom Transportation Germany GmbH
[2] Alstom Switzerland Ltd.

Abstract

The health condition of bogie components has a direct influence on passenger ride comfort, running stability, and train reliability. Track condition plays also a crucial role in wheel/rail interaction and fatigue of bogie components. From the aspect of maintenance cost, bogies and track are among the main cost drivers. Condition-Based Maintenance (CBM) has the potential to enhance the reliability and to reduce the maintenance cost. To implement CBM for bogie components and track, condition monitoring is required. This provides a great landing platform for data science and Artificial Intelligence (AI) techniques. This paper presents some application examples of data science and AI in condition monitoring for bogie components and track, including axle box temperature monitoring; actuator temperature prediction; gearbox and motor fault detection on a wheel lathe; permanent onboard monitoring for gearbox and motor; trackside acoustic monitoring; and track condition monitoring using in-service trains. These examples cover the topics of extracting insights from data and the adoption of rule-based methods or Machine Learning (ML) for fault detection and diagnosis. As an outlook, the extension to predictive maintenance and the importance of Machine Learning Operations (MLOps) are discussed.

Keyworts: condition monitoring; bogie; track; data science; artificial intelligence

Session 10
Yan Niu, Dietmar Kraft, Kai Heimann

1 Introduction

The health condition of bogie components has a direct influence on passenger ride comfort, running stability, and train reliability. In addition, track condition plays also a crucial role in wheel/rail interaction and the fatigue of bogie components. In case an in-service failure happens to a bogie component or a track section, it may potentially cause unexpected train delay, significant economic losses, and in some cases even catastrophic accidents. From the aspect of Life Cycle Costs (LCC), maintenance costs represent a considerable percentage [1]. If the maintenance activities are scheduled at fixed intervals (e.g. based on time or distance), bogies and track are often the main cost drivers. Bogie maintenance accounts for 25-30% of the maintenance costs for the entire train [2], while track maintenance (excluding switch) is about 25% of the total maintenance costs for a line [3]. To remain competitive, railway companies face increasing economic pressure to reduce these maintenance costs.

In 2017, The Economist published a famous article “The world’s most valuable resource is no longer oil, but data”. In that article, data is referred to as “the oil of the digital era” [4]. Also in 2017, McKinsey shared a study titled “The rail sector’s changing maintenance game – How rail operators and rail OEMs can benefit from digital maintenance opportunities”. According to McKinsey, Condition-Based Maintenance (CBM) has the potential to offer at least 10-15% reduction in maintenance costs, with predictive maintenance contributing additional 10% savings [5]. In the following years, Artificial Intelligence (AI) has been demonstrating its transformative potential in infusing every industry. In 2022, Alstom released a white paper, looking at the potential of AI in the transport sector, and predictive maintenance is among the use cases [6]. In 2024, International Union of Railways (UIC) unveiled a report, “The journey toward AI-enabled railway companies”. According to UIC’s observation, rolling stock predictive maintenance has enabled a 15% increase in reliability and a 20% reduction in maintenance cost, and rail infrastructure predictive maintenance can drive a 15-25% reduction in unplanned downtime and a 15-30% reduction in maintenance costs [7].

According the definition in ISO 13372, CBM is the type of maintenance which is governed by condition monitoring programs [8]. Based on the condition data from the asset, CBM infers if there is an impending failure. Predictive maintenance is a specific case of CBM, as it predicts the Remaining Useful Life (RUL) until failure occurs if no action is taken [1]. To implement CBM or predictive maintenance for bogies and track, condition monitoring is needed. This actually provides a great landing platform for data science and AI techniques. This paper presents some application examples of data science

and AI in condition monitoring for bogie components and track. These examples cover the topics of extracting insights from data and the adoption of rule-based methods or machine learning for fault detection and diagnosis. Before diving into those application examples, a brief review is given on data science, AI, and condition monitoring.

2 A brief review on some key concepts

Data science is the discipline of extracting knowledge and insights from data. The key steps of a data science project are listed in Figure 1. During the data analysis, there are often several iterations to get good insights from the data. In some cases, AI techniques are also adopted to assist the data analysis.

AI encompasses a wide array of tools, e.g. rule-based systems, machine learning, deep learning, etc. [9] [10]. Figure 2 illustrates the relationship between data science and AI, and there is a clear overlap among data science, AI, Machine Learning (ML), and deep learning. Machine Learning is the "field of study that gives computers the ability to learn without being explicitly programmed" [10]. Figure 3 presents the family of Machine Learning which has three main categories, including supervised learning, unsupervised learning, and reinforcement learning. Under supervised learning, two subcategories exist: regression and classification. Regression outputs numerical values, while classification delivers categorical values, e.g. healthy or faulty. The life cycle of a machine learning project is illustrated in Figure 4.

Condition monitoring is referred to as the process of "acquisition and processing of information and data that indicate the state of a machine over time" [8]. The workflow of condition monitoring can be described by the flow chart in Figure 5. By comparing Figure 5 with Figure 1 and Figure 4, it is interestingly observable that condition monitoring actually serves as a great landing platform for data science and AI.



Figure 1 : Key steps of a data science project [11]

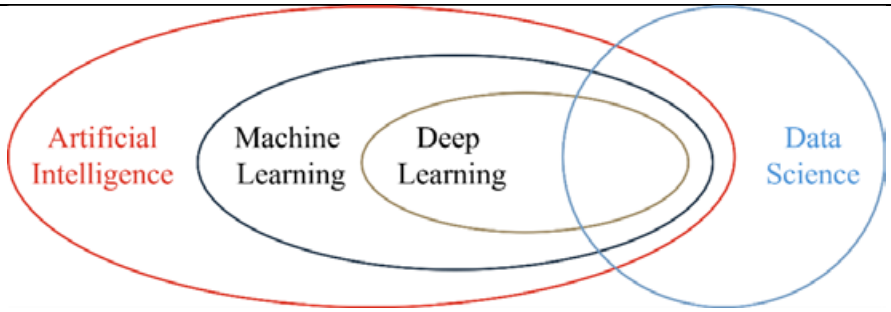


Figure 2 : Relationship between data science and AI [11]

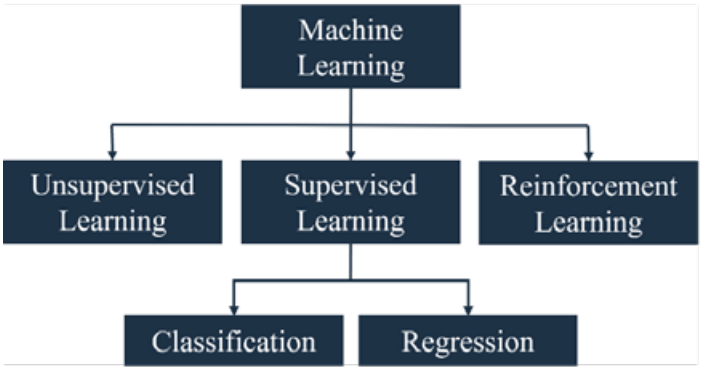


Figure 3 : Categories of machine learning [10]

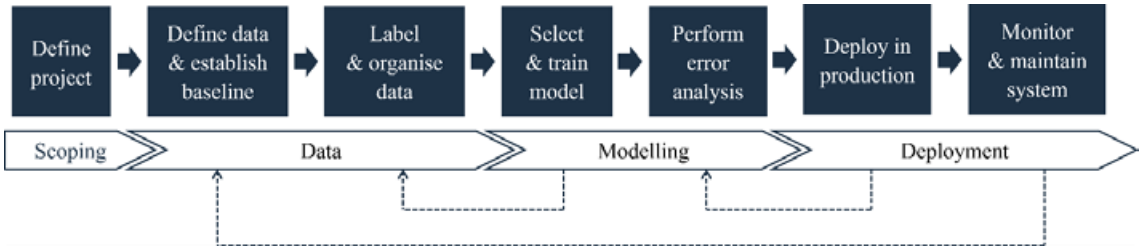


Figure 4 : Machine learning project life cycle [12]

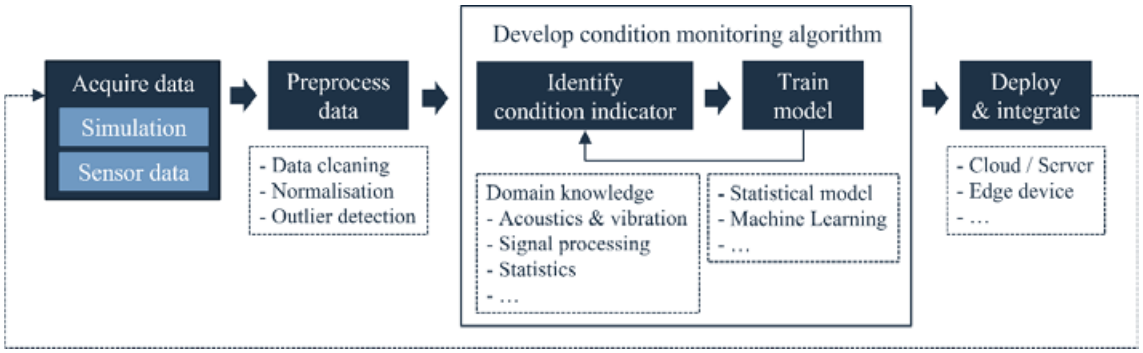


Figure 5 : Condition monitoring workflow (based on [13])

3 Temperature-based applications

During a medical check-up, body temperature belongs to one of the key indicators of our health condition. A high body temperature, also known as a fever, is often a sign that our body is fighting against an infection or inflammation. In machine condition monitoring, temperature is also an important physical parameter. This section presents two application examples using temperature data.

3.1 Axle box temperature monitoring

For interoperable rolling stock with a maximum design speed higher than or equal to 250 km/h, it is mandatory to install onboard axle box condition monitoring, e.g. using temperature sensors [14] [15]. Figure 6 presents in color the maximum temperature of each axle box across a fleet of high-speed trains within two years. From the color pattern in Figure 6, it is noticed that the axle boxes on motor bogies in general have experienced slightly higher temperature than those on trailer bogies. To identify the axle bearings that are about to fail, one method is to set temperature limits as alarm levels [15]. This is a rule-based method which also belongs to AI, but not machine learning. Another possible method is to build a digital twin to predict the axle box temperature, e.g. using a deep learning algorithm called Long Short-Term Memory (LSTM) network [16]. A significant deviation between the predicted and the measured temperature indicates an imminent bearing failure. In this case, it is an application of the regression technique in Figure 3.

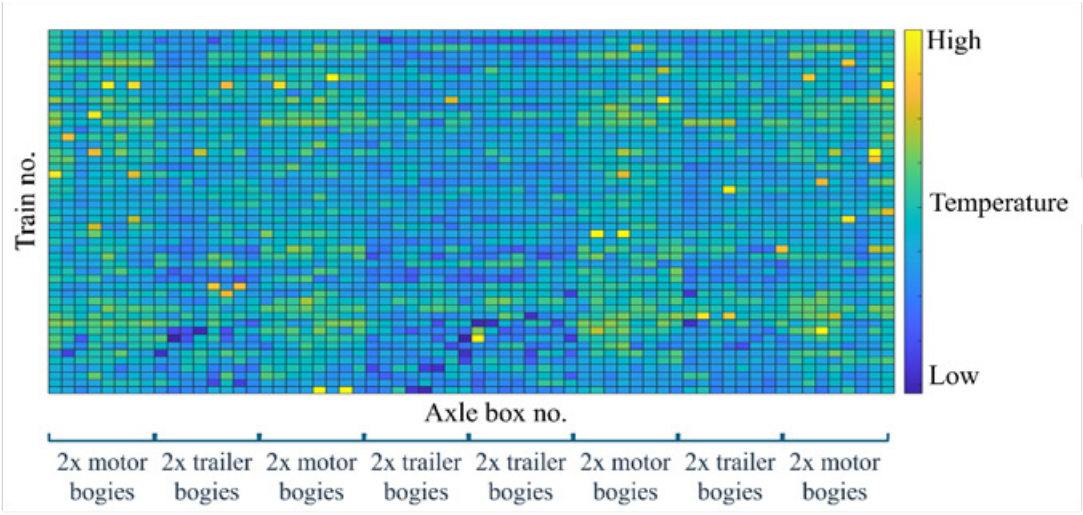


Figure 6 : Maximum temperature for each axle box across a fleet of high-speed trains within two years

3.2 Actuator temperature prediction

In standard railway vehicles, mechatronics has already entered into the level of application in bogies [17] [18]. Actuators, as a fundamental component in mechatronic systems, are responsible of executing the control action on connected bogie components [19]. Each type of actuators has a specified operating temperature range. Exceeding the temperature limits may influence the actuator performance and in turn the control effectiveness. Thus, actuators are often integrated with sensors for temperature monitoring. To evaluate if the temperature of a type of actuators will stay within limits, one solution is to perform track tests using prototype mechatronic bogies. However, it is sometimes difficult to cover a certain track section which may potentially bring extra valuable information. In such a scenario, a digital twin of the actuator can be utilized to predict the temperature. The basic idea is described in Figure 7. An example is provided in Figure 8, where the predicted temperature in general follows the measured temperature. This is also an application of regression in Figure 3.

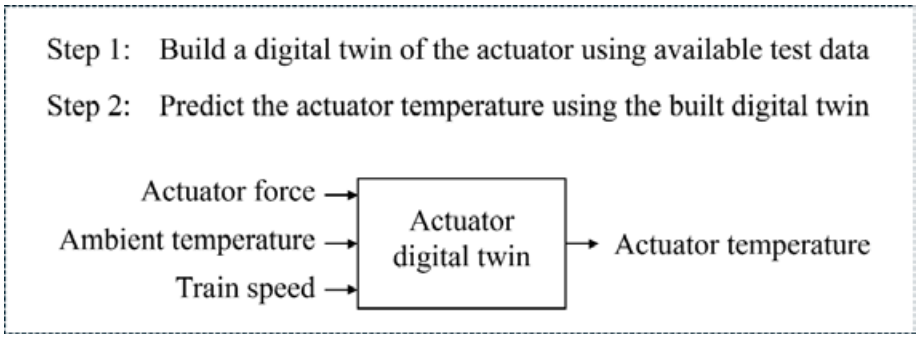


Figure 7 : Actuator temperature prediction - basic idea

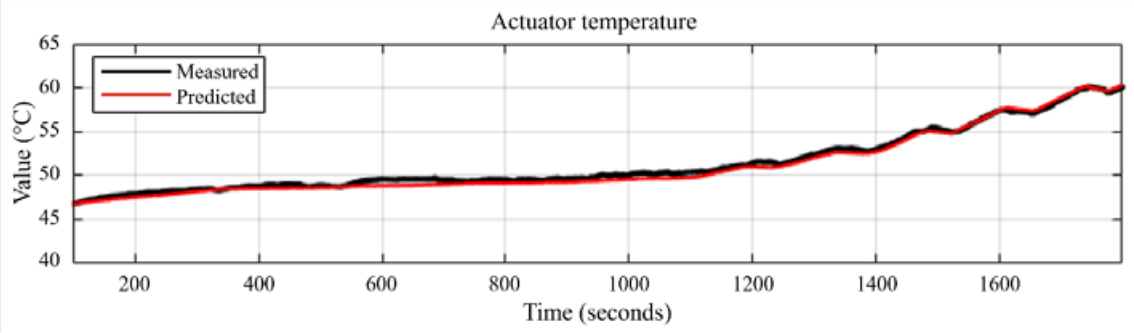


Figure 8 : Actuator temperature prediction - an example

4 Vibration-based applications

A defect or degradation within a structure may lead to variations in its vibration characteristics. This means that the health condition of a mechanical structure can be evaluated based on its vibration [20]. In machine condition monitoring, vibration analysis is a central topic and it is also considered as the most prevalent method [21] [22]. This section focuses on the application of vibration-based methods. When a bogie is running on a track section or on a test rig, the bogie components will vibrate as a response to the excitation. At the same time, the air particles surrounding the bogie components are forced to move at the same frequencies. Sound waves are formed in this process and propagate through longitudinal wave motion of air particles [23]. This means that the vibration of the bogie components can either be measured directly using e.g. accelerometers or be sensed indirectly using e.g. microphones. Through vibration analysis, it is then possible to extract the vibration characteristic information and perform further evaluation on the health condition of the bogie components and the track. The basic idea is illustrated in Figure 9.

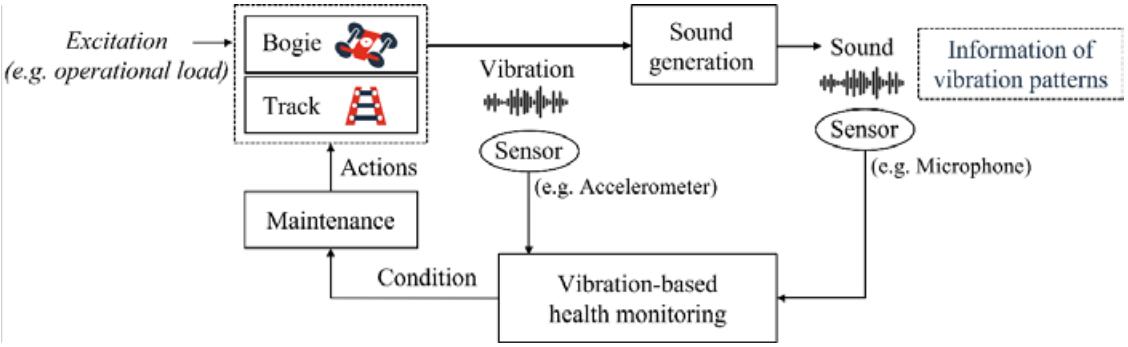


Figure 9 : Vibration-based condition monitoring for bogie components and track - basic idea

4.1 Gearbox and motor fault detection on a wheel lathe

Wheel lathes are among the most important machining tools in rolling stock maintenance. When a motor bogie is put on a wheel lathe for reprofiling work, the wheelset is driven by the traction rollers of the wheel lathe, and the connected gearbox and motor are driven inversely by the wheelset. Directly after the wheel reprofiling work, the wheel tread is in a good condition, and the traction wheels of the wheel lathe can be kept running for a short duration. This provides an opportunity for performing vibration-based fault detection for gearbox and motor, e.g. using a wireless accelerometer. In corporation with an urban rail operator, this method has been applied for a fleet of metro trains for motor

bearing fault detection since 2021, and expanded for gearbox bearing fault detection since 2024. Figure 10 shows some photos from this application. The fault detection is realized by comparing the calculated characteristic value with a defined threshold, as illustrated in Figure 11. This is a rule-based method which belongs to AI, but not machine learning. As the real conditions of the bearings are known from maintenance feedback, the results can be categorized into one of the four groups in Table 1. Then the threshold is fine-tuned accordingly. If it is a FP (i.e. false alarm), increase the threshold; if it is a FN, reduce the threshold; and in case it is a TP or a TN, keep the threshold unchanged.



Figure 10 : Gearbox and motor bearing fault detection on a wheel lathe – an application example: (a) a train on a wheel lathe; (b) a wireless accelerometer on motor housing; (c) a detected faulty gearbox bearing

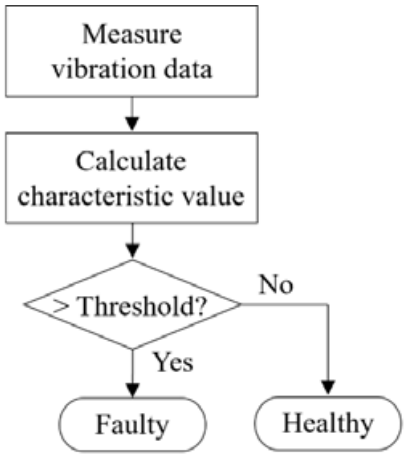


Figure 11 : Gearbox and motor fault detection on a wheel lathe – principle

Table 1: Possible results from fault detection and maintenance feedback (this table is generated based on [24])

		Results from fault detection	
		Faulty	Healthy
Results from maintenance feedback	Faulty	True Positive (TP)	False Negative (FN)
	Healthy	False Positive (FP)	True Negative (TN)

4.2 Permanent onboard condition monitoring for gearbox and motor

During the train operation, there are a variety of dynamic scenarios, e.g. different operating modes (acceleration, cruising, coasting, braking), different track sections (straight lines or curves), etc. This causes the load zone in gearboxes and motors often not always to be at the same position. Permanent onboard condition monitoring is capable of covering different dynamic scenarios. Furthermore, it also enables monitoring gearboxes and motors in a preferred frequency, e.g. everyday. These features actually increase the chance to detect defects or degradation in an early stage. If it is a brand new train or a component replacement has just been performed, it is often reasonable to assume the gearboxes or motors are in a good condition, and this period may serve as a reference phase, allowing a threshold to be set based on the statistical distribution of the vibration data. As the train is put in service, the freshly acquired vibration data are compared with that threshold. In case an unusual number of values go beyond the limit, it is often a sign of a defect or degradation. The principle is illustrated in Figure 12. This approach has long been adopted in process engineering for Statistical Process Control (SPC) [25]. It is an approach to anomaly detection which falls under unsupervised learning in Figure 3.

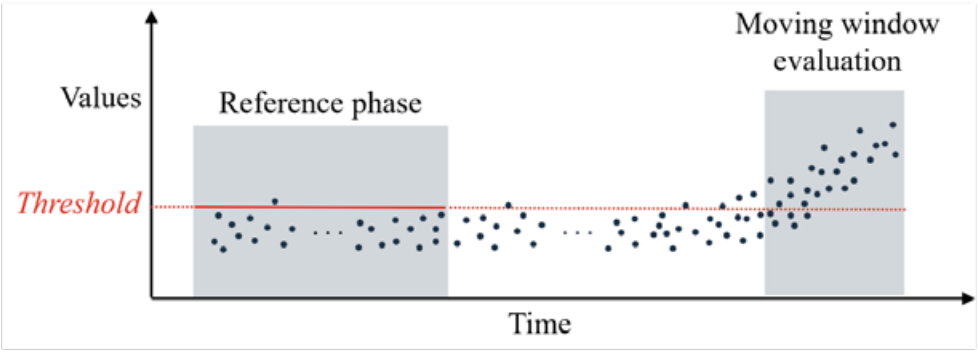


Figure 12 : Fault detection based on statistical process control - principle

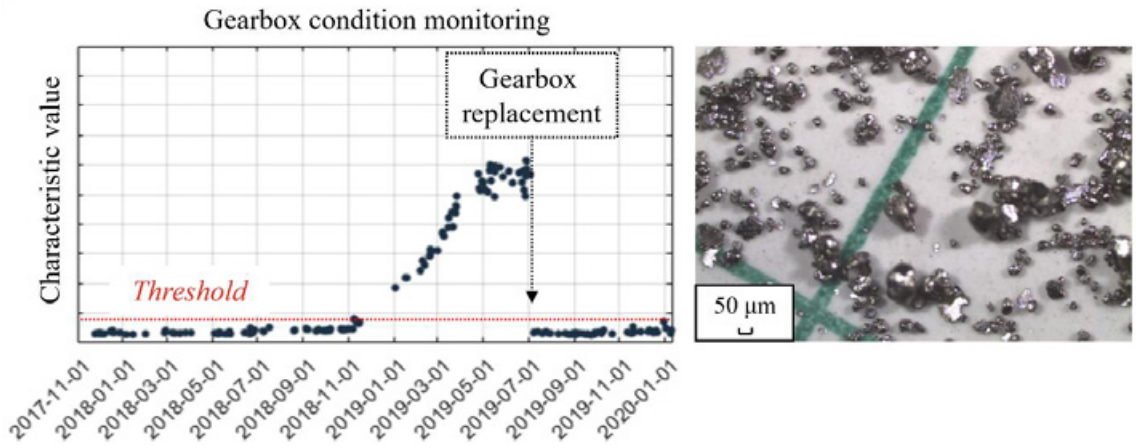


Figure 13 : An example of detected gearbox condition degradation: (a) time history of the vibration characteristic values; (b) a microscopic image of the gearbox oil sample

Figure 13 shows an application example. The characteristic values of the acceleration data from a gearbox exhibited a clear upward trend, ultimately surpassing the threshold. Following a gearbox replacement, the characteristic values decreased to within the normal range. The microscopic image of the gearbox oil sample confirms that the gearbox oil was contaminated.

4.3 Trackside acoustic monitoring

Sounds from passing trains carry not only the rhythm from the Doppler effect but also the valuable health condition information of bogie components. According to a cost-benefit analysis report from Rail Safety and Standards Board (RSSB) in UK, trackside acoustic monitoring is highly likely to deliver significant benefits for train owners and maintainers [26]. Alstom's Mechatronics team has developed a pilot solution for trackside acoustic monitoring, and this pilot solution has been deployed at regular time intervals on the networks of two urban rail operators [27].

In partnership with an operator, Alstom has developed a permanent trackside acoustic monitoring system for axle bearing fault detection. Figure 14 provides a photo from the monitoring system. Every time when a train passes by the measurement area, the sound waves emitted from the passing axle boxes are captured by the two microphone arrays. The recorded data are then post-processed, and the health condition of the axle bearings are evaluated using a trained ML model, as illustrated in Figure 15. This is an application of the classification technique from Figure 3.



Figure 14 : A photo from a permanent trackside acoustic monitoring system

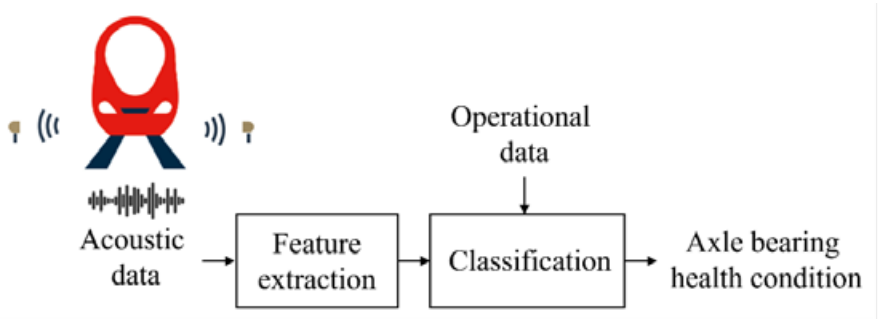


Figure 15 : Axle bearing fault detection using trackside acoustic monitoring - process

4.4 Track condition monitoring using in-service trains

Track irregularities can cause high dynamic forces which have a direct influence on wheel-rail interaction, ride comfort, and the fatigue of bogie components. By analysing onboard acceleration data, it is possible to gain insights into the condition of track [28]. A cost-efficient approach is to integrate monitoring systems on in-service trains, so that they can serve as a moving platform for monitoring the track condition during passenger service. In corporation with an urban rail operator, two trains have been instrumented with a track condition monitoring system. Figure 16 shows different types of sensors that are installed on an in-service train.

One use case is corrugation detection using Axle Box Acceleration (ABA) data. An example is provided in Figure 17. With the ABA data, track sections are classified if corrugation is detected. This is again an application of the classification technique from Figure 3. The related track section is localised with the help of the GPS data.

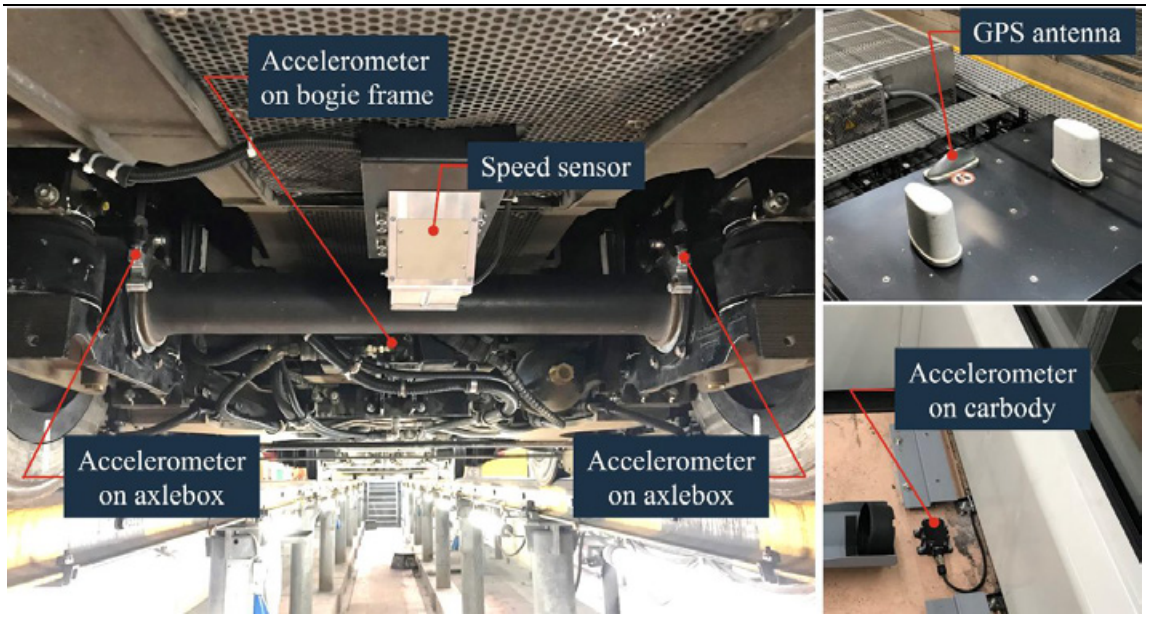


Figure 16: Track condition monitoring – sensors installed on an in-service train

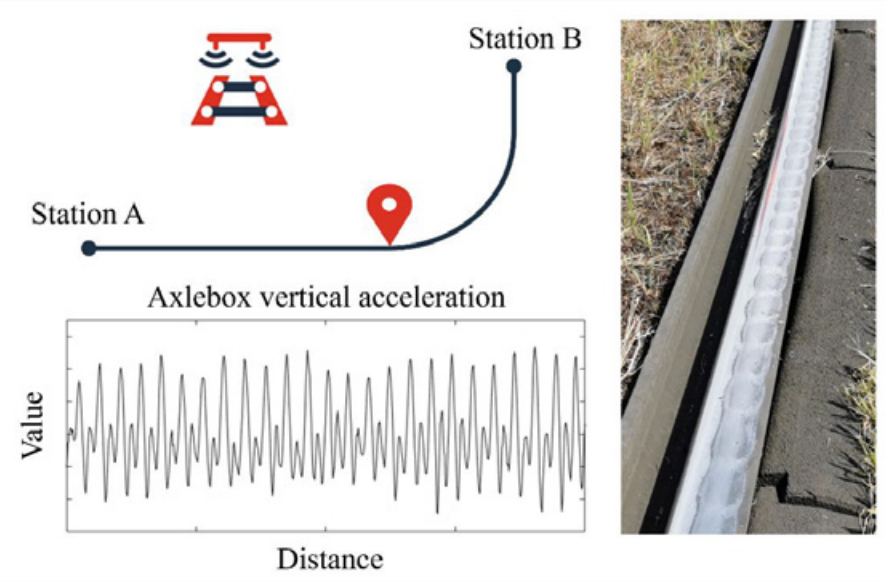


Figure 17 : Corrugation detection using axle box acceleration data – an example

5 Outlook

As a specific case of CBM, predictive maintenance estimates also the RUL, enabling a further reduction of the LCC. A reliable prediction of the RUL is crucial for a successful implementation of predictive maintenance. Leveraging the condition monitoring platform, AI can effectively contribute to achieving this goal.

Machine Learning Operations (MLOps) is the standardisation and streamlining of ML life cycle management [29]. It combines ML development with system operation (Ops) to ensure efficient and reliable deployment and maintenance of ML models. As it comes to the deployment phase of an ML model, MLOps becomes critical. Some sensors may not function properly, a data drift may occur due to a change in the operational condition or a degradation of the infrastructure, and so on. It is necessary to monitor the performance of the ML model regularly and update the ML model if there is a need. Adopting an MLOps mindset is key to unlocking the full potential of AI.

The journey of landing data science and AI is not over yet. There are more breathtaking views awaiting beyond the horizon.

Literature

[1] P. Dersin, A. Alessi, O. Fink, B. Lamoureux and M. Brahimi, "Prognostics and Health Management in Railways," in *Handbook of RAMS in Railway Systems: Theory and Practice*, Q. Mahboob and E. Zio, Eds., Boca Raton, FL, USA, CRC Press, Taylor & Francis Group, 2018, pp. 79-98.

[2] E. Kwasnicki, “Fahrwerk,” in *Handbuch Schienenfahrzeuge: Entwicklung, Produktion, Instandhaltung*, C. Schindler, Ed., Hamburg, DVV Media Group GmbH, 2014, pp. 349-371.

[3] E. Lay and R. Rensing, “Weichen,” in *Handbuch Eisenbahninfrastruktur*, L. Fendrich and W. Fengler, Eds., Berlin, Springer-Verlag GmbH, 2019, pp. 249-326.

[4] "The world's most valuable resource is no longer oil, but data," *The Economist,* 2017.

[5] S. Stern, A. Behrendt, E. Eisenschmidt, S. Reimig, L. Schirmers and I. Schwerdt, "The rail sector's changing maintenance game: How rail operators and rail OEMs can benefit from digital maintenance opportunities," McKinsey & Company, 2017.

[6] *Alstom white paper: Artificial Intelligence across Alstom,* Alstom, 2022.

[7] *The journey toward AI-enabled railway companies,* International Union of Railways (UIC), 2024.

[8] *ISO 13372: Condition Monitoring and Diagnostics of Machines - Vocabulary,* International Organization for Standardization, 2012.

[9] K. Warwick, Artificial Intelligence - The Basics, Routledge, 2012.

[10] A. Ng, "Machine Learning Specializaton," Stanford Online, DeepLearning.AI.

[11] A. Ng, "AI for Everyone," DeepLearning.AI.

[12] A. Ng, "Machine Learning in Production," DeepLearning.AI.

[13] "Condition Monitoring," MathWorks, [Online]. Available: https://www.mathworks.com/discovery/condition-monitoring.html.

[14] *EN 15437-1 Railway applications - Axlebox condition monitoring - Interface and design requirements - Part 1: Trackside equipment and rolling stock axlebox,* The European Committee for Standardization (CEN), 2022.

[15] *EN 15437-2 Railway applications - Axlebox condition monitoring - Interface and design requirements - Part 2: Performance and design requirements of on-board systems for temperature monitoring,* The European Committee for Standardization (CEN), 2022.

[16] M. Tesar and S. Leineweber, "Digitaler Zwilling: Höhere Verfügbarkeit für Radsatzlager," *EI - Der Eisenbahningenieur,* no. 2, pp. 28-31, 2025.

[17] M. Breemeersch und V. Brundisch, „Das mechatronische Fahrwerk - aktuelle Entwicklungen," *ZEVrail,* Bd. 138, Nr. 3, pp. 90-98, 2014.

[18] D. Kraft, "Mechatronik und Zustandsüberwachung in Schienenfahrzeug-Fahrwerken," in *The 2nd International Railway Symposium Aachen*, Aachen, Germany, 2019.

[19] M. Spiryagin, S. Bruni, C. Bosomworth, P. wolfs and C. Cole, Rail Vehicle Mechatronics, Boca Raton, FL, USA: CRC Press, 2022.

[20] C.-P. Fritzen, "Vibration-Based Techniques for Structural Health Monitoring," in *Structural Health Monitoring*, D. Balageas, C. Fritzen and A. Güemes, Eds., London, UK, ISTE Ltd., 2006, pp. 45-224.

[21] R. B. Randall, Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications, West Sussex, UK: John Wiley & Sons, Ltd., 2011.

[22] J. Kolerus and E. Becker, Condition Monitoring und Instandhaltungsmanagement, Tübingen: expert verlag, 2022.

[23] F. A. Everest and K. C. Pohlmann, Master Handbook of Acoustics, The McGraw-Hill Companies, 2009.

[24] G. James, D. Witten, T. Hastie and R. Tibshirani, An introduction to statistical learning - with applications in R, Springer, 2013.

[25] C. R. Farrar and K. Worden, Structural Health Monitoring: A Machine Learning Perspective, West Sussex: John Wiley & Sons, Ltd., 2013.

[26] A. Bevan and P. Shackleton, "Comparison of options for acoustic monitoring of axle bearings: Cost-benefit analysis," Rail Safety and Standards Board Ltc., 2023.

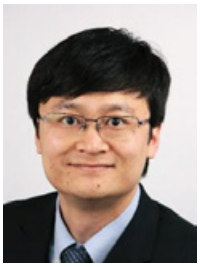
[27] Y. Niu, D. Kraft, C. Manns and M. Breemeersch, "Cost-Efficient Solutions for Bogie Health Monitoring," *ZEVrail,* vol. 147, no. 4, pp. 128-133, 2023.

[28] "DIN EN 13848-1: Bahnanwendungen - Oberbau - Gleislagequalität - Teil 1: Beschleunigung der Gleisgeometrie," Deutsches Institut für Normung e.V., 2019.

[29] M. Treveil and Dataiku, Introducing MLOps, O'Reilly Media, Inc., 2020.

Session 10
Yan Niu, Dietmar Kraft, Kai Heimann

Authors



Niu, Yan

Bachelor degree in Mechanical Engineering and Automation from Shanghai Jiao Tong University (SJTU) in China in 2006. Master degree in Mechatronics and PhD degree in Mechanical Engineering from University of Siegen in Germany in 2009 and 2018. Research Assistant with focus on Structural Health Monitoring (SHM) at Institute of Mechanics and Control Engineering - Mechatronics at University of Siegen from 2009 to 2013. CBM Data Engineer at Bombardier Transportation GmbH from 2013 to 2021. Data Engineer at Alstom since 2021.



Kraft, Dietmar

Studies of Mechanical Engineering with specialisation on railway technology at Technical University of Aachen (RWTH) until 1992. Research Assistant with focus on railway mechatronics at Institute of Railway Vehicles and Conveyor technique (IFS) at RWTH Aachen. PhD degree in 1998. Since 1998 different functions in bogie development at Alstom Switzerland (resp. predecessor companies) including vehicle dynamics, testing, safety, project and bid engineering, product management, mechatronic projects. Since 2018 product manager mechatronics.



Heimann, Kai

Bachelor's and Master's degrees in Mechanical Engineering from the University of Siegen, Germany, in 2012 and 2013. Executive MBA from the University of Siegen Business School in 2019. From 2014 to 2021, held engineering roles at Bombardier Transportation GmbH with a focus on technical project execution in bogie development for regional and high-speed trains. Since 2021, working at ALSTOM in roles related to mechanical design, R&D program management, and innovation. Currently leading a team of engineers in R&D, with responsibilities spanning strategic technology development, innovation and European-funded research projects.

IRSA 2025

SESSION 11

FAHRZEUG UND FAHRWEG

VEHICLE AND CHASSIS

Session 11
Stefan Lipinski, Gustav Hellings, Christian Schindler

Optimierung eines selbststeuernden Einzelachsfahrwerks mittels eines genetischen Algorithmus

5. International Railway Symposium Aachen

Lipinski, Stefan[1], Hellings, Gustav[1], Univ.-Prof. Dr.-Ing Schindler, Christian[1]

[1]Institut für Schienenfahrzeuge der RWTH Aachen

Zusammenfassung

Dieses Paper beschreibt die Optimierung des Fahrwerks des Aachener Rail Shuttle (ARS) mittels des genetischen Algorithmus NSGA-II. Ziel der Optimierung war es, das Fahrwerk so auszulegen, dass ein Kompromiss aus stabiler Geradeausfahrt und verschleißarmer Bogenfahrt erreicht wird.

Dazu werden im Rahmen von Voruntersuchungen zunächst die zu optimierenden Parameter, die zur Bewertung heran zu ziehenden Zielgrößen und die einzusetzenden Fahrszenarien definiert.

Für die Optimierung erzeugt der genetische Algorithmus iterativ Parametersets auf Basis ihrer „Fitness“, die ihrerseits aus den zuvor festgelegten Zielgrößen bestimmt wird. Die Ermittlung der Zielgrößen eines Parametersets erfolgt mit Hilfe von Simulationen des Mehrkörpermodells des ARS.

Der Algorithmus ist in der Lage, das Laufverhalten sowohl bezüglich Geradeaus- als auch Bogenfahrt zu verbessern. Der vorgeschlagene Parametersatz weist eine gegenüber dem Ausgangszustand verringerte Querfedersteifigkeit bei gleichzeitig erhöhter Dämpfersteifigkeit auf. Der Rechenzeitaufwand für eine Optimierung beträgt ca. achteinhalb Stunden.

Keywords: Einzelachsfahrwerk; Lauftechnik; passive Steuerung; genetischer Algorithmus; NSGA-II

1 Einleitung

Der Aachener Rail Shuttle (ARS) ist ein leichter, batteriebetriebener Schienenbus der dritten Generation, der fahrerlos mit erhöhter Taktdichte auf ländlichen Strecken eingesetzt werden soll, um den Schienenpersonennahverkehr auf dem Land attraktiver zu gestalten [1]. Der ARS besteht nicht wie üblich aus einem selbstragenden, röhrenförmigen Wagenkasten, sondern aus einem Chassis ähnlich dem Aufbau eines Flachgüterwagens, auf dem über die Sekundärfederstufe die Leichtbaufahrgastzelle abgestützt ist. Als Teil des Leichtbaukonzepts wurden anstelle von herkömmlichen Drehgestellen Einzelachsfahrwerke vorgesehen, wobei der große Achsabstand von 9,4 m kombiniert mit den üblicherweise bogenreichen Trassierungen von ländlichen Strecken eine Steuerung der Fahrwerke notwendig macht. Für den ARS wurde aus diesem Grund ein selbststeuerndes Radsatz-Einzelachsfahrwerk entwickelt, das von den Führungskräften im Rad-Schiene-Kontakt gelenkt wird.

1.1 Fahrwerkkonstruktion

Das untersuchte Fahrwerk ist ein angetriebenes Einzelachsfahrwerk mit einem Radsatz, der gegenüber dem umschließenden Fahrwerkrahmen rotatorisch gelagert ist, siehe Abb. 1. Gegenüber dem Chassis kann der Fahrwerkrahmen über die in blau dargestellten Primärfedern vertikal federn. Die insgesamt vier in grün dargestellten, unter einem

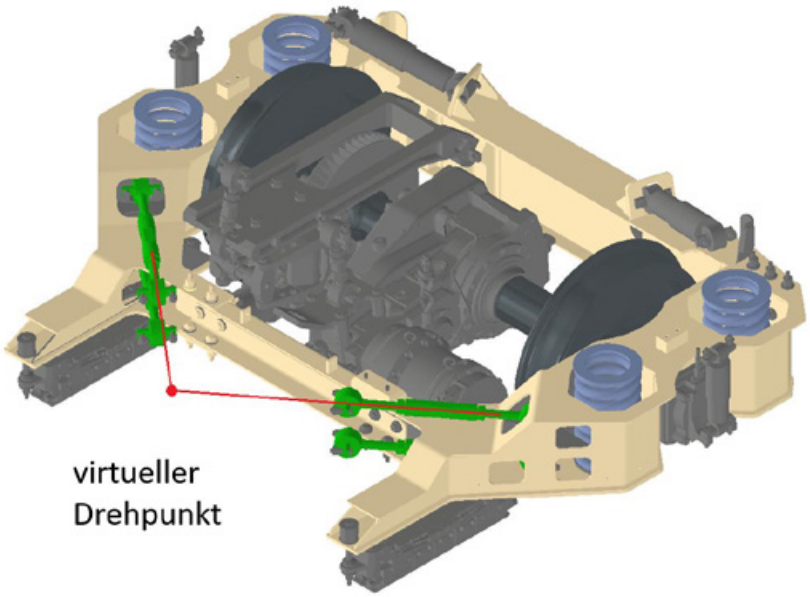


Abb. 1: Isometrische Ansicht des untersuchten Fahrwerks mit virtuellem Drehpunkt

Winkel bezüglich der Hochachse verbauten Schräglenker verbinden über Schraubverbindungen den Fahrwerkrahmen mit dem Chassis. Durch die Gummi-Metall-Elemente in den Lenkeraugen lassen die Lenker Hubbewegungen mit bis zu 10° Rotation und Schwenkbewegungen mit bis zu 6° Auslenkung zu.

Vertikale Relativbewegungen zwischen Fahrwerkrahmen und Chassis werden durch die am Achslager verbauten Vertikaldämpfer, laterale Bewegungen durch die am vorderen Querbalken angebrachten Schlingerdämpfer gedämpft. Die Bremsvorrichtung und das Getriebe werden jeweils über Drehmomentstützen im Chassis aufgehängt. Weiterhin sind tief aufgehängte Magnetschienenbremsen verbaut.

Die flexiblen Gummi-Metall-Elemente der Lenker ermöglichen Freiheitsgrade für Hub- und Rollbewegungen zwischen Fahrwerk und Chassis, um die Realisierung einer entgleisungssicheren Primärfederung zu ermöglichen.

Da die Längslenker außerdem unter einem Winkel um die Hochachse verbaut sind, entsteht ein virtueller Drehpunkt, um den das Fahrwerk bei Einwirkung einer Querkraft seitlich unter dem Chassis ausschwenken kann, siehe Abb. 1. Der Drehpunkt für diese Bewegung ergibt sich dabei aus dem Schnittpunkt der Lenkerachsen, der aufgrund der Anordnung der Lenker gegenüber dem Radsatz in Richtung Fahrzeugmitte verschoben ist. Weiterhin handelt sich bei der Bewegung um eine zusammengesetzte Translation und Rotation, da sich der Drehpunkt bedingt durch die Verschiebung der Lenkerachsen beim Ausschwenken ebenfalls bewegt.

1.2 Selbstlenkung durch Spurführung

Die schwenkbare Anbindung des Fahrwerks an das Chassis ermöglicht es, dass die Spurführungskräfte das Fahrwerk bei einer Bogenfahrt selbstständig zum Einlenken bringen, da sie einen Hebelarm gegenüber dem virtuellen Drehpunkt aufweisen. Dabei wird das Fahrwerk so lange weiter ausgelenkt, bis die Primärfedern das entsprechende Gegenmoment erzeugen.

Da alle Antriebs- und Bremskräfte an beiden Rädern gleichermaßen auftreten, wirken sie parallel, aber mit entgegengesetzten Hebelarmen auf den Fahrwerkrahmen, sodass weder das Antreiben noch das Bremsen ein Ausdrehmoment erzeugt. Auf diese Weise werden die auf das Fahrwerk wirkenden Längskräfte von den Querkräften entkoppelt, die Längskräfte beeinflussen auch im Bogen nicht den Einlenkwinkel des Fahrwerks.

Die Selbstlenkung durch Spurführung ist jederzeit gewünscht, da das Fahrwerk dabei stets in Richtung niedriger Schräglaufwinkel strebt.

1.3 Auslenkung durch Querbeschleunigung

Neben der gewünschten Selbstlenkung tritt zusätzlich situationsabhängige Auslenkung durch Querbeschleunigung auf, da die Zentrifugalkräfte, die auf das Chassis und die Fahrgastzelle wirken, nur durch die Querfederwirkung der Primärfedern und die Lenker übertragen werden können. Auch in diesem Fall entsteht ein Auslenkmoment, da die Primärfedern einen Hebelarm gegenüber dem virtuellen Drehpunkt aufweisen. Das Ausmaß der Auslenkung durch Querbeschleunigung hängt von der Federsteifigkeit und der Positionierung der Lenker, der Federn und zu einem kleinen Teil der Drehmomentstützen von Bremse und Getriebe ab.

Die Auslenkung durch Querbeschleunigung kann vorteilig sein, um in engen Bögen einen zusätzlichen Beitrag zum Einlenkwinkel zu liefern, den die Spurführungskräfte allein nicht aufbringen können. Gleichzeitig führt dies aber in weiteren Bögen dazu, dass das Fahrwerk stärker als notwendig einlenkt. Zusätzlich kann es bei langsamer Fahrt in überhöhten Bögen dazu führen, dass sich das Fahrwerk antiradial einstellt. Die Auslenkung durch Querbeschleunigung muss daher auf ein sinnvolles Maß begrenzt werden.

2 Methodik und Voruntersuchungen

2.1 Optimierung mittels NSGA-II und Mehrkörpersimulation

Die Optimierung erfolgt auf Grundlage des sogenannten Non-dominated Sorting Genetic Algorithm (NSGA-II). NSGA-II ist ein multikriterieller, evolutionärer Algorithmus, der für ein Problem mit mehreren Zielgrößen einen Satz von Lösungen erzeugt, wobei jede der Lösungen eine der Zielgrößen nur verbessern kann, indem sie andere Zielgrößen verschlechtert [2]. Da in der Lauftechnik die stabile Geradeausfahrt und die verschleißfreie Bogenfahrt gegensätzliche Anforderungen an das Fahrwerk stellen [3], eignet sich NSGA-II, um eine Fahrwerksauslegung zu finden, die einen sinnvollen Kompromiss dieser Eigenschaften bietet.

NSGA-II nutzt angelehnt an die Evolutionstheorie die Mechanismen der Paarung, Mutation und Selektion, um auf Basis der Fitness der Individuen einer Generation neue Generationen zu erzeugen, die das Optimierungsproblem besser lösen als die jeweils vorausgegangene [2]. Angewandt auf die Fahrwerksoptimierung stellt jeweils ein Satz an zu untersuchenden Fahrwerksparametern ein Individuum dar und die Fitness dieses Individuums wird mittels seiner Laufeigenschaften in Gerade und Bogen bewertet.

Um die Bewertung der Laufeigenschaften zu ermöglichen, werden die Parameter eines Individuums an ein Mehrkörpersimulationsmodell des Fahrzeugs übergeben. Es werden Fahrten auf vorab festgelegten Strecken mittels Zeitschrittintegration simuliert und die Simulationsergebnisse bezüglich der definierten Zielgrößen ausgewertet.

Die Optimierung beginnt mit der Erzeugung von zufälligen Parametern für die Startindividuen, deren Fitness mit Hilfe der Mehrkörpersimulationen bewertet wird. Anschließend kombiniert der Algorithmus die Parameter der Individuen mit der besten Fitness (Paarung) und verändert zusätzlich bei einem Teil der Individuen Parameter auf zufälliger Basis (Mutation). Auf diese Weise werden iterativ neue Generationen an Individuen erzeugt, wobei die Beibehaltung von Individuen mit besserer Fitness zu einer Konvergenz in Richtung gutes Laufverhalten führt (Selektion). Zusätzlich versucht NSGA-II einen Mindestabstand zwischen den Ergebnissen der Individuen einzuhalten, um eine Konvergenz gegen lokale Minima zu vermeiden. [2]

Für die Fahrwerksoptimierung wurden insgesamt 50 Generationen mit je 20 Individuen ausgewertet.

2.2 Auswahl der Parameter

Für die Erzeugung und Bewertung der Individuen mittels NSGA-II ist es notwendig, vorab die zu optimierenden Parameter festzulegen. Die Auswahl der Parameter muss dabei so erfolgen, dass eine Variierung eines Parameters auch zur Veränderung mindestens einer der Zielgrößen führt. Gleichzeitig müssen ausreichend Parameter betrachtet werden, sodass jede Zielgröße durch mindestens einen Parameter beeinflusst wird.

Für die Wahl der Parameter werden zunächst alle Bauteile betrachtet, die Einfluss auf das Laufverhalten nehmen können. Da die Komplexität des Problems jedoch mit der Anzahl an variierten Parametern steigt, wird durch Voruntersuchungen analysiert, welche Parameter den größten Einfluss auf das Laufverhalten nehmen. Wenn zwei Parameter den gleichen Effekt auf das Laufverhalten nehmen, wird nur der Einflussreichere in die Betrachtung aufgenommen, um die Komplexität weiter zu senken.

Die Komponenten, die laterale oder longitudinale Kräfte zwischen Fahrwerk und Chassis erzeugen oder übertragen sind die Primärfedern durch ihre Querfederwirkung, die Schlingerdämpfer, sowie die Gummi-Metall-Lager der Längslenker und der Drehmomentstützen von Getriebe und Bremse.

Die primären Schraubenfedern erzeugen bei Querversatz zwischen den Anbindungspunkten an Chassis und Fahrwerk eine Rückstellkraft. Diese wirkt einer Auslenkung jederzeit entgegen und weist einen Hebelarm gegenüber dem virtuellen Drehpunkt auf. Dadurch dient sie als Zentrierung für das Fahrwerk. Da die Rückstellkraft durch die radiale Federsteifigkeit bestimmt wird, wird letztere als Parameter für die Optimierung aufgenommen.

Die in Schienenfahrzeugen eingesetzten Schlingerdämpfer sind häufig als nichtlineare Dämpfer ausgeführt, die hohe Dämpferkräfte bereits bei niedrigen Dämpfergeschwindigkeiten erzeugen, während sie bei höheren Dämpfergeschwindigkeiten über eine abgeflachte Kennlinie verfügen [4]. Dies soll eine effektive Dämpfung des Wellenlaufs in der Geraden ermöglichen, ohne gleichzeitig die Ausdrehbewegung bei Bogenein- und -ausfahrt zu stark zu dämpfen.

Da die Schlingerdämpfer am untersuchten Fahrwerk an der dem Drehpunkt gegenüberliegenden Seite des Fahrwerks angebracht sind und damit den größten Hebelarm aller untersuchten Bauteile aufweisen, wurden im ARS auch Dämpfer mit nichtlinearer Kennlinie verbaut. Die Dämpfersteifigkeit der ersten Stufe hat einen maßgeblichen Einfluss auf die Stabilität des Fahrwerks und wird daher als Optimierungsparameter aufgenommen. Da außerdem eine sinnvolle Umschaltgeschwindigkeit zwischen starker und schwacher Dämpfung von der Gesamtauslegung des Fahrwerks abhängt, wird diese ebenfalls optimiert.

Neben den Schraubenfedern erzeugen auch alle Gummi-Metall-Lager Rückstellkräfte, da sie bei einer Ausdrehbewegung kardanisch belastet werden. Aufgrund der kleinen Torsionswinkel liegen diese Kräfte jedoch um eine Größenordnung niedriger als die Rückstellkräfte der Schraubenfedern. Dies gilt ebenso für den Vergleich der Dämpfungskräfte des Schlingerdämpfers mit denen der Gummi-Metall-Lager. Daher wird die Feder-Dämpfer-Charakteristik der Lager in der Optimierung nicht betrachtet.

Da darüber hinaus beim Ausdrehen des Fahrwerks die Anbindungspunkte aller Bauteile näherungsweise entlang einer Kreisbahn um den virtuellen Drehpunkt verschoben werden, beeinflusst auch die konstruktive Position dieser Punkte sowohl die wirkenden Kräfte als auch Drehmomente. Da eine Verschiebung der Anbindungspunkte eines Bauteils jedoch die gleiche Auswirkung wie die Änderung des jeweiligen Bauteilparameters bewirkt, werden die Positionen dieser Punkte nicht betrachtet.

Da die Selbststeuerung des Fahrwerkkonzepts darauf beruht, dass der virtuelle Drehpunkt einen Hebelarm gegenüber der Radsatzachse aufweist, soll neben den genannten Bauteilparametern auch dieser Hebelarm optimiert werden. Dazu werden die

Anbindungspunkte der Längslenker am Chassis in Querrichtung verschoben, sodass sich der Einbauwinkel um die Hochachse ändert. Da der virtuelle Drehpunkt im Schnittpunkt der Lenkerachsen liegt, führt eine Erhöhung des Einbauwinkels somit zu einer Senkung des Hebelarms zwischen Drehpunkt und Radsatzachse. Im Rahmen der Optimierung dient der Einbauwinkel als veränderlicher Parameter.

2.3 Auswahl der Zielgrößen

Neben den zu optimierenden Parametern müssen auch die Zielgrößen, anhand derer die Fitness des Laufverhaltens bewertet wird, vorab festgelegt werden. Hierfür können alle Größen gewählt werden, die eine quantifizierbare Aussage über die Erfüllung der gewünschten Eigenschaften bezüglich Laufverhalten zulassen. Auch hier steigt die Komplexität des Problems mit steigender Anzahl der Zielgrößen. Da außerdem jede Zielgröße einen Anteil an der Gesamtfitness trägt, sinkt der Einfluss einer einzelnen Zielgröße mit steigender Zielgrößenzahl. Daher sollen nur so viele Zielgrößen wie nötig eingeführt werden.

Der Zweck des selbstgesteuerten Fahrwerks ist es, ein Einlenken im Bogen zu ermöglichen, um den Verschleiß durch auftretenden Schlupf zu minimieren. Aus diesem Grund bietet sich die kumulierte Verschleißarbeit W_v bei einer Bogenfahrt als Bewertungsgröße an. Sie errechnet sich aus

$$W_V = \sum_{t=0}^{t_{end}} (T_x \cdot s_x + T_y \cdot s_y) \cdot v_{ref} \cdot \Delta t$$

Dabei sind T_x und T_y die Kraftschlusskräfte im Radkontakt, s_x und s_y sind die Schlüpfe entlang der jeweiligen Richtung, v_{ref} ist die Referenzgeschwindigkeit des Schlupfes und Δt ist die Zeitschrittweite der Integration. Für die Verschleißbewertung B_v werden die Verschleißarbeiten der jeweiligen Bögen mit dem Gewichtungsfaktor G_R des Bogens, der sich aus dem Anteil des Bogenradius an allen Bögen der Strecke ergibt, multipliziert und aufsummiert:

$$B_v = \sum W_{v,R} \cdot G_R$$

Das Fahrwerk lenkt, wie in Kapitel 1 beschrieben nicht nur durch die Spurführungskräfte, sondern auch durch die Zentrifugalkraft aus. Da das Ausmaß und auch die Richtung dieses Lenkeffekts jedoch nur von der aktuell wirkenden Querbeschleunigung abhängt, kann im Allgemeinen nicht davon ausgegangen werden, dass dieser Lenkeffekt das Fahrverhalten immer verbessert. Aus diesem Grund soll neben dem Verschleiß zusätzlich der im Bogen verbleibende Schräglaufwinkel bewertet werden. Dazu wird für Bögen mit den Radien 600 m bis 264 m jeweils der Absolutwert des im konstanten Bogen

verbleibenden Schräglaufwinkels per Simulation bestimmt, mit der Häufigkeit des jeweiligen Bogenradius gewichtet und anschließend aufsummiert. Diese gewichtete Summe von Schräglaufwinkeln stellt die Bewertungsgröße für den Algorithmus dar und muss minimiert werden:

$$B_\alpha = \sum \alpha_i \cdot G_R$$

Der Vorteil dieser Bewertungsgröße liegt darin, dass das vorteilige Verhalten bei engen und das nachteilige Verhalten bei weiten Bögen gleichermaßen implizit abgedeckt wird. Ebenso wird das Zusammenwirken von Einlenkung durch Spurführung und Einlenkung durch Zentripetalkraft korrekt erfasst und gemeinsam bewertet.

In der Geradeausfahrt muss die Stabilität bei Höchstgeschwindigkeit sichergestellt werden, um ein Entgleisen zu verhindern. Zur Bewertung der Stabilität dient der Median der Amplitude der Radsatzquerauslenkung über 5 Sekunden des Grenzzyklus, mit dem das Fahrzeug nach einmaliger halbcosinusförmiger Anregung auf idealem Gleis schlingert.

$$B_S = median(\hat{y}_{RS})_{t=5s}$$

Da ein stabiler Grenzzyklus mit Radaufstandspunkten innerhalb der Fahrfläche kein erhöhtes Risiko gegenüber einem vollständig abklingenden Zyklus darstellt, wird die Betrachtung der Geradeausfahrt außerdem ergänzt durch die Bewertung des Fahrkomforts bei Fahrt mit Höchstgeschwindigkeit auf einem mit Gleislageunebenheiten behafteten Gleis. Als Bewertungsgröße dient der durchschnittliche Kontinuierliche Komfort in lateraler Richtung entsprechend der EN 12299 [5] bei einer Fahrt über eine Minute.

$$B_K = \frac{1}{12} \cdot \sum_{t=60s} C_{Cy}$$

Weiterhin können die Bewertungsgrößen in NSGA-II unterschiedlich gewichtet werden. Dies wurde genutzt, um einerseits zwischen Haupt- und Nebenbewertungsgrößen zu unterscheiden und andererseits, um die Gegebenheiten der Strecke zu berücksichtigen.

Der Verschleiß bei Bogenfahrt sowie die Radsatzauslenkung in der Geraden sind weit verbreitete Bewertungsgrößen, die sich auch über unterschiedliche Fahrzeuge hinweg sinnvoll vergleichen lassen. Daher werden diese im Rahmen dieser Untersuchung als Hauptbewertungsgrößen eingesetzt und im Algorithmus mit doppelter Gewichtung versehen. Der gewichtete Schräglaufwinkel und der Fahrgastkomfort sollen die Hauptbewertungsgrößen ergänzen und eine an das untersuchte Fahrzeug angepasste

Optimierung ermöglichen, weshalb diese als Nebenbewertungsgrößen im Algorithmus nur mit einfacher Gewichtung versehen werden.

Auf der zur Optimierung genutzten Strecke Düren-Heimbach liegt das Verhältnis aus gerader Strecke zu Bögen bei 1:1,6. Deshalb werden die Gewichtungen von Verschleiß und gewichtetem Schräglaufwinkel im Bogen zusätzlich um den Faktor 1,6 erhöht.

Insgesamt ergeben sich damit die Gewichtungsfaktoren anhand Tabelle 1

Tabelle 1: Gewichtungsfaktoren der einzelnen Zielgrößen

Gewichteter Verschleiß B_V	Gewichteter Schräglaufwinkel B_α	Radsatzauslenkung B_S	Querkomfort B_K
3,2	1,6	2,0	1,0

2.4 Auswahl der Streckenparameter

Da die Bewertung der Fitness der Individuen auf Basis von Zeitschrittintegrationsrechnungen eines Mehrkörpersystems vorgenommen wird, hat die Auswahl der Strecken und der dabei gefahrenen Geschwindigkeit ebenfalls einen wesentlichen Einfluss auf das Ergebnis. Daher müssen die Fahrszenarien so gewählt werden, dass der Einfluss der Optimierungsparameter auf die definierten Zielgrößen möglichst groß wird, da der Algorithmus dann am schnellsten gegen die optimalen Lösungen konvergiert.

Für die Stabilitätsuntersuchung wurde als Fahrszenario eine Geradeausfahrt mit 110 km/h auf einem ideal geradem, ungestörten Gleis mit einmaliger, halbcosinusförmiger Anregung gewählt [4], wobei die Fahrgeschwindigkeit 110 % der projektierten Höchstgeschwindigkeit des Fahrzeugs entspricht.

Da die Untersuchung des Komforts auf einem idealen Gleis nicht zielführend ist, wurde außerdem eine weitere Geradeausfahrt bei 110 km/h, jedoch auf einem fehlerbehafteten Gleis – entsprechend den Vorgaben aus DIN EN 14363 [6] für Geschwindigkeiten unterhalb von 120 km/h – vorgesehen. Die Bewertung des Komforts erfolgt für eine Fahrzeit von einer Minute.

Für den Verschleiß und den verbleibenden Schräglaufwinkel wurden insgesamt sechs Simulationsszenarien von Bogenfahrten vorgesehen. Alle Bogenfahrten beginnen auf einer Geraden, die über einen Übergangsbogen in einen Bogen mit festgelegtem Radius führt. Die eingesetzten Radien wurden äquidistant zwischen 600 m und 180 m

Bogenradius gewählt und betragen 600 m, 516 m, 432 m, 348 m, 264 m und 180 m. Die Geschwindigkeit ist in allen Simulationen so gewählt, dass sich eine unausgeglichene Querbeschleunigung von 1 m/s^2 ergibt.

Um die Gewichtungen in NSGA-II sinnvoll zu wählen, wurde die Optimierung beispielhaft für die Strecke Düren-Heimbach durchgeführt. Die Gewichtungen zwischen den Zielgrößen für Geradeausfahrt und Bogenfahrt wurden entsprechend der Häufigkeiten von gerader Strecke zu bogiger Strecke angepasst, und die Gewichtungen der verbleibenden Schräglaufwinkel wurden anhand der Verteilung der jeweiligen Bogenradien umgesetzt.

3 Ergebnisse

Abb. 2 zeigt die relativen Bewertungen von Geradeaus- und Bogenfahrt für die Individuen der finalen Generation. Jeder Punkt im Diagramm entspricht den Zielgrößen eines bestimmten Individuums. Für die Darstellbarkeit der Ergebnisse wurden die Zielgrößen jeweils auf die Werte des momentan im ARS verbauten Parametersatzes normiert und entsprechend ihrer Gewichtung zusammengefasst. Somit bedeutet ein Wert kleiner als 1 eine Verbesserung gegenüber dem Ist-Zustand, während ein Wert über 1 eine Verschlechterung angibt.

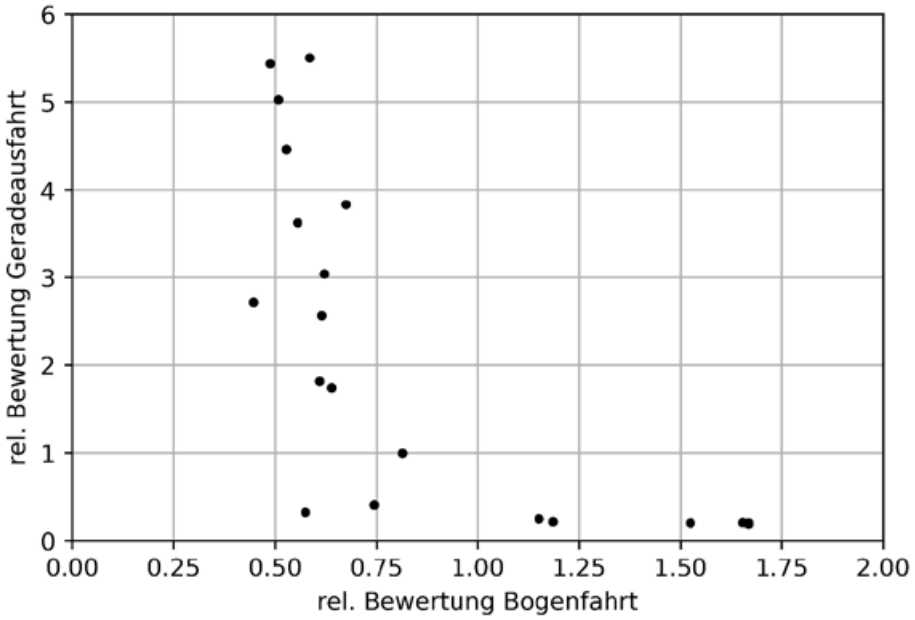


Abb. 2: relative Bewertungen der Individuen der finalen Generation; normiert auf Bewertung der ARS-Konfiguration

Die relative Bewertung der Bogenfahrt liegt für 14 der 20 Individuen unterhalb von 1, die relative Bewertung der Geradeausfahrt liegt für acht Individuen unterhalb von 1. Bei zwei Individuen konnten sowohl die Bogenfahrt als auch die Geradeausfahrt verbessert werden.

Als beste Kompromisslösung ergibt sich das Individuum, das sowohl in der Bogen- als auch in der Geradeausfahrt die größte Verbesserung erreicht hat. Tabelle 2 zeigt die gefundenen Parameter sowie die ursprünglichen Parameter zum Vergleich.

Tabelle 2: Parametervergleich zwischen gefundener Lösung und ARS-Konfiguration

Parametersatz	c_{Feder}	$d_{Dämpfer}$	$v_{schalt,Dämpfer}$	α_{Lenker}
	[N/mm]	[Ns/mm]	[m/s]	[°]
Individuum 5	67	999	0,11	61,5
ARS	171	470	0,02	50

Abb. 3 zeigt die relativen Bewertungen für die einzelnen Zielgrößen. Bezüglich des gewichteten Schräglaufwinkels erreichen alle Individuen eine Verbesserung gegenüber der derzeit verbauten Konfiguration. Ebenso verbessert die Mehrheit der gefundenen Parametersätze die im Bogen auftretende Verschleißarbeit. Der Querkomfort nach EN 12299 fällt bei der Hälfte der Individuen besser aus. Die Radsatzauslenkung ist bei 8 Individuen kleiner als im Ist-Zustand.

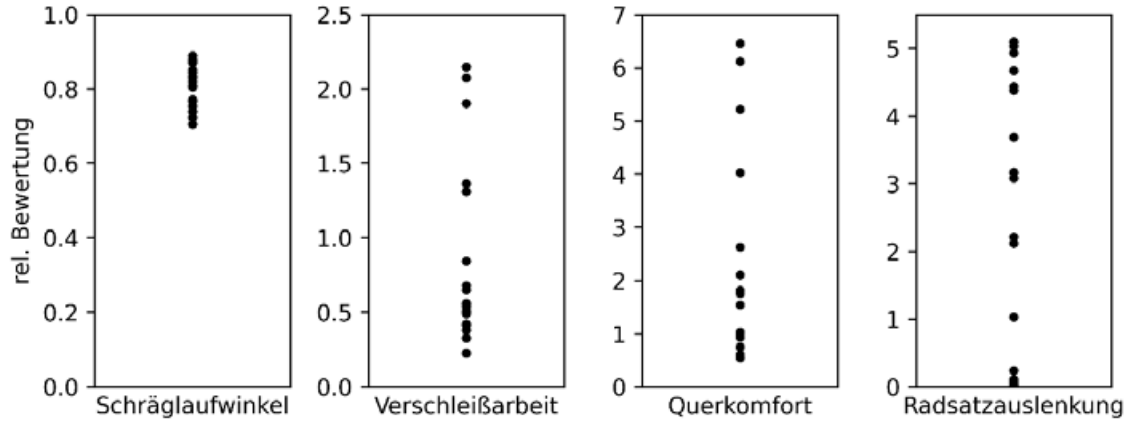


Abb. 3: Zielgrößen der Optimierung für die Individuen der finalen Generation; normiert auf Zielgrößen der ARS-Konfiguration

Abb. 4 zeigt die Verteilungen der Parameter der Individuen der finalen Generation. Auch hier wurden die Parameter jeweils auf die momentan im ARS verbauten Bauteile normiert.

Die durch NSGA-II gefundenen Federsteifigkeiten sind bei allen Individuen niedriger als beim ARS und bewegen sich in einem Bereich zwischen 35 % und 55 % der ursprünglichen Federsteifigkeit. Die Dämpfersteifigkeit liegt bei sechs Individuen höher, beim Rest tiefer als beim ARS. Die Umschaltgeschwindigkeit des zweistufigen Dämpfers liegt für 18 Individuen höher als beim ARS, ein großer Teil der Werte verteilt sich außerdem über einen breiten Bereich zwischen Faktor 4 und Faktor 9. Der Lenkerwinkel teilt sich in zwei Gruppen, wobei 16 Individuen einen deutlich niedrigeren Winkel, vier Individuen einen deutlich größeren als den Ursprungswinkel aufweisen.

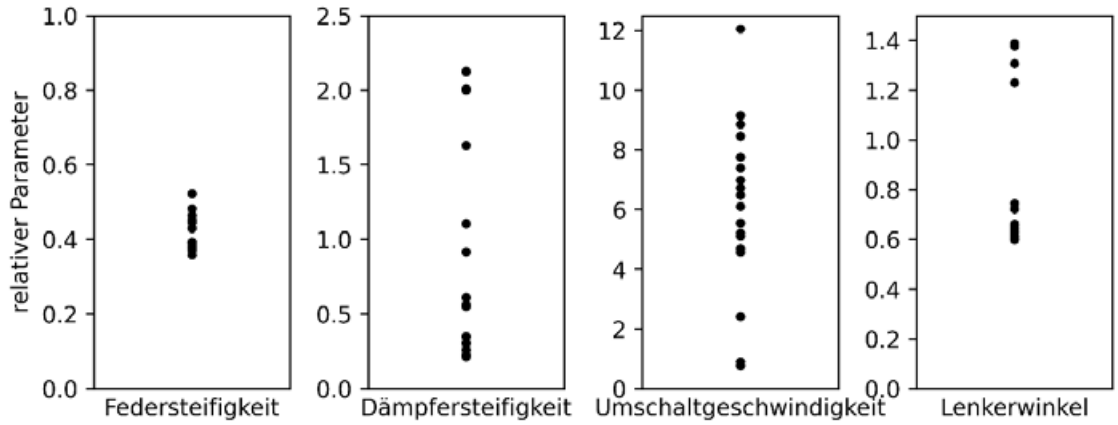


Abb. 4: Einzelparameter der Individuen der finalen Generation; normiert auf jeweilige Parameter der ARS-Konfiguration

4 Diskussion

4.1 Interpretation der Ergebnisse

Die in Abb. 2 gezeigten Ergebnisse der ermittelten Individuen liegen alle in einem Bereich, bei dem entweder die Geradeaus- oder die Bogenfahrt oder beides verbessert werden konnte. Dies zeigt, dass die momentane Auslegung des ARS, beispielhaft für die Strecke Düren-Heimbach demonstriert, noch nicht optimal ist. Insbesondere bei der Bogenfahrt können noch bessere Laufeigenschaften erzielt werden, da hier 14 der Individuen besser bewertet wurden. Die gewichteten Schräglaufwinkel liegen bei allen

Individuen niedriger, ebenso weisen alle Individuen eine niedrigere Federsteifigkeit als die ursprüngliche auf. Hieraus lässt sich ableiten, dass die momentan verbaute Feder eine zu hohe Querfedersteifigkeit aufweist und das Potential der Selbstlenkung des Fahrwerks dadurch noch nicht vollständig ausgeschöpft wird.

Bei der Geradeausfahrt wurden acht gefundene Individuen besser bewertet als die ursprüngliche Konfiguration. Abb. 5 zeigt die relativen Bewertungen für den Querkomfort und die Radsatzauslenkung der Individuen der finalen Generation. Wenn ein Individuum einen besseren Querkomfort bietet, ist auch die Stabilität dieses Individuums besser. Hieraus lässt sich schließen, dass die Individuen, bei denen die Zielgrößen der Geradeausfahrt nicht verbessert werden konnten, auch tatsächlich keine gleichwertige Kompromisslösung zu den restlichen Individuen darstellen.

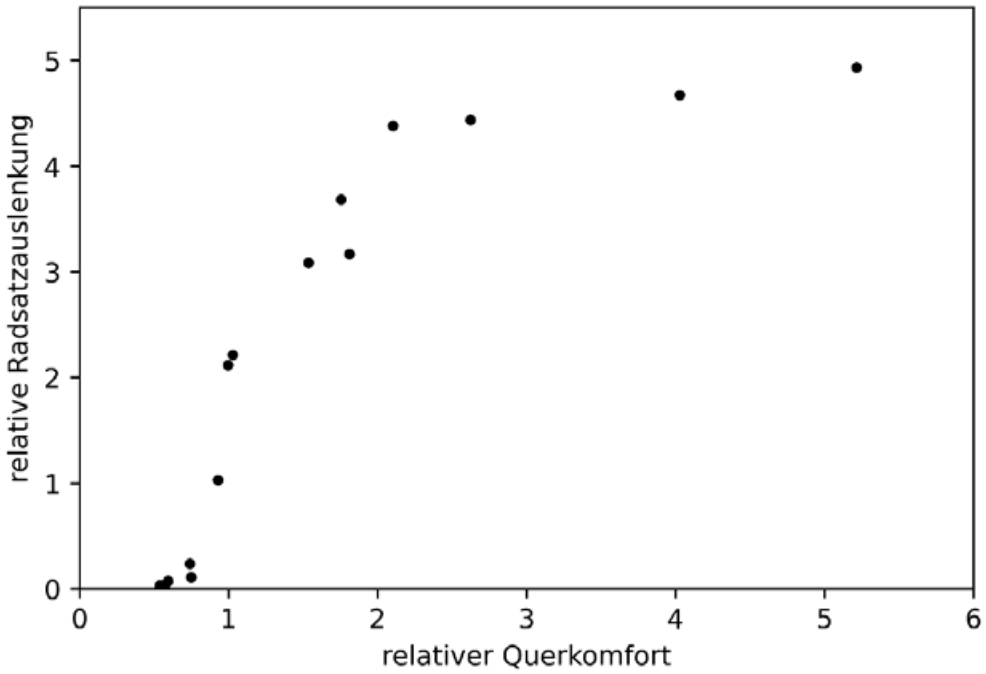


Abb. 5: relative Radsatzauslenkung in Abhängigkeit des relativen Querkomforts für die Individuen der finalen Generation, jeweils normiert auf Radsatzauslenkung und Querkomfort der ARS-Konfiguration

Wie in Kapitel 2 beschrieben, wählt NSGA-II die Individuen nicht rein auf Basis ihrer Fitness aus, sondern versucht zusätzlich einen Mindestabstand zwischen den gefundenen Lösungen einzuhalten, um eine Konvergenz gegen lokale Minima zu verhindern. Dies hat bei der Fahrwerkoptimierung dazu geführt, dass auch Individuen mit zu niedriger

Dämpfung beibehalten wurden, da sie aus Sicht der Bogenfahrt dennoch als beste Lösungen eingestuft werden.

4.2 Umsetzbarkeit der Ergebnisse

Da die Optimierung auf Basis von numerischer Simulation eines Mehrkörpermodells erfolgt ist, prüft der Algorithmus nicht, ob die ausgewählten Parameter eines Individuums auch tatsächlich konstruktiv umsetzbar sind.

Die Querfedersteifigkeit liegt für alle Individuen unterhalb der derzeit im ARS verbauten Federsteifigkeit in einem Intervall von ca. 62 bis 90 N/mm. Bei einer derzeit verbauten Vertikalfedersteifigkeit von 375 N/mm müsste eine Feder mit kleinerem Windungsdurchmesser bei größerer Federlänge verbaut werden, um die niedrigen Querfedersteifigkeiten bei gleichbleibender Vertikalsteifigkeit zu erreichen. Hier muss überprüft werden, ob es aus Bauraumsicht möglich und ggf. sinnvoll ist, auch die Vertikalfedersteifigkeit zu senken und größere Federwege in der Primärstufe zu akzeptieren.

Die mit NSGA-II bestimmte Dämpfersteifigkeit liegt mehr als doppelt so hoch wie die derzeit im ARS verbaute. Aufgrund des niedrigen Gewichts des Fahrzeugs wurden bei der Entwicklung Dämpfer mit mittlerer Steifigkeit gewählt. Daher ist es möglich, die Dämpfersteifigkeit an die Ergebnisse des Algorithmus anzupassen. Da die mit NSGA-II ermittelten Umschaltgeschwindigkeiten jedoch größtenteils deutlich höher als die des verbauten Dämpfers ausfallen, kann auch der Einsatz eines linearen, nicht schaltenden Dämpfers in Betracht gezogen werden. Dafür muss geprüft werden, ob der Dämpfer bei Umschaltgeschwindigkeiten oberhalb von 0,1 m/s bei den simulierten Szenarien überhaupt seine Umschaltgeschwindigkeit erreicht. Wenn dies nicht der Fall ist, ist auch der Verbau eines linearen Dämpfers eine Option.

Der Einbauwinkel der Lenker ist mit 61,5° größer als bei der derzeitigen Konstruktion. Hier muss geprüft werden, ob die Festigkeit des Fahrwerkrahmens und des Chassisquerträgers ausreicht, da die Lenkerkräfte, die für die Übertragung der Längskräfte aus dem Antreiben und Bremsen notwendig sind, über $1/\cos(\alpha_{Lenker})$ skalieren und damit mit steigendem Einbauwinkel immer größer werden.

Weiterhin war es notwendig, die Gewichtungen für den Algorithmus festzulegen. Daher wurde die Optimierung am Beispiel der Strecke Düren-Heimbach durchgeführt, sodass die ermittelten Parameter speziell für diese Strecke sinnvoll sind. Es ist jedoch ohne Änderung der Methodik möglich, die Optimierung für andere Strecken durchzuführen,

indem die Gewichtungen entsprechend den Rahmenbedingungen der gewünschten Strecke angepasst werden.

4.3 Anwendung eines genetischen Algorithmus zur Fahrwerksoptimierung

Der Vorteil der Anwendung eines genetischen Algorithmus liegt darin, dass eine Vielzahl von Parametern und (gegenläufigen) Zielgrößen gleichzeitig untersucht werden kann, um einen sinnvollen Kompromiss zu finden. Durch eine passende Wahl insbesondere der Zielgrößen kann eine schnelle Konvergenz gegen optimale Lösungen erreicht werden, sodass im Vergleich zu Monte-Carlo-Simulationen eine große Menge an Parameterkombinationen nicht überprüft und damit Rechenzeit gespart wird.

Die Wahl des genetischen Algorithmus spielt im Rahmen dieser Untersuchung nur eine untergeordnete Rolle. Ein Hauptvorteil von NSGA-II gegenüber anderen multikriteriellen, evolutionären Algorithmen liegt in der geringeren Rechenkomplexität und damit höheren Rechengeschwindigkeit [2]. Da hier die Bewertung der Individuen anhand von numerischer Zeitschrittintegration erfolgt, übersteigt die Simulationsdauer die reine Rechendauer des Algorithmus um mehrere Größenordnungen, sodass letztere keinen relevanten Einfluss auf die Gesamtdauer einer Optimierung nimmt.

5 Fazit

Das Ziel dieser Untersuchung war es, das selbststeuernde Einzelachsfahrwerk des ARS bezüglich seiner Laufeigenschaften in Geradeaus- und Bogenfahrt zu optimieren. Hierzu wurde der genetische Algorithmus NSGA-II für die iterative Erzeugung von Parametersätzen verwendet, während Mehrkörpersimulation eingesetzt wurde, um die jeweiligen Parametersätze bezüglich ihrer Laufeigenschaften zu analysieren und bewerten.

Die vorgeschlagene Methode ist in der Lage, das Laufverhalten bezüglich Rad-Schiene-Verschleißes und Schräglaufwinkel bei Bogenfahrt und bezüglich Stabilität und Fahrgastkomfort bei Geradeausfahrt zu verbessern.

Der gefundene Parametersatz besitzt eine etwa 60 % schwächere Querfedersteifigkeit der Primärfeder, während die Dämpfersteifigkeit 110 % höher als die des momentan im ARS

verbauten Parametersatzes liegt. Inwieweit diese Anpassungen konstruktiv umsetzbar sind, muss Gegenstand von weiteren Untersuchungen sein.

Die Optimierung wurde beispielhaft an der Strecke Düren-Heimbach durchgeführt, indem die Gewichtungen in NSGA-II entsprechend ermittelt wurden. Durch Anpassung dieser Gewichtungen ist es ohne weiteren Aufwand möglich, das Fahrzeug auch für andere Strecken zu optimieren. Der Rechenaufwand für die vollständige Optimierung betrug etwa 8 Stunden und 33 Minuten.

Da sich der Demonstrator des ARS derzeit noch in der Inbetriebnahme befindet, war es noch nicht möglich, Testfahrten durchzuführen. Aus diesem Grund konnte das Mehrkörpermodell des ARS zum jetzigen Zeitpunkt noch nicht verifiziert werden, weshalb die Ergebnisse dieser Untersuchung ebenfalls noch einer Verifizierung bedürfen.

Literatur

[1] C. Schindler, „The Aachen Rail Shuttle ARS - Autonomous and energy self-sufficient feeder transport," *Journal of Rail Transport Planning & Management,* 2022.

[2] K. Deb, A. Pratap, S. Agarwal und T. Meyarivan, „A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation, vol. 6, no. 2,* pp. pp. 182-197, 2002.

[3] K. Knothe und S. Stichel, Schienenfahrzeugdynamik, Heidelberg: Springer Berlin, 2003.

[4] O. Polach, M. Berg und S. Iwnicki, „Kap. 12 Simulation," in *Handbook of Railway Vehicle Dynamics*, Taylor & Francis Group, 2006, pp. 360-417.

[5] „DIN EN 12299 - Bahnanwendungen - Fahrkomfort für Fahrgäste - Messung und Auswertung," 2025.

[6] „DIN EN 14363 - Bahnanwendungen - Versuche und Simulationen für die Zulassung der fahrtechnischen Eigenschaften von Eisenbahnfahrzeugen - Fahrverhalten und stationäre Versuche," 2022.

Session 11
Stefan Lipinski, Gustav Hellings, Christian Schindler

Autoren



Lipinski, Stefan

Stefan Lipinski, M.Sc

Wissenschaftlicher Mitarbeiter am Institut für Schienenfahrzeuge der RWTH Aachen University (ifs)

Seffenter Weg 8, 52074 Aachen



Hellings, Gustav

Gustav Hellings, B.Sc

Studentische Hilfskraft am Institut für Schienenfahrzeuge der RWTH Aachen University (ifs)

Seffenter Weg 8, 52074 Aachen



Schindler, Christian

Univ.-Prof. Dr.-Ing Christian Schindler

Leiter des Instituts für Schienenfahrzeuge der RWTH Aachen University (ifs)

Seffenter Weg 8, 52074 Aachen

Session 11
Oliver Hagemann, Martin Kuhlins, Leon Beckmann

Herausforderungen des Engineerings für das Redesign/Refurbishment der S-Bahn Rheinland (BR422 und BR423)

Hagemann, Oliver; Kuhlins, Martin; Beckmann, Leon

DB Systemtechnik GmbH, Kompetenzzentrum HGV und ET,
Breitenbachstr. 69, 47809 Krefeld

Zusammenfassung

Die DB Regio AG hat am 22.04.2024 den Anschlussvertrag[1] für den weiteren S-Bahn-Betrieb der Linien im Großraum Köln bis 2032 gewonnen. In diesem Zusammenhang müssen die insgesamt 99 Fahrzeuge der Baureihen 422 und 423 bis Ende 2026 modernisiert sein. Mit der Konstruktion, Nachweisführung und Zulassung des Redesigns ist die DB Systemtechnik GmbH beauftragt. Das Engineering wird dabei durch die DB Systemtechnik GmbH aus Krefeld umgesetzt. An dieser Stelle berichtet das Engineering-Team von seinen praktischen Erfahrungen aus dem Umbau von Schienenfahrzeugen. Die Fertigung der Fahrzeuge findet durch die Talbot Services GmbH in Aachen zusammen mit DB Regio statt. Die modernisierten Musterbaufahrzeuge sind dem Aufgabenträger im Juni 2025 erfolgreich vorgestellt worden.

Keywords: **Redesign/Refurbishment/Modernisierung, Engineering, ET422, ET423, DB Systemtechnik GmbH, DB Regio AG**

1 Redesign ET422 und ET423 der S-Bahn Rheinland

Die DB Regio AG hat am 22.04.2024 den Anschlussvertrag[1] für den weiteren Betrieb der Linien im Großraum Köln gewonnen und wird daher bis 2032 die Verkehrsleistung erbringen. In diesem Zusammenhang müssen die insgesamt 99 Fahrzeuge der Baureihen

[1] vgl. https://www.deutschebahn.com/de/presse/presse-regional/pr-duesseldorf-de/presseinformationen-regional/Neuer-Vertrag-fuer-die-S-Bahn-Koeln-bringt-preisgekroente-Zuege-auf-die-Schiene--12833854: Stand 30.03.2025

422 und 423 bis Ende 2026 im Fahrgastraum modernisiert und nach dem 4. Eisenbahnpaket zugelassen werden.[2] Die Fahrzeuge erhalten zur Steigerung des Fahrgastkomforts u.a. neue Monitore im Deckenbereich, ein WLAN-System, 230 V-Fahrgastraum-Steckdosen mit USB-Lademöglichkeit sowie LED-Technik in der Fahrgastraumbeleuchtung. Ebenso erhalten die Fahrzeuge Änderungen im Design durch den Austausch der Sitzpolster und einer neuen Außen-Lackierung. Diese erstmalige Projektkonstellation aus DB Regio, DB Systemtechnik und Talbot Services ist für alle Beteiligten mit neuen Erfahrungen verbunden gewesen. Dabei sorgt die sehr hohe Präsenz der DB Systemtechnik vor Ort in Aachen, dass der Umbau der Fahrzeuge im ersten Quartal 2025 zielgerichtet mit DB Regio und Talbot Services umgesetzt werden kann. Am 02.07.2025 berichtet Steve Pryce, Gesamtprojektleiter für das Redesign, dass am 29.06.2025 die Musterbaufahrzeuge dem Aufgabenträger erfolgreich vorgestellt wurden.[3]

2 DB Systemtechnik GmbH

Die DB Systemtechnik GmbH, eine Tochtergesellschaft der Deutschen Bahn AG, ist der europaweit führende Dienstleister im Eisenbahnsektor für Ingenieur- und Prüfdienstleistungen[4]. Sie bietet Produkte für das Eisenbahnwesen, angefangen bei der Planung und Beratung für die Infrastruktur über die Planung und Optimierung von Werkstätten zur Instandhaltung der Schienenfahrzeuge bis hin zu Dienstleistungen direkt am Fahrzeug an. Hierzu gehören u.a. die Zulassung, Mess- und Prüfdienstleistungen, Unfallinstandsetzung und Modernisierung. Der Bereich Fahrzeuge (TVE2) der DBST besteht aus 7 Kompetenzzentren mit über 180 Mitarbeitenden in Deutschland. Er bietet ein breites Portfolio wie z.B. die Fertigungsbegleitung von Neufahrzeugen, die Stör- und Verfügbarkeitsanalyse, die Konstruktion bei Unfallinstandsetzungen und die Durchführung von Redesigns. Das Redesign bietet die Möglichkeit, bestehende Fahrzeuge wirtschaftlich zu modernisieren und damit auch eine Verlängerung der Lebensdauer des Fahrzeugs zu erwirken. Die Fahrzeugmodernisierung für die Fahrzeuge der S-Bahn Rheinland wird durch das Kompetenzzentrum HGV und ET (TVE21) in Krefeld geplant, konstruiert, nachgewiesen und durch enge Begleitung der Musterbauten mit zur Zulassung geführt.

[2] vgl. https://wir.gorheinland.com/angebot/die-s-bahn-rheinland-kommt/: Stand 30.03.2025
[3] Vgl. https://www.linkedin.com/posts/stevepryce_philipvonessen-robertdaesrfeld-mariuskrabiell-activity-7345861913228050432-J1Tq?utm_source=share&utm_medium=member_ios&rcm=CoAAEU4qcEBobeDJgThWC5p-KawIIBjikwabd0: Stand 07.07.2025
[4] Vgl. https://www.db-systemtechnik.de/dbst-de: Stand 07.07.2025

3 Herausforderungen und Lösungswege aus Sicht des Engineerings

3.1 3D-Modellierung der Einbauräume

Beschreibung

Die technische Dokumentation von Einbauräumen bei Bestandsfahrzeugen ist nicht immer lückenlos vorhanden. In der Regel liegen keine 3D-Modelle vor und manchmal sind auch keine vollständigen Zeichnungen verfügbar. Zudem entspricht der tatsächliche Einbauraum nicht immer den vorliegenden Unterlagen, was selbst bei guter Datenlage stets eine Überprüfung am Fahrzeug notwendig macht. Erst im Anschluss dieser Bestandsaufnahme am Fahrzeug kann die für den Umbau benötigte 3D-Umgebung im CAD-System erstellt werden.

Herausforderungen

Die Entwicklung von mechanischen Neuteilen konzentriert sich überwiegend auf Metallkonstruktionen. Die Herausforderung liegt somit nicht in den hohen technischen Anforderungen an die einzelnen Neuteile. Vielmehr stellt die mechanische Integration neuer Betriebsmittel in eine Bauraumumgebung die Herausforderung dar. Teilweise sind Züge der gleichen Baureihe unterschiedlich ausgestattet oder gefertigt. Dies bringt konstruktive Anforderungen mit sich. In dem hier vorgestellten Redesign befinden sich zudem zwei verschiedene Baureihen parallel im Umbau. Ziel ist, nach Möglichkeit eine gemeinsame, wirtschaftliche Konstruktionslösung zu finden. Allerdings steigert dieser Umstand die Anforderungen an die Konstruktion.

Lösungswege

Im Vorfeld an die Konstruktion ist eine umfangreiche Foto- und Videodokumentation unerlässlich. Ein konsequenter Einsatz robuster Arbeitsmittel wie Gliedermaßstab, Gimbal-Kamera und Dummies aus Pappe und 3D-Druck-Prototypen haben sich bewährt. Eine präzise Vorbereitung zur Bauraumaufnahme sorgt dafür, dass das Fahrzeug nur für kurze Zeit für das Engineering benötigt wird, um die notwendigen Informationen für die Erstellung der 3D-Bauraumumgebung zu erhalten. Nach der Bauraumaufnahme ist eine Einbaukonstruktion möglichst flexibel auszugestalten. Die Konstruktion ist so auszulegen, dass sie bei verschiedenen Gegebenheiten (verschieden ausgestatteten Fahrzeugen) und bei möglichen Planungsänderungen in der näheren Umgebung des Einbaus weiterhin einsetzbar bleibt.

Abbildung 1 zeigt beispielhaft, wie sich komplexe Einbauraum-Umgebungen mit einfachen Mitteln wirtschaftlich und effektiv analysieren lassen.



Abbildung 1: Einfache und effiziente Bauraumaufnahme des zu modellierenden Einbauraums

3.2 Einbauräume für elektrische Leitungen und Betriebsmittel

Beschreibung

Im Rahmen dieses Redesigns werden pro Fahrzeug ca. 5000 m neue Elektroleitungen verlegt. Dafür werden sowohl vorhandene Kabelkanäle als auch neu zu schaffende Verlegewege verwendet.

Herausforderungen

Die elektrische Bestandsdokumentation, bestehend aus Stromlaufplan, Verbindungstabellen und Bundtabellen mit dazugehörigen Bundplänen (Zeichnungen) reicht häufig nicht aus, um eine verlässliche Aussage über die freien Kapazitäten der vorhandenen Kabelkanäle treffen zu können. Ebenso geht aus dieser Dokumentation nicht immer hervor, welche Leitungskategorien (z.B. Energie und Daten) durch den Hersteller wie verlegt sind. Beispielhaft für eine solche Herausforderung steht hier die Integration der Fahrgastraumsteckdosen, da neue Leitungswege geplant, gefertigt und gemäß Regelwerk nachgewiesen werden müssen. Zusätzlich wird das Engineering an vorhandenen Steckverbindern, Wagenübergängen und Kupplungen vor weitere Herausforderungen gestellt. Hier wird die Dokumentation geprüft und am Fahrzeug verifiziert, ob geeignete Reserveleitungen vorhanden sind, welche kostengünstig genutzt werden können. Des Weiteren zeigt die Erfahrung, dass in den Steckverbindern die zur Verfügung stehende Kapazität für Erweiterungen oft nicht ausreichend ist. Das Engineering sorgt in der Konstruktion für die Bereitstellung ausreichender zusätzlicher Leitungs- und Steckplatzkapazitäten. Hierbei werden Kundenwünsche für Schaffung weiterer Reservekapazitäten für zukünftige Modernisierungen berücksichtigt.

Lösungswege

Am Beispiel der Fahrzeugintegration der Fahrgastraumsteckdosen hat das Engineering zunächst mit der Prüfung der Dokumentation und der Verifikation der Ist-Situation in beiden Fahrzeugen begonnen. Dazu werden notwendige Seitenverkleidungen entfernt und mögliche Isolierungen freigelegt. Anschließend wird vor Ort die mögliche Leitungsverlegung unter Berücksichtigung der Gegebenheiten festgelegt und danach mittels Zeichnungen und elektrischer Planunterlagen für die Fertigung dokumentiert.



Abbildung 2:Leitungsverlegung der Fahrgastraumsteckdosen

3.3 Rechtzeitige Bereitstellung von Lieferanten-Dokumentation für Gesamtkonstruktion

Beschreibung

Neue Komponenten werden in früher Projektphase mit kundenspezifischen Ausschreibungskriterien durch den jeweiligen Projekteinkauf bestellt. Infolge der Ausschreibung werden häufig Komponenten angeboten, die erst nach Auftragseingang fertig entwickelt werden, um allen Anforderungen zu entsprechen.

Herausforderung

Nach der Beauftragung des Engineerings startet der Redesign-Entwicklungsprozess. Zu Beginn steht hierbei die Konzeptionsphase für die einzelnen Arbeitspakete, die im Laufe der Zeit vom Grobkonzept über das Feinkonzept in die Konstruktion führt. Als ein Beispiel stellt der Wechselrichter für die Fahrgastraumsteckdosen eine besondere Herausforderung für das Engineering dar. Dieser Wechselrichter basiert auf einem bestehenden Gerät, das für die konkrete Anwendung lediglich intern elektrisch modifiziert werden muss. Damit allerdings die mechanische Konstruktion bereits starten kann, hat der Hersteller nach Auftragserteilung die Vorgaben für die mechanische Konstruktion geliefert. Während des Einbauengineerings hat der Hersteller jedoch die Geometrie des Gehäuses ohne vorherige Rücksprache so abgeändert, dass die bisherige Konstruktion nicht mehr verwendet werden kann. Dies führt zu einer Überarbeitung, der bereits abgeschlossenen Konstruktion gemäß ursprünglicher Herstellervorgabe.




Generell wird das Engineering durch unvollständige Informationen in Datenblättern, Komponentenbeschreibungen etc. vor Herausforderungen gestellt, die vermeidbar wären. Hierzu zählen z.B. die Informationen zu Gewichtsangaben, Steckerbelegungen, Leistungsaufnahmen, Wartungshinweisen, Prüfnormen oder auch Umweltbedingungen.

Lösungswege

Aus Sicht des Engineerings müssen gemäß Pflichtenheft zu Beginn der Konzeptions- und Konstruktionsphase die Vorgaben der Hersteller vorliegen. Diese dürfen während der Phase des Engineerings nicht einseitig geändert werden. Daher ist es essenziell, dass das Engineering den Einkaufsprozess vom Angebot bis zum Abschluss der Zulassung begleitet und frühzeitig mit den Lieferanten in den Dialog geht.

3.4 Modernisierung Fahrgastinformationssystem

Beschreibung

Die Modernisierung des Fahrgastinformationssystems (FIS) oder Reisenden-informationssystems (RIS) ist Teil nahezu jedes Redesigns. Neben Änderungen des Designs fallen diese Änderungen dem Fahrgast sofort ins Auge. Dazu gehören oft neue Fahrtzielanzeigen zur Außenseite sowie Monitore im Fahrgastraum.

Herausforderung

Die Herausforderungen für das Engineering beim Arbeitspaket FIS sind vielfältig, da dieses im Regelfall die umfangreichste Aufgabe ist. Hierzu zählen die Konzeption und Sicherstellung der Integrationsfähigkeit in das Fahrzeug, die Begleitung der Beschaffung der Komponenten, die Konstruktion und die Begleitung des Musterbaufahrzeugs. All dies erfordert auf Seiten des Engineerings einen sehr hohen Abstimmungsaufwand mit allen Beteiligten, damit die Integration gemäß Kundenwunsch erfolgt.

Zu diesen Herausforderungen bei der Umrüstung des FIS sticht dieses Redesign mit weiteren, neuen Herausforderungen beim Engineering hervor:

- Zeitgleiche Entwicklung mit möglichst gleicher FIS-Ausrüstung und FIS-Konstruktion für zwei verschiedene Baureihen in kurzer Zeit
- FIS-Modernisierung soll bereits die Vorrüstung für kommende FIS-Funktionserweiterungen beinhalten
- FIS-System befindet sich beim Hersteller noch in finaler Entwicklung

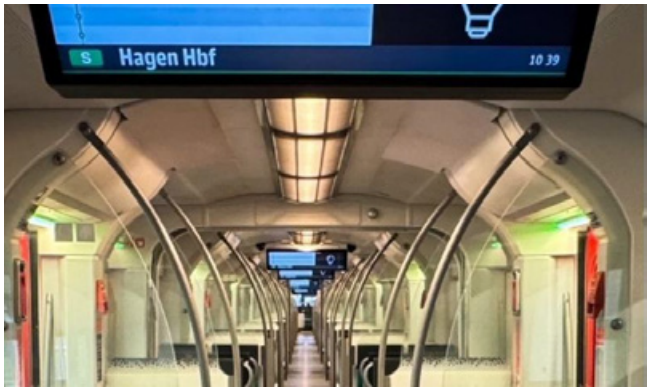


Abbildung 3: Neue Fahrgastmonitore in den Fahrzeugen der S-Bahn Rheinland

Lösungswege

Das Engineering hat sich bei diesem Umbau mit dem Hersteller des Systems sehr eng abgestimmt. Eigens dafür ist innerhalb des Engineerings ein kleines Team für dieses Arbeitspaket gebildet worden, um kurzfristig die Entwicklung beim Hersteller zu unterstützen. Zeitgleich sind von diesem Team die benötigten Vorgaben an Konstruktion und Fertigung übermittelt worden, damit alle Schnittstellen zueinander passen. Dieses pragmatische Vorgehen hat dafür gesorgt, dass innerhalb der äußerst kurzen Entwicklungszeit beim Hersteller gleichzeitig das Material beschafft, die Konstruktion entwickelt und schlussendlich der parallellaufende Musterbau der Fahrzeuge erfolgreich durchgeführt wurde. Am Ende ermöglicht dies den Fahrgästen ein modernes FIS mit Darstellung von Live-Informationen während der Fahrt.

3.5 Piktogramme und Anschriften

Beschreibung

Die Fahrzeuge der S-Bahn Rheinland erhalten im Rahmen des Redesigns neue Piktogramme und Anschriften. Sie lassen sich in betriebliche und nicht betriebliche

Kennzeichnungen unterteilen. Zu den betrieblichen Piktogrammen und Anschriften gehören insbesondere die am Langträger angebrachten Kennzeichnungen. Diese sind umfangreich normativ vorgegeben. Nicht betriebliche Kennzeichnungen werden teilweise durch die TSI-PRM und teilweise durch Kundenwünsche bestimmt.

Herausforderungen

Die Herausforderung des Engineerings liegt in einem sehr hohen Abstimmungsaufwand, um alle Anforderungen des Kunden, des Designers, des Aufgabenträgers, des Marketings, des Piktogramm-Lieferanten, der Normung und den Gegebenheiten am Fahrzeug zu erfüllen. Der zeitliche Anteil dieser Abstimmung am gesamten Arbeitspaket ist größer als derjenige der reinen Zeichnungserstellung. Des Weiteren ist eine vorausgehende Recherche in Regelwerken notwendig. Der zeitliche Aufwand hierfür hat in den vergangenen Jahren wesentlich zugenommen.
Ein Beispiel für eine notwendige Änderung und die Bedeutung von Piktogrammen für Redesignmaßnahmen ist in Abbildung 4 dargestellt. Nach Norm dürfen in einem Infofeld maximal fünf Piktogramme dargestellt sein. Daher mussten in diesem Fall Vorrang-Sitzplätze im Zug verlegt werden. Eigentlich sollten sich diese PRM-Plätze neben den Rollstuhlplätzen befinden, doch dies hätte zu mehr als fünf Piktogrammen geführt. Ein Entfall optionaler Piktogramme wäre denkbar gewesen, wurde aber vom Kunden nicht gewünscht. Zudem musste das Infofeld mit weißer Schrift auf transparentem Hintergrund geändert werden. Der notwendige Kontrast war nicht gegeben, sodass schließlich eine nachtblaue Schrift auf weißem Hintergrund gewählt wurde.



Abbildung 4: Beispielpiktogramm „Infofeld 2. Klasse“ (v.l.n.r.: blau auf weiß mit Rollstuhlplatz, weiß auf transparent, blau auf weiß mit Vorrangplatz)

Lösungswege

Zur weitgehend reibungsfreien Bearbeitung des Arbeitspakets und zur Vermeidung von aufwändigen Nacharbeiten, Projektverzögerungen und Problemen bei der Zulassung ist eine disziplinierte Dokumentation und eine umfangreiche Betreuung des Arbeitspakets über den gesamten Redesignprozess hinweg essenziell. Das beginnt bei der Bestandsaufnahme der Fahrzeuge vor Umbau mittels Videos und Fotos, geht über die Kommunikation mit Kunde, Umbauer und Lieferant und endet bei einer detaillierten technischen Zeichnung. Es haben sich Design Reviews (CDR, IDR, FDR) zur verbindlichen Absprache mit dem Kunden und eine enge Vorort-Betreuung während des Klebeprozesses bewährt. Die Zeichnungserstellung erfordert deshalb neben Zeichnungskompetenz tiefes Regelwerkswissen und eine hohe Kommunikationsfähigkeit des Verantwortlichen.

3.6 Nachweisführung gemäß 4. Eisenbahnpaket

Beschreibung

Im Rahmen der Fahrzeugmodernisierung wird neben der elektrischen und mechanischen Konstruktion durch das Engineering auch die Dokumentation für die Nachweisführung als Teil der Zulassung erstellt. Diese Dokumente dienen zur Bestätigung der Einhaltung des geltenden Regelwerks. Wichtige Nachweise sind hierbei z.B. zu Masseänderungen, zum Brandschutz und zur elektromagnetischen Verträglichkeit (EMV) vorzulegen. Diese drei Punkte sind insofern besonders, da ohne ihre unbedenkliche Bewertung das Fahrzeug nicht in Betrieb gehen kann.

Herausforderungen

Jeder notwendige Nachweis im Rahmen der Fahrzeugmodernisierung stellt das Engineering vor eigene Herausforderungen. Stellvertretend werden hier die Herausforderungen für Masseänderungen, Brandschutz und EMV erläutert.

Änderungen am Fahrzeug, ziehen im Regelfall Masseänderungen nach sich, welche nach TSI und NNTR nachgewiesen sein müssen, um die Fahrzeugsicherheit zu gewährleisten. Masseänderungen werden vor Umbau mittels Massenbilanz-Prognose prognostiziert und nach Umbau mittels Massenbilanz-Validierung nachgewiesen. Dazu dient vor Umbau die Ausgangsdokumentation des Halters. Sie beinhaltet u.a. das Fahrzeug-Layout und Informationen zu Masseschwerpunkten. Diese liegen oft nicht nach aktuellem Regelwerk vor. Deshalb sind heutige Regelwerks-Anforderungen mit dieser Dokumentation, die möglicherweise 15 Jahre und älter ist, nicht immer reibungslos zu erfüllen. Zum

Zeitpunkt der Auslieferung des Fahrzeugs gab es u.U. andere Vorschriften. Somit fehlen benötigte Informationen. Die Basis für den Umbau ist stets die aktuelle Eingangs-Verwiegung des Musterbaufahrzeugs. Diese sollte mit der gleichen Waage durchgeführt werden wie am Ende die Ausgangsverwiegung nach Umbau.

Neben der Nachweisführung für die Masseänderungen am Fahrzeug ist die Bewertung des Brandschutzes verpflichtend. Hierbei werden neue Komponenten, die brandschutzkonforme Konstruktion sowie die weiterhin bestehenden Fluchtmöglichkeiten bewertet. Die Herausforderungen für das Engineering liegen überwiegend in der Norm EN45545-2. Sie besagt, dass jede neu in das Fahrzeug eingebrachte Komponente der gültigen Brandschutznorm für Schienenfahrzeuge genügen muss. Dazu wird jede Komponente in eine Brandschutz-Liste (FCIL) mit ihren notwendigen Nachweisen aufgelistet und bewertet. In der Praxis liegt die Herausforderung nicht in der Methodik, sondern in der rechtzeitigen Bereitstellung dieser Nachweise. Oft ist es so, dass Komponenten gemäß Kundenwunsch verbaut werden, die neu auf dem Markt sind oder auch erstmalig in ein Schienenfahrzeug eingebaut werden sollen. Zu diesem Zeitpunkt liegen durch den Lieferanten meist noch keine gültigen Nachweise vor. Es kommt auch vor, dass die Nachweise von nicht akkreditierten Prüflaboren erstellt oder sie so alt sind, dass sie nicht mehr verwendet werden können. Dann müssen während der oft kurzen Projektlaufzeit die entsprechenden Nachweise durch den Lieferanten neu erbracht werden. Die überwiegende Mehrheit der Komponenten wird in kurzer Zeit erfolgreich bewertet, wohingegen der geringe verbleibende Anteil den Großteil der Zeit benötigt.

Weiterhin bildet die Nachweisführung für die EMV einen wichtigen Bestandteil im Zulassungsprozess. Dabei ist die Verwendung von EMV-konformen elektrischen Betriebsmitteln, wie z.B. dem Router eines WLAN-Systems, dem Fahrgastinformationsdisplay oder dem Wechselrichter für 230 V-Fahrgastraum-Steckdosen, unerlässlich, um unerwünschte Störaussendungen zu vermeiden und die Störfestigkeit der Geräte sicherzustellen. Das System Bahn, sei es das Schienenfahrzeug selbst oder die Eisenbahn-Infrastruktur, jedoch auch der Fahrgast selbst dürfen nicht beeinträchtigt werden. Die Herausforderungen liegen in der Nachweisführung. Sie ist zweistufig aufgebaut und fordert einen EMV-Plan sowie die EMV-Prüfberichte von akkreditierten Prüflaboren (EN50121, EMV06). Im zweiten Schritt nach Umbau folgt meist die EMV-Messung des betriebsbereiten Fahrzeugs. In diesem Projekt sticht besonders der Wechselrichter für die Fahrgastraum-Steckdosen hervor, weil für diesen nach über 3 Monaten nach Auftragsvergabe noch keine EMV-Nachweise vorlagen. Ohne diese Nachweise kann das Fahrzeug nicht betrieben werden.

Lösungswege

Für die drei stellvertretenden Herausforderungen Masseänderungen, Brandschutz und EMV sind durch das Engineering folgende Lösungswege gewählt worden.

Masseänderungen: Auf Grundlage der projektbezogenen Fahrzeugverwiegungen werden im Rahmen des Massemanagements die erforderlichen Parameter gemäß der gültigen TSI und NNTR berechnet. Dabei begleitet die DB Systemtechnik GmbH mit Ihrer umfangreichen Kompetenz den Halter bei der Erfüllung der für die Zulassung notwendigen Anforderungen, um Fahrzeuge jeglicher Art modernisieren zu können. Dies beginnt bereits bei der Bereitstellung der Fahrzeugverwiegungen und wird fortgeführt über technische Klärungen zu Bremse, Drehgestell, Wagenkasten, Fahrdynamik und Radsätzen. Dadurch wird die Interoperabilität des Fahrzeugs im Einklang mit dem Regelwerk sichergestellt.

Brandschutz: Das Engineering berät den Auftraggeber bei der Komponentenauswahl und wählt bewährte Lösungen aus dem Schienenfahrzeugsektor. Sie sind meist einfach zu beschaffen und haben eine bekannte Nachweisdokumentation, um das Redesign zügig durchzuführen. Weiterhin begleitet das Engineering den Einkaufsprozess und zeigt bereits in früher Projektphase die notwendigen Nachweise an, damit diese parallel zur Konstruktion erstellt bzw. zur Verfügung gestellt werden können. Für den Fall, dass es keinen ausreichenden Nachweis gibt, besteht die Möglichkeit, durch das eigene akkreditierte Brandschutzlabor der DBST kurzfristig den notwendigen Nachweis zu erhalten.

EMV: Das Engineering unterstützt den Projekteinkauf durch die Lastenhefterstellung und anschließende Bewertung der Angebote, damit auch die notwendigen Nachweise eingekauft werden. Darüber hinaus fließt die langjährige Erfahrung des Engineerings bei der Auswahl von bewährten Komponenten bis zur abschließenden EMV-Messung stets mit ein. Bei der EMV-Messung ist eine geschickte Wahl des Ortes sinnvoll, um äußere störende EMV-Einflüsse auf die Messung auszuschließen. Falls erforderlich, kann durch das DBST-eigene akkreditierte EMV-Prüflabor die Nachweisführung unterstützt werden. So können Komponenten oder auch, wie in diesem Umbau, das gesamte Fahrzeug auf die normativen Anforderungen hin geprüft und verifiziert werden.

4 Danksagung

Wir bedanken uns bei der Region NRW der DB Regio AG und der Talbot Services GmbH für die gute Zusammenarbeit in dieser erstmaligen Projektkonstellation und freuen uns auf weitere gemeinsame Projekte.

5 Abkürzungsverzeichnis

3D	dreidimensional
BR	Baureihe
CAD	Computer-Aided Design
DB	Deutsche Bahn
DBST	DB Systemtechnik GmbH
CDR	Conceptual Design Review
EMV	Elektromagnetische Verträglichkeit
ET	Elektrischer Triebzug/Elektrische Triebzüge
FDR	Final Design Review
FIS	Fahrgastinformationssystem
FCIL	Fire Safety List
HGV	Hochgeschwindigkeitsverkehr
IDR	Intermediate Design Review
LED	Light Emitting Diode
NNTR	Notifizierte Nationale Technische Regeln
RIS	Reisendeninformationssystem
TSI	Technische Spezifikation für die Interoperabilität
TSI-PRM	Zugänglichkeit für Menschen mit Behinderung und Menschen mit eingeschränkter Mobilität
PRM	Persons with Reduced Mobility
WLAN	Wireless Local Area Network

Session 11
Oliver Hagemann, Martin Kuhlins, Leon Beckmann

Autoren



Hagemann, Oliver

Oliver Hagemann hat an der Technischen Universität Braunschweig Maschinenbau studiert. Nach dem Studium startete er 1996 bei der Siemens Verkehrstechnik in Erlangen. Nach mehreren Aufgaben und Stationen bei Siemens war er zuletzt Leiter des Eisenbahntestzentrums in Wegberg-Wildenrath („PCW"). Dort verantwortete er die drei Themengebiete (Bahn-) Infrastruktur, die akkreditierte Prüfstelle und das Eisenbahnverkehrsunternehmen (EVU). Seit dem 01.01.2019 ist er Leiter des Kompetenzzentrums HGV und ET mit Sitz in Krefeld der DB Systemtechnik GmbH.



Kuhlins, Martin

Martin Kuhlins hat an der Technischen Hochschule Mittelhessen Elektrotechnik studiert und das Studium 2016 als Master of Science abgeschlossen. Anschließend startete er am 07.11.2016 bei der DB Systemtechnik GmbH in Krefeld als Konstrukteur Elektrotechnik. Dort hat er u.a. die LZB-Fahrzeugintegration im Rahmen des Redesigns ET420 S-Bahn München verantwortet. Seit dem 01.09.2021 ist er Leiter des Kompetenzzentrums Redesign, Umbau und Sanierung ET in Krefeld und verantwortet mit seinem Team die Umsetzung von Redesigns/Refurbishments.



Beckmann, Leon

Leon Beckmann studierte in den Fachrichtungen Maschinenbau und Energietechnik an der FH Aachen und der Universität Hannover. Nach Abschluss mit dem Grad Master of Science startete er zum 01.08.2023 bei der DB Systemtechnik GmbH in Krefeld als Konstrukteur Mechanik. Seitdem ist er am Projekt Redesign S-Bahn Köln beteiligt. Die Tätigkeitsschwerpunkte umfassen Konstruktion, Dokumentation und Umbaubegleitung am Fertigungsstandort.

Session 11
Margit Olbrich, Toni Schnell

Digitales Instandhaltungsregelwerk bei der Deutschen Bahn, Personenverkehr

Know-how für Profis am Zug

Olbrich, Margit[1], Schnell, Toni[2]

[1]DB Systemtechnik GmbH
[2]Inxire GmbH

Zusammenfassung

Über die Unternehmen DB Fernverkehr AG und DB Regio AG transportiert die DB AG täglich hunderttausende Reisende in ihren ICE-, IC- und Regionalbahnen sowie S-Bahnen. Die Fahrzeuginstandhaltung sorgt dafür, dass die Züge geprüft, gewartet und instandgesetzt werden, um die Sicherheit, Zuverlässigkeit und den Komfort der Reisenden zu gewährleisten. Die Erstellung, Bearbeitung und Verteilung der Instandhaltungsregelwerke erfolgt digital über die webbasierte Anwendung REFI auf Basis einer standardisierten XML-Datenstruktur. Dies ermöglicht die effektive Modernisierung der Arbeitsprozesse und treibt die Digitalisierungsstrategie und grüne Transformation der DB AG voran. Neben Arbeitsanweisungen und Instandhaltungshandbüchern wurden Prüfprozesse zur Qualitätssicherung eingeführt, um eine fachlich richtige Darstellung sicherzustellen. Die Nutzung einer einheitlichen XML-Datenstruktur ermöglicht einen effizienten Datenaustausch und eine gesteigerte Transparenz bei der Erstellung und Verteilung der Instandhaltungsregelwerke. Die papierlose Bereitstellung der Regelwerke auf Tablets reduziert Kosten und Ressourcen, während gleichzeitig die Umwelt geschont wird. Die Integration von Augmented Reality in die Instandhaltungsprozesse bietet das Potenzial, die Ausführung der Arbeiten zu optimieren. Durch den Einsatz von Künstlicher Intelligenz auf Basis der strukturierten Regelwerksinhalte können Instandhaltungsteams schnellere und präzisere Hilfestellungen erhalten. Die kontinuierliche Weiterentwicklung der REFI-Anwendung positioniert sie als zentrale Plattform für die intelligente Instandhaltung der Zukunft.

Keywords: Instandhaltungsregelwerke; Instandhaltung von Schienenfahrzeugen; Digitalisierung: webbasiertes Redaktionssystem

Session 11
Margit Olbrich, Toni Schnell

1 Einleitung

Die Deutsche Bahn Personenverkehr mit den Unternehmen DB Fernverkehr AG und DB Regio AG befördert mit einer heterogenen Fahrzeugflotte täglich hunderttausende Reisende in ihren ICE-, IC- und Regionalbahnen sowie S-Bahnen. Der Bereich Instandhaltung sorgt dafür, dass die Fahrzeuge so geprüft, gewartet, instandgesetzt und für den Einsatz vorbereitet werden, dass die Reisenden sicher, zuverlässig und komfortabel an ihrem Ziel ankommen. Als Grundlage für die Arbeiten der Fahrzeuginstandhaltung bei DB Personenverkehr wird in einem webbasierten Redaktionssystem die Erstellung der Arbeitsbeschreibungen als Instandhaltungsregelwerke bis zur Verteilung und Kenntnisnahme dieser, DV-technisch durchgeführt und dokumentiert.

Die DB Systemtechnik GmbH und die Firma Inxire GmbH unterstützen die DB Personenverkehr dabei, die Lösungen für die Instandhaltungsregelwerke innovative und digital zu entwickeln und damit das Zukunftspotential für eine effektivere Fahrzeuginstandhaltung zu erschließen.

Was diese digitalen Instandhaltungsregelwerke für die Instandhaltung von Schienenfahrzeugen der DB AG konkret bedeuten und welches Zukunftspotential in ihnen steckt, wird in diesem Artikel beschrieben.

2 Regulatorische Vorgabe

Die zugrundeliegenden Vorgaben, dass Fahrzeuge und deren Komponenten instand zu halten sind, damit sie während ihres Einsatzes den gesetzlichen Bestimmungen entsprechen, sind geregelt im Allgemeines Eisenbahngesetz (AEG) und der Eisenbahn-Bau- und Betriebsordnung (EBO).

„Gemäß § 4a AEG sind die Eisenbahnen und Halter von Eisenbahnfahrzeugen für die Instandhaltung ihrer Eisenbahnfahrzeuge zuständig. Sie können die Aufgaben der für die Instandhaltung zuständige Stelle (ECM) selber wahrnehmen oder per Vertrag auf einen Dritten übertragen [1].“

Nach § 4 EBO müssen Fahrzeuge so beschaffen sein, dass diese den Anforderungen der Sicherheit und Ordnung genügen und nach § 32 (2) EBO sind Fahrzeuge planmäßig, wiederkehrend zu untersuchen.

„Jede für die Instandhaltung zuständige Stelle ist verpflichtet, die zur Instandhaltung übernommenen Fahrzeuge in betriebssicherem Zustand zu halten. Sie hat dazu ein Instandhaltungssystem einzurichten und darüber Aufzeichnungen zu führen. Für jedes Fahrzeug sind Instandhaltungsunterlagen zu führen und für die gesamte Dauer der Verwendbarkeit des Fahrzeugs aufzubewahren [1].“

Auf Basis dieser Vorgaben hat die DB AG im Rahmen eines breit angelegten Programms die technische Qualitätssicherung grundlegend strukturiert und den Aufbau eines Qualitätsmanagements vorangetrieben. Das 1999 gegründete Projekt QSI (Qualifizierung Systeme der Instandhaltung) gehörte dazu [2].

Aus Sicht der Arbeitsprozesse bei der Fahrzeuginstandhaltung ist die höchstmögliche Rechtssicherheit für Mitarbeiter:innen der Fokus. Über verschiedene Teilprojekte wurden aus dieser Sichtweise folgende Ziele verfolgt:

- „Umfassende und lückenlose Beherrschung der Instandhaltungsprozesse“
- „Transparenz aller Vorgänge und Entscheidungen“
- „Dokumentation zur Nachweisführung der Arbeitserledigung und von Festlegungen im erforderlichen Umfang“
- „Höchstmögliche Rechtssicherheit für Prozessverantwortliche und Mitarbeiter:innen - die rechtliche Bedeutung der Arbeitspapiere liegt vor allem in deren Beweisfunktion“
- „Doppelprüfungen im sicherheitsrelevanten Bereich und bei hoher Fehlerwahrscheinlichkeit“
- „Halter von Eisenbahnfahrzeugen sind stets und umfassend für die Instandhaltung ihrer Fahrzeuge verantwortlich“

„Fahrzeugsicherheit, Verfügbarkeit und Verkehrstauglichkeit sind die Faktoren, an denen die Qualität der Instandhaltung von Eisenbahnfahrzeugen gemessen wird. Dabei sind es die Instandhaltungskosten, die über Erfolg oder Misserfolg eines Eisenbahnverkehrsunternehmens mitentscheiden. Durch die auf den ersten Blick gegenläufigen Forderungen nach hoher Qualität und niedrigen Kosten wird ein Kräftefeld aufgebaut, in dem sich die für die Instandhaltung Verantwortlichen bewegen müssen, um ein Gesamtoptimum für das Eisenbahnunternehmen zu erreichen [3].“

2.1 Anforderungen an das digitale Instandhaltungsregelwerk

Die Notwendigkeit für ein Instandhaltungsregelwerk ergibt sich aus [4]

- Gesetzliche Verantwortung der Eisenbahnen, wie in Kapitel 1 beschrieben
- Notwendigkeit zur fachlichen Anleitung der Instandhaltungsmitarbeiter:innen
- Sicherung und Weiterentwicklung des Know-how und der Erfahrungen auf dem Gebiet der Fahrzeuginstandhaltung

Um den Anforderungen der verschiedenen Eisenbahnverkehrsunternehmen (EVU) der DB AG an die Fahrzeuginstandhaltung gerecht zu werden, wurden folgende Muss-Anforderungen an das Fahrzeuginstandhaltungsregelwerk festgelegt:

- Vollständigkeit
- Richtigkeit, Eindeutigkeit und Aktualität
- Anwendergerechte Gestaltung und Regelungstiefe
- Verfügbarkeit beim Anwender / der Anwenderin
- Aufbewahrung der Instandhaltungsunterlagen für jedes Fahrzeug über die gesamte Lebensdauer

2.2 Zielstellung zur Digitalisierung der Instandhaltungsregelwerke für Schienenfahrzeuge

Das Ziel zur Digitalisierung der Instandhaltungsregelwerke für Schienenfahrzeuge war es, eine IT-Anwendung zu entwickeln, mit der die Qualität der Instandhaltung durch klare und eindeutige Vorgaben zu einer „berechenbaren“ bzw. „kontrollierbaren“ Größe wird.

Unter diesen Maßgaben erstellt, bearbeitet und verteilt die DB AG ihre Instandhaltungsregelwerke digital über die webbasierte Anwendung REFI, welches auf Basis einer standardisierten XML-Datenstruktur arbeitet.

Durch diese standardisierten Datenstrukturen können moderne Ausgabetechnologien eingesetzt und Arbeitsprozesse kontinuierlich modernisiert werden. So werden die Digitalisierungsstrategie und die grüne Transformation der DB AG für die Fahrzeuginstandhaltung vorangetrieben.

Um für alle Fahrzeuginstandhaltungsprozesse die wichtigsten Informationen zur Verfügung zu stellen, wurde sich zu Projektbeginn auf zwei Dokumenttypen konzentriert, die Arbeitsanweisung (AA) und das Instandhaltungshandbuch (IHB).

Daraus entwickelte sich das Regelwerk für die Fahrzeuginstandhaltung, in folgendem Zusammenhang:

- Die Basis der Instandhaltungsvorgaben bilden Arbeitsanweisungen für alle Instandhaltungsarbeiten. Die Arbeitsanweisungen sind fachliche Anleitungen der Arbeiten. Zudem dienen diese als Schulungsbasis für die Mitarbeiter:innen der Fahrzeuginstandhaltung.
- Die Arbeitsanweisungen werden in Instandhaltungshandbüchern zusammengefasst.
- Jedes Handbuch beinhaltet nur die für die Fahrzeugbaureihe bzw. -bauart notwendigen Anweisungen.
- Die Handbücher sind in einer für die Instandhaltungsmitarbeiter:innen verständlicher Form erstellt.

Neben der Erstellung und Bearbeitung des Instandhaltungsregelwerks wurden mehrere Prüfprozesse nach Mehraugenprinzip zur Qualitätssicherung eingeführt, die eine fachlich richtige und vollständige Darstellung aller Instandhaltungsarbeiten sicherstellen.

3 Einsatz eines Redaktionssystems zur Bearbeitung des Instandhaltungsregelwerks

Die Anwendung REFI (Regelwerke Fahrzeuginstandhaltung), welche bei der DB für den Personenverkehr eingesetzt wird, verfügt über ein integriertes Redaktionssystem, welches viele Funktionen zur inhaltlichen Bearbeitung beinhaltet. Für alle Dokumente gibt es mehrstufige Prozesse zur Qualitätssicherung (QS), bestehend aus Lesungen und Prüfungen der Anwendbarkeit in den Werkstätten sowie Freigaben und Inkraftsetzungen. Ersteller der Dokumente sind interne Mitarbeiter der DB AG oder beauftragte technische Redakteure, die hierfür Zugriff auf die Anwendung erhalten.

3.1 Erstellung des digitalen Instandhaltungsregelwerks auf Basis einer einheitlichen Strukturdefinition

Das Regelwerk von DB-Personenverkehr beinhaltet alle Vorgaben für die ordnungsgemäße Instandhaltung. Das Kernelement der Vorgaben ist die Arbeitsanweisung, die Auswirkung auf folgende Aspekte hat:

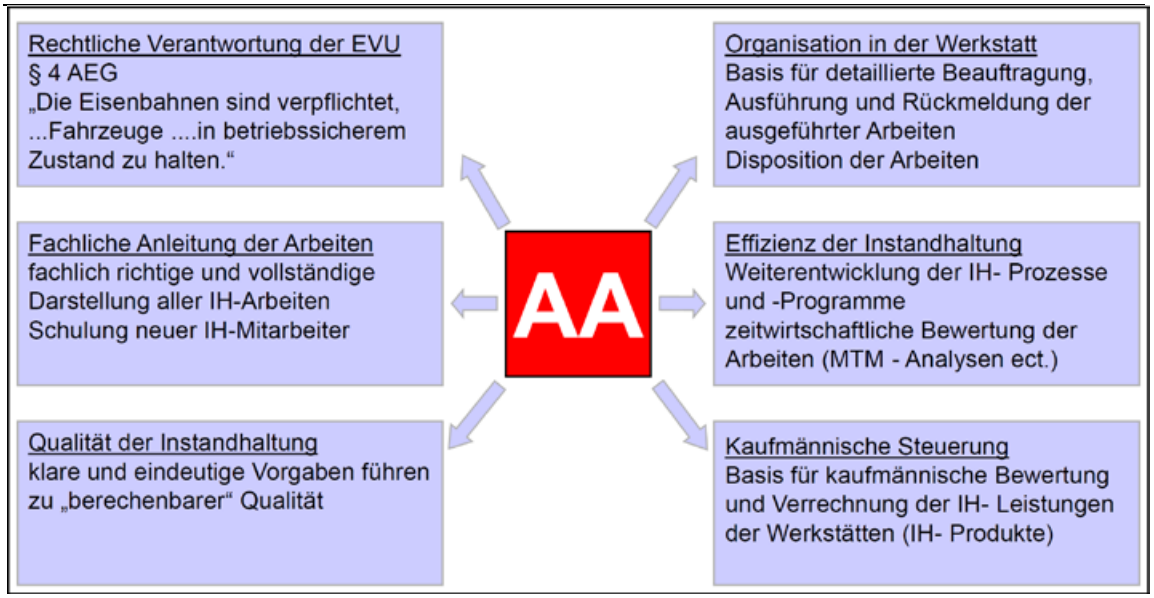


Abbildung 1: Die Arbeitsanweisung als Kernelement und Know-how Träger

In einer Arbeitsanweisung wird für ein Fahrzeug bzw. einer oder mehrerer Baureihen die auszuführenden Arbeiten an einem Fahrzeugteil oder einer Komponente des Fahrzeuges beschrieben.

Diese Beschreibung ist per Definition aus den Konzernrichtlinien in mehreren Informationsteilen aufgebaut und wird heute im Standard von allen Fahrzeugherstellern benutzt.

Eine Arbeitsbeschreibung besteht aus:

- Einem Kopfbereich, in dem Informationen beschrieben sind, die für die gesamte Arbeitsanweisung gelten
- Haupttätigkeiten, mit:
    - dem Arbeitsinhalt
    - den Qualitätsanforderungen
    - zusätzliche Hinweisen
    - ggf. weiteren Unterarbeitsschritten
- Verzeichnissen zu:
    - Grenzmaßen, die bei den auszuführenden Arbeiten zu beachten sind
    - Werkzeuge, die bei den auszuführenden Arbeiten zu verwenden sind
    - Materialien, die bei den auszuführenden Arbeiten einzusetzen oder zu tauschen sind
    - Mess- und Prüfmittel, die bei den auszuführenden Arbeiten einzusetzen sind

Die folgende Darstellung zeigt die Datenstruktur einer Arbeitsanweisung bei der DB AG:

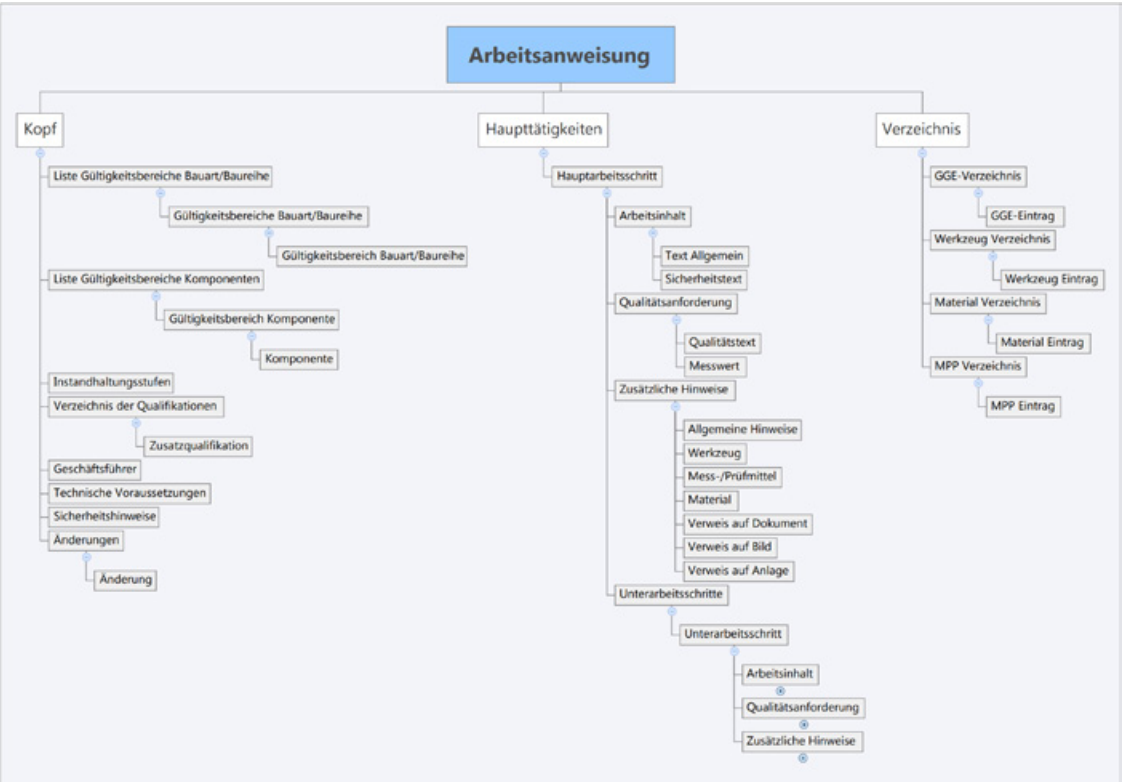


Abbildung 2: Datenstruktur einer Arbeitsanweisung bei der DB AG

3.1 Einheitliche Datenstruktur über XML

Um immer eine gleiche, einheitliche Datenstruktur zu nutzen, hat man sich bereits 2002 dazu entschieden, diese Datenstruktur auf Basis eines Schemas für die Auszeichnungssprache XML (Extensible Markup Language) zu definieren.

Der Vorteil der XML-basierten Verarbeitung von Informationen ist, dass einzelne Informationsobjekte über vielfältige Ausgabe-Interfaces den Endanwendern zur Verfügung gestellt werden können.

Grundsätzlich kann XML zum Beschreiben, Speichern und Austauschen von Daten genutzt werden [5].

- Durch den strukturierten Aufbau des XML-Dokuments mit den Elementen als kleinste Einheit, lassen sich bedarfsgerecht Darstellungen je nach Anwendungsfall z.B. als Arbeitsanweisung, Materialliste, Meß- und Prüfmittelübersicht generieren.

Anders gesagt, die strukturierten Daten lassen sich in jeglichen textuellen Aufbau zusammensetzen. Dies gilt ebenso für den Austausch der Daten mit Anwendungen, Menschen und Netzwerken.

- Da für jedes Fahrzeug die Instandhaltungsunterlagen entsprechend lange aufzubewahren sind [6], leitet sich aus dem Fahrzeuglebenszyklus ein ähnlicher Zyklus für die Unterlagen ab. In diesem Zusammenhang ergibt sich XML als idealer Begleiter für die digitale Vorhaltung der Dokumente unabhängig von proprietären Dokumentformaten oder Softwarelösung mit einem geringen Lebenszyklus.
- Durch den Einsatz des Datenschemas für XML, XSD – XML Schema Definition, wird die Struktur des XML-Dokuments und die einzusetzenden Dateitypen eindeutig definiert [7]. Durch entsprechende Schemavalidierung von übergebenen XML-Dokumenten, z.B. den Arbeitsanweisungen, wird eine hohe Qualität sichergestellt, da u.a. bei Importen die Gültigkeit, Vollständigkeit und Wohlgeformtheit geprüft wird.

Folgend ein Auszug aus dem Datenschema (XSD - XML Schema Definition) einer Arbeitsanweisung bei der DB AG:

```
<xsd:complexType name="ArbeitsschrittTyp">
    <xsd:sequence>
        <xsd:element name="Nummer"                type="xsd:string"/>
        <xsd:element name="ChangedFlag"           type="FlagTyp"                minOccurs="0"/>
        <xsd:element name="Sicherheitshinweis"    type="TextTyp"                minOccurs="0"/>
        <xsd:element name="Arbeitsinhalt"         type="TextTyp"                minOccurs="0"/>
        <xsd:element name="Qualitaetsanforderung" type="QualitaetsanforderungTyp" minOccurs="0"/>
        <xsd:element name="ZusaetzlicherHinweis"  type="ZusaetzlicherHinweisTyp" minOccurs="0"/>
    </xsd:sequence>
</xsd:complexType>
<xsd:complexType name="ZusaetzlicherHinweisTyp">
    <xsd:choice minOccurs="0" maxOccurs="unbounded">
        <xsd:element name="AllgHinweis"       type="TextTyp"/>
        <xsd:element name="Werkzeug"          type="BezeichnungTyp"/>
        <xsd:element name="MessPruefmittel"   type="BezeichnungTyp"/>
        <xsd:element name="Material"          type="BezeichnungTyp"/>
        <xsd:element name="AaVerweis">
```

Abbildung 3: Datenschema (XSD) einer Arbeitsanweisung bei der DB AG

Diese Struktur ist bis heute stabil und der Austausch zwischen unterschiedlichen Anwendungen wird dadurch real praktiziert.

Einige Beispiel aus der Praxis dafür sind:

- Seit 2010 wird dieser Datenaustausch aktiv bei Fahrzeugbeschaffungsprojekten durchgeführt, in dem die Fahrzeughersteller die Beschreibungen der Instandhaltungsarbeiten auf Basis eines von der DB AG vorgegebenen XML-Schemas (XSD) erstellen können.

Die vertraglich vereinbarten, zugelieferten XML-Dateien der Arbeitsanweisungen werden anschließend mittels einer Importfunktionalität in das DB-eigene Redaktionssystem - Regelwerkerstellung Fahrzeuginstandhaltung (REFI) - übertragen.

- Auf dieser Basis der XSD konnte eine Vereinheitlichung aller Regelwerke der DB AG auf strukturierte Datenhaltung in unterschiedlichen Systemen (z.B. DM-Cube – DB-Cargo, DB Netz, REFI – DB-Fernverkehr & DB-Regio, …) erreicht werden.
- Die Migration von Regelwerksunterlagen z.B. auf Grund einer Softwareabkündigung kann auf Basis der eindeutigen Definition einfach realisiert werden.
- Die Regelwerkserstellung und -verteilung erfolgt immer in der gleichen Datenstruktur, wodurch klare und eindeutige Vorgaben in einer berechenbaren bzw. kontrollierbaren Qualität entstehen. Die Instandhalter:innen aus unterschiedlichen Werken und unterschiedlichen Unternehmensbereichen erhalten Vorgaben in bekannten Strukturen, der Einarbeitungsaufwand wird somit spürbar minimiert.

3.2 Auswirkungen des digitalen Instandhaltungsregelwerks auf die Regelwerkerstellungs-, -verteilungs- und Instandhaltungsprozesse

Neben dem unmittelbaren Vorteil zum Datenaustausch über die XML-basierte Datenstruktur entstehen weitere Synergien im Zusammenhang mit der Digitalisierung der Instandhaltung. So können die Prozesse der Regelwerkserstellung, -verteilung und der Instandhaltung mit der geeigneten IT-Lösung REFI effektiv und effizient gestaltet werden.

Nachfolgend weitere Beispiele aus der Praxis, welche durch die digitale Verarbeitung entstanden sind:

- Digitale Umsetzung von mehrstufigen Zeichnungs- bzw. Freigabeprozessen durch Workflowfunktionen, bei denen zudem der Entstehungs- und Qualitätssicherungsprozess der Instandhaltungsregelwerke nachvollziehbar dokumentiert werden. Daraus resultiert eine hohe Transparenz zum Entstehungsprozess der Instandhaltungsregelwerke und damit analog eine hohe Rechtssicherheit.

- Revisionssichere, digitale Vorhaltung der Instandhaltungsvorgaben. Dadurch ist eine zentrale Aufbewahrung der Instandhaltungsregelwerke möglich, welche im Bedarfsfall effizient abgerufen werden können.
- Papierlose Bereitstellung der Instandhaltungsvorgaben durch verschiedene Ausgabemöglichkeiten als PDF oder HTML auf Tablets der Instandhalter:innen. Damit wird der aufwendige, papiergebundene Verteilungsprozess ersetzt, was sowohl Kosten & Ressourcen minimiert als auch einen Beitrag für die Umwelt leistet. Zudem können die benötigten Informationen zielgerichtet an die ausführenden Mitarbeiter:innen digital verteilt werden, der „Gang zum Regelwerksschrank“ entfällt.
- Durch die standardisierte Form der Anleitung zu den Arbeiten (Arbeitsanweisungen) kann in vielen Werkstätten, die Fahrzeuge der DB AG instand halten, ein Testat "Qualifizierung Systeme der Instandhaltung" festgestellt werden, das bestätigt, dass die Prozessorganisation für die Fahrzeuginstandhaltung korrekt und wirkungsvoll gelebt wird.

4 Ausblick

Der Schritt, die digitalen, in XML-Elementen strukturierten Informationen auf moderne Ausgabegeräte, wie Tablets, den Instandhalter:innen direkt an den Fahrzeugen zur Verfügung zu stellen, ist bereits Realität in den Werken von DB-Fernverkehr und DB - Regio und gehört zur aktuellen Digitalisierungsstrategie der DB AG.

Arbeitsprozesse in den Werken von DB Fernverkehr und DB Regio, wie die digitale, örtliche Verteilung und Bestätigung über den Erhalt der Regelwerksinformationen, werden bereits gelebt. Zudem befindet sich ein weiteres Projekt in der Pilotphase, welches die Regelwerkssuche mit intelligenten Strategien erweitert, um den Instandhalter:innen eine noch bedarfsgerechtere Ausgabe zum angefragten Informationsinhalt bereitzustellen.

Die Inhalte der Instandhaltungsregelwerke in ein VR-Framework zu integrieren, kann die Lücke über „Augmented Education“ in den Werkstätten der DB AG von der digitalen Erstellung der Instandhaltungsregelwerke bis zur qualitätsgesicherten Ausführung der Arbeiten durch die Instandhalter:innen schließen.

Bereits im Jahr 2016 wurde ein Proof of Concept (PoC) für den Einsatz von Augmented Reality (AR) in REFI realisiert.

Als Hardware kam dabei die Microsoft HoloLens zum Einsatz — ein Mixed-Reality-Headset, das digitale Inhalte direkt in die reale Umgebung projiziert und so eine immersive Verschmelzung von physischer und virtueller Welt ermöglicht.

Auf Softwareseite wurde das Framework Unity verwendet, das sich besonders für die Entwicklung interaktiver 3D- und Mixed-Reality-Anwendungen eignet.

Ein zentrales Ziel des PoC war es, die bereits vorhandenen Inhalte aus den digitalen Regelwerken auf die HoloLens zu übertragen und dort für eine kontextsensitive Anzeige sowie Navigation verfügbar zu machen. Dies konnte technisch ohne größere Hürden umgesetzt werden: Die strukturierten Daten ließen sich problemlos transformieren und in der Brille visualisieren.

Aus dem PoC ergaben sich folgende wesentliche Erkenntnisse:

- Regelwerksinhalte konnten erfolgreich auf der Brille dargestellt werden, wodurch Monteure oder Auszubildende direkten Zugriff auf relevante Vorschriften und Anleitungen am Einsatzort erhielten.
- Es zeigte sich jedoch, dass zu wenige 3D-Modelle und Bildmaterialien vorhanden waren, um eine wirklich reichhaltige Augmentierung der realen Umgebung zu ermöglichen. Dadurch blieb das AR-Erlebnis visuell eingeschränkt und oft noch textbasiert.
- Die eingesetzte Hardware wies zudem einige technische und ergonomische Limitierungen auf: Das Sichtfeld der HoloLens war relativ klein, die Brille selbst vergleichsweise schwer und nach längerem Tragen unangenehm. Damit war der Dauereinsatz in praktischen Wartungs- oder Schulungssituationen noch nicht optimal.

Gleichzeitig wurde deutlich, dass die Inhalte der Regelwerke in dieser Form hervorragend für Ausbildungszwecke geeignet sind. Die Erkenntnisse sind in der strategischen Planung und Weiterentwicklung eingeflossen.

Auch können KI-gestützte Lösungen die Instandhaltung wesentlich verbessern, indem sie auf die digitalen Instandhaltungsregelwerke zugreifen und diese effizient auswerten.

Mit dem Einzug leistungsfähiger Künstlicher Intelligenz und insbesondere der Large Language Models (LLMs) ergeben sich auch für die REFI-Anwendung völlig neue Möglichkeiten. Die Inhalte der Regelwerke sind in der REFI-Datenbank bereits klar strukturiert und mit semantischen Beziehungen angereichert. Dadurch sind sie „KI-ready“ und können von Sprachmodellen deutlich präziser interpretiert und verarbeitet werden, als es bei unstrukturierten Formaten wie etwa klassischen PDFs der Fall wäre.

Auf dieser Basis lässt sich beispielsweise eine KI-gestützte Suche mit dem RAG-Ansatz (Retrieval-Augmented Generation) einfach realisieren. Bei diesem Verfahren wird das vorhandene Sprachmodell durch externe, domänenspezifische Wissensbestände ergänzt. Das bedeutet, dass die KI nicht nur auf ihr trainiertes Sprachverständnis zurückgreift, sondern zusätzlich gezielt relevante Regelwerksinhalte abruft und einbezieht. So können Antworten aktueller, spezifischer und faktenbasierter generiert werden.

Dies eröffnet in der Praxis die Möglichkeit, Instandhaltungsteams schnelle, kontextsensitive Hilfestellungen zu liefern — direkt basierend auf den jeweils gültigen technischen Regelwerken. Dadurch lassen sich Fehler vermeiden, Wartungszeiten verkürzen und die Qualität der Instandhaltung insgesamt steigern.

5 Schlussteil

Die kontinuierliche Weiterentwicklung der REFI-Anwendung hat gezeigt, dass sie nicht nur aktuelle Anforderungen an eine moderne Instandhaltungsausführung erfüllt, sondern auch eine solide Basis für künftige technologische Entwicklungen bietet. Damit entwickelt sich REFI von einer Datenbank für digitale Regelwerkserstellung zu einer zentralen Plattform für die intelligente, qualitätsgesicherte und zukunftsorientierte Ausführung der Instandhaltungsarbeiten. Sie trägt nicht nur zur Effizienzsteigerung und Fehlervermeidung im heutigen Instandhaltungsbetrieb bei, sondern schafft auch die Voraussetzungen, um den Instandhaltungsbetrieb von Fahrzeugen mit hochautomatisierten und autonomen Funktionen nachhaltig sicherzustellen.

Literatur

[1] https://www.eba.bund.de/DE/Themen/Fahrzeuge/Instandhaltung/instandhaltung_node.html, abgerufen am 28.06.2022

[2] Pergande H.-G.: Projekt QSI - Neue Dokumentation für die Instandhaltung. Der Eisenbahningenieur, Ausgabe Juli 2001

[3] Pergande H.-G.: Neuorientierung der Fahrzeuginstandhaltung: Projekt QSI. Der Eisenbahningenieur, Ausgabe Dezember 2000

[4] Ackermann, H.: Instandhaltungs– Management– Systeme DB AG, mehr als nur technische Herausforderung. Aachen, 10.12.2013, Folie 10ff. http://www.dmg-berlin.info, abgerufen am 03.06.2022

[5] Quin, L.: XML Essentials. https://www.w3.org/standards/xml/core, abgerufen am 29.06.2022

[6] https://www.eba.bund.de/DE/Themen/Fahrzeuge/Instandhaltung/instandhaltung_node.html, abgerufen am 28.06.2022

[7] https://www.w3.org/standards/xml/schema.html, abgerufen am 30.06.2022

Abbildungen

[1] Die Arbeitsanweisung als Kernelement und Know-how Träger, Ackermann, H.: Instandhaltungs– Management– Systeme DB AG, mehr als nur technische Herausforderung. Aachen, 10.12.2013, Folie 10ff. http://www.dmg-berlin.info, abgerufen am 03.06.2022

[2] Datenstruktur einer Arbeitsanweisung bei der DB AG, Schnell, Toni: Arbeitsdarstellung aus dem Projekt REFI der DB AG

[3] Datenschema (XSD) einer Arbeitsanweisung bei der DB AG, Arbeitsdarstellung aus dem Projekt REFI der DB AG

Session 11
Margit Olbrich, Toni Schnell

Autoren



Olbrich, Margit

Dipl. .-Inf.-Wirtin (FH) Margit Olbrich (58), Studium Informations- und Wissensmanagement an der FH Darmstadt. Seit 2001 Projektleiterin für IT-Anwendungen bei der DB AG und seit 2008 die Anwendungsverantwortliche für die Anwendung REFI (Regelwerkserstellung Fahrzeuginstandhaltung) und deren Voranwendung QSI5 bei der Deutschen Bahn, DB-Systemtechnik, Abteilung TT.TVE43.

Anschrift: DB-Systemtechnik, TT.TVE 43, Weilburger Straße 22, 60326 Frankfurt a. Main, Deutschland

E-Mail: Margit.Olbrich@deutschebahn.com



Schnell, Toni

Geschäftsführer der inxire GmbH

Anschrift: inxire GmbH, Hanauer Landstraße 293 A, 60314 Frankfurt am Main, Deutschland / Germany

E-Mail: toni.schnell@inxire.com

IRSA 2025

SESSION 12

SICHERHEIT + BAHNÜBERGANG

SAFETY + LEVEL CROSSING

Session 12
Carolin Lang, Lukas Leicht, Patrick Ziesel, Aaron Zdarsky, Martin Cichon

Konzepte zur netzunabhängigen Energieversorgung für eine alternative Sicherung von Bahnübergängen

Lang, Carolin; Leicht, Lukas; Ziesel, Patrick; Zdarsky, Aaron und Cichon, Martin

Institut für Fahrzeugsystemtechnik, Karlsruher Institut für Technologie

Zusammenfassung

Auf Nebenstrecken bleiben Bahnübergänge häufig nichttechnisch gesichert, da die Installation von herkömmlichen Sicherungsanlagen kostspielig ist. Ein wesentlicher Kostentreiber sind dabei die Erdarbeiten, die im Zuge eines Netzanschlusses und der Verlegung der Sensorik anfallen.

Um die Kosten für eine technische Sicherung des Bahnübergangs zu senken und Erdarbeiten einzusparen, wird in dieser Untersuchung auf einen Netzanschluss verzichtet und die benötigte Energie vor Ort erzeugt. Dafür werden verschiedene netzunabhängige Konzepte analysiert und hinsichtlich ihrer Eignung bewertet.

Ein Konzept ist eine Energieversorgung mittels Photovoltaikmodulen in Kombination mit einem Batteriespeicher. Außerdem wird noch eine Brennstoffzelle mit Photovoltaikunterstützung untersucht.

Ein weiteres Konzept ist die induktive Energieübertragung. Passiert ein Fahrzeug einen Bahnübergang, so kann eine am Fahrzeug integrierte Primärspule kontaktlos eine im Fahrweg verbaute Sekundärspule speisen. Dieser induzierte Strom kann über Batterien gepuffert und dem Übergang zur Verfügung gestellt werden.

Keywords: Netzunabhängige Energieversorgung, Sicherung von Bahnübergängen, Photovoltaik, Brennstoffzelle, Induktive Energieübertragung

1 Einleitung

Aufgrund der steigenden Nachfrage nach klimafreundlichem und nachhaltigem Verkehr rückt der Eisenbahnverkehr in den Fokus. Um diese steigende Nachfrage bedienen zu können, besteht die Möglichkeit, die Infrastruktur – darunter auch Nebenstrecken – auszubauen. [1]

Gerade auf Nebenstrecken gibt es Bahnübergänge, die nicht technisch gesichert sind. Gemäß Eisenbahnbau- & Betriebsordnung darf auf eine technische Sicherung verzichtet werden, wenn der Bahnübergang wenig frequentiert ist. Um trotzdem eine ausreichende Sicherheit zu gewährleisten, müssen querende Züge abhängig von den Sichtbedingungen ihre Geschwindigkeit reduzieren und ihre Ankunft gegebenenfalls mittels Pfeifen ankündigen. [2]

Diese Sicherheitsvorkehrungen bergen einige Nachteile. Zum einen steigt der Energieverbrauch der passierenden Züge. Zum anderen steigt die Beschwerdelage durch die erhöhten Lärmemissionen der Fahrzeuge [3].

Eine technische Nachsicherung ist mit finanziellen Hürden verbunden. Ein wesentlicher Kostentreiber sind Erdarbeiten, die im Zuge eines Netzanschlusses und der Verlegung der Sensorik anfallen. [4]

Im Zuge des Forschungsprojekts ATTRACTIVE wird ein kostengünstiger Ansatz untersucht, um solche Bahnübergänge nachzusichern. Ein Bestandteil davon ist eine netzunabhängige Energieversorgung, um die Kosten für die genannten Erdarbeiten zu reduzieren. [5]

Deshalb werden in dieser Untersuchung folgende Energieversorgungslösungen betrachtet:

- Photovoltaikmodule mit Batteriespeicher
- Brennstoffzelle in Kombination mit Photovoltaikmodul
- Dynamisch induktiv geladene Batteriespeicher

2 Grundlagen und Stand der Technik

2.1 Bahnübergänge

Bahnübergänge werden nach § 11 der Eisenbahn-Bau- und Betriebsordnung (EBO) als „[…] höhengleiche Kreuzungen von Eisenbahnen mit Straßen, Wegen und Plätzen. […]“ definiert [2].

Bahnübergänge lassen sich gemäß [6] in zwei Hauptgruppen unterteilen: Technisch gesicherte und nichttechnisch gesicherte Übergänge. Wann eine technische Sicherung notwendig ist, hängt vom täglichen kreuzenden Straßenverkehrsaufkommen ab und ist ebenfalls in der EBO festgelegt [2].

Bei Bahnübergängen mit technischer Sicherung werden Straßenverkehrsteilnehmende mithilfe von technischen Anlagen audiovisuell – also durch Schranken, Lichtzeichen oder akustische Signale – auf ein sich näherndes Schienenfahrzeug aufmerksam gemacht [6]. Bahnübergänge ohne technische Sicherung werden lediglich durch das Aufstellen von Andreaskreuzen gesichert, welche die Vorfahrt der Schienenfahrzeuge anzeigen. [6], [7]

Ein zentraler Nachteil nicht technisch gesicherter Bahnübergänge besteht darin, dass Züge diese häufig nur mit verminderter Geschwindigkeit passieren dürfen [6].

Technische Anforderungen werden bei bundeseigenen Strecken durch die Technischen Richtlinien der Deutschen Bahn geregelt. Im Besonderen gilt für Bahnübergänge die DB-Richtlinie 815 „Bahnübergänge planen und Instand halten“. Diese wurde vom Eisenbahn-Bundesamt (EBA) als technisches Regelwerk anerkannt. [8]

Für nichtbundeseigene Bahnen gelten für die technische Anforderungen der Bahnübergangssicherung die Vorschriften in der Bahnübergangsvorschrift für nichtbundeseigene Eisenbahnen (BÜV-NE). In der BÜV-NE werden zwar Anforderungen an die Auslegung der Stromversorgung gegeben, dies gilt jedoch ausschließlich für eine Stromversorgung mit Netzanschluss. Die Versorgung eines Sicherungssystems ohne Netzanschluss wird in der Vorschrift nicht weiter berücksichtigt. [9]

2.2 Photovoltaik

Photovoltaik nutzt den photoelektrischen Effekt aus, bei dem Licht Elektronen in einem Halbleitermaterial löst. Durch eine p-n-Übergangsschicht entsteht dabei ein elektrisches Feld, das die freien Elektronen in eine Richtung lenkt und so elektrischen Strom erzeugt. Auf diese Weise lässt sich also Lichtenergie direkt in elektrische Energie umwandeln.

Das auf ein Photovoltaikmodul einfallende Licht kann in drei Strahlungsarten unterteilt werden. *Direktstrahlung* bezeichnet den Anteil der Sonnenstrahlung, der die Erdatmosphäre ungehindert durchdringt und direkt auf das Modul trifft. Ein weiterer Teil der Sonnenstrahlung wird durch Moleküle und Partikel in der Atmosphäre gestreut. Der Anteil, der nach dieser Streuung das Modul erreicht, wird als *Diffusstrahlung* bezeichnet.

Reflektierte Strahlung entsteht, wenn Sonnenlicht zunächst von Oberflächen wie dem Erdboden oder Gebäuden reflektiert wird und anschließend auf das Modul trifft.

Die Zusammensetzung dieser drei Strahlungsarten hängt vom geografischen Standort, der Tages- und Jahreszeit sowie von der Ausrichtung und Neigung des Moduls ab.

Photovoltaikmodule können nur einen Teil der einfallenden Strahlungsenergie in elektrische Energie umwandeln. Der Wirkungsgrad liegt bei handelsüblichen Silizium-Solarmodulen typischerweise zwischen $15 - 24\,\%$. Neben dem Wirkungsgrad beeinflusst insbesondere die Modulfläche maßgeblich den erzielbaren Energieertrag.

Zum Vergleich verschiedener Photovoltaikmodule wird die Nennleistung in der Einheit *Watt-Peak* (Wp) angegeben. Sie bezeichnet die elektrische Leistung, die ein Modul unter festgelegten Standard-Testbedingungen erbringen kann. Diese lauten:

- Bestrahlungsstärke: $1000\ \frac{\mathrm{W}}{\mathrm{m}^2}$
- Temperatur des Solarmoduls: $25\ °\mathrm{C}$
- Lichtspektrum: AM 1,5

Das Lichtspektrum AM 1,5 (Air Mass) beschreibt die spektrale Verteilung der Sonnenstrahlung nach dem Eintritt in die Erdatmosphäre. Bei einem Wert von 1,5 legt das einfallende Licht aufgrund des Sonnenstands den 1,5-fachen Weg zur Erdoberfläche zurück im Vergleich zur Wegstrecke im Zenit. [10]

2.3 Induktive Energieübertragung

Bei der induktiven Energieübertragung erzeugt eine Primärspule durch Wechselstrom ein zeitlich veränderliches Magnetfeld. Dieses koppelt magnetisch mit einer Sekundärspule, in der gemäß Faradayschem Induktionsgesetz eine Spannung induziert wird. Die Effizienz der Übertragung hängt dabei wesentlich von Stromstärke, Frequenz und der Kopplung der Spulen ab. Die Energieübertragung erfolgt kontaktlos über das Magnetfeld zwischen der Sender- und der Empfangsspule. [11]

Ein induktives Energieübertragungssystem erfordert neben Primär- und Sekundärspule einen Umrichter auf der Primärseite zur Erzeugung hochfrequenter Wechselströme. Auf der Sekundärseite wandelt ein Gleichrichter die induzierte Spannung in Gleichstrom, der zum Laden einer angeschlossenen Batterie genutzt wird. [12]

2.4 Stand der Forschung

Es existiert bereits Forschung bezüglich netzunabhängiger Energieversorgung von Bahnübergängen. In einer Veröffentlichung aus dem Jahr 2021 von der Cardiff University zusammen mit Network Rail in Großbritannien werden unterschiedliche Energy-Harvesting-Möglichkeiten miteinander verglichen. Darin soll eine Lichtsignalanlage elektrisch versorgt werden, um einen technisch nicht gesicherten Bahnübergang nachzusichern. Die erste vorgestellte Lösung beinhaltet ein System mit Solarmodulen, einer Methanol-Brennstoffzelle sowie Blei-Säure-Batterien. Das zweite System nutzt zwei Solarmodule, eine Windturbine und eine Natrium-Nickel-Chlorid-Batterie. Die Aktivierung der Lichtsignalanlage erfolgt über konventionelle Achszähler. Es wird mit einem Energiebedarf von etwa 1,3 kWh/Tag gerechnet. [13]

Im Jahr 2024 wird von der Tashkent State Transport University in Usbekistan ein Energieversorgungskonzept bestehend aus Photovoltaik in Kombination mit einem Energiespeicher zur Versorgung von Eisenbahnsignalen vorgesellt. Der Fokus lag hierbei auf einer höheren Zuverlässigkeit der Energieversorgung. Die Leistungsfähigkeit der Photovoltaiksysteme wurde für zwei Regionen in Usbekistan analysiert, wobei sowohl ein autarker als auch ein netzgekoppelter Betrieb betrachtet wurde. Es wird von einem Leistungsbedarf von 600 W für die Signale ausgegangen und einem Energiespeicher mit einer Kapazität von 3 kWh. Für die Solarmodule ist eine Fläche von 0,5 m^2 vorgesehen, womit die Autoren eine jährliche Ausbeute von etwa 938 kWh erwarten. [14]

Außerdem werden mit dem Industrial Area Crossing Signal System von O´Conner Engineering Inc. bereits Bahnübergangssignale für Industriegebiete in den USA kommerziell vertrieben. Diese werden mit Hilfe von Solarmodulen in Kombination mit einer Batterie mit Energie versorgt. Die dort eingesetzten Solarmodule erbringen eine Leistung von 90 W bis 140 W. Die verwendeten Batterien sind mit einer Kapazität von 110 Ah angegeben. Die Angabe zur Klemmenspannung fehlt, jedoch werden andere Komponenten des gleichen Systems mit 12 V Gleichspannung betrieben. [15]

Die Forschung zur induktiven Energieübertragung konzentriert sich vor allem auf Straßenfahrzeuge, die durch den Fahrweg geladen werden. Unterschieden wird zwischen stationärem Laden im Stand und dynamischem Laden während der Fahrt [16]. Für leichte Nutzfahrzeuge existieren Standards wie SAE J2954 sowie IEC 61980-1 bis -3 [12]. Für den Schienenverkehr liegen bislang keine normativen Vorgaben vor.

Die benötigte fahrzeugseitige Infrastruktur wird durch Ladepads mit einer Leistung von bis zu 20 kW realisiert, wobei die Anzahl durch den Bauraum begrenzt ist [17].

Mehrere Forschungsprojekte evaluieren die technische Umsetzbarkeit induktiver Energieübertragung im Straßenverkehr. Im Projekt FABRIC wurden dynamische Ladevorgänge bei unterschiedlichen Fahrgeschwindigkeiten untersucht – mit Wirkungsgraden von 62 % bei 30 km/h und 60 % bei 50 km/h. [18]

In Busan (Südkorea) wurde das quasi-dynamische Laden einer Straßenbahn mit Gummibereifung erprobt. Auf bis zu 250 m langen Abschnitten konnten Übertragungsleistungen bis 1 MW realisiert werden. Die Effizienz variiert jedoch stark mit der übertragenen Leistung – oberhalb von 350 kW steigt sie auf über 80 % bei einem Spulenabstand von 60 mm. [19]

3 Methodik

3.1 Ort

Als Untersuchungsort für die exemplarische Analyse verschiedener Energy-Harvesting-Methoden wird ein Bahnübergang in Siegsdorf (Bayern) nahe des Chiemsees gewählt. Konkret handelt es sich um den Übergang „Höpfling“, dessen geographische Lage in Abbildung 1 nachvollzogen werden kann.



Abbildung 1: Bahnübergang Höpfling

Die Bahnstrecke ist eine eingleisige Nebenstrecke, die Traunstein, Siegsdorf und Ruhpolding miteinander verbindet. Der Übergang befindet sich zwischen einer Brücke über den Fluss „Weiße Traun“ und einer T-Kreuzung zur Ruhpoldinger Straße. Südlich des Übergangs befindet sich die Haltestelle „Höpfling“, an der Züge aus beiden Richtungen halten.

3.2 Betrachtete Bahnübergangssicherung

Der in Abschnitt 3.1 vorgestellte Bahnübergang ist technisch nicht gesichert. Für eine technische Sicherung sind verschiedene Komponenten erforderlich, für deren Energieverbräuche von folgenden Werten ausgegangen wird:

Tabelle 1: Energieverbräuche der betrachteten Komponenten

Komponente	Verbrauch im Leerlauf	Verbrauch im Betrieb
Steuereinheit tagsüber	28,7 W	32,6 W
Steuereinheit nachts	4,5 W	–
Master-Lichtsignalanlage	6,6 W	7,8 W
Slave-Lichtsignalanlagen	$3 \cdot 4{,}2\ \mathrm{W} = 12{,}6\ \mathrm{W}$	$3 \cdot 5{,}4\ \mathrm{W} = 16{,}2\ \mathrm{W}$

Bei den angegebenen Verbräuchen handelt es sich um eigens durchgeführte Messungen von prototypischen Komponenten, die den Autoren im Rahmen dieser Untersuchung zur Verfügung stehen. Das in dieser Untersuchung betrachtete Sicherungssystem basiert auf einer funkbasierten Kommunikation zwischen einer fahrzeugseitigen Einheit und einer ortsfesten Steuereinheit. Nähert sich ein Zug dem Bahnübergang, so löst die Steuereinheit die Master-Lichtsignalanlage aus und schaltet sie nach erfolgter Überfahrt wieder ab. Zur vollständigen Absicherung des Übergangs werden zusätzlich 3 Slave-Lichtsignalanlagen eingesetzt, die das Signal der Master-Einheit über Funk synchron übernehmen. Zur Reduzierung des Energieverbrauchs verfügt die Steuereinheit über einen Ruhemodus, der abends nach Überfahrt des letzten Zugs aktiv wird und morgens vor Überfahrt des ersten Zugs wieder abgeschaltet wird.

3.3 Betrachteter Energiebedarf

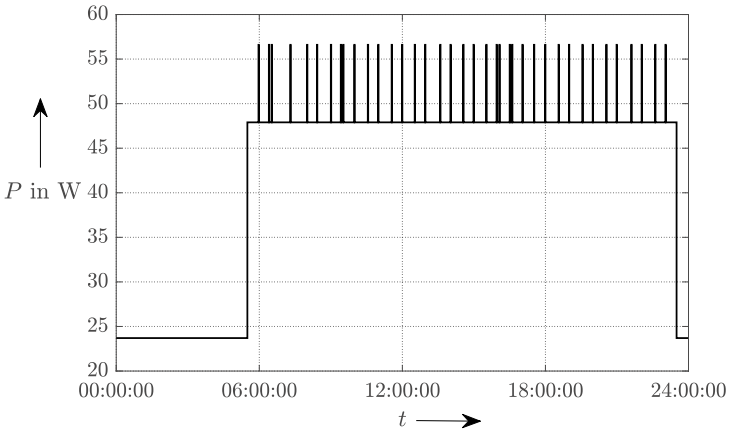


Abbildung 2: Lastprofil

Da die Energieverbräuche im Leerlauf und im Betrieb sowie tagsüber und nachts unterschiedlich sind, müssen plausible Annahmen getroffen werden, wann und wie lange die Bahnübergangssicherung ausgelöst wird und in den Ruhezustand geht. Zur Ermittlung des täglichen Energiebedarfs wurde das Kursbuch [20], [21] herangezogen, aus dem die Ankunftszeiten der Züge an der nahegelegenen Haltestelle „Höpfling“ abgeleitet wurden. Aus jenem Kursbuch und den

aufgeführten Energieverbräuchen ergibt sich das in Abbildung 2 gezeigte Lastprofil über einen Tag. Es wird fahrplanbedingt angenommen, dass die Steuereinheit zwischen 05:30 Uhr und 23:30 Uhr aktiv ist und sich außerhalb dieser Zeiten im Ruhezustand befindet. Das entspricht einem Zeitraum von einer halben Stunde vor Passieren des ersten Zugs und einer halben Stunde nach Passieren des letzten Zugs. Darüber hinaus wird angenommen, dass jeder passierende Zug die Bahnübergangssicherung für 1 Minute aktiv hält. Es wird vereinfachend davon ausgegangen, dass das dargestellte Lastprofil für alle Tage des Jahres repräsentativ ist.

4 Simulation verschiedener Energy-Harvesting-Ansätze

4.1 Photovoltaikmodule mit Batteriespeicher

Zur Versorgung des Übergangs mit Photovoltaikmodulen und Batteriespeichern steht im Rahmen dieser Untersuchung das in Abbildung 3 dargestellte System zur Verfügung.

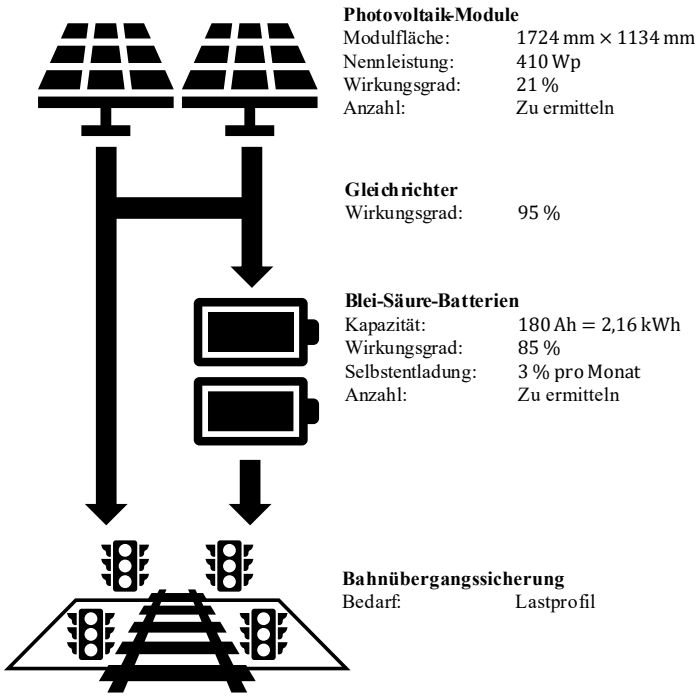


Abbildung 3: Schematischer Aufbau der Photovoltaikanlage

Die technischen Eigenschaften der eingesetzten Komponenten basieren auf Herstellerdatenblättern und wurden durch Literaturwerte ergänzt.

Bei den Wirkungsgraden des Gleichrichters und der Blei-Säure-Batterie wurde sich an [20] orientiert. Sie wurden so gewählt, dass sie zu den darin angegebenen Bereichen der Zellwirkungsgrade und Gesamtwirkungsgrade inklusive Umrichter passen. Der Wert für die monatliche Selbstentladung stammt ebenfalls aus einem in [20] angegebenen Bereich.

Zur adäquaten Simulation der elektrischen Leistung, die die Photovoltaikmodule zur Verfügung stellen, wird auf historische Einstrahlungswerte zurückgegriffen. Dazu eignen

sich Einstrahlungsdaten, die der Deutsche Wetterdienst auf seiner Homepage zur Verfügung stellt [21]. Dort sind Datensätze zur Global- und Diffusstrahlung hinterlegt, die von den verschiedenen Wetterstationen in Deutschland gemessen wurden.

Eine Analyse der Daten ergibt, dass an der nächstgelegenen Wetterstation in den Wintermonaten von 2010 bis 2011 die geringste Einstrahlung im Vergleich zu allen anderen verfügbaren Jahren zu verzeichnen war. Dieser Zeitraum dient daher im Folgenden als Vergleichszeitraum, auf den die Photovoltaikanlage ausgelegt wird.

Aufgrund des signifikanten Einflusses der Modulorientierung auf den Energieertrag wurden folgende Parameter festgelegt:

- Der Anstellwinkel beträgt $\beta = 40°$
- Die Module sind nach Süden ausgerichtet, also beträgt $\varepsilon = 180°$

Außerdem wird davon ausgegangen, dass die Module zu keiner Zeit verschattet sind. Bezüglich des Anstellwinkels sei erwähnt, dass eine steilere Anstellung in den Wintermonaten mit tiefstehender Sonne zu einer größeren Ausbeute von Direktstrahlung führt. Andererseits nimmt mit steileren Winkeln auch die Ausbeute von Diffusstrahlung ab. Vorsimulationen ergaben für dieses Szenario, dass die Gesamtausbeute bei einem Anstellwinkel von 40° höher ist als bei steileren Winkeln.

Aus den genannten historischen Einstrahlungsdaten werden die Anteile der Direktstrahlung, Diffusstrahlung und reflektierter Strahlung berechnet. Unter Berücksichtigung des Anstellwinkels, der Ausrichtung und des Wirkungsgrades der Module kann dann ein zeitabhängiges Profil elektrischer Leistung berechnet werden, die die Module zur Verfügung stellen.

Ziel der Simulation ist es, eine Kombination aus den Anzahlen n_{Module} und $n_{\text{Batterien}}$ zu finden, die bei dem gewählten Zeitraum für die solare Einstrahlung noch genügend elektrische Leistung zur sicher-

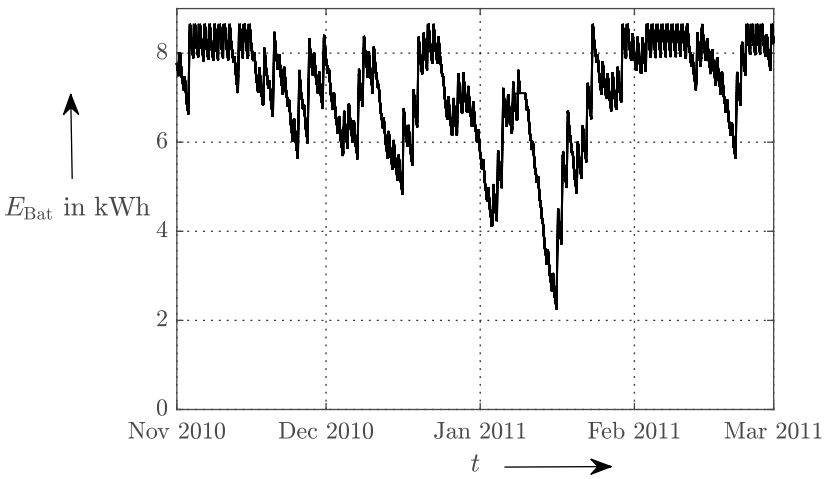


Abbildung 4: Verlauf des Ladezustands der Batterien

en Versorgung der Bahnübergangssicherung bereitstellt.

Für die Kombination $n_{\text{Module}} = 4$ und $n_{\text{Batterien}} = 4$, ergibt sich der in Abbildung 4 gezeigte Verlauf des Ladezustands der Batterien, wenn von einem initialen Ladestand von 90 % ausgegangen wird. In Phasen geringer solarer Einstrahlung reicht die durch die Photovoltaikmodule erzeugte Energie nicht aus, um den Bedarf vollständig zu decken. In diesen Zeiten wird aus der Batterie ausgespeist, wodurch der Ladestand teilweise deutlich abfällt. Allerdings ist der Ladestand stets positiv und außerhalb der in [20] angegebenen Entladetiefe von 70 bis 80 %.

Durch den ungewöhnlich dunklen Vergleichszeitraum kann die Anlage mit dieser Kombination als ausreichend dimensioniert betrachtet werden.

4.2 Brennstoffzelle mit Photovoltaikunterstützung

Zur netzunabhängigen Versorgung des Bahnübergangs steht im Rahmen dieser Untersuchung das in Abbildung 5 gezeigte Brennstoffzellensystem zur Verfügung.

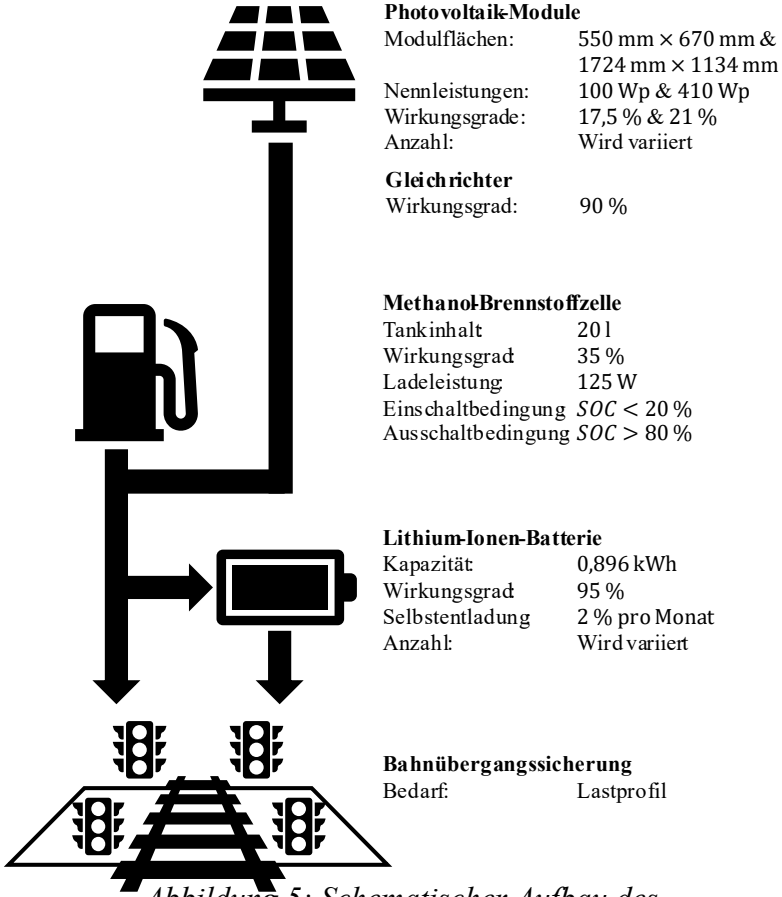


Abbildung 5: Schematischer Aufbau des Brennstoffzellensystems mit Photovoltaikunterstützung

Die technischen Eigenschaften der eingesetzten Komponenten basieren auf Herstellerdatenblättern und wurden durch Literaturwerte ergänzt.

Bei den Wirkungsgraden des Gleichrichters und der Lithium-Ionen-Batterie wurde sich an [20] orientiert. Sie wurden so gewählt, dass sie zu den darin angegebenen Bereichen der Zellwirkungsgrade und

Gesamtwirkungsgrade inklusive Umrichter passen. Der Wert für die monatliche Selbstentladung stammt ebenfalls aus einem in [20] angegebenen Bereich.

Ziel der Simulation ist die Bestimmung des Autarkiegrads des Systems unter den gleichen Betriebsbedingungen wie in Abschnitt 4.1. Der Methanoltank ist zu Beginn der Simulation voll und wird wiederaufgefüllt, sobald sein Füllstand unter 10 % beziehungsweise 2 l fällt. In Abhängigkeit der installierten PV-Modulleistung P_{ges} und der Anzahl $n_{\mathrm{Batterien}}$ der Lithium-Ionen-Batterien ergibt sich Autarkiegrad A, der in Abbildung 6 gezeigt ist.

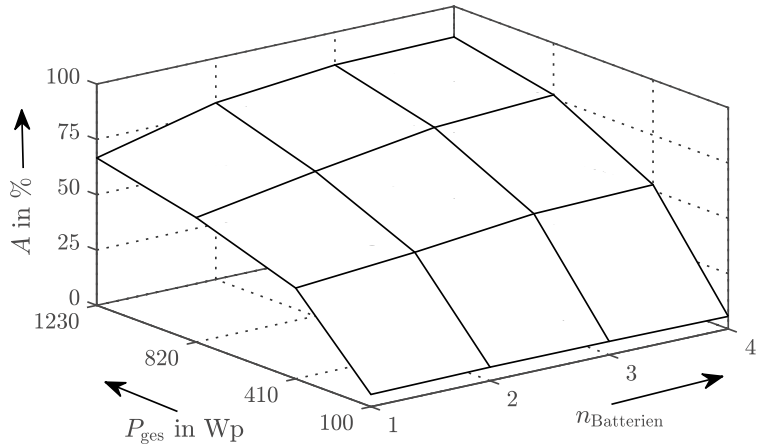


Abbildung 6: Autarkiegrad des Brennstoffzellensystems

Es ist erkennbar, dass der Autarkiegrad anfangs schnell steigt mit zunehmender Modulleistung und Batterieanzahl, für volle Autarkie jedoch überproportional mehr Module und Batterien benötigt werden. Das Nachfüllintervall für den Methanoltank ist umgekehrt proportional zum Autarkiegrad. Bei 5 % werden im betrachteten Zeitraum etwa 5 Tankinhalte benötigt, während bei 80 % etwa 1 Tankinhalt ausreicht.

4.3 Induktive Energieübertragung durch Schienenfahrzeuge

Für jedes Schienenfahrzeug wird der Einsatz von fünf Ladepads angenommen, wodurch eine Übertragungsleistung von 100 kW erreicht wird. Die übertragbare Energie ist das Produkt aus Übertragungsleistung und Ladezeit und hängt damit von der Leistungsfähigkeit der Ladeinfrastruktur und der Verweildauer des Fahrzeugs über den Ladepads ab. Mit sinkender Fahrzeuggeschwindigkeit erhöht sich also die übertragbare Energie bei konstanter Übertragungsleistung.

Die angenommene Übertragungseffizienz liegt bei 75 %. Dieser Wert ist etwas niedriger als der aus Südkorea, siehe Abschnitt 2.4, was auf einen größeren Luftspalt und den Einfluss stählerner Radsätze auf das Magnetfeld zurückzuführen ist. Bei einer angenommenen Durchfahrtsgeschwindigkeit von 30 km/h ergibt sich eine erforderliche

Ladestrecke von 13,83 m, um die angestrebte Energiemenge zu übertragen. Pro Ladevorgang können so 46,08 Wh übertragen werden.

In Abbildung 7 ist der zeitliche Verlauf des Batterieladezustands bei einer auf 14 m aufgerundeten Ladestrecke gezeigt. Dargestellt sind zwei Szenarien: 30 km/h bei Halt an der Station sowie 60 km/h bei Durchfahrt. Im letztgenannten Szenario ist die verfügbare Energie nicht ausreichend, um den Bahnübergang zuverlässig zu versorgen.

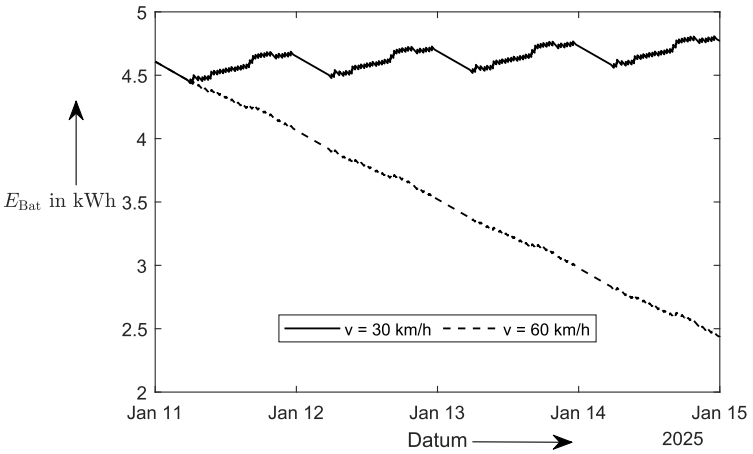


Abbildung 7 - Energieübertragung durch Schienenfahrzeuge mit einer Ladestreckenlänge von 14 m

4.4 Induktive Energieübertagung durch Straßenfahrzeuge

Aufgrund des begrenzten Bauraums wird für Straßenfahrzeuge von einem einzelnen Ladepad mit einer Übertragungsleistung von 20 kW pro Fahrzeug ausgegangen [18]. Für die Berechnung werden Durchschnittswerte verwendet. Bei einer angenommenen Fahrgeschwindigkeit von 30 km/h – bedingt durch die Lage des Bahnübergangs nahe der Kreuzung – und einer durchschnittlichen Effizienz von 62 % (basierend auf dem Projekt FABRIC) ergibt sich ein realistisches Energieszenario.

Die verfügbare Ladestrecke am Bahnübergang Höpfling ist durch ihre Lage auf einer Brücke auf 30 m begrenzt. Laut Verkehrskonzept beträgt die tägliche Verkehrsbelastung dort rund 400 Fahrzeuge [22], wovon 4 % als batterieelektrisch angenommen werden – leicht über dem deutschlandweiten Anteil von 3,3 % (Stand Januar 2025) [23], da künftig von einer Steigerung des Anteils ausgegangen wird. Es wird optimistisch angenommen, dass sämtliche batterieelektrischen Fahrzeuge über eine kompatible induktive Ladeinfrastruktur verfügen.

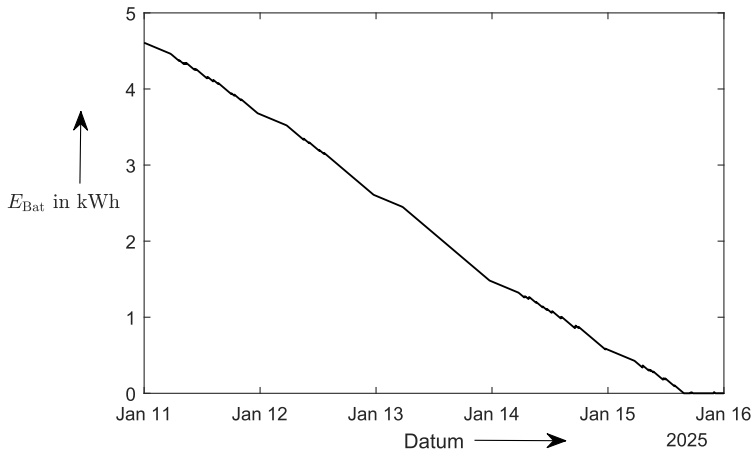


Abbildung 8 – Energieübertragung durch Straßenfahrzeuge mit einer Ladestreckenlänge von 30 m

In Abbildung 8 ist der Energiegehalt der Batterie über sechs Tage gezeigt. Die 30 m lange Ladestrecke reicht unter den getroffenen Annahmen nicht aus, um den Bahnübergang dauerhaft zu versorgen – der Speicher entleert sich innerhalb weniger Tage. Um den Energiebedarf im betrachteten Lastfall (vier Lichtzeichen, eine Steuereinheit) zu decken, wäre eine Ladestrecke von mindestens 161,1 m erforderlich. Bei einem steigenden Anteil an ladenden Fahrzeugen kann diese Länge perspektivisch entsprechend reduziert werden.

5 Zusammenfassung

Die Untersuchung zeigt, dass eine Photovoltaikanlage mit Batteriespeicher grundsätzlich geeignet ist, um einen Bahnübergang netzunabhängig mit Energie zu versorgen. Für einen sicheren Betrieb muss jedoch auch in Phasen außergewöhnlich geringer Sonneneinstrahlung eine ausreichende Energieversorgung gewährleistet sein. Daraus ergibt sich ein hoher Dimensionierungsbedarf für Modulfläche und Speicherkapazität. Diesem kann entgegengewirkt werden, indem der Energieverbrauch der verwendeten Komponenten optimiert wird. Einsparpotential liegt unter anderem bei der Steuereinheit, die den größten Energiebedarf aufweist.

Ein hybrides System aus Brennstoffzelle und Photovoltaikunterstützung stellt ebenfalls eine geeignete Lösung dar. Durch die Kombination beider Technologien kann die Dimensionierung der Photovoltaikkomponenten reduziert werden. Der regelmäßige Nachfüllbedarf des Methanoltanks stellt jedoch einen betrieblichen Aufwand dar. Dieser kann durch größere Photovoltaikmodule und Batteriespeicher reduziert werden. Insgesamt ermöglicht dieser Ansatz einen kompakteren und flexibleren Systemaufbau.

Die induktive Energieübertragung durch Schienenfahrzeuge ist technisch realisierbar, erfordert jedoch eine flächendeckende Ausstattung der Fahrzeuge mit entsprechender

Ladeinfrastruktur. Auf Streckenabschnitten mit vielen Fahrzeugen kann das daher die Wirtschaftlichkeit beeinträchtigen.

Die induktive Energieübertragung durch Straßenfahrzeuge ist nur unter sehr spezifischen Voraussetzungen sinnvoll. Neben einer ausreichend hohen Anzahl kompatibler Fahrzeuge sind lange Ladestrecken erforderlich, was die bauliche Umsetzung und Wirtschaftlichkeit erschweren. Zudem ist die Akzeptanz seitens der Fahrzeughalter kritisch zu betrachten, da ein wirtschaftlicher Anreiz zur Nachrüstung notwendig wäre. Insgesamt erscheint dieser Ansatz derzeit nur eingeschränkt praktikabel.

Unabhängig vom gewählten Versorgungskonzept sind rechtliche Rahmenbedingungen zu schaffen, da der netzunabhängige Betrieb von Bahnübergangssicherungen gemäß [9] in Deutschland bislang nicht vorgesehen ist. Eine Anpassung der Vorschriften ist erforderlich, um innovative Versorgungslösungen rechtlich abzusichern und deren Einsatz zu ermöglichen.

Danksagung

Diese Arbeit wurde im Rahmen des Projekts ATTRACTIVE (FKZ 01F1190A), gefördert durch das Bundesministerium für Digitales und Verkehr, durchgeführt.

Literatur

[1] „Neu- und Ausbau Infrastruktur“. Zugegriffen: 10. Juli 2025. [Online]. Verfügbar unter: https://www.bmv.de/SharedDocs/DE/Artikel/E/schiene-neubau-ausbau-erhalt/neubau-ausbau-infrastruktur-bedarfsplan.html

[2] „§ 11 EBO - Einzelnorm“. Zugegriffen: 6. Dezember 2024. [Online]. Verfügbar unter: https://www.gesetze-im-internet.de/ebo/__11.html

[3] von N. Häußinger, „Wenn das Pfeifen nervt ...“, Tegernseerstimme. Zugegriffen: 10. Juli 2025. [Online]. Verfügbar unter: https://tegernseerstimme.de/wenn-das-pfeifen-nervt/

[4] J. Oelsner und J. Bruder, „Analyse Optimierungspotential Bahnübergangssicherung“, EBA Forschungsbericht, Apr. 2019. Zugegriffen: 26. November 2024. [Online]. Verfügbar unter: https://www.dzsf.bund.de/SharedDocs/Downloads/DZSF/Veroeffentlichungen/Forschungsberichte/2019/EBA_Forschungsbericht_2019-04_Anlagen.pdf?__blob=publicationFile&v=5

[5] „Autonomous Train TRansponder Alerting Cross Track Intercepting VEhicles - ATTRACTIVE“. Zugegriffen: 11. Juli 2025. [Online]. Verfügbar unter: https://www.bmv.de/SharedDocs/DE/Artikel/DG/mfund-projekte/attractive.html

[6] E. Schöne, „Bahnübergänge“, in *Handbuch Eisenbahninfrastruktur*, L. Fendrich und W. Fengler, Hrsg., Berlin, Heidelberg: Springer, 2019, S. 673–709. doi: 10.1007/978-3-662-56062-4_13.

[7] *Straßenverkehrs-Ordnung (StVO)*. 2024.

[8] „EBA - Homepage - BÜ-Leitfaden“. Zugegriffen: 4. Juli 2025. [Online]. Verfügbar unter: https://www.eba.bund.de/SharedDocs/Downloads/DE/Infrastruktur/Bahnuebergaenge/21_bue_leitfaden.html

[9] F. Christ *u. a.*, *VDV BÜV- NE:Vorschrift für die Sicherung der Bahnübergänge bei nichtbundeseigenen Eisenbahnen*, VDV-Schrift 363, August 2020.

[10] K. Mertens, *Photovoltaik: Lehrbuch zu Grundlagen, Technologie und Praxis*, 7. Auflage. München: Hanser, 2025. doi: 10.3139/9783446480674.

[11] A. Triviño-Cabrera, J. M. González-González, und J. A. Aguado, *Wireless Power Transfer for Electric Vehicles: Foundations and Design Approach*. in Power Systems. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-26706-3.

[12] D. Barth, „Mathematische Methoden für die effiziente Auslegung und den Betrieb induktiver Batterieladesysteme“. Zugegriffen: 12. November 2024. [Online]. Verfügbar unter: https://publikationen.bibliothek.kit.edu/1000157668

[13] A. Jeyanathan, S. Chew, D. Stevenson, und L. Cipcigan, „Whole Life Costing of Off- Grid Power Supplies for Rail Level Crossing Safety Systems“, in *2021 56th International Universities Power Engineering Conference (UPEC)*, Aug. 2021, S. 1–6. doi: 10.1109/UPEC50034.2021.9548187.

[14] A. V. Ablayeva, N. V. Yaronova, und O. O. Ruzimov, „Ensuring a Reliable Power Supply for Railway Traffic Lights Using Renewable Energy Sources“, in *2024 International Ural Conference on Electrical Power Engineering (UralCon)*, Sep. 2024, S. 600–604. doi: 10.1109/UralCon62137.2024.10718877.

[15] „Railroad Crossing Signals“. Zugegriffen: 10. Juli 2025. [Online]. Verfügbar unter: https://www.oconnerengineering.com/Railroad-Crossing-Signals.html

[16] Mauro Feliziani, Tommaso Campi, Silvano Cruciani, und Francesca Maradei, *Wireless Power Transfer for E-Mobility : Fundamentals and Design Guidelines for Wireless Charging of Electric Vehicles*. Amsterdam: Academic Press, 2023. [Online]. Verfügbar unter: http://www.redi-bw.de/db/ebsco.php/search.ebscohost.com/login.aspx%3fdirect%3dtrue%26db%3dnlebk%26AN%3d3451950%26site%3dehost-live

[17] S. Lee, J. Huh, C. Park, N.-S. Choi, G.-H. Cho, und C.-T. Rim, „On-Line Electric Vehicle using inductive power transfer system“, in *2010 IEEE Energy Conversion Congress and Exposition*, Sep. 2010, S. 1598–1601. doi: 10.1109/ECCE.2010.5618092.

[18] „Final Report Summary - FABRIC (FeAsiBility analysis and development of on-Road chargIng solutions for future electric vehiCles) | FP7“, CORDIS | European Commission. Zugegriffen: 27. März 2025. [Online]. Verfügbar unter: https://cordis.europa.eu/project/id/605405/reporting

[19] G. Lee, M. Y. Kim, S.-G. Lee, und J. H. Kim, „Operational Verification of Semidynamic Wireless Power Transfer in Light-Rail Transit Systems“, *IEEE Transactions on Transportation Electrification*, Bd. 11, Nr. 1, S. 348–358, Feb. 2025, doi: 10.1109/TTE.2024.3390569.

[20] T. Aundrup *u. a.*, „Batteriespeicher in der Nieder- und Mittelspannungsebene“, VDE Verband der Elektrotechnik, Mai 2015.

[21] „Index of /climate_environment/CDC/observations_germany/climate/10_minutes/solar/historical/“. Zugegriffen: 10. Juli 2025. [Online]. Verfügbar unter: https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/solar/historical/

[22] PLANUNGSGESELLSCHAFT STADT-LAND-VERKEHR GmbH, „Verkehrsgutachten Siegsdorf“, München, Feb. 2017. Zugegriffen: 4. März 2025. [Online]. Verfügbar unter: https://www.rathaus-siegsdorf.de/fileadmin/Dateien/PDF-Dateien/Verkehrsgutachten_gesamt.o.pdf

[23] „Kraftfahrt-Bundesamt - Umwelt“. Zugegriffen: 19. März 2025. [Online]. Verfügbar unter: https://www.kba.de/DE/Statistik/Fahrzeuge/Umwelt/umwelt_node.html;jsessionid=AAD8E5EF4385DDB5115717CDEADD4DEB.live11313

Session 12

Carolin Lang, Lukas Leicht, Patrick Ziesel, Aaron Zdarsky, Martin Cichon

Autoren



Carolin Lang, M. Sc.
Wissenschaftliche Mitarbeiterin, Institutsteil Bahnsystemtechnik
Institut für Fahrzeugsystemtechnik, KIT
Rintheimer Querallee 2, 76131 Karlsruhe



Lukas Leicht, M. Sc.
Wissenschaftlicher Mitarbeiter, Institutsteil Bahnsystemtechnik
Institut für Fahrzeugsystemtechnik, KIT
Rintheimer Querallee 2, 76131 Karlsruhe



Patrick Ziesel, M. Sc.
Wissenschaftlicher Mitarbeiter, Institutsteil Bahnsystemtechnik
Institut für Fahrzeugsystemtechnik, KIT
Rintheimer Querallee 2, 76131 Karlsruhe



Aaron Zdarsky, B. Sc.
Studentischer Mitarbeiter, Institutsteil Bahnsystemtechnik
Institut für Fahrzeugsystemtechnik, KIT
Rintheimer Querallee 2, 76131 Karlsruhe



Prof. Dr.-Ing. Martin Cichon
Institutsleiter Institutsteil Bahnsystemtechnik
Institut für Fahrzeugsystemtechnik, KIT
Rintheimer Querallee 2, 76131 Karlsruhe

Session 12
Matthias Jelinski, Walter Schmitz, Stefan Pickl

Security-Management: Risiko-orientierte Bewertung von physischen und Cyber-Angriffen auf das Eisenbahnsystem

Jelinski, Matthias[1], Schmitz, Walter[2], Pickl, Stefan[3]

[1]IVE - Ingenieurgesellschaft für Verkehrs- und -Eisenbahnwesen mbH, Hannover
[2]CreaLab GmbH, Feldkirchen-Westerham
[3]Institut für Theoretische Informatik, Mathematik und Operations Research, Universität der Bundeswehr München

Zusammenfassung

Die Eisenbahn wird u. a. in Deutschland zu den kritischen Infrastrukturen gezählt. Offene Eisenbahnsysteme bieten exponierte bzw. attraktive Ziele für Angriffe. Detaillierte Angriffspotenziale sind bisher nur in den Teil- und Untersystemen hinreichend gut untersucht. Auf Basis eines Fallbeispiels wird eine risiko-orientierte und szenario-basierte Methode für eine ganzheitliche Bewertung von Angriffen in physischer, Cyber- oder hybrider Form vorgestellt. Mittels der Methode können zielführende Sicherheitsmaßnahmen in einem Security-Management festgelegt werden.

Keywords: Security; Angriffssicherheit; kritische Infrastruktur (KRITIS), Digitalisierung; Resilienz; Angriffspotenzial; hybride Bedrohung; Vulnerabilität; Vignette; Szenario; Risiko

1 Einleitung

Für Bahnsysteme ist schon seit jeher ein hoher Technisierungsgrad im Vergleich zu anderen technischen Systemen der jeweiligen Zeit charakterisierend. Eine zunehmende Digitalisierung und Vernetzung ist der aktuelle Ausdruck davon. Gleichzeitig ist die Eisenbahn ein zentraler Verkehrsträger und wird u. a. in Deutschland zu den kritischen Infrastrukturen (KRITIS) gezählt. Sie bietet exponierte bzw. attraktive Ziele für physische Angriffe, für Cyber-Angriffe oder – in ihrer Kombination – für hybride Angriffe. Detail-

lierte Angriffspotenziale sind nur in den Teil- und Untersystemen hinreichend gut untersucht. Eine systematische und strukturierte Ableitung von Sicherheitsanforderungen im Sinne von *security*-Anforderungen für das deutsche Eisenbahnsystem in seiner Gesamtheit fehlte bislang. Im Jahr 2021 hat das Deutsche Zentrum für Schienenverkehrsforschung (DZSF) das Forschungsprojekt „Identifikation bestehender Angriffspotentiale für das System Bahn“ bzw. „REAVRS – Revealing Existing Attack Vulnerabilities in the Rail System“ vergeben.

Ein Ergebnis des DZSF-Forschungsprojekts [1] ist eine risiko-orientierte und szenariobasierte Methode für eine ganzheitliche Bewertung von Angriffen in physischer, Cyber- oder hybrider Form. Mittels der REAVRS-Methode können zielführende Sicherheitsmaßnahmen in einem Security-Management festgelegt werden.

2 Ermittlung bestehender Angriffspotenziale

Angriffspotential wird als ein Maß für den notwendigen Aufwand eines bestimmten Angriffs auf ein bestimmtes (Evaluierungs-) Objekt verstanden – angelehnt an die Definition in der technischen Richtlinie TR-03147 [2] des Bundesamtes für Sicherheit in der Informationstechnik (BSI). Das Objekt bzw. die Objekte sind Bestandteile des Eisenbahnsystems. Für die Ermittlung der Angriffspotenziale erfolgen die Abbildung des Eisenbahnsystems sowie die Systematisierung von Bedrohungen.

2.1 Abbildung des Eisenbahnsystems

Als Grundlage wird die Definition des europäischen Eisenbahnsystems genutzt. Sie ist in der Interoperabilitätsrichtlinie RL EU 2016/797 [3] und insbesondere den Technischen Spezifikationen Interoperabilität (TSI) angegeben. Wesentliche Charakteristika sind:

- Unterscheidung von strukturellen Teilsystemen – bspw. Eisenbahninfrastruktur, Energie, Fahrzeuge – gegenüber funktionellen Teilsystemen – bspw. Instandhaltung, Betriebssteuerung oder Telematikanwendungen
- Definition zweier eigenständigen Teilsysteme für streckenseitige und fahrzeugseitige Leit- und Sicherungstechnik – Abgrenzung einerseits untereinander sowie andererseits von den Teilsystemen Infrastruktur und Fahrzeuge

Zusätzlich erfolgt die Definition eines Teilsystems „Digitale Systeme im Eisenbahnwesen“ losgelöst von den TSI, dafür mit Bezug auf das IT-Grundschutz-Kompendium [4] des BSI.

Für die weitere Abbildung werden Konzepte der Systemhierarchie bspw. gemäß DIN EN 50126 [5] angewandt. Das betrachtete Gesamtsystem wird schichtweise in weitere Teil- oder Untersysteme und schließlich in Komponenten untergliedert. Den Teil- und Untersystemen sowie Komponenten können wiederum Eigenschaften, Funktionen und Schnittstellen untereinander zugeordnet werden.

In dieser Abbildung des Eisenbahnsystems können Teil- oder Untersysteme oder Komponenten als mögliche Angriffspunkte bzw. als generische Orte eines Angriffs angegeben werden.

2.2 Systematisierung von Bedrohungen und Differenzierung von Tätertypen sowie Tatabsicht

Als Grundlage für die Systematisierung der Bedrohungen sind umfangreiche Recherchen erfolgt zu physischen Angriffen auf bzw. in Bahnsystemen weltweit und weitreichend in die Vergangenheit einerseits [6], [7] und zu Cyber-Angriffen andererseits [8].

Eine Bedrohung liegt vor, wenn es einen potentiellen Täter gibt, der sowohl die Absicht als auch die Fähigkeit hat, einen Angriff durchzuführen. Die Fähigkeit beschreibt, dass der Täter über ein geeignetes Angriffsmittel verfügt, das bei dem Angriffspunkt im Eisenbahnsystem eingesetzt werden kann. Ein Angriffsmittel muss nicht zwingend eine Waffe sein. Für eine generische Beschreibung werden die Angriffsmittel in acht Kategorien für physische Angriffe und fünf Kategorien für Cyber-Angriffe eingeteilt (Abbildung 1). Das schließt eine Kombination von physischen und Cyber-Angriffsmitteln nicht aus.

Physische Angriffe

Mechanische Mittel	
P1	Gegen Infrastruktur, Hardware
P2	Gegen Zug/Züge von außen
(Unkonventionelle) Spreng- und/oder Brandvorrichtung	
P3	Gegen Eisenbahn-Infrastruktur
P4	Gegen Zug/Züge von außen
P5	Im Zug/in Zügen
Waffen gegen Person/Personen	
P6	CBRN-Waffe (chemisch, biologisch, radiologisch und nuklear)
P7	Handfeuerwaffen
P8	Schlagwerkzeuge (Messer, Beile, Stöcke)

Cyber-Angriffe

Angriffsmittel zur initialen Kompromittierung	
C11	**Social Engineering** (Phishing, E-Mail-Spoofing, Cross-Site-Scripting)
C12	**Direkter Angriff auf IT-System** (Angriff auf passwortgeschützte Komponente, Exploits, IP-Spoofing, DNS-Spoofing, Man-In-The-Middle, Replay-Angriff)
C13	**Infizierte oder manipulierte Hardware**
Angriffsmittel für Folgeangriffe	
C20	Ransomware, Viren, Trojaner, Würmer, Pharming, Bots, Botnetze, Injection, Drive-By-Downloads, Defacement, Cross-Site-Request-Forgery, Replay-Angriff
Angriffsmittel für die Störung des Systems	
C30	Denial-of-Service-Angriff, DDoS-Angriff

Abbildung 1: Systematisierung von Angriffsmitteln

Anhand der Tatabsichten werden Tätertypen in drei Hauptgruppen eingeteilt. Zwischen physischen oder Cyber-Angriffen wird dabei nicht unterschieden:

- Kriminelle, Cyber-Kriminelle (K): Übergeordnete Tatabsicht ist das Streben nach Gewinn, insbesondere nach finanziellem Gewinn, kann aber auch persönlicher Gewinn ohne finanziellen Vorteil des Täters sein. Charakterisierend sind die Profitorientierung und nachfolgend die Erzeugung von Sachschaden. Personenschäden werden in der Regel nicht angestrebt, evtl. aber billigend in Kauf genommen.
- Terroristen, Cyber-Terroristen (Ter): Übergeordnete Tatabsicht ist die Verbreitung von Angst und Schrecken als Druckmittel für das Durchsetzen politischer, religiöser oder ideologischer Ziele. Charakterisierend sind eine hohe Gewaltbereitschaft und die Erzeugung maximaler Sach- und Personenschäden. Eine Profitorientierung ist dem untergeordnet.
- Staatlich geförderte Angreifer, Cyber-Angreifer und Angreifer mit staatlichem Hintergrund (StA): Übergeordnete Tatabsicht ist das Erreichen politischer, wirtschaftlicher oder militärischer Ziele. Hierzu gehören z. B. Spionage oder Sabotage bspw. für eine Destabilisierung. Charakterisierend sind die nahezu unbegrenzt zur Verfügung stehenden Mittel und Ressourcen der Angreifer.

Aus der weitergehenden Differenzierung der drei Hauptgruppen ergeben sich 14 verschiedene Tätertypen mit bestimmten Tatabsichten. Bspw. erfolgen die Unterscheidung von Einzel- oder Gruppentäter oder die Berücksichtigung besonderer Merkmale wie bspw. Innentäter oder Selbstmordattentäter. Alle Tätergruppen begehen im rechtlichen Sinn Straftaten oder rechtswidrige Taten für das Erreichen ihrer Tatabsichten.

2.3 Beschreibung von Angriffspotenzialen als Vignette

Angriffspotenziale können durch die drei Kenngrößen Tätertyp, Angriffspunkt und Angriffsmittel beschrieben werden. Die kombinierte Angabe

- eines Angreifers mit Bezug auf die Tatabsicht eines Tätertyps,
- eines für die Tatabsicht als geeignet angesehenen Angriffsmittels und
- eines für das Angriffsmittel möglichen Angriffspunkts im Eisenbahnsystem

wird als Vignette bezeichnet. Vignetten werden u. a. durch Abstraktion bzw. Generalisierung von spezifischen Gegebenheiten historischer Angriffe gebildet und generieren so eine breite Datenbasis für Analysen von Angriffspotenzialen. Wird anschließend eine Vignette wieder in ein reales Umfeld überführt und wird der Angriffspfad mit dort im-

plementierten Sicherheitsmaßnahmen beschrieben, entsteht aus der Vignette ein Angriffsszenario. In einem Angriffsszenario können wiederum auch mehrere Vignetten eingehen. Insbesondere Cyber-Angriffe, die aus einer inertialen Kompromittierung und einem weiteren Folgeangriff bestehen, können so abgebildet werden. Auch hybride Angriffe lassen sich durch mehrere Vignetten modellieren.

Ein intelligenter Täter wird immer das für seine Tatabsicht und zum ausgewählten Angriffspunkt geeignete Angriffsmittel einsetzen. Hieraus können bestimmte Kombinationen von Tätertypen und Angriffsmittel ausgeschlossen werden. Auch im Vergleich zum Eisenbahnsystem ist eine Anwendung bestimmter Angriffsmittel in bestimmten Teil- bzw. Untersystemen nicht möglich oder wird ausgeschlossen. Bspw. wird der Einsatz von mechanischen Mittel gegen Infrastruktur (P1) in bestimmten Teilbereichen eines Bahnhofs als Vandalismus angesehen und nicht als ein Angriff auf das gesamte Eisenbahnsystem verstanden.

Werden Einzel-Vignetten betrachtet, können aus den Gegenüberstellungen Tätertypen – Angriffsmitteln und Angriffsmitteln – Angriffspunkte im Eisenbahnsystem vereinfachend 19 strukturelle Teil- bzw. Untersysteme als Angriffspunkte für physische Angriffe und 18 Angriffspunkte der digitalen Systeme im Eisenbahnwesen für Cyber-Angriffe identifiziert werden. Hieraus ergeben sich 330 Vignetten für physische Angriffe und 238 Vignetten für Cyber-Angriffe. Wird der Detailierungsgrad der Abbildung – insbesondere des Eisenbahnsystems erhöht – steigt die Anzahl Vignetten schnell stark an.

3 Modellbildung und Entwicklung der REAVRS-Methode

Ausgehend von einer Vignette kann ein Angriffsszenario nach dem Sanduhr-Modell entwickelt werden. Dieses Vorgehen wird für die REAVRS-Methode übernommen (Abbildung 2).

3.1 Ganzheitlicher Risikoansatz

Aufbauend auf dem Risikomanagementprozess nach DIN ISO 31000 [9] wird ein ganzheitlicher Risikoansatz zum Schutz der kritischen Infrastruktur Eisenbahn entwickelt (Abbildung 3). Dieser Ansatz umfasst die Bedrohungs- und Vulnerabilitätsmodelle aus der Ursachenanalyse, das Schadensmodell aus der Konsequenzenanalyse sowie eine Risikobewertungsmethode, die in Kombination mit einer Strategie zum Umgang mit Ungenauigkeit und Ungewissheit auch die Ableitung von Empfehlungen zum Schutz der kritischen Infrastruktur Eisenbahn unterstützt.

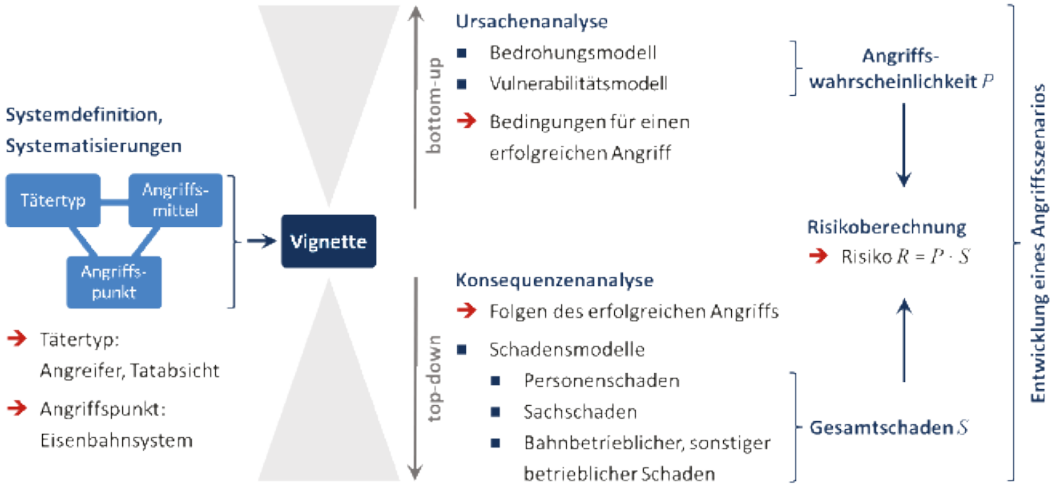


Abbildung 2: Vorgehen und Modellübersicht der REAVRS-Methode

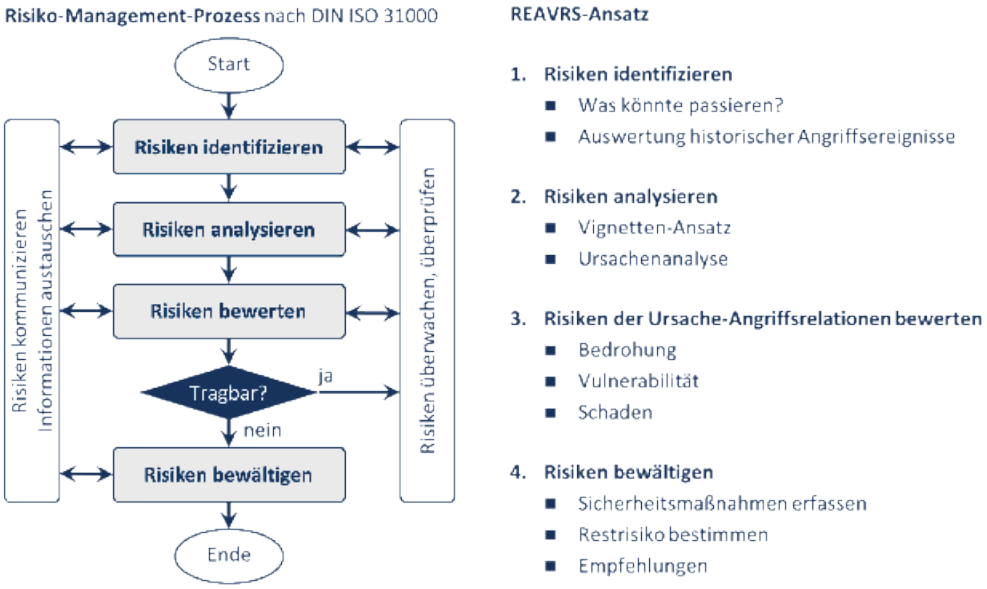


Abbildung 3: Risikomanagementprozess nach DIN ISO 31000

Vignetten stellen Angriffsereignisse mit unbekannter Angriffswahrscheinlichkeit dar und beschreiben somit ungewisse Risikosituationen. Die Herausforderung besteht darin, eine geeignete Methode zur Bewertung ungewisser Risikosituationen bereitzustellen und das Dunkelfeld ungewisser Angriffsereignisse aufzuhellen (Abbildung 4).

Ausgehend von einer Vignette werden Schätzwerte für die Einflussgrößen *Bedrohung*, *Vulnerabilität* u. a. ermittelt. Aus den Größen *Bedrohung* und *Vulnerabilität* wird die Größe *Angriffswahrscheinlichkeit* abgeleitet. Die Teilschäden werden zur Größe *Gesamtschaden* zusammengefasst und schließlich werden *Angriffswahrscheinlichkeit* und *Ge-*

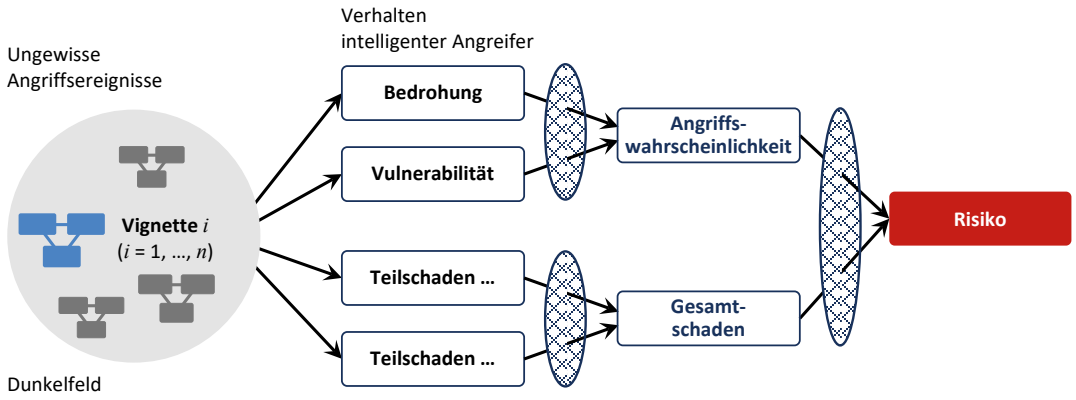


Abbildung 4: Risikobewertungsmethode bei Ungewissheit

samtschaden zur Größe *Risiko* aggregiert. Dabei werden unterschiedliche Aggregationsfunktionen genutzt. So soll bspw. die Aggregationsfunktion zur Berechnung des Risikos das Produkt von Eintrittswahrscheinlichkeit und Schaden darstellen. Da es sich hier um Größen unterschiedlicher Dimensionen (Personenschaden in Anzahl Tote und Verletzte, Sachschaden in Geldeinheiten) oder um unscharfe Größen handelt, die nur in qualitativen Kategorien (wie z. B. geringe oder hohe Bedrohung) angegeben werden können, stellt sich die Frage, wie das Produkt solch unterschiedlicher Größen bestimmt werden soll. Zur Lösung dieser Frage wird die Theorie unscharfer Mengen (engl.: fuzzy sets theory) herangezogen. Für jede der qualitativ unterschiedlichen Größen wird eine Skalierung von 1 bis 5 festgelegt, die angibt, ob es sich um einen sehr geringen, geringen, mittleren, hohen oder sehr hohen Wert handelt. Das heißt, exakte Werte werden mit Hilfe der fuzzy sets in unscharfe Werte übergeführt.

Die Bestimmung des Risikos erfolgt vereinfachend in Baumstrukturen (Abbildung 5). Ausgangspunkt ist der Knoten *Risiko*, der sich auf der Auflösungsebene 1 befindet. Alle anderen Größen werden im Rahmen eines top-down-Prozesses auf den darunter liegenden Auflösungsebenen erfasst. Die Baumstruktur repräsentiert die Auffächerung bzw. Verfeinerung der Größe *Risiko* über mehrere Auflösungsebenen hinweg.

Interpretiert man die Kanten der Baumstruktur als funktionale Beziehung zwischen zwei Größen, dann gilt:

- *Risiko* ist eine Funktion von *Angriffswahrscheinlichkeit* und *Gesamtschaden*.
- *Angriffswahrscheinlichkeit* ist eine Funktion von *Bedrohung* und *Vulnerabilität*.
- *Gesamtschaden* ist eine Funktion der Teilschäden.

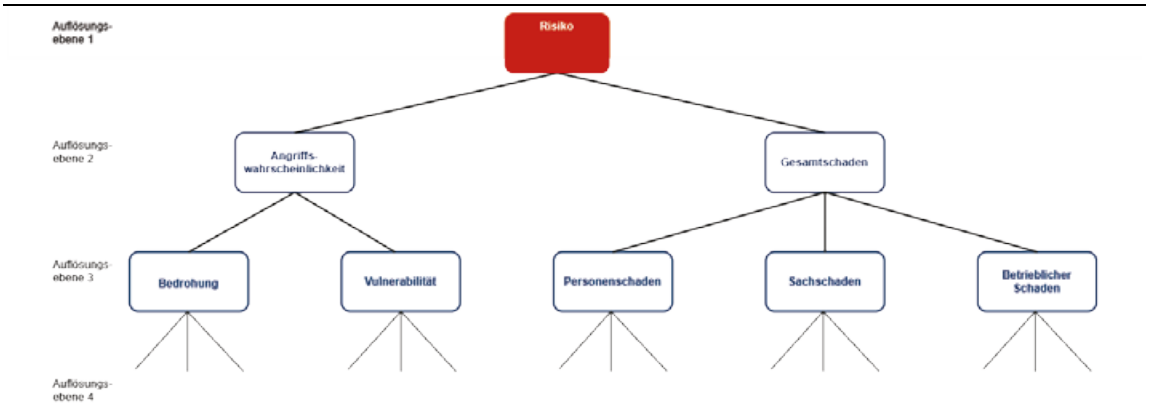


Abbildung 5: Baumstruktur zur Erfassung des Risikos eines Angriffs

Bei der vorliegenden Unschärfe der Beziehungen kann die Baumstruktur auch zur Beschreibung von Arbeitshypothesen genutzt werden wie:

- Das Risiko ist umso größer, je größer die Angriffswahrscheinlichkeit und je größer der Schaden ist.
- Die Angriffswahrscheinlichkeit ist umso größer, je größer die Bedrohung und je größer die Vulnerabilität ist.
- Der Gesamtschaden ist umso größer, je größer die Teilschäden sind.

Dieser top-down-Verfeinerungsprozess wird solange fortgesetzt bis eine Auflösungsebene erreicht ist, auf der nur noch einfache Größen anzutreffen sind. So wird bspw. angenommen, dass „Anzahl Tote“ und „Anzahl Verletzte“ den Personenschaden hinreichend genau erfassen, also nicht mehr weiter zu Leicht- und Schwerverletzten aufgelöst werden müssen. Demgegenüber stellen die Bestimmungsgrößen Bedrohung und Vulnerabilität komplexe Größen dar, deren Ursachen bzw. Einflussfaktoren noch in Bedrohungs- und Vulnerabilitätsmodell zu erfassen sind.

3.2 Konkretisierung der Vignettenbeschreibung <Ter-E/P5/Pz-ET>

Vignetten werden durch Angabe der Id in der Form <Tätertyp/Angriffsmittel/Angriffspunkt> bezeichnet. Am Beispiel der Vignette <Ter-E/P5/Pz-ET> wird die Entwicklung sowie die Anwendung der REAVRS-Methode gezeigt. Dabei wird das Angriffsszenario „Angriff mit einer Rucksackbombe“ in einem ICE in einem Eisenbahntunnel konstruiert und die Vignettenbeschreibung konkretisiert (Abbildung 6).

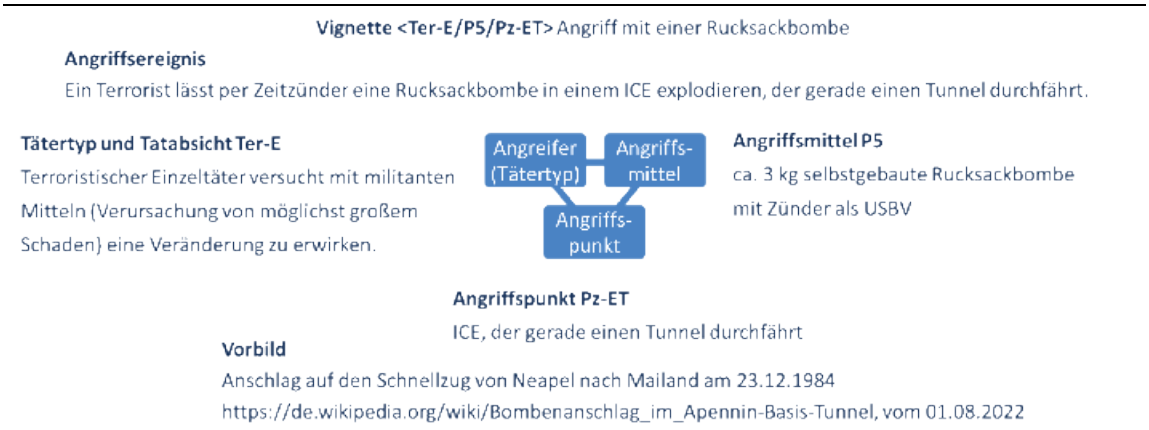


Abbildung 6: Konkretisierung der Vignettenbeschreibung <Ter-E/P5/Pz-ET>

3.1 Ursachenanalyse

Im Sanduhr-Modell bottom-up erfolgt die Ursachenanalyse, aus der die Bedingungen für einen erfolgreichen Angriff abgeleitet werden. Hierzu werden ein Bedrohungsmodell und ein Vulnerabilitätsmodell entwickelt. Grundlage sind die Kenngrößen, die anhand eines Fischgräten-Diagramm nach ISHIKAWA identifiziert werden (Abbildung 7 für Vignette <Ter-E/P5/Pz-ET>).

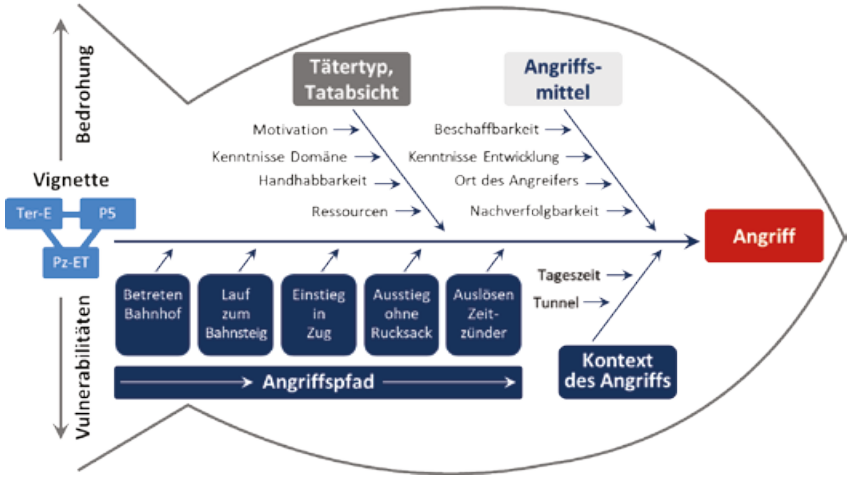


Abbildung 7: Ursachenanalyse mittels Fischgräten-Diagramm

Über der horizontalen Hauptgräte werden die Kenngrößen der Bedrohung für Tätertyp, Tatabsicht und Angriffsmittel angegeben. Unterhalb der Hauptgräte wird die Vulnerabilität anhand eines Angriffspfads abgeleitet. Im Angriffspfad werden alle chronologischen Teilereignisse oder Phasen des Angriffs und die dabei betroffenen bzw. vom Angreifer durchlaufene oder genutzte Untersysteme im Eisenbahnsystem berücksichtigt.

Für die Vignette <Ter-E/P5/Pz-ET> sind alle im Angriffspfad vom Angreifer durlaufenden Untersysteme öffentlich bzw. für Reisende zugänglich. Für Vignetten anderer Angriffe sind im Angriffspfad Zugänglichkeiten zu bestimmten Untersystemen bzw. für Cyber-Angriffe entsprechende Zonen und Conduits zu berücksichtigen.

3.2 Konsequenzenanalyse und Schadensmodell

Im Sanduhr-Modell top-down erfolgt die Konsequenzenanalyse, aus der die Folgen eines erfolgreichen Angriffs abgeleitet werden. Hierzu wird ein Schadensmodell entwickelt, in dem vier Schadensarten unterschieden werden:

- Personenschaden: Aus den Rechercheergebnissen erfolgter Angriffe und aus zur betrachteten Vignette vergleichbaren Ereignissen wird der Personenschaden abgeleitet. Dabei werden Tote und Verletzte unterschieden und ins Verhältnis zu den betroffenen Personen im Eisenbahnsystem gesetzt.
- Sachschaden: Der monetäre Schaden an strukturellen Teilsystemen kann bspw. auf Grundlage des Kostenkennwertekatalogs der Deutschen Bahn abgeleitet werden.
- Bahnbetrieblicher Schaden: Für die Abbildung von Störungen des funktionellen Teilsystems Betriebssteuerung und Verkehrsführung insbesondere durch Cyber-Angriffe werden die von einer Sperrung oder Störung eines Streckenabschnitts betroffene Anzahl Züge im Verhältnis zur Gesamtzahl bzw. Gesamtnetzleistung gesetzt. Das deutsche Eisenbahnnetz wird dafür im Netz-Evaluationsmodell ‚NEMO' [10], [11] mit einem beispielhaften Fahrplan für einen Deutschlandtakt abgebildet.
- Sonstiger betrieblicher Schaden: Schäden an den strukturellen Teilsystemen – insbesondere Infrastruktur und Fahrzeug – führen zu Störungen der funktionellen Teilsysteme oder Angriffe – insbesondere Cyber-Angriffe – führen zu direkten Störungen funktioneller Teilsysteme. Für die Modellbildung erfolgt vereinfachend eine direkte qualitative Abschätzung in den Schwergradkategorien entsprechend [5].

Für die Vignette <Ter-E/P5/Pz-ET> wird der Personenschaden gegenüber den Verhältnismäßigkeiten des Vorbilds – Anschlag auf den Schnellzug von Neapel nach Mailand am 23.12.1984 – aufgrund der konzentrierten Wirkung einer USBV in einem Tunnel erhöht.

Als konkreter Angriffsort wird ein Eisenbahntunnel auf der Schnellfahrstrecke Hannover – Würzburg, angenommen. Für die Ableitung des bahnbetrieblichen Schadens wird eine Sperrung zwischen Kassel und Göttingen in beiden Richtungen berücksichtigt sowie eine Umleitung via Eichenberg. Diese Annahmen sind Basis für die Ableitung des Gesamtschadens der Vignette <Ter-E/P5/Pz-ET> (Abbildung 8).

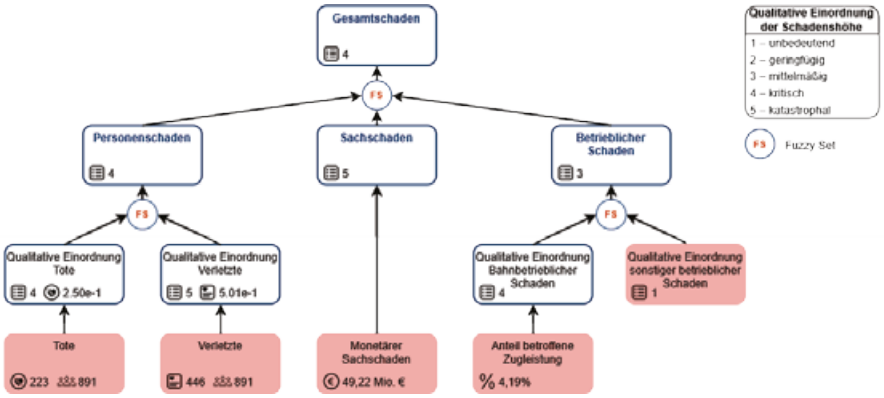


Abbildung 8: Ableitung des Gesamtschadens, Vignette <Ter-E/P5/Pz-ET>

3.3 Bedrohungsmodell

Das Bedrohungsmodell berücksichtigt die in der Ursachenanalyse identifizierten Kenngrößen. Leitfrage ist, wie die Größe *Bedrohung* von den einzelnen Kenngrößen beeinflusst wird. Dabei wird von der Grundannahme ausgegangen, dass Bedrohung als Funktion der Größen *Angriffsabsicht* und *Angriffsfähigkeit* aufgefasst werden kann. Das heißt, eine Bedrohung liegt dann vor, wenn es einen Angreifer gibt, der sowohl die Fähigkeit als auch die Absicht hat anzugreifen. Ausgehend von dieser Grundannahme werden weitere Arbeitshypothesen abgeleitet und in einem hierarchisch geordneten Bedrohungsbaum abgebildet (Abbildung 9).

Die Angriffsabsicht wird als Funktion der Motivation des Angreifers sowie der Erpressbarkeit des Angegriffenen abgebildet:

- Je motivierter der Angreifer ist, desto entschlossener verfolgt er seine Absicht.
- Je erpressbarer der Angegriffene wirkt, desto mehr fühlt sich der Angreifer in seiner Angriffsabsicht bestärkt.

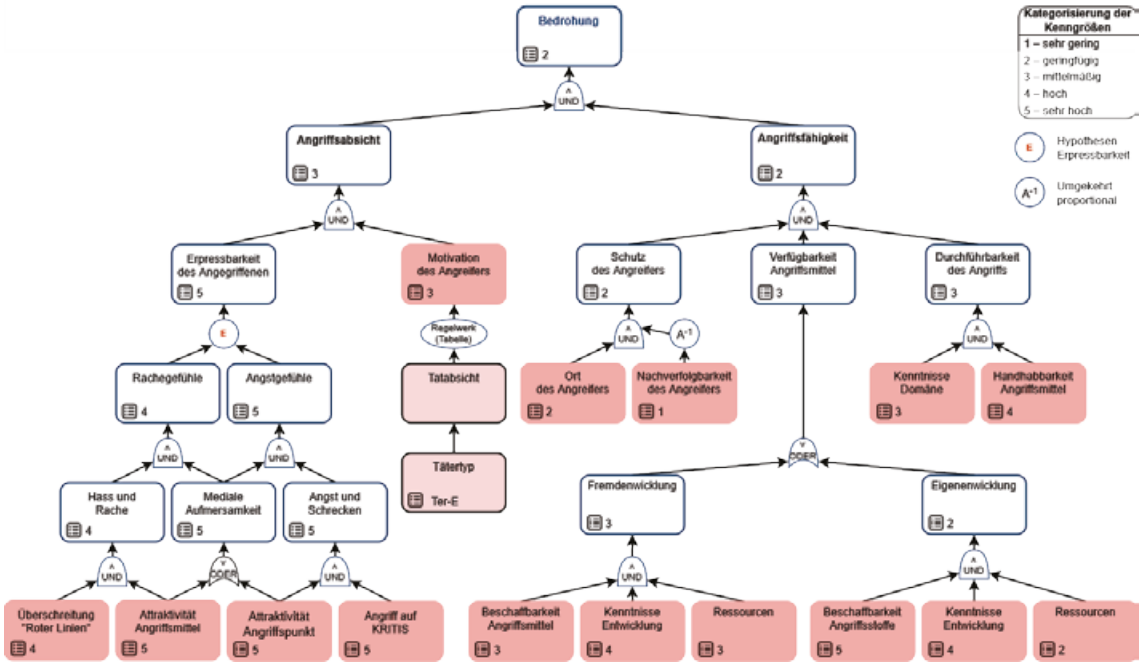


Abbildung 9: Bedrohungsbaum, Vignette <Ter-E/P5/Pz-ET>

Die Angriffsfähigkeit wird mit folgenden Voraussetzungen und Arbeitshypothesen beschrieben:

- Der Angreifer ist hinreichend geschützt.
- Der Angreifer verfügt über geeignete Angriffsmittel.
- Der Angriff ist durchführbar bzw. nicht unmöglich.

Damit ist die Angriffsfähigkeit eines Täters umso höher einzuschätzen, je größer der Eigenschutz des Täters ist, je größer die Verfügbarkeit des Angriffsmittels ist und je einfacher bzw. weniger komplex der Angriff durchgeführt werden kann.

3.4 Vulnerabilitätsmodell

Das deutsche Eisenbahnsystem ist ein offenes System. Auf Zugangskontrollen zu den Bahnhöfen, den Bahnsteigen und den Zügen wird weitgehend verzichtet. Damit ist die Konsequenz verbunden, dass das Eisenbahnsystem mit vielen Fahrgästen, komplexen Infrastruktureinrichtungen und Zügen zahlreiche, durchaus leicht erreichbare Angriffspunkte bietet, die aber das Eisenbahnsystem noch nicht vulnerabel bzw. verwundbar machen müssen. Vulnerabilität entsteht erst dann, wenn eine vom Angreifer ausnutzbare Schwachstelle existiert [12].

Das Vulnerabilitätsmodell baut auf Vignetten auf, deren Angriffspfade in Phasen eingeteilt sind. Solange keine Sicherheitsmaßnahmen entlang des Angriffspfades installiert sind, kann ein Angreifer immer wieder auf diesem Weg unkontrolliert seinen vorgesehenen Angriffspunkt erreichen und dort ungehindert den geplanten Angriff vorbereiten und durchführen. Das heißt, der ungeschützte Angriffspfad stellt eine Schwachstelle dar, die das Eisenbahnsystem verwundbar macht. Neben Schwachstellen sind auch Erfolgsaussichten von denkbaren Sicherheitsmaßnahmen zu analysieren. Dazu wird festgelegt, wo bzw. in welcher Phase welche Sicherheitsmaßnahmen eingesetzt werden sollen. Anschließend wird in einem Simulationsexperiment analysiert, wie erfolgreich die so positionierten Sicherheitsmaßnahmen den Angriffspfad überwachen. Dazu werden Ereignisse simuliert, die sich aus den Aktivitäten des Angreifers sowie der Sicherheitsmaßnahmen ergeben. Ein Entdeckungsereignis kann bspw. eintreten, wenn eine verdächtige Person oder ein gefährliches Gut auf dem Weg zum Angriffspunkt in den Sensitivitätsbereich von Sicherheitsmaßnahmen (Kameras, Sensoren, Patrouillen, Kontrollen) gelangt. Ob dann tatsächlich eine Entdeckung vorliegt oder nicht, wird über die Entdeckungswahrscheinlichkeiten der vor Ort aktivierten Sicherheitsmaßnahmen bestimmt.

Die funktionale Beziehung zwischen den Größen *Vulnerabilität*, *Abwehrerfolg*, *Entdeckung* und *Gefahrenabwehr* kann wie folgt als Hypothese formuliert werden: Je größer der Abwehrerfolg, desto geringer ist die Vulnerabilität. Das heißt: Je effizienter die Sicherheitsmaßnahmen entlang des Angriffspfades eingesetzt werden, desto erfolgreicher ist die Gefahrenabwehr oder desto geringer ist die Vulnerabilität.

Die Wahrscheinlichkeit des Abwehrerfolgs wird mit Hilfe des Tools ‚ScenSim' [13] durch eine stochastische Simulation bestimmt. Der ermittelte Schätzwert *Abwehrerfolg* wird anschließend in die entsprechende qualitative Vulnerabilitätskategorie übergeführt.

4 Beispielrechnung

4.1 Risikobestimmung

Für die Bestimmung des Risiko werden zahlreiche weitere Annahmen getroffen, auf deren Basis die qualitativen Eingangsgrößen für die *Bedrohung*, die *Vulnerabilität* und den *Gesamtschaden* bestimmt werden. Die Bestimmung der Bedrohung (Abbildung 9) und die Ableitung des Gesamtschadens (Abbildung 8) für die Vignette <Ter-E/P5/Pz-ET> wurden schon angegeben. Das Risiko für Vignette <Ter-E/P5/Pz-ET> ergibt sich mit der Kategorie 4 als hoch (Abbildung 10).

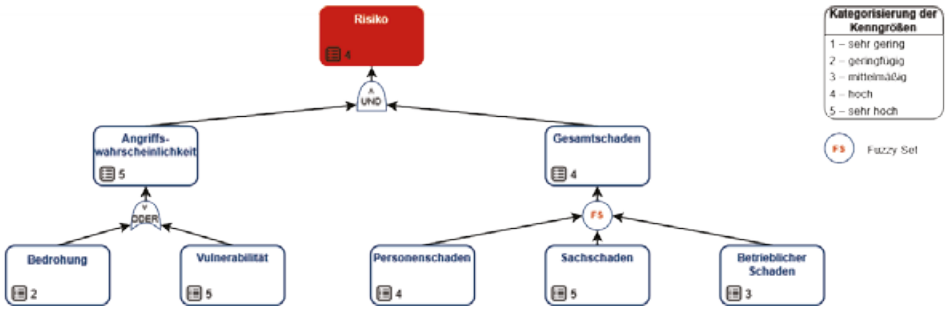


Abbildung 10: Bestimmung des Risikos, Vignette <Ter-E/P5/Pz-ET>

4.2 Risikovergleich und Wirkung von Sicherheitsmaßnahmen

Mit der REAVRS-Methode können einerseits die Wahrscheinlichkeit und die Auswirkung von Angriffen abgeschätzt werden. Andererseits können Maßnahmen zur Vermeidung oder Verringerung der Risiken berücksichtigt und ihre Wirkung durch Risikovergleich dargestellt werden.

Ein Angriffsszenario wird dazu in Phasen bzw. Teilstrecken eingeteilt, in denen geprüft wird, welche Sicherheitsmaßnahmen vor Ort bzw. in den Untersystemen verfügbar sind, welche Sicherheitsmaßnahmen den Angreifer oder das Angriffsmittel entdecken könnten und welche Ereignisse dabei ausgelöst werden könnten. Betrachtet man Sicherheitsmaßnahmen als weitere Kenngröße eines Angriffsszenarios, dann lässt sich jede Konfliktsituation als erweiterte Vignette in der Form <Angreifer/Angriffsmittel/Angriffspunkt/Sicherheitsmaßnahme> darstellen. Eine so erweiterte Vignette ist einer weiteren Risikoanalyse zu unterziehen, in der die erforderlichen Eingangsdaten abgeleitet werden.

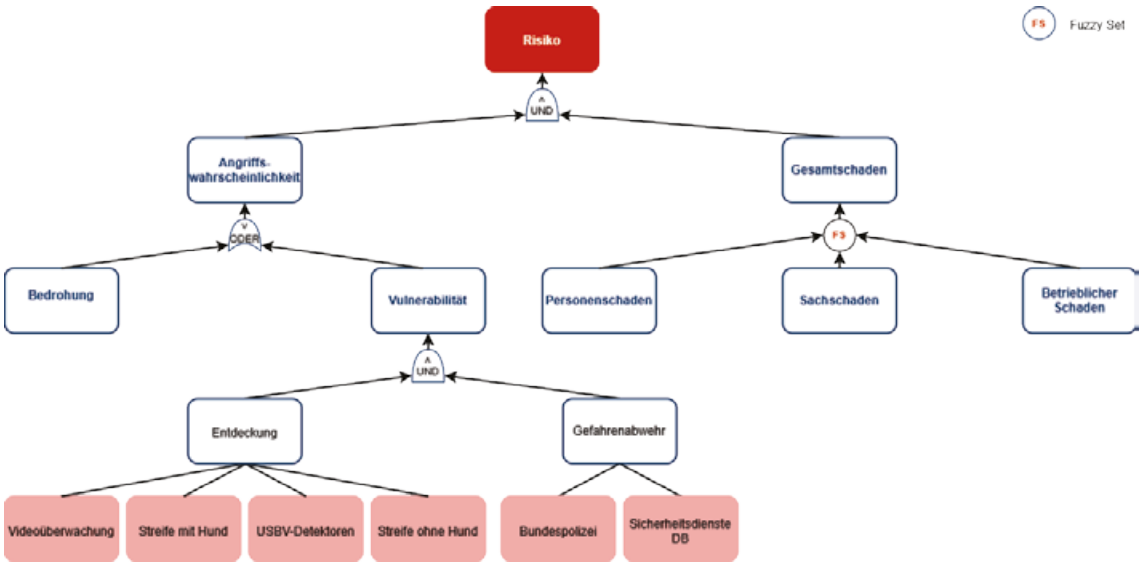


Abbildung 11: Einfluss von Sicherheitsmaßnahmen auf Vulnerabilität und Risiko

Zur Visualisierung des Einflusses von Präventionsmaßnahmen wird Vulnerabilität in die beiden Prozesse *Entdeckung* und *Gefahrenabwehr* aufgeteilt (Abbildung 11), denen wiederum je nach Verfügbarkeit und Bedarf spezielle Sicherheitsmaßnahmen zugeordnet werden. Diese Maßnahmen werden hinsichtlich ihrer Erfolgsaussichten bewertet. Voraussetzung für deren Erfolg ist aber, dass die ergriffenen Maßnahmen die Gefahr rechtzeitig entdecken.

Am Beispiel der Vignette <Ter-E/P5/Pz-ET> kann gezeigt werden, in welchem Umfang Vulnerabilität und Risiko durch Sicherheitsmaßnahmen verringert werden können. Hierzu wird eine Risikomatrix genutzt (Abbildung 12). So könnte bspw. eine lückenhafte durch eine lückenlose Videoüberwachung verbessert und eine zusätzliche Sprengstoffdetektion durch Spürhund oder Sensoren bzw. Detektoren ergänzt werden.

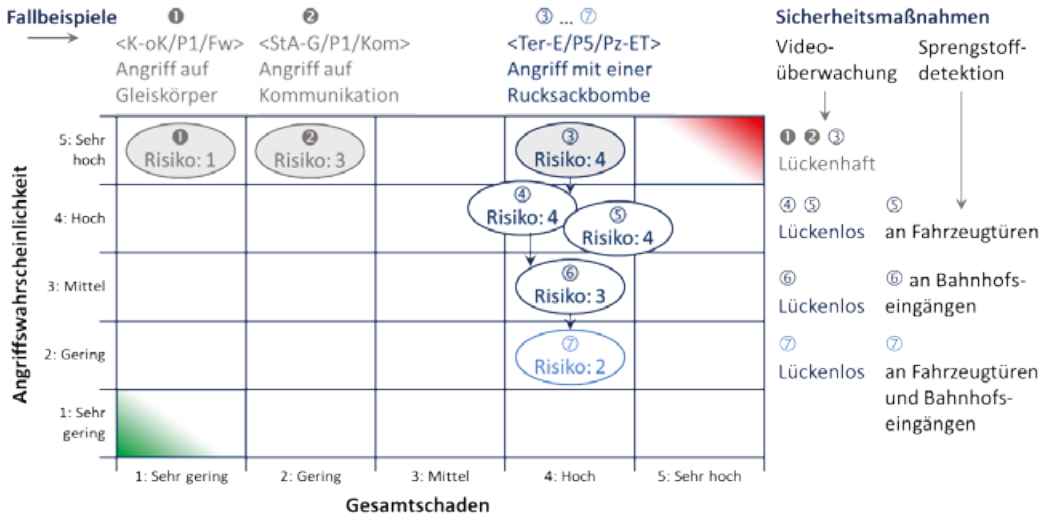


Abbildung 12: Risikovergleich für unterschiedliche Fallbeispiele und Sicherheitsmaßnahmen

Die lückenhafte Videoüberwachung führt bei allen Fallbeispielen zur Einstufung einer Angriffswahrscheinlichkeit als sehr hoch (1, 2, 3). Durch einen Ausbau von einer lückenhaften Videoüberwachung zu einer lückenlosen in den Bahnhöfen und in den Fahrzeugen kann die Angriffswahrscheinlichkeit zwar um eine Kategorie gesenkt werden. Eine Risikominderung durch Einordnung in eine geringere Kategorie gelingt aber noch nicht (4). Durch die zusätzliche Berücksichtigung einer (lückenhaften) Sprengstoffdetektion an den Fahrzeugtüren kann keine weitere Änderung von Angriffswahrscheinlichkeit, Gesamtschaden und Risiko erreicht werden (5). Eine Minderung des Risikos ergibt sich aber bei einer zusätzlichen Sprengstoffdetektion an den Bahnhofseingängen (6). Eine vollständige

lückenlose Videoüberwachung in Kombination mit Sprengstoffdetektoren an den Fahrzeugtüren und an den Bahnhofseingängen führt nochmals zu genau einer geringeren Kategorie der Angriffswahrscheinlichkeit und des Risikos (7).

5 Schlussfolgerungen und Empfehlungen

Auf Basis eines Fallbeispiels wird die Anwendung der REAVRS-Methode gezeigt. Durch Risikovergleiche können einerseits zielführende Sicherheitsmaßnahmen angegeben werden. Andererseits werden hieraus allgemeine Empfehlungen für die Risikobehandlung bei Angriffspotentialen abgeleitet. In der Gesamtheit bildet die REAVRS-Methode einen Managementprozess ab – insbesondere für ein Security-Management im Eisenbahnwesen auf hoher Systemebene und für die Begegnung jeglicher Angriffe.

Die Schlussfolgerungen basieren ausschließlich auf den im DZSF-Projekt beispielhaft untersuchten Vignetten und durchgeführten Variationsrechnungen. Um allgemeingültige Aussagen machen zu können, müssten weitere Vignetten untersucht werden. Dennoch können auf der Basis der gewonnenen Erkenntnisse erste Antworten auf die nachfolgenden allgemeingültigen Fragen gegeben werden:

Haben Cyber-Angriffe oder physische Angriffe auf das Eisenbahnsystem ein höheres Risiko?

- Grundsätzlich können mit der REAVRS-Methode konkrete Vignetten oder Angriffsszenarien physischer Angriffe wie auch von Cyber-Angriffen untersucht werden. Allerdings kann auch festgestellt werden, dass attraktive Angriffspunkte für physische Angriffe in einem offenen System vorhanden sind. Geeignete Angriffsmittel dafür sind leicht verfügbar. So können aus physischen Angriffen insbesondere hohe Personenschäden folgen.
- Für Cyber-Angriffe sind entsprechende IT-Kenntnisse erforderlich. Andererseits können Angriffe von entfernten Orten erfolgen. Cyber-Angriffe auf digitale Systeme können zu einem hohen bahnbetrieblichen Schaden oder zu wirtschaftlichen Schäden vor allem der Eisenbahnverkehrsunternehmen führen. Das gilt insbesondere für IT-Systeme auf die viele Personen zugreifen können. Hierzu zählen nicht nur Informations- und Vertriebskanäle für Reisende, sondern auch Dispositionssysteme für Personale. Für Personenschäden müssen Cyber-Angriffe die in der Leit- und Sicherungstechnik immanente Sicherungslogik – die auch in den digitalen Systemen abgebildet ist – umgehen. Dazu sind Domäne-Kenntnisse erforderlich. Der für einen Angreifer erforderliche Aufwand wird im Bedrohungsmo-

dell bewertet. Durch die Sicherungslogik werden Angriffe erschwert, da kompromittierte Systemkomponenten aufgrund der im Eisenbahnsystem üblichen Absicherungen gegen Fehler durch andere Komponenten erkannt oder ausgeglichen werden. Es liegt ein Fall von *safety-through-security* vor.

In welche Bereiche sollte vorrangig investiert werden?

- Verhaltenserkennung (z. B. durch Videoüberwachung) und Detektion von Angriffsmitteln (z. B. durch Gepäckkontrolle) sollten nicht gegeneinander ausgespielt werden. In einem durch eine große Anzahl unterschiedlicher Vignetten gebildeten Szenario ist nur ein Netzwerk sich gegenseitig ergänzender Sicherheitsmaßnahmen wirksam.

Wo sollten Kontrollen durchgeführt werden?

- Generell sollten potentielle Angreifer schon an den Zugängen zum Eisenbahnsystem identifiziert und neutralisiert werden. Der Verbund der Sicherheitssysteme zum Schutz des Eisenbahnsystems gegen Angriffe beginnt deshalb spätestens an den Systemzugangspunkten und wirkt im Verbund sich ergänzender Sicherheitsmaßnahmen im gesamten System. Da es keine hundertprozentige Sicherheit gibt, haben Kontrollen hinter dem Peripherieschutz (z. B. auf dem Bahnsteig oder beim Einstieg in den Zug) ihre Berechtigung.

Peripherieschutz oder Objektschutz?

- Der Überwachung aller Zugänge des Eisenbahnsystems an den Systemzugangspunkten kommt eine große Bedeutung zu. Dabei stellt sich die Frage nach einer Überwachung entweder am Bahnhofseingang oder an den Zug-Türen in dieser Absolutheit nicht. Da bei allen Sicherheitsmaßnahmen ein Restrisiko besteht, ist ein Systemverbund sich ergänzender, aber dadurch auch redundanter Maßnahmen denkbar. Die Analyse einer Vielzahl unterschiedlicher Vignetten, die in ihrer Gesamtheit die Bedrohungslage aufspannen, bestimmt letztlich die Dimension des Sicherheitsverbundes, in dem technische Maßnahmen und taktisches Vorgehen zusammenspielen.

Einzelmaßnahmen oder im Verbund?

- Um das Risiko erkennbar zu verringern, müssen vor allem die Angriffsmittel mit ihren Bestandteilen (z. B. fester Sprengstoff, Flüssigsprengstoff, Eisengegenstände wie z. B. Messer) identifiziert und entdeckt werden. Das heißt, die darauf

ansprechenden Detektoren müssen in den Mix von Sicherheitsmaßnahmen aufgenommen werden. Dabei sollten Schwächen einzelner Sicherheitsmaßnahmen durch Stärken anderer Maßnahmen kompensiert werden. Dazu können bspw. bauliche Maßnahmen getroffen werden, die gewährleisten, dass die Angreifer und Angriffsmittel nahe genug an die Detektoren herangeführt werden.

Literatur

[1] Pickl, Stefan; Buchetmann, Brigitte; Alonso Villota, Marina; Loreth, Teresa; Jelinski, Matthias; Schmitz, Walter; Neubecker, Karl Adolf; Siefer, Thomas; Stute, Friedrich: *Identifikation bestehender Angriffspotentiale für das System Bahn. REAVRS - Revealing Existing Attack Vulnerabilities in the Rail System.* Deutsches Zentrum für Schienenverkehrsforschung (DZSF), Schlussbericht, 18.08.2024

[2] Technische Richtlinie TR-03147: *Vertrauensniveaubewertung von Verfahren zur Identitätsprüfung natürlicher Personen.* Version 1.0.6 vom 28.12.2021, Bundesamt für Sicherheit in der Informationstechnik, https://www.bsi.bund.de/dok/TR-03147

[3] Richtlinie (EU) 2016/797 des Europäischen Parlaments und des Rates vom 16. Mai 2016 über die Interoperabilität des Eisenbahnsystems in der Europäischen Union. (Neufassung, Interoperabilitätsrichtlinie), Abl. L 138 vom 26.05.2016, S. 44, Konsolidierter Text: Abl. L 165, S. 27 vom 28.05.2020, ELI: http://data.europa.eu/eli/dir/2016/797/2020-05-28

[4] Schildt, Holger; Alberts, Katrin; Förster, Stefanie; Hoffmann, Brigitte; Nißing, Fabian; Welticke, Jessika: *IT-Grundschutz-Kompendium.* Edition 2022, Bundesamt für Sicherheit in der Informationstechnik, Bonn https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/IT_Grundschutz_Kompendium_Edition2022.pdf?__blob=publicationFile&v=3#download=1 vom 12.10.2022

[5] DIN EN 50126; VDE 0115-103: *Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS).* Teil 1: Generischer RAMS Prozess. Ausgabe Oktober 2018, deutsche Fassung EN 50 126-1: 2017, Teil 2: Systembezogene Sicherheitsmethodik. Ausgabe Oktober 2018, deutsche Fassung EN 50 126-2:2017

[6] Iffländer, Lukas; Loreth, Teresa; Alonso Villota, Marina; Pickl, Stefan; Schmitz, Walter; Neubecker, Karl Adolf: *Physical Attacks on the Railway System.* Congress in Computer Science, Computer Engineering & Applied Computing (CSCE), Las Vegas, NV, USA, 2023, doi: 10.1109/CSCE60160.2023.00391

[7] Iffländer, Lukas; Buder, Thomas: *Physische Angriffe auf das Bahnsystem – ein historischer Rückblick*. ETR - Eisenbahntechnische Rundschau, 2023, NR. 7+8

[8] Iffländer, Lukas; Buder, Thomas, Buchetmann, Brigitte: *Cyberangriffe auf das Bahnsystem – Was bisher geschah*. EI-Der Eisenbahningenieur, August 2023

[9] DIN ISO 31000: *Risikomanagement - Leitlinien (ISO 31000:2018)*. Ausgabedatum 2018-10

[10] Sewcyk, Bernd: *Makroskopische Abbildung des Eisenbahnbetriebs in Modellen zur langfristigen Infrastrukturplanung.* Wissenschaftliche Arbeit Nr. 61 des Instituts für Verkehrswesen, Eisbahnbau und -betrieb der Universität Hannover, Eurailpress Tezlaff-Hestra, Hamburg

[11] Nemo: *Langfristplanung der Schieneninfrastruktur*. IVE mbH, Hannover, https://ivembh.de/softwareprodukte/infrastruktur/nemo/ vom 05.07.2025

[12] IEC 62443-3-3:2019 + AC:2019: *Industrial communication networks - Network and system security*. Part 3-3: System security requirements and security levels. Publication date 2020-01

[13] Scenario Simulation ScenSim: *Eine stochastische Ereignissimulation zur Erfassung der Verwundbarkeit eines offenen Systems*. CreaLab GmbH, Feldkirchen-Westerham, https://www.crealab-gmbh.de/risk/unsere-risikomanagement-tools/#scensim vom 05.07.2025

Session 12
Matthias Jelinski, Walter Schmitz, Stefan Pickl

Autoren-Daten



Dr.-Ing. Matthias Jelinski

Matthias Jelinski hat Bauingenieurwesen mit der Vertiefung Eisenbahnbau und -betrieb an der Universität Hannover studiert und dort auch promoviert. Er war wissenschaftlicher Mitarbeiter am Institut für Verkehrswesen, Eisenbahnbau und -betrieb (IVE) und ist heute Senior Consulter bei der Ingenieurgesellschaft für Verkehrs- und Eisenbahnwesen (IVE mbH) in Hannover. Hier verantwortet er Sicherheitsuntersuchungen sowie die Bearbeitung von Forschungsaufträgen.

E-Mail: matthias.jelinski@ivembh.de



Dipl.-Math. Dipl.-Phys. Walter Schmitz

Walter Schmitz hat Mathematik und Physik an der Universität des Saarlandes studiert. Bis 2008 war er Ressortleiter strategische Planung zu Konfliktanalyse, Strukturanalyse, Logistik und LCC bei der Industrieanlagen-Betriebsgesellschaft mbH (IABG), Ottobrunn. Seit 2008 ist er als selbständiger Berater für den Schutz kritischer Infrastrukturen, Risiko- und Krisenmanagement für die CreaLab GmbH, tätig.

E-Mail: w.schmitz@crealab-gmbh.de



Prof. Dr. Stefan Pickl

Stefan Wolfgang Pickl studierte Mathematik, Theoretische Elektrotechnik und Philosophie an der TU Darmstadt (Diplom 1993, ERASMUS-Stipendiat an der EPFL Lausanne), 1998 Promotion (TU Darmstadt), 2004 Ruf auf die C4-Professur für Operations Research an der Universität der Bundeswehr München und 2004/05 Habilitation (Universität zu Köln)

E-Mail: stefan.pickl@unibw.de

Session 12
Shota Miyoshi, Makoto Yagi, Daisuke Yamaguchi, Tomonori Hasegawa

Development and field experiment of the safety assisting system for passive level crossings

Miyoshi, Shota[1], Yagi, Makoto[1], Yamaguchi, Daisuke[1] and Hasegawa, Tomonori[1]

[1]National Traffic Safety and Environment Laboratory, Japan

Abstract

This study develops a level crossing safety assisting system to improve the safety of passive level crossings by assisting the safety-oriented actions of both level crossing users and train drivers as a provisional measure until full safety measures such as crossing barriers are implemented. The system consists of bidirectional communication between the level crossing and the train. The ground device informs the level crossing users of the approaching train, and the onboard device informs the train driver of the operating status of the ground device. This paper reports the results of field experiments to verify the communication and assisting information notification functions.

Keywords: Level crossing; Safety; Safety assisting system; General-purpose technologies

1 Introduction

Ensuring safety is crucial for the rail transport. Level crossings are points where rail and road traffic intersect, making it particularly challenging to ensure safety. Among level crossings, the passive level crossings, which has neither warning devices nor barriers, has a particularly high potential for accidents, and thus the improvement of safety is required[11]. Level crossing protection devices are expensive with their installation and maintenance,

because of the high reliability required to ensure that they warn level crossing users of the approach of the train. Therefore, it is difficult to install such measures in rural railways in a severe economic condition, and many passive level crossings remain in these areas. As a result, there is a demand for low-cost and effective safety improvement measures for the passive level crossings.

As low-cost safety improvement measures not relying on active devices for train detection, the coloring of pathways, installation of bypass fence[3], installation of more visible warning signs[8] and warnings through sound or displays based on the detection of level crossing users and road vehicles are being implemented not only in Japan but also world-wide [7, 9]. Even more effective are the implementation of simple manual gates to pause the movement of level crossing users [10]. These methods are effective in terms of focusing the attention of level crossing users at the level crossing[4]. However, at passive level crossings with limited sight and sound of approaching trains, an active system for train detection is required to notify level crossing users of the approach of a train[11].

As precedents of low-cost systems that inform level crossing users of the approach of the train, there exist a system that detects trains approaching level crossings by installing axle sensors and radars on the ground[12], and a system that transmits the position of the train to the ground by detecting the train's position by onboard Global Navigation Satellite System (GNSS) and one-way communication using Wireless Fidelity (Wi-Fi) from the train to the level crossing[6]. These systems also use Internet of Things (IoT) devices as processing units. These precedents aim to ensure the safety of level crossings as protection devices at a low cost using such general-purpose technologies. However, the use of general-purpose technologies has the higher possibility of equipment failure and lacks fail-safe characteristics. It leads to challenges in the reliability of the device's operation and its transition to a safe state. This makes it difficult to ensure the safety of level crossings by means of devices with using general-purpose technologies.

For this reason, the authors develop the device as assisting the safety-oriented actions of those involved at the level crossing rather than the protection device. This may reduce the reliability requirements for the device's operation and could allow for a low-cost configuration that does not require safety certifications like the Safety Integrity Level (SIL).

The safety assisting system aims to assist the safety-oriented actions of both level crossing

users and railways at level crossings, and thus propose a low-cost safety assisting system utilizing general technologies based on train position detection using GNSS and bidirectional communication. The safety assisting system not only notifies level crossing users of an approaching train by transmitting the train position from the train to the level crossing, but also includes a function to provide feedback on the system's operational status to the train driver through communication from the level crossing to the train. This enables the train driver to take safety-oriented actions based on the operational status of the devices. By implementing assist for respective safety-oriented actions of the level crossing users and the train driver, the safety assisting system is considered to be capable of achieving a certain degree of safety at a low cost.

This paper reports the concept, specification and field experiment of the safety assisting system for the level crossing users and the train driver.

2 Situation analysis of level crossings and accidents on the level crossings

In Japan, level crossings are classified into three types as shown in Table 1, based on the presence or absence of equipment such as barriers and warning devices[2, 5]. Table 1 indicates that there are many passive level crossing that have neither barriers nor warning devices especially in regional railways consisting of small to medium-sized private railways and tramways. A comparison of the installation of level crossings with Germany shows that the total number of level crossing and the number of level crossing per railway line length is higher in Japan, while the total number of passive level crossing is lower in Japan.

In terms of level crossing accidents, Table 2(a) shows that Japan has a low number of accidents per level crossing while a high number of accidental causalities per level crossing [1, 2, 5]. Furthermore, Table 2(b) shows the difference of the incidence of accidents by level crossing type. In Japan, the incidence of accidents at passive level crossings is higher than at other types of level crossing. In Germany, on the other hand, the incidence of significant accidents is almost equal at active level crossing with barriers and passive level crossings.

In Japan, this situation of level crossing accidents has led to the basic policy for safety measures of abolishing all passive crossings or installing the full protection devices consisting

of barriers and warning devices. Under this policy, the number of passive level crossings has been reduced over the years. However, as the abolition of level crossings progresses, it has become increasingly difficult to coordinate with local community regarding further abolition of level crossings due to problems such as longer alternative routes bring greater inconvenience for level crossing users. The installation of protection devices requires not only the high initial cost, but also running costs and continuous effort to maintain the functionality of the protection devices of the active level crossings.

Table 1: Number of level crossings installed by type as of 2023[2, 5].

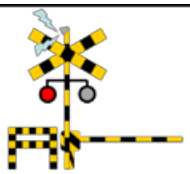
	 Active level crossing with barriers (Automatic or manual)	 Active level crossing with only automatic warning	 Passive level crossing	Total
All of Japan	29,422 (91%)	582 (2%)	2,367 (7%)	32,371
Small private railways and tramways of Japan	6,293 (82%)	211 (3%)	1,192 (15%)	7,696
All of Germany (except tramways)	9,662 (68%)	529 (4%)	3,841 (28%)	13,503

Table 2: The situation of level crossings and level crossing accidents in 2023[1, 2, 5].

(a) Overall situation at the level crossings.

	Total length of railways [km]	Total number of level crossings	Total number of accidents on level crossings	Average number accidents per crossing	Total number of casualties on level crossings	Average number of casualties per crossing
Japan *	27,733	32,371	257	7.9×10^{-3}	164	5.1×10^{-3}
Germany	39,892	13,503	158	11.7×10^{-3}	49	3.7×10^{-3}

* All numbers include the tramways defined in the Act on Rail Tracks (Act No. 76 of 1921) with total length 525 km and 448 crossings.

(b) Average number of accidents per level crossings by level crossing type.

	Active level crossing with barriers (Automatic or manual)	Active level crossing with only automatic warning	Passive level crossing
Japan	7.7×10^{-3}	3.4×10^{-3}	11.0×10^{-3}
Germany*	3.4×10^{-3}	13.2×10^{-3}	3.4×10^{-3}

* Only significant accidents included (total 51 significant accidents out of 158 accidents).

Regarding level crossing accidents in Japan, 51% of accidents are caused by level crossing users entering immediately before the train passes at the level crossing[5]. Additionally, the incidence of such accidents is lower for active level crossings with barriers and active level crossings with only automatic warning, compared to passive level crossing. This indicates that the warning device has a preventive effect on accidents caused by level crossing users entering immediately before the train passes at the level crossing[7] . Therefore, a device that warns level crossing users of the approach of the train can be said to be effective in improving the safety of passive level crossings.

3 Development of the safety assisting system for passive level crossings

3.1 Concept of the safety assisting system for passive level crossings

Based on the situation analysis stated in section 2, the authors propose a safety assisting system for passive level crossings that assists both level crossing users and train drivers by providing assisting information to improve the safety of passive level crossings.

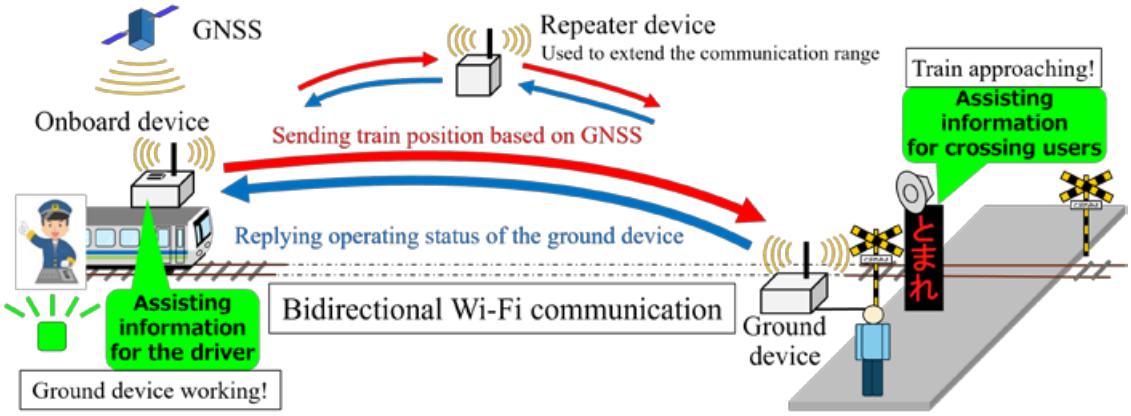


Figure 1: Overview of the configuration and operation of the safety assisting system for passive level crossings.

Figure 1 shows the configuration of the safety assisting system. It uses bidirectional wireless communication to transmit the train position information from the onboard device to the ground device installed at the level crossing. The ground device informs level crossing users of the approaching train when a train is approaching the crossing.

Additionally, the ground device replies to the onboard device to confirm that the informing action by the ground device is functioning. The onboard device informs the train driver whether the ground device functions properly based on the reply. This bidirectional notification action enables to improve level crossing safety by not only assisting level crossing users at level crossings but also train drivers, thus integrating safety checks by level crossing users, train drivers' driving behavior when passing level crossings and the safety assisting system.

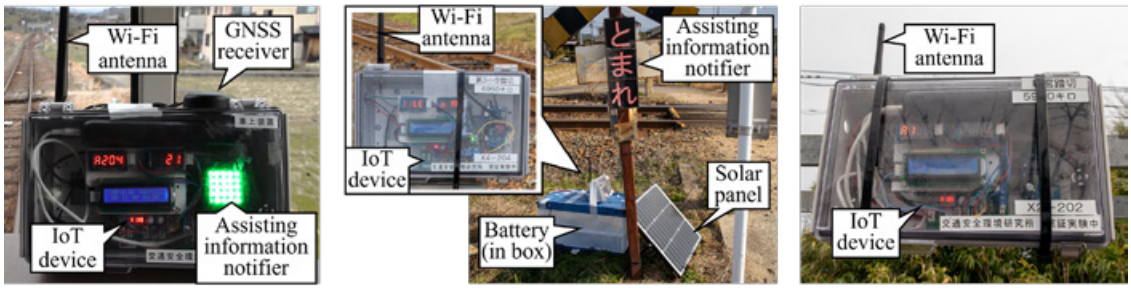
The safety assisting system should be positioned as being different from railway crossing protection devices, and it assumes the possibility of malfunction or non-operation. This assumption allows the safety assisting system to be low-cost with using general-purpose technologies and will contribute to the early realization of improved level crossing safety.

Passive level crossings originally have a risk that level crossings with limited sight and sound of approaching trains are recognized as a defect of the crossing legally. The installation of the safety assisting system at a passive level crossing is expected to reduce the risk of accidents at the level crossing in most situations where it operates normally. However, there is a risk that accidents may occur anew in cases of malfunction or non-operation of the safety assisting system, and that the railway operator will be sued regarding the inadequate maintenance of the devices. For the operation of the safety assisting system, it is of course necessary to reduce the risk of level crossing accidents occurring at the level crossing. Moreover, it is also required that the safety assisting system does not increase the risk of railway operators being accused of inadequate maintenance of their devices. For this purpose, it may also be necessary to share the concept of the safety assisting system among railway operators and potential level crossing users in the region.

3.2 Technical specification of the safety assisting system for passive level crossings

The design specification of the safety assisting system consists of general-purpose technologies such as GNSS, IoT devices, and Wi-Fi. Figure 2(a) shows the onboard device. It obtains information about the current position and speed of the train by using the GNSS receiver and transmits to the ground device and repeater devices via Wi-Fi communication as shown in Figure 1. In addition to this, if the train is approaching a upcoming level

crossing but the operating status of the ground device at that level crossing is not known, the onboard device inform the train driver of a warning about the failure of the ground device. In the prototype device shown in Figure 2(a), the "Assisting information notifier" indicates the operating status of the ground device.



(a) Onboard device (b) Ground device (c) Repeater device

Figure 2: Devices constituting the safety assisting system for passive level crossings.

Figure 2(b) shows the ground device. It informs level crossing users of the approach of the train by blinking light and sound using information about the approaching train received from the onboard device or repeater device when the train approaches to the level crossing. It also returns its operational status to the onboard device as shown in Figure 1. In addition, regarding the securing of power supply, it is assumed that supplying power from external sources is difficult in areas where passive level crossings exist. Therefore, the authors aim to use solar panels and batteries to operate the ground device through daytime charging. The output power of the solar panels and the energy capacity of the batteries need to be determined based on factors such as the energy consumption per operation, the number of operations of the ground device per day, and the percentage of sunshine.

Figure 2(c) shows the repeater device. It is installed on the ground as required when the radio line-of-sight distance around the level crossing is short. It retransmits the information received from the onboard and the ground device to nearby devices as shown in Figure 1. This is designed to clear communication failure caused by long distance or obstacles in Wi-Fi communication.

4 Field experiment

4.1 Evaluation items of the field experiment and composition of the experimental system

Field experiment was conducted to confirm the operation of the wireless communication and assisting information notification functions of the safety assisting system. In this experiment, the following items were evaluated in detail concerning the possibility of information transmission via Wi-Fi communication and the notification of assisting information in conjunction with the approach of trains at level crossings:

- The established communication range between the ground device, repeater device and onboard device in the field experiment
- The start position of notifying the assisting information

The self-supply of the power of the ground device through solar panels and batteries was not tested, as the amount of power generated by the solar panels in cloudy weather was lower than expected and insufficient at the time of the field experiment.

Figure 3 shows the setup of the devices for the field experiment. The ground device was placed at the level crossing at the position of the red point in Figure 3. The repeater devices were placed at three positions of white points in the direction towards Station B from the crossing to verify the feasibility of communication with using the repeater devices. The route from the level crossing towards Station A and to repeater device 2 towards Station B is straight. In contrast, the route between repeater devices 2 and 3 is a curve with bushes along the route, causing the reduced radio line-of-sight distance.

4.2 Results of the experiment

4.2.1 The established communication range between the ground device, repeater device and onboard device

Figure 4 shows the established communication range between the ground device, repeater device and onboard device as a result of the field experiment.

Regarding the extension in communication range by the repeater device, communication

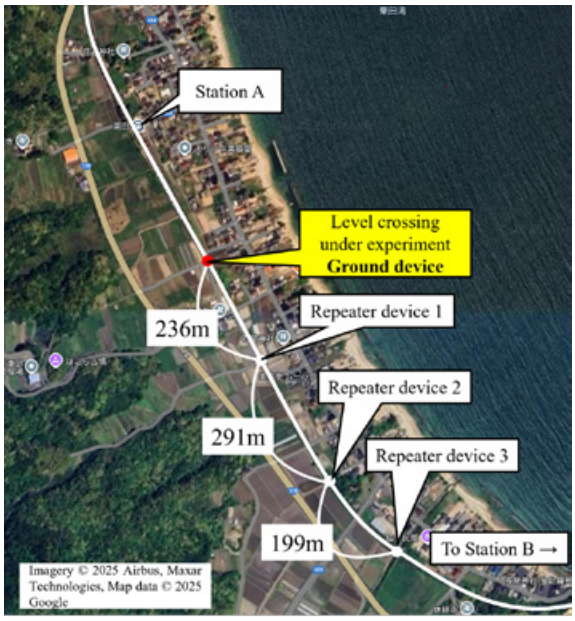


Figure 3: Configuration of the position of the ground and repeater devices in field experiment.

was successfully established between the ground device and the onboard device through repeater devices 1 and 2, within a range of 699 m from the level crossing. However, due to the installation position in curved section with bushes along the curve, no communication was established between repeater devices 2 and 3, and thus the effect of the repeater device 3 in extending communication range could not be achieved. Nevertheless, since communication was successfully established between repeater devices 1 and 2, it can be said that the function itself of the repeater devices have been implemented without problems.

From the above, it is verified that the safety assisting system can accommodate any route shapes through the proper installation of the repeater devices. It is a practical consideration to adjust the installation positions to ensure the communication between the repeater devices unlike the repeater device 3 of the field experiment.

4.2.2 The start position of notifying the assisting information

The start position of notifying the assisting information were set to 300 m before the level crossing for trains approaching from Station A, and 600 m before the level crossing for trains approaching from Station B. These values were set by taking into account the train's

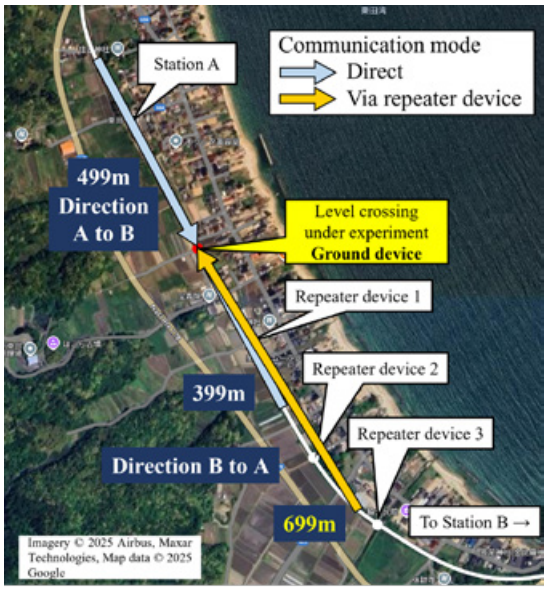


Figure 4: Distance of communication established between the ground device, repeater device and onboard device in the field experiment.

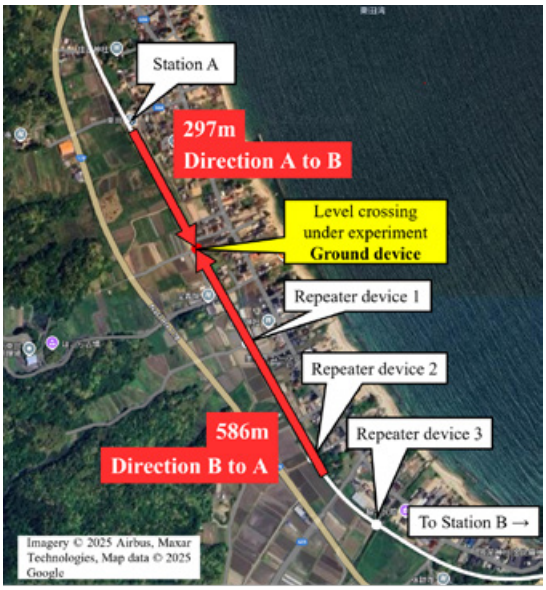


Figure 5: The start position of notifying the information in the field experiment.

operating speed and the requirement for the assisting information notification starting at least 30 seconds before the train passes at the level crossing.

The field experiments were conducted four times for each direction. Averaging the test results, as shown in Figure 5, for trains approaching from Station A, the assisting information notification started 297 m before the crossing, while for trains approaching from Station B, it started 586 m before the crossing. The variation in values was not significant. The start position of notifying the assisting information were delayed by around 10 metres compared to the set value. Considering the speed of the train, this delay can be assessed as a time delay of around one second. The time delay is considered to have been caused by train movement during satellite positioning and data transmission at the onboard device, as well as during data transmission via repeater devices.

From the above, it is verified that the safety assisting system can notify the level crossing users and the train driver at the desired timing before the train approaches to the level crossing. It is a practical consideration to adjust the timing of the assisting information notification by taking this time delay into account when setting the starting position of the assisting information notification.

5 Conclusion

This paper describes the development of the safety assisting system for passive level crossing and the confirmation of its basic operations through field experiment. As a result, it was shown that the basic operation of the system, consisting of wireless communication between the ground device and the onboard device, as well as the notification of assisting information, could be carried out without any problems.

Measures for the passive level crossing are required to be addressed promptly. The specifications of the safety assisting system, including the method of notifying the assisting information, are being continuously studied in collaboration with railway operators to ensure that it is suited to the regions to be implemented. The authors are also currently considering introducing self-sufficiency in power supply. Long-term operational on-site tests are planned for the future, and research and development will continue to the implementation of this system to the society.

Acknowledgments

We would like to thank WILLER TRAINS, Inc. and Kitakinki Tango Railway Company for providing the opportunity of field experiment and technical advices.

Literature

[1] Bundesstelle für Eisenbahnunfalluntersuchung: Jahresbericht 2023. (2023) https://www.eisenbahn-unfalluntersuchung.de/SharedDocs/Downloads/BEU/Jahresberichte/Jahresbericht_2023.pdf

[2] Eisenbahn-Bundesamt: Bericht des Eisenbahn-Bundesamts Berichtsjahr 2023. (2024) https://www.eba.bund.de/SharedDocs/Downloads/DE/Allgemeines/Sicherheitsberichte/sicherheitsbericht_2023.pdf

[3] GRABUŠIĆ, Silvestar ; BARIĆ, Danijela: Multi-criteria analysis of technical solutions for increasing level crossing safety: case study level crossing Poznanovec in Croatia. In: *Transportation Research Procedia* 64 (2022), pp. 248–256

[4] GRIPPENKOVEN, Jan: Wahrnehmung und Verhalten am Bahnübergang. In: *Deine Bahn 02/2017* (2017), pp. 10–15

[5] Ministry of Land, Infrastructure, Transport and Tourism: Information on safety in rail

transport. (FY 2023) (in Japanese) https://www.mlit.go.jp/tetudo/tetudo_fr8_000055.html

[6] MIRZA, Ghulam F. ; JUNEJO, Yahya S. ; SHAIKH, Faisal K. ; CHOWDHRY, Bhawani S. ; SHAH, Ali A.: Smart Solution for Accidents Avoidance on Railway Level Crossings. In: *International Conference on Robotics and Automation in Industry (ICRAI)* (2023)

[7] MIYOSHI, Shota: Activities on the technical verification of assisting devices and addressing social issues on the class 4 level crossings. In: *Proceedings of the Symposium of National Traffic Safety and Environment Laboratory*. (2024), pp. 59–69 (in Japanese)

[8] NGUYEN, Ngoc Anh T. ; TRUONG, Long T. ; SKARBEZ, Richard: Improving nighttime visibility and safety at passive railway level crossings: New designs incorporating photoluminescent markings and signs. In: *Traffic Injury Prevention* (2025), pp. 1–8

[9] OELSNER, Jenny ; BUDER, Jens: Analyse Optimierungspotential Bahnübergangssicherung. In: *EBA Forschungsbericht 2019-4* (2019)

[10] SUGIMOTO, Yasuhiko: Safety measures at class 4 level crossings (introduction of level crossing gates). In: *West Japan Railway vompany Technical Information Magazine* (2022), No. 46, pp. 4–6 (in Japanese)

[11] The Accident Prevention Analyst Official, Assistant to Director for Management Secreatriat, JAPAN TRANSPORT SAFETY BOARD: The JTSB Digest No. 31 (2019)

[12] WULLEMS, Chris ; BAKER, Ross ; BEH, Gregory ; UPTON, Pete ; WAYTH, Rod: Trialling Low-cost Level Crossing Warning Devices in Australia. In: *10th World Congress on Railway Research (WCRR)* (2013)

Session 12
Shota Miyoshi, Makoto Yagi, Daisuke Yamaguchi, Tomonori Hasegawa

Authors



Miyoshi, Shota
Ph.D. in Engineering
Researcher at the National Traffic Safety and Environment Laboratory
contact: s-miyoshi@ntsel.go.jp



Yagi, Makoto
Bachlor of Engineering
Visiting researcher at the National Traffic Safety and Environment Laboratory
contact: m-yagi@ntsel.go.jp



Yamaguchi, Daisuke
Ph.D. in Engineering
Principal researcher at the National Traffic Safety and Environment Laboratory
contact: d-yama@ntsel.go.jp



Hasegawa, Tomonori
Ph.D. in Engineering
Executive researcher at the National Traffic Safety and Environment Laboratory
contact: hasegawa@ntsel.go.jp

IRSA 2025

SESSION 13

INFRASTRUKTURZUSTAND I

INFRASTRUCTURE CONDITION I

Session 13
Kira Zschiesche, Johannes Pelz, Alexander Reiterer

KI-Anwendungen zur regelmäßigen Überwachung der Eisenbahninfrastruktur

Zschiesche, Kira[1], Pelz, Johannes[1], Reiterer, Alexander [1,2]

[1]Fraunhofer Institute for Physical Measurement Techniques IPM, 79110 Freiburg, Germany

[2]Department of Sustainable Systems Engineering INATECH, Albert-Ludwigs-University Freiburg, Germany

Zusammenfassung

Die Verwendung von Künstlicher Intelligenz (KI) zur Überwachung der Eisenbahninfrastruktur gewinnt zunehmend an Bedeutung. KI-Technologien, insbesondere im Bereich des Deep Learning, ermöglichen eine effiziente Analyse großer Datenmengen. Diese Technologien verbessern die Wartung und Inspektion der Bahninfrastruktur, indem sie die Erkennung von Defekten und Anomalien automatisieren. Die regelmäßige Überwachung ist notwendig, um die Auswirkungen von erhöhten Belastungen und extremen Wetterbedingungen zu bewältigen. Verschiedene Datensätze werden genutzt, um KI-Modelle zu trainieren und zu validieren. Diese Modelle sind in der Lage, verschiedene Infrastrukturkomponenten wie Gleise, Oberleitungen und Bauwerke zu analysieren. Eine Herausforderung liegt in der Datenqualität und der Notwendigkeit, repräsentative Datensätze zu erstellen. Dennoch zeigt die Automatisierung durch KI großes Potenzial, um die Effizienz und Sicherheit im Schienentransport langfristig zu verbessern. Diese Arbeit soll eine systematische Übersicht über bestehende Datensätze und Anwendungsbeispiele von KI im Bereich der Infrastrukturüberwachung geben. Die weltweiten Datensätze zeigen eine starke Heterogenität und fehlende Standards. Zugleich werden praxisnahe Lösungsansätze und damit die potenziellen Möglichkeiten von KI aufgezeigt.

Keywords: Eisenbahninfrastruktur, Deep Learning, Künstliche Intelligenz, Wartung, Inspektion, Datensätze

1 Einleitung

Der Schienentransport von Passagieren und Gütern bietet eine hohe Kapazität und Geschwindigkeiten, einen geringen Energieverbrauch und ist umweltfreundlich. Für eine moderne Wirtschaft ist ein effizientes und funktionierendes Transportsystem unerlässlich [1]. Längere Störungen oder Ausfälle können erhebliche Auswirkungen auf die Gesellschaft und die Betreiber haben. Die Infrastruktur ist hohen Belastungen ausgesetzt, die durch ein steigendes Transportvolumen, schwerere Fahrzeuge, höhere Geschwindigkeiten und extreme Wetterereignisse verursacht werden. Daher sind Wartungsmaßnahmen notwendig, um den Schienentransport effizient und störungsfrei zu gestalten. Eine regelmäßige oder kontinuierliche Überwachung der Infrastruktur spielt hierbei eine zentrale Rolle.

Um die Auswirkungen der Belastungen auf die Infrastruktur erfassen und beurteilen zu können, wurden verschiedene Messmethoden und -systeme entwickelt, darunter manuelle Systeme wie Trolleys, beispielsweise das Trimble GEDO Vorsys [2] sowie automatisierte Messzüge wie der „Doctor Yellow“ in Japan [3] und der IRIS320 in Frankreich. Diese Züge sind mit speziellen Sensoren ausgestattet, die Inspektionen bei hohen Geschwindigkeiten ermöglichen. Viele der verwendeten Sensoren erzeugen große Datenmengen, wie etwa Punktwolken durch LiDAR (Light Detection and Ranging) oder Bilder von Kamerasystemen. Da die optische Messsensorik einen wesentlichen Teil der Infrastrukturdatenerfassung im Eisenbahnbereich ausmacht, wird im Folgenden ein besonderer Fokus auf diesen Aspekt gelegt. Die Auswertung der großen Datenmengen erfolgt häufig noch manuell unter Einsatz von Experten. Eine automatisierte Analyse der Inspektionsdaten kann jedoch die benötigte Zeit und den Einsatz von Personal reduzieren. Künstliche Intelligenz (KI) ermöglicht es, die Zeit und Kosten für die Erkennung von Defekten oder Anomalien zu senken. KI-Anwendungen werden zunehmend in der Bahnindustrie eingesetzt, vor allem im Bereich des Gleissystems, bei dem sie einen Anteil von 77 % an Veröffentlichungen hat, während andere Bereiche wie Bauwerke (14 %) und Oberleitungssysteme (4 %) weniger Beachtung finden [4]. Die Voraussetzung für den Einsatz von KI, insbesondere bei Deep Learning (DL), sind großen Mengen an Daten, die für das Lernen zur Verfügung stehen müssen.

2 Künstliche Intelligenz

Künstliche Intelligenz, als Teilgebiet der Informatik, imitiert menschliche kognitive Fähigkeiten, indem sie versucht aus Eingabedaten Informationen zu gewinnen. Folgt die KI dabei programmierten Abläufen, wie Algorithmen und statistischen Methoden, um aus

den Daten zu lernen und Vorhersagen oder Entscheidungen zu treffen, sprechen wir von Maschinellem Lernen (ML). Dabei lernt der Algorithmus durch Wiederholungen eine Aufgabe selbstständig zu lösen, ohne dass ein Lösungsweg vorher definiert wird. Durch die Verfügbarkeit von großen Datenmengen und hohen Rechenleistungen sind in den vergangenen Jahren große Fortschritte gemacht worden. Die Auswahl geeigneter Trainingsdaten ist dabei entscheidend für den erfolgreichen Einsatz von KI-Anwendungen. Die KI kann nur aus den Daten lernen, welche sie zur Verfügung gestellt bekommt. Repräsentieren diese Daten nicht die Gesamtheit der Situationen oder Zustände, mit denen die KI im Einsatz konfrontiert wird, kann es zu fehlerhaften Aussagen kommen. Um dem entgegenzuwirken, wird versucht, die KI abstrakt genug zu halten, um auch neue Daten korrekt verarbeiten zu können und andererseits die KI nicht zu einfach zu modellieren, sodass die zugrunde liegenden Muster nicht mehr erfasst werden. Um die auftretenden Unsicherheiten beurteilen zu können, haben sich verschiedenen Metriken wie Accuracy, Precision, Recall oder F1-Score etabliert [5]. Besonders in sicherheitskritischen Bereichen wie zum Beispiel bei der Bewertung von systemrelevanten Infrastrukturen muss ein besonderes Augenmerk auf die Zuverlässigkeit und Robustheit der Entscheidungen gelegt werden. Zudem kommt in Zukunft der Fähigkeit der KI, ihre eigenen Schlussfolgerungen zu erklären, eine hohe Bedeutung zu – Forschungen in diesem Bereich laufen unter dem Begriff der „explainable artificial intelligence (XAI)“ [6, 7].

Es gibt vielfältige Anwendungsgebiete für den Einsatz von KI: Kategorisierung von Bildern für das autonome Fahren, Erkennen und Interpretieren von verbaler Sprache für Sprachassistenzen oder Mustererkennung, um Ereignisfolgen zu erkennen. Deep Learning (DL), eine Unterkategorie des ML, hat sich im Bereich der Bildanalyse, Videoverarbeitung und Punktwolkenanalyse hervorgetan. Dabei basieren die Algorithmen auf künstlichen neuronalen Netzen, welche aus mehreren Schichten und Knoten bestehen. Jede dieser Schichten kann komplexe Muster in den Daten erkennen. Einen allgemeinen Überblick über den Einsatz von DL im Bahnsektor geben [8, 9]. Klassische Analysewerkzeuge des DL sind dabei die Klassifikation, Objektdetektion und Segmentierung (siehe Abbildung 1). Bei der Klassifikation werden Objekte basierend auf Bilddaten in bestimmte Kategorien eingeteilt. Für jedes einzelne Bild kann eine Klasse mit einer bestimmten Wahrscheinlichkeit bestimmt werden. Bei der Detektion werden einzelne Objekte identifiziert und lokalisiert. Dabei werden in den Bildern Rechtecke, sogenannte Bounding Boxes, eingezeichnet, die zu erkennende Objekte umrahmen und ihnen die jeweilige Klasse zuordnen. Bei der Segmentierung wird jeder Datenpunkt einer bestimmten Grundklasse zugeordnet.

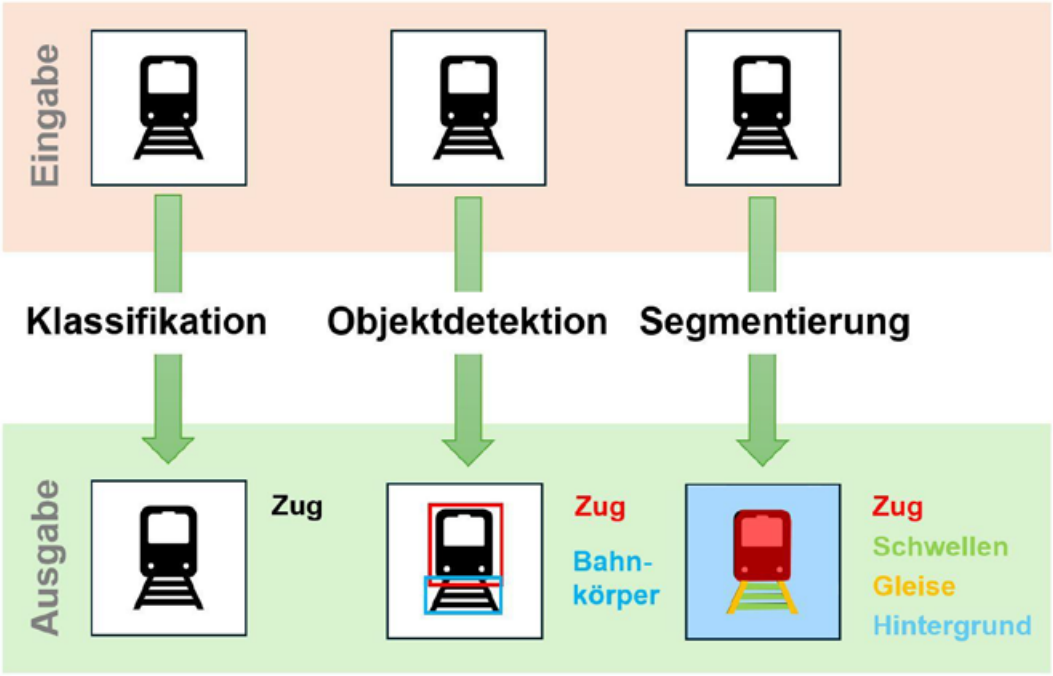


Abbildung 1 Deep Learning. Unterschieden wird zwischen der Klassifikation, der Objektdetektion und der Segmentierung.

3 KI-Anwendungen zur Überwachung der Infrastruktur

Eine regelmäßige und kontinuierliche Überwachung der Infrastruktur spielt eine wichtige Rolle, um die Funktionsfähigkeit des Schienentransports langfristig aufrechtzuerhalten. Im Rahmen der regelmäßigen Überwachung der Eisenbahninfrastruktur sind mehrere Bereiche von besonderer Bedeutung, da sie entscheidend zur Sicherheit und Effizienz des gesamten Systems beitragen. Dazu zählt insbesondere der Gleisbereich, der aus mehreren Komponenten besteht, wie etwa den Schwellen, die die Schienen stabilisieren, den Weichen, die den Zugverkehr auf unterschiedliche Gleise umleiten, sowie den Gleisen selbst, die die Hauptfahrbahn bilden. Darüber hinaus spielt der Schotter eine essenzielle Rolle, da er das Gleisbett stabilisiert und die Befestigungen sichert, die die Schienen an ihrem Platz halten. Ein weiterer wichtiger Aspekt der Infrastrukturüberwachung ist das Oberleitungssystem. Dieses umfasst die Masten, die die Oberleitungen tragen, die Isolatoren, die elektrische Trennungen gewährleisten, und die Ober- sowie Fahrleitungen, die für die Stromversorgung der elektrischen Züge verantwortlich sind. Auch die Befestigungen innerhalb dieses Systems sind entscheidend, um die Stabilität und Sicherheit der Oberleitungen zu gewährleisten. Zusätzlich sind Brücken und Tunnel bedeutende Infrastrukturkomponenten, die sowohl für die Überquerung von Hindernissen

als auch für unterirdische Durchgänge für Züge notwendig sind. Schließlich ist der Lichtraum, der den vertikalen und horizontalen Raum definiert, den ein Zug beim Fahren einnehmen darf, ebenfalls ein wichtiger Faktor. Zusätzlich sind Informationen über Signalanlagen, Bahnsteige und Lärmschutzwände weitere Komponenten, die für den Betrieb und die Sicherheit eine wichtige Rolle spielen.

Ein häufiges Anwendungsbeispiel ist die Erkennung von Hindernissen im Gleisbereich. In Bildern und Videos [10, 11], in einer Kombination von realen Bildern und synthetisch erzeugten Bildern [12] oder in Kombination von Punktwolken und Bildern [13] werden unterschiedliche Ansätze verfolgt. Auch wenn die Erkennung von Hindernissen keine Infrastruktur darstellt, sind Hindernisse doch auch immer in Bezug zur Infrastruktur zu betrachten. Die frühzeitige Erkennung von Hindernissen auf den Gleisen, wie z. B. Personen, Tiere oder Objekte, ist essenziell, um Unfälle und Kollisionen zu vermeiden, setzt aber eine Definition des Gefahrenraums voraus. Dies schützt sowohl die Fahrgäste als auch das Personal und verringert das Risiko von Sachschäden der Infrastruktur. Durch die zunehmende Automatisierung im Bahnbetrieb wird diese Anwendung immer wichtiger, auch im Hinblick auf ATO (Automatic Train Operation). Die Erkennung von Hindernissen kann dazu beitragen, den Wartungsbedarf der Infrastruktur zu bewerten, indem sie auf potenzielle Probleme hinweist, die behoben werden müssen, bevor sie zu größeren Störungen führen.

3.1 Methodik der Literaturrecherche

Die im Mai und Juni 2025 durchgeführte Recherche nach relevanten Artikeln und Veröffentlichungen fokussierte sich auf Datensätze und Anwendungen zur Infrastrukturüberwachung im Eisenbahnsektor. Die Recherche erfolgte in den Datenbanken Fraunhofer E-Lib, UB Freiburg, Google Scholar und Elicit. Es wurden ausschließlich Datensätze mit optischer Messtechnik berücksichtigt, mit Fokus auf Deep-Learning-Methoden und aussagekräftige Datensätze aus dem Zeitraum 2015 bis 2024 – unabhängig vom Herkunftsland. Die Methodik folgt einem mehrstufigen Ansatz: Zunächst wurde ein „Top-Down-Approach“ mit schlagwortbasierter Suche und bibliografischem Schneeballsystem durchgeführt. Anschließend folgte ein „Bottom-Up-Approach“ über direkte Datenbankrecherchen, um weitere oder aktuellere Publikationen zu erfassen. Ergänzend kam eine KI-gestützte Suche mit Tools wie Perplexity und Elicit zum Einsatz, basierend auf natürlichsprachlichen Eingaben. Abschließend wurden alle Publikationen nach verwendeter Sensorik kategorisiert.

3.2 Datensätze für KI-Anwendungen

Der Einsatz von KI und die damit verbundenen Datensätze gewinnen zunehmend an Bedeutung. Diese Datensätze, die aus einer Vielzahl von unterschiedlichen Quellen stammen, ermöglichen es, Muster und Anomalien in der Infrastruktur zu erkennen. Durch die Analyse dieser Daten können präzisere Vorhersagen über Wartungsbedarf oder Sicherheitsrisiken getroffen werden. Die Implementierung von KI-gestützten Lösungen auf Basis dieser Datensätze kann nicht nur die Effizienz des Schienentransports verbessern, sondern trägt auch zur Erhöhung der Sicherheit und Zuverlässigkeit bei, indem potenzielle Probleme frühzeitig identifiziert und behoben werden können. Ein zentrales Risiko ist dabei die Datenqualität. Ungenaue, unvollständige oder veraltete Daten können zu falschen Analysen und Entscheidungen führen. Nicht repräsentative Datensätze führen zu voreingenommenen Ergebnissen. Modelle, die zu stark auf spezifischen Datensätzen trainiert werden, neigen dazu, in der Praxis zu versagen, wenn sie auf unbekannten Daten angewendet werden. International betrachtet stellt die Variabilität der Infrastruktur eine Herausforderung dar. Unterschiedliche technische Standards und Bauweisen in verschiedenen Ländern führen zu Herausforderungen bei der Datenintegration und Analyse. Bahnsteige zum Beispiel unterscheiden sich stark in Höhe, Länge und Zugangssystemen. In Europa variieren die Bahnsteighöhen zwischen 30 cm und 76 cm, während Japan und Südkorea meist hohe Bahnsteige um 110 cm nutzen. Die USA zeigen große Unterschiede je nach Betreiber. Weltweit existieren über 30 verschiedene Spurweiten, die sich historisch, technisch und strategisch entwickelt haben. Die am weitesten verbreitete ist die Normalspur mit 1.435 mm, die in Europa, Nordamerika, China und Australien genutzt wird. In Russland und ehemaligen Sowjetstaaten ist die Breitspur mit 1.520 mm üblich, während Spanien und Indien lange auf 1.676 mm setzten.[14] Zudem besteht die Gefahr von Bias, wenn Datensätze aus bestimmten Regionen oder Ländern nicht repräsentativ sind. Dies kann zu verzerrten Ergebnissen bei internationalen Analysen führen. Diese Probleme erfordern eine enge Zusammenarbeit zwischen internationalen Organisationen, eine gemeinsame Entwicklung von Standards und eine sorgfältige Planung, um die Herausforderungen der Internationalität im Bereich der Eisenbahndatensätze zu bewältigen und die Vorteile eines global vernetzten Schienentransports zu realisieren. Zudem ist die Verfügbarkeit von öffentlich zugänglichen Datensätzen begrenzt. Mit ihrem Forschungsbericht „Aufbereitung von Datensätzen für die Anwendung des automatisierten Fahrens im Eisenbahnbetrieb“ [15] hat das Deutsche Zentrum für Schienenverkehrsforschung (DZSF) Annotations- und Qualitätssicherungsprozesse erstellt, welche die

Herausforderungen und Lösungsansätze aufzeigen, um die Nutzung und Qualität offener Daten zu verbessern.

Verschiedene Ansätze haben sich aufgrund der Komplexität der Lerndaten auf die Verwendung synthetischer Daten konzentriert. So generierte [16] 14.615 annotierte Bilder, um Fremdkörper wie Plastikbeutel oder Luftballons in Oberleitungen zu detektieren. Zunächst wurden Bilder von Oberleitungen gesammelt und Anomalien mit Photoshop erstellt. Dann generierte ChatGPT textuelle Beschreibungen, die in Bilder mittels Stable Diffusion (Deep-Learning-Text-zu-Bild-Generator) umgewandelt wurden. Schließlich wurden die Bilder nachbearbeitet, um die Auflösung zu verbessern. Um Trainingsdaten für die semantische Segmentierung von Punktwolken zur Erkennung von Bahnübergängen zu erhalten, haben [17] reale und synthetische Punktwolken kombiniert. Zunächst wurden reale Punktwolken von Bahnübergängen mithilfe mobiler Laserscanner erfasst. Anschließend wurden synthetische Punktwolken generiert, indem sie 3D-Modelle von Bahnübergängen und deren Umgebung in einer simulierten Umgebung erstellten. Durch die Kombination konnte die Vielfalt der dargestellten Szenen erhöht werden.

In [18, 19] wird der Datensatz RailGoerl24 vorgestellt. Er umfasst eine Sammlung von 12.205 RGB-Bildern und 384 Mio. Punkten einer 3D-LiDAR-Punktwolke für die automatische Personenerkennung im Bahnbereich. Aufgenommen wurde der Datensatz im Görlitzer Eisenbahntestzentrum, um die Entwicklung fahrerloser Züge zu unterstützen. Er umfasst 33.556 Box-Annotationen für die Objektklasse Person. Der Datensatz Open Sensor Data for Rail 2023 (OSDaR23) bietet ebenfalls eine umfassende Multisensoraufnahme der Eisenbahnumgebung. Hierbei kommt ein Schienenfahrzeug zum Einsatz, das mit verschiedenen Sensoren ausgestattet ist, darunter LiDAR, RGB-Kameras, eine Infrarotkamera sowie ein Radar. Der Datensatz umfasst 22 verschiedene Klassen und ist ebenfalls auf die Anwendung des ATO ausgelegt [20–22].

Eine Übersicht einiger Datensätze findet sich im Anhang.

3.3 Anwendungen mit Bilddaten

Bilddatensätze spielen eine zentrale Rolle in der Entwicklung und dem Training von ML-Modellen, da sie es ermöglichen, visuelle Informationen zu analysieren und Muster zu erkennen. Diese Daten stammen aus verschiedenen Quellensystemen, wie zum Beispiel spezieller Messkameras an Trägerplattformen wie Zügen oder UAVs (Unmanned Aerial Vehicle). Ein effektiver Bilddatensatz muss sorgfältig annotiert und qualitativ hochwertig sein, um die Genauigkeit der ML-Modelle zu gewährleisten. Durch die Verwendung von

Bilddatensätzen können KI-Anwendungen zur Überwachung und Wartung von Bahninfrastrukturen beitragen.

Die Erkennung von Gleisen ist eine wichtige Anwendung im Bahnsektor. Für Aufgaben wie die Erkennung von Hindernissen oder die gezielte Suche nach Schäden muss je nach Aufnahmesituation und verwendeter Sensorik zunächst das Gleis im Bild gefunden werden. Um dies zu adressieren, wurde der Datensatz Rail-DB mit 7.432 Bildpaaren und Annotationen unter unterschiedlichen Bedingungen generiert [23]. Zu den größten Herausforderungen bei der Erkennung von Gleisen gehören die Vielfalt der Beleuchtungssituationen, Schienenwechsel und die Anforderungen bei hohen Geschwindigkeiten [24]. Li & Peng zeigen mit ihrem zeilenbasierten Framework eine Segmentierungsmethode für Gleise mit einer Accuracy von 92,77 % [25]. Der Bereich des autonomen Fahrens fokussiert sich stark auf die Anwendung für Straßenfahrzeuge und lässt den Bereich des schienengebundenen Verkehrs meist außen vor. Der öffentliche Datensatz RailSem19 dagegen umfasst 8.500 annotierte Sequenzen aus der Perspektive von Zügen und Straßenbahnen, darunter über 1.000 Beispiele mit Bahnübergängen und 1.200 Straßenbahnszenen, um diese Lücke des Forschungsfeldes aktiv zu bedienen [26, 27].

Um Bilddaten nur vom Gleisbereich zu erhalten, werden Kameras oder Kamerasysteme unterhalb des Trägerfahrzeuges installiert und die Bilder während der Fahrt aufgezeichnet. So können Auffälligkeiten wie Schienenoberflächenfehler, z. B. Risse oder Riffelungen, aber auch fehlende Befestigungen detektiert werden.

Arain et al. haben einen Bilddatensatz erstellt, speziell für das Klassifizieren und Detektieren von Gleisoberflächenanomalien [28, 29]. Der Datensatz enthält dabei Bilder von Rillen, Fugen, Rissen, Splittern, Schälungen, Abplatzungen und Squats, jedoch in einer sehr ungleichen Verteilung an Bildern. RailSet befasst sich ebenfalls mit der Erkennung von Anomalien im Gleis [30]. Der Datensatz beinhaltet 6.600 manuell annotierte Bilder normaler Situationen und 1.100 Bilder von Schienenfehlern wie Lochanomalien und Gleisunterbrechungen. Um den Datensatz repräsentativ zu gestalten, wurden synthetische Bilder mittels des Deep-Learning-Algorithmus StyleMapGAN generiert. Diese Daten sollen die Entwicklung autonomer Züge unterstützen, die in der Lage sind, Gleisschäden zu erkennen. Auch [31, 32] beschäftigen sich mit der Erkennung von Anomalien im Gleis. Ein aus mehreren Robotern bestehendes System soll Oberflächenfehler wie Risse, Squats, Rillen und Rost erkennen. Das System verwendet Ultraschallsensoren in Kombination mit Bildverarbeitung mittels OpenCV und DL-Algorithmen zur Klassifizierung der erkannten Oberflächenfehler. Es wurde eine

Accuracy von 90,30 % erreicht. Um bessere Ergebnisse bei der Klassifizierung von Oberflächendefekten in Schienen zu erreichen, nutzen [33] Merkmalsextraktionen und -kombinationen. Durch Kontrastveränderungen werden Beleuchtungsprobleme auf der Schienenoberfläche behoben. Der vorgeschlagene Ansatz erreichte einen Klassifizierungserfolg von 97,10 %.

Um fehlende Klammern im Befestigungssystem automatisiert zu bestimmen, haben [34] RGB-Bilder während einer Fahrt mit 20 km/h aufgezeichnet. Sie unterscheiden zwischen einem intakten Befestigungssystem, einem, bei dem eine Klammer fehlt und einem, bei dem beide Klammern fehlen. Die Ergebnisse mit unterschiedlichen Netzwerkarchitekturen zeigten über 94 % Accuracy in der Testphase.

Bahnschwellen sind sicherheitskritische Komponenten. Aufgrund häufiger hochintensiver Belastungen durch die Interaktion zwischen Zug und Gleis sowie widriger Umgebungsbedingungen können sich z. B. verschiedene Rissmuster auf den Betonbahnschwellen entwickeln, die deren Haltbarkeit und Lebensdauer verringern. Mit der Erkennung von Defekten an Schwellen und Befestigungselementen in Bildern beschäftigt sich [35, 36]. Zwei unterschiedliche Erkennungsaufgaben werden dabei für die Schwellen definiert: die Objekterkennung (gutes, beschädigtes oder fehlendes Befestigungsmittel) und die semantische Segmentierung (von Splittern, bröckelnden Betonschwellen und anderen Materialklassen). Der verwendete Datensatz umfasst 137 km Gleis mit Bounding Boxes von 203.287 Schwellen. Die Gesamtzahl der defekten Befestigungselemente im Testsatz betrug 1.087 (0,13 % aller Befestigungselemente), einschließlich 22 vollständig fehlender Befestigungselemente und 1.065 gebrochener Befestigungselemente. 2.524 Schwellen wurden als nicht inspizierbar definiert. Das Ergebnis zeigte nur eine fehlerhafte Detektion. [37] untersucht die Anwendung von KI zur Identifikation und Klassifizierung von Rissen in Bahnschwellen in drei verschiedenen Typen: longitudinal, transversal und schräg, basierend auf ihren spezifischen Risswinkeln. Der verwendete Datensatz wurde dafür unter Laborbedingungen erfasst. Es wurden künstliche Risse (kleinste Breite: 1 mm) auf den Schwellen erzeugt, die an beliebigen Stellen und in beliebiger Ausrichtung auf der Oberseite der Schwelle angebracht wurden. Mittels Augmentierungstechniken konnte so ein Datensatz von 283 Bildern generiert werden. Nahezu 100 % der Rissverläufe konnten korrekt bestimmt werden.

Da das Oberleitungssystem entgegen dem Gleisbereich oberhalb der Fahrzeuge verortet ist, wird dieses in der Regel durch Sensorik erfasst, welche auf oder vor dem Fahrzeug befestigt ist. Für die Position des Fahrdrahtes werden häufig LiDAR-Sensoren eingesetzt,

für die Abnutzung des Fahrdrahtes oder einzelner Komponenten des Oberleitungssystems häufiger kamerabasierte Systeme. Das Oberleitungssystem besteht dabei aus vielen einzelnen Komponenten, an die unterschiedliche Anforderungen gestellt werden können. Chen et al. erproben die Detektion von Befestigungsklemmen [38], Kang et al. detektieren und klassifizieren Bilddaten zur Erkennung von Oberflächenfehlern bei Isolatoren [39], während Zhong et al. Befestigungspins in einem dreistufigen automatisierten Inspektionssystem untersuchen [40]. Alle drei Anwendungsbeispiele zeigen den erfolgreichen Einsatz von KI-Methoden zur Inspektion der Oberleitungselemente. Ob die Bild-Datensätze der Öffentlichkeit zur Verfügung stehen, ist nicht bekannt.

FRSign stellt einen Datensatz zur Erkennung von Bahnverkehrslichtsignalen in Frankreich dar. Die Aufnahmen wurden von fahrenden Zügen gemacht und umfassen manuelle Annotationen. Der veröffentlichte Datensatz enthält über 100.000 Bilder von sechs Typen französischer Signale und deren Farbkombinationen zur Klassifikation [41, 42]. Ebenfalls mit Signalen beschäftigt sich der Datensatz GERALD [43]. Der Datensatz umfasst 5.000 Bilder aus einer Vielzahl von Bahnszenen sowie 30.000 Annotationen der häufigsten Typen deutscher Hauptbahnsignale und ist für Forschungs- und nicht-kommerzielle Zwecke frei verfügbar. Der gezeigte Ansatz demonstriert den erfolgreichen Einsatz von KI zur Erkennung von Bahnsignalen. Die erreichte Accuracy reicht jedoch wahrscheinlich nicht aus, um in einem Betriebsszenario wie ATO eingesetzt zu werden [44].

Um die Integrität von Tunneln oder anderer Bauwerke zu gewährleisten, werden Untersuchungen zur frühzeitigen Erkennung von Rissen oder Materialermüdungen durchgeführt. Um einen Datensatz zur automatisierten Risserkennung mittels KI in Tunneln zu generieren, nutzen Song et al. mehrere seitlich und nach oben ausgerichtete Kameras. Der verwendete Datensatz enthält 15.718 händisch annotierte Bilder (512 x 512 Pixel) mit Rissen für die Segmentierung. Die Ergebnisse zeigen, dass der Ansatz eine Risssegmentierung in nahezu Echtzeit durchführen kann [45]. Deutlich erschwert wird die Segmentierung durch Verschmutzungen oder feuchte Stellen an der Tunnelwand.

3.4 Anwendungen mit Punktwolken

Das Verstehen von Eisenbahnszenen hat für verschiedene Anwendungen eine hohe Bedeutung, z. B. für autonome Züge, Building Information Modeling (BIM) oder digitale Zwillinge. Dabei wird die automatische semantische Segmentierung der Punktwolken als

bevorzugte Methodik eingesetzt, beispielhaft genannt seien [46–49]. Punktwolken ermöglichen es, dreidimensionale Informationen zu analysieren und komplexe Strukturen zu erkennen. Komplexe und unregelmäßige Formen können erfasst werden, die sonst möglicherweise übersehen worden wären. Sie können leicht in Geoinformationssysteme (GIS) integriert werden, um umfassende Analysen und Visualisierungen von Infrastrukturdaten zu ermöglichen. Punktwolken haben darüber hinaus die spezifische Eigenart, dass die Messwerte und deren Verteilung in der Wolke sensorabhängig sind. Kellner et al. zeigen dieses Phänomen durch den Vergleich von echten und synthetischen Punktwolken einer Brücke. Eine auf eine Punktwolke trainierte KI kann nur bedingt für verschiedene Sensoren generalisieren und wird unterschiedliche Ergebnisse für verschiedene Punktwolken des gleichen Objekts liefern [50]. Ebenfalls sind viele KI-Modelle anfällig dafür, dass sie den räumlichen Kontext in Punktwolken auf simple geometrische Zusammenhänge reduzieren, was das Erkennen von Objekten erschwert [51].

RailCloud-HdF stellt einen großen LiDAR-Punktwolken-Datensatz dar, welcher der Segmentierung von Bahnszenen dient [52]. Die Annotation erfolgte manuell für die Klassen Erdoberfläche, Vegetation, Gebäude, Masten, Gleise, Oberleitungen, Schranken an Bahnübergängen und Strukturen (z. B. Tunnel, Brücken). Rail3D umfasst drei verschiedene Eisenbahnumgebungen in Ungarn, Frankreich und Belgien, die ein breites Spektrum von Eisenbahnanlagen und -bedingungen abbilden [53, 54]. Der Datensatz umfasst die Klassen Erdoberfläche, Vegetation, Gleis, Masten, verschiedene Oberleitungen, Signale, Zäune, größere Bahninfrastrukturen und Gebäude. Jiang et al. stellen ebenfalls einen Datensatz für die semantische Segmentierung zur Verfügung [55, 56]. Der Datensatz RailPC umfasst rund 3 Milliarden Punkte, die über 25 km Schienen in städtischen und bergigen Umgebungen gesammelt wurden und mit 16 spezifischen Klassen annotiert sind. Die Ergebnisse der Tests mit verschiedenen Deep-Learning-Modellen zeigen, dass bestehende Methoden in der Segmentierung von Bahninfrastruktur wie Gleisen und Masten erheblichen Verbesserungsbedarf haben. Der RailPC-Datensatz ist öffentlich zugänglich und wird kontinuierlich aktualisiert [56].

Speziell für die Segmentierung von Tunnelverkleidungen ist der Datensatz Seg2Tunnel erstellt worden [57]. Der Datensatz umfasst 1.300 Tunnelringe, die aus fünf verschiedenen Tunneln stammen und die Anwendung von 3D-Deep-Learning-Techniken zur automatischen Segmentierung ermöglicht. Die automatisierte Berechnung von Deformationen basierend auf den segmentierten Punktwolken zeigt eine geringe durchschnittliche relative Fehlerquote. Insgesamt bietet die Studie eine Grundlage für

zukünftige Forschungsarbeiten im Bereich der automatisierten Tunnelinspektion und -analyse.

Die semantische Segmentierung des Oberleitungssystems stellt einen weiteren Schritt hin zur Digitalisierung der Eisenbahninfrastruktur dar. [58, 59] beschäftigen sich mit der Segmentierung von Oberleitungssystemen bzw. mit dem Tragwerk oberhalb der Bahnstrecke des niederländischen Eisenbahnsystems. Der Datensatz enthält auf 800 m Strecke 14 Klassen und ist öffentlich zugänglich. Er weist jedoch ein großes Ungleichgewicht der einzelnen Klassen auf. Mit der Erfassung des Oberleitungssystems befassen sich auch [60]. Der Datensatz umfasst acht Kategorien (Masten, Isolatoren, Ausleger, Tragseil, Fahrdraht, Stützrohr, Hänger, Fahrdrahtseitenhalter) und die experimentellen Ergebnisse zeigen, dass die vorgeschlagene DL-Methode eine hohe Accuracy bei der Erkennung dieser Komponenten erzielt.

4 Zusammenfassung und Diskussion

Zusammenfassend lässt sich sagen, dass eine KI-Auswertung basierend auf 2D-Bilddaten schon heute zu verwertbaren Ergebnissen führen kann. Die Objekterkennung von Rissen in Schwellen [37] zeigt mit einer Accuracy von 96,9 % eine sehr hohe Erkennungsrate auf, während die Segmentierung von Rissen in Tunnelwänden [45] mit einer Accuracy von 80 % zwar niedriger ist, aber dafür sogar echtzeitfähig. Auch Oberflächendefekte auf den Schienen [31–33] können mit 90,30 % - 97,10 % erkannt werden. Für die Erkennung von Signalen existieren mit GERALD [43, 44] und FRSign [42] große Datensätze, welche für das Training von zuverlässigen KI-Modellen entscheidend sind.

Das Beispiel der Schwellenerkennung [35, 36] zeigt eindrücklich, dass selbst eine vermeintlich hohe Accuracy für zukünftige Anwendungen im sicherheitskritischen Bereich unzureichend sein wird. So käme ein naives Modell, welche alle Schwellen als „intakt“ vorhersagt, auf eine 99,87 % Trefferquote in einem Datensatz mit 99,87 % intakter Schwellkörper, aber dennoch würden 0,13 % oder 264 defekte Schwellkörper nicht richtig erkannt. Für den sicherheitskritischen Bahnbereich spielen also andere Metriken wie der Recall eine zusätzliche entscheidende Rolle. Dieser beschreibt die Anzahl der tatsächlich erkannten Defekte. Die Erkennung aller Defekte wird bei sicherheitsrelevanten Strukturen angestrebt. Um in diesen Bereich zu gelangen, sind größere Datensätze notwendig als die hier vorgestellten.

Die Generalisierung über verschiedene Sensoren stellt KI-Modelle, welche direkt auf Punktwolken arbeiten, aktuell noch vor Herausforderungen. Die Aufbereitung der Daten

bei Punktwolken ist allein durch die schiere Datenmenge sehr aufwendig. Synthetisch generierte Daten können ein Weg sein, diese Kosten zu minimieren.

Es zeigt sich darüber hinaus, dass es keinen öffentlichen Datensatz von erforderlicher Größe und Internationalität gibt, mit dem sich einzelne Problemstellungen für verschiedene Infrastrukturen lösen ließen. Die geringe Diversität der Datensätze erschwert es, allgemeingültige Aussagen über die Übertragbarkeit und Verwendbarkeit der entwickelten Lösungen zu treffen. Beispielsweise werden verschiedene Punktwolken des gesamten Zugumfelds aus Frankreich [61], China [55, 56], Spanien [46, 52, 62], Belgien und Ungarn [53], Signale aus Deutschland [44] und Frankreich [41] und länderspezifische Architekturen wie beispielsweise Oberleitungsbögen aus den Niederlanden [58] als Datengrundlage verwendet. Der RailSem19 Datensatz sticht hier hervor, welcher Bilddaten aus 38 verschiedenen Ländern beinhaltet [26, 27] und auch der RailSet Datensatz [30–32] wurde in verschiedenen Ländern erstellt.

Besonders interessant könnte der noch nicht veröffentlichte Rail-Bench [63] Datensatz werden. Hier wird in Aussicht gestellt, einen Benchmarkdatensatz zur Verfügung zu stellen, welcher die Vergleichbarkeit der KI-Modelle für Multisensorsysteme drastisch erhöhen würde.

Der Einsatz von KI zur regelmäßigen Infrastrukturüberwachung zeigt großes Potenzial in der Automatisierung. Routineaufgaben können effizient und schneller durchgeführt werden. Wartungsstrategien und frühzeitige Problemerkennung können Kosten und Personalaufwand reduzieren. Durch den Einsatz einer Echtzeit-Datenanalyse können Entscheidungen zeitnah getroffen werden und eine sofortige Reaktion erfolgen. Der hohe anfängliche Investitionsaufwand zur Implementierung eines KI-Ansatzes, die Komplexität der Entwicklung und Wartung von KI-Systemen und fehlende Flexibilität, sich an Situation und Änderungen anzupassen, wirken sich zurzeit noch negativ auf die praktische Anwendung aus.

Die Datenabhängigkeit, sowohl in Quantität als auch in Qualität, stellt eine Herausforderung dar. Ohne die benötigten Mengen an hochwertigen Daten kann es zu ungenauen oder fehlerhaften Analysen kommen. Datensätze, gewonnen durch verschiedene Sensoren (Kamera-, LiDAR-Sensoren), bieten unterschiedliche Vor- und Nachteile. Punktwolken bieten den Vorteil der dreidimensionalen Information von Objekten im Raum. Durch die Tiefeninformation kann die Position des Objektes im Raum einfacher und direkter bestimmt werden. Gleichzeitig sind sie weniger anfällig für Veränderungen der Lichtverhältnisse, Schatten und Reflexionen als Bilder. Ihr Einsatz im Bereich der Infrastrukturüberwachung fokussiert sich hauptsächlich auf das

semantische Segmentieren. Bilddaten können sehr hohe Auflösungen haben und bieten Informationen über Texturen, Farben und Oberflächenmerkmale. Für den Einsatz zur Objekterkennung und -klassifizierung bietet dies einen Vorteil. Die Verwendung von synthetischen Daten führt möglicherweise zu einem mangelhaften Kontext im Lernprozess eines KI-Systems. Die Realität der Umgebung wird in ihrer Komplexität nicht vollständig dargestellt und nicht alle Variation und Unregelmäßigkeiten werden ausreichend abgebildet. Dies kann die Robustheit des Modells beeinträchtigen.

Angesichts dessen, dass es sich bei Deep Learning als Technologie für die Bahninfrastrukturüberwachung um ein recht junges Forschungsfeld handelt, lässt sich unter Betrachtung der Jahreszahlen der gefundenen Datensätze von einer recht hohen Dynamik sprechen. Es ist davon auszugehen, dass sich diese die nächsten Jahre weiter erhöht.

Zukünftig kann XAI helfen, die Modelle besser zu verstehen und für bahnspezifische Probleme zu optimieren. Durch immer leistungsfähigere Hardware wird es in Zukunft möglich sein, die in den Studien vorgestellten Technologien echtzeitfähig auf Mess- oder Regelzügen einzusetzen. Eine kosteneffiziente, kontinuierliche und reproduzierbare Infrastrukturüberwachung wird so möglich.

Annex

Die folgenden Tabellen zeigen eine Übersicht über verschiedene Datensätze, erfasst durch optische Sensorik oder synthetisch erzeugt, zur regelmäßigen Überwachung der Eisenbahninfrastruktur für KI-Anwendungen. Zunächst werden Bilddatensätze, gefolgt von Punktwolken sowie Daten aus Multisensorsystemen und synthetisch erzeugte Datensätze aufgezählt.

Session 13
Kira Zschiesche, Johannes Pelz, Alexander Reiterer

Tabelle 1: Datensätze für KI-Anwendungen zur Infrastrukturüberwachung mittels Kameras

Datensatz	Jahr	Infrastrukturobjekt	Daten Bilder	Annotiert	Öffentlich zugänglich	Aufgabe*	Herkunfts-land	Quelle
Finding railway fasteners in image data - ProRail	2020	Schienenbefestigung	47	ja	unklar	O	NL	[64]
Image data of insulation joints - ProRail	2019	Isolierter Schienenstoß	28,2 Tsd	ja	unklar	O	NL	[65]
Image data of spark erosion	2021	Funkenerosion an Schienenstoß	0,3 Tsd	ja	unklar	O	NL	[66]
RailSem19	2019	Gesamtes Zugumfeld	8,5 Tsd	ja	ja	S	38 Länder	[26, 27]
The GERALD Dataset	2023	Lichtsignale	5 Tsd	ja	ja	O	DE	[44]
FRSign: A Large-Scale Traffic Light Dataset for Autonomous Trains	2020	Lichtsignale	105 Tsd	ja	ja	O	FR	[42]
RAilWay PEdestrian Dataset (RAWPED) Dataset	2020	Passanten	26 Tsd	ja	ja	O	k. A.	[11]
Railway Obstacle Dataset (RAILOD)	2022	Hindernisse auf der Strecke	4,7 Tsd	ja	ja	O	CN	[10]

An Investigation of Railway Fastener Detection	2021	Schienenbefestigung	3 Tsd	ja	auf Anfrage	K	SE	[34]
Real-Time Tunnel Crack Analysis System via Deep Learning	2019	Tunnelwand	15,7 Tsd	ja	unklar	S	CN	[45]
ESRORAD dataset	2022	Szenen aus dem Straßenverkehr inklusive Bahnumfeld	102,7Tsd	ja	unklar	O	FR	[12]
RailEye3D	2021	Personenerkennung im Bahnumfeld	10 Bildsequenzen	ja	auf Anfrage	O	AT	[67, 68]
Rail-DB	2022	Gleiserkennung	7,4 Tsd	ja	auf Anfrage	S	CN	[23, 25]
RailSet: A Unique Dataset for Railway Anomaly Detection	2022	Defekte / Anomalien im Fahrweg	7,7 Tsd	ja	ja	S	Verschiedene Länder	[30]
RailVID: A Dataset for Rail Environment Semantic	2022	Gesamtes Zugumfeld	1,1 Tsd	ja	ja	S	CN	[69, 70]
rail-hp8ij: Rail Computer Vision Project	2022	Signale und Personen	0,6 Tsd	ja	ja	O	k. A.	[71]
Railway track surface faults dataset	2024	Defekte auf der Gleisoberfläche	5,2 Tsd	ja	ja	K	PK	[28, 29]

Crack detection dataset	2019	Rost und Risse in Gleisen	0,6 Tsd	ja	ja	K	k. A.	[32]
RailDefectDetection	2021	Defekte auf der Gleisoberfläche	1,8 Tsd	ja	ja	O, K	TR	[72]

Tabelle 2: Datensätze für KI-Anwendungen zur Infrastrukturüberwachung mittels LiDAR

Datensatz	Jahr	Infrastrukturobjekt	Daten	Annotiert	Öffentlich zugänglich	Aufgabe*	Herkunfts-land	Quelle
MOMIT: 3D Point Cloud of a railway slope	2020	Gesamtes Zugumfeld	110 Mio Punkte	nein	ja	k. A.	FR	[61]
RailPC: A large-scale railway point cloud semantic segmentation dataset	2024	Gesamtes Zugumfeld	2 Mrd Punkte / 25 km	ja	auf Anfrage	S	CN	[56]
Labelled high resolution point cloud dataset of 15 catenary arches in the Netherlands	2022	Oberleitungsbögen	15 Mio Punkte	ja	ja	S	NL	[58]
Annotated mobile laser scans of the Dutch railway environment	2021	Gesamtes Zugumfeld	9,4 km	ja	nein	S	NL	[49]
Rail3D: Multi-Context Point Cloud Dataset	2024	Gesamtes Zugumfeld	288 Mio Punkte / 5,8 km	ja	ja	S	BE, FR, HU	[53]

Automated Recognition of Railroad Infrastructure in Rural Areas from LIDAR Data	2015	Gesamtes Zugumfeld	12,5 Mio Punkte / 0,55 km	ja	nein	S	AT	[73]
WHU-Railway3D	2024	Gesamtes Zugumfeld	4,6 Mrd Punkte / 30 km	ja	auf Anfrage	S	CN	[74]
SemanticRail3D	2021	Gesamtes Zugumfeld	2,8 Mrd Punkte / 90 km	ja	ja	S	ES	[46, 62]
Seg2Tunnel	2024	Tunnel	821 Mio Punkte / 1,8 km	ja	auf Anfrage	S	CN	[57]

Tabelle 3: Datensätze für KI-Anwendungen zur Infrastrukturüberwachung mittels Multisensorsystemen

Datensatz	Jahr	Infrastrukturobjekt	Daten	Annotiert	Öffentlich zugänglich	Aufgabe*	Herkunfts-land	Quelle
Rail-Bench: Benchmarksuite für die Automatisierung des Schienenverkehrs	k. A.	Gesamtes Zugumfeld	k. A.	ja	ja	k. A.	k. A.	[63]
RailCloud-HdF	2024	Gesamtes Zugumfeld	8 Mrd Punkte / 267 km Track	ja	auf Anfrage	S	FR	[52]
Open Sensor Data for Rail 2023 (OSDaR23)	2023	Hindernisse im Zugumfeld	1534 frames	ja	ja	O	DE	[20]

RailGoerl24: Görlitz Rail Test Center CV Dataset 2024	2024	Personenerkennung im Bahnumfeld	12,2 Tsd Bilder / 383 Mio Punkte	ja	ja	O	DE	[18, 19]

Tabelle 4: Datensätze für KI-Anwendungen zur Infrastrukturüberwachung mittels synthetisch generierter Daten

Datensatz	Jahr	Infrastrukturobjekt	Daten	Annotiert	Öffentlich zugänglich	Aufgabe*	Herkunfts-land	Quelle
RailFOD23: A dataset for foreign object detection on railroad transmission lines	2024	Fremdkörper an Oberleitungen	14.6 Tsd synth. Bilder	ja	ja	O	CN	[16, 75]
Synthetic Railway Level Crossing Dataset (SRLC)	2021	Bahnübergänge	32 Mio Punkte	ja	ja	S	SE	[17]
Syntethic Annotated Railway Dataset (SARD)	2022	Gesamtes Zugumfeld	7,5 Tsd Punkte	ja	ja	O	IT	[76]
Nuage de points 3D des infrastructures ferroviaires	2017	Gesamtes Zugumfeld	2 km	nein	ja	k. A.	FR	[77]

* K = Klassifikation, O = Objektdetektion, S = Segmentierung, k. A. = keine Angabe

5 Literatur

1. Mattsson L-G, Jenelius E (2015) Vulnerability and resilience of transport systems – A discussion of recent research. Transportation Research Part A: Policy and Practice 81:16–34. doi:10.1016/j.tra.2015.06.002
2. Trimble Inc. (2022) GEDO Vorsys. pre-measurment for tamping
3. Hisa T, Kanaya M, Sakai M, Hamaoka K (2012) Rail and contact line inspection technology for safe and reliable railway traffic. Hitachi Review 61(7)
4. Phusakulkajorn W, Núñez A, Wang H, Jamshidi A, Zoeteman A, Ripke B, Dollevoet R, Schutter B de, Li Z (2023) Artificial intelligence in railway infrastructure: current research, challenges, and future opportunities. Intelligent Transportation Infrastructure 2. doi:10.1093/iti/liad016
5. Szeliski R (2021) Computer Vision: Algorithms and Applications, 2nd Edition
6. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, Guidotti R, Del Ser J, Díaz-Rodríguez N, Herrera F (2023) Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion 99:101805. doi:10.1016/j.inffus.2023.101805
7. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 58:82–115. doi:10.1016/j.inffus.2019.12.012
8. Di Summa M, Griseta ME, Mosca N, Patruno C, Nitti M, Renò V, Stella E (2023) A Review on Deep Learning Techniques for Railway Infrastructure Monitoring. IEEE Access 11:114638–114661. doi:10.1109/ACCESS.2023.3309814
9. Oh K, Yoo M, Jin N, Ko J, Seo J, Joo H, Ko M (2022) A Review of Deep Learning Applications for Railway Safety. Applied Sciences 12(20):10572. doi:10.3390/app122010572
10. Guan L, Jia L, Xie Z, Yin C (2022) A Lightweight Framework for Obstacle Detection in the Railway Image Based on Fast Region Proposal and Improved YOLO-Tiny Network. IEEE Trans. Instrum. Meas. 71:1–16. doi:10.1109/TIM.2022.3150584
11. Toprak T, Belenlioglu B, Aydin B, Guzelis C, Selver MA (2020) Conditional Weighted Ensemble of Transferred Models for Camera Based Onboard Pedestrian Detection in Railway Driver Support Systems. RAWPED Dataset. IEEE Trans. Veh. Technol.:1. doi:10.1109/TVT.2020.2983825

12. Khemmar R, Mauri A, Dulompont C, Gajula J, Vauchey V, Haddad M, Boutteau R (2022) Road and Railway Smart Mobility: A High-Definition Ground Truth Hybrid Dataset. Sensors 22(10):3922. doi:10.3390/s22103922
13. Bai R, Wu Z, Xu T (2024) A Lightweight Camera and Lidar Fusion Framework for Railway Transit Obstacle Detection CACML 2024. Association for Computing Machinery, Erscheinungsort nicht ermittelbar, S 303–308
14. Menius R (2024) Bahnbau und Bahninfrastruktur. Ein Leitfaden zu bahnbezogenen Infrastrukturthemen, 11. Aufl. Springer Vieweg, Wiesbaden, Heidelberg
15. Kaden K-U, Schwalbe K, Scheithauer M (2023) Aufbereitung von Datensätzen für Anwendungen des automatisierten Fahrens im Eisenbahnbetrieb. Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt
16. Chen Z, Yang J, Feng Z, Zhu H (2024) RailFOD23: A dataset for foreign object detection on railroad transmission lines. Sci Data 11(1):72. doi:10.1038/s41597-024-02918-9
17. Uggla G, Horemuz M (2021) Towards synthesized training data for semantic segmentation of mobile laser scanning point clouds: Generating level crossings from real and synthetic point cloud samples. Automation in Construction 130:103839. doi:10.1016/j.autcon.2021.103839
18. Tagiew R, Wunderlich I, Sastuba M, Seitz S (2025) RailGoerl24: Görlitz Rail Test Center CV Dataset 2024 IEEE Engineering Reliable Autonomous Systems (ERAS)
19. Tagiew R, Wunderlich I, Zanitzer P, Sastuba M, Knoll C, Göller K, Amjad H, Seitz S (2025) Görlitz Rail Test Center CV Dataset 2024 (RailGoerl24). https://data.fid-move.de/dataset/railgoerl24. Zugegriffen: 11. Juni 2025
20. Tagiew R, Klasek P, Tilly R, Köppel M, Denzler P, Neumaier P, Klockau T, Boekhoff M, Schwalbe K (2024) OSDaR23: Open Sensor Data for Rail 2023. 8th International Conference on Robotics and Automation Engineering (ICRAE), Singapore, Singapore:270–276. doi:10.1109/ICRAE59816.2023.10458449
21. Tilly R, Neumaier P, Tagiew R, Schwalbe K, Klasek P, Klockau T, Denzler P, Köppel M (2023) Offener Multisensordatensatz für die Entwicklung der Umfeldwahrnehmung beim vollautomatischen Fahren. Eisenbahntechnische Rundschau April 2023(4)
22. Tilly R, Neumaier P, Schwalbe K, Klasek P, Tagiew R, Denzler (2023) Open Sensor Data for Rail 2023. https://data.fid-move.de/dataset/osdar23. Zugegriffen: 11. Juni 2025
23. Li X, Peng X (2022) Rail-DB: dataset. https://github.com/Sampson-Lee/Rail-Detection. Zugegriffen: 11. Juni 2025

24. Wang Y, Wang L, Hu YH, Qiu J (2019) RailNet: A Segmentation Network for Railroad Detection. IEEE Access 7:143772–143779. doi:10.1109/ACCESS.2019.2945633
25. Li X, Peng X (2022) Rail Detection: An Efficient Row-based Network and A New Benchmark MM '22: Proceedings of the 30th ACM International Conference on Multimed. Association for Computing Machinery, New York, United States, S 6455–6463
26. Zendel O, Murschitz M, Zeilinger M, Steininger D, Abbasi S, Beleznai C (2025) RailSem19 Dataset. https://wilddash.cc/railsem19. Zugegriffen: 27. Mai 2025
27. Zendel O, Murschitz M, Zeilinger M, Steininger D, Abbasi S, Beleznai C (2019) RailSem19: A Dataset for Semantic Rail Scene Understanding Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), S 0
28. Arain A, Mehran S, Shaikh MZ, Kumar D, Chowdhry BS, Hussain T (2024) Railway track surface faults dataset. Data in Brief 52:110050. doi:10.1016/j.dib.2024.110050
29. Arain A, Mehran S, Shaikh MZ, Kumar D, Hussain T, Chowdhry BS (2022) Railway Track Surface Faults Dataset. Mehran University of Engineering and Technology. https://data.mendeley.com/datasets/8hxtgyyxrw/2. Zugegriffen: 12. Juni 2025
30. Zouaoui A, Mahtani A, Hadded MA, Ambellouis S, Boonaert J, Wannous H (2022) RailSet: A Unique Dataset for Railway Anomaly Detection. In: IEEE (Hrsg) 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS), S 1–6
31. Iyer S, Velmurugan T, Gandomi AH, Noor Mohammed V, Saravanan K, Nandakumar S (2021) Structural health monitoring of railway tracks using IoT-based multi-robot system. Neural Comput & Applic 33(11):5897–5915. doi:10.1007/s00521-020-05366-9
32. Iyer S (2025) GitHub - kriiyer/Dataset: Dataset for crack and rust detection. https://github.com/kriiyer/Dataset/tree/master. Zugegriffen: 27. Juni 2025
33. Aydin I, Akin E, Karakose M (2021) Defect classification based on deep features for railway tracks in sustainable transportation. Applied Soft Computing 111:107706. doi:10.1016/j.asoc.2021.107706
34. Chandran P, Asber J, Thiery F, Odelius J, Rantatalo M (2021) An Investigation of Railway Fastener Detection Using Image Processing and Augmented Deep Learning. Sustainability 13(21):12051. doi:10.3390/su132112051

35. Gibert X, Patel VM, Chellappa R (2015) Deep Multi-task Learning for Railway Track Inspection. IEEE Transactions on Intelligent Transportation Systems (18):153–164. doi:10.1109/TITS.2016.2568758
36. Gibert X, Patel VM, Chellappa R (2015) Robust Fastener Detection for Autonomous Visual Railway Track Inspection 2015 IEEE Winter Conference on Applications of Computer Vision. IEEE, S 694–701
37. Jamshidi A, Faghih-Roohi S, Hajizadeh S, Núñez A, Babuska R, Dollevoet R, Li Z, Schutter B de (2017) A Big Data Analysis Approach for Rail Failure Risk Assessment. Risk Analysis 37(8):1495–1507. doi:10.1111/risa.12836
38. Chen J, Liu Z, Wang H, Nunez A, Han Z (2018) Automatic Defect Detection of Fasteners on the Catenary Support Device Using Deep Convolutional Neural Network IEEE Transactions on Instrumentation and Measurement, Bd 67, S 257–269
39. Kang G, Gao S, Yu L, Zhang D (2019) Deep Architecture for High-Speed Railway Insulator Surface Defect Detection: Denoising Autoencoder With Multitask Learning. IEEE Trans. Instrum. Meas. 68(8):2679–2690. doi:10.1109/TIM.2018.2868490
40. Zhong J, Liu Z, Han Z, Han Y, Zhang W (2019) A CNN-Based Defect Inspection Method for Catenary Split Pins in High-Speed Railway. IEEE Trans. Instrum. Meas. 68(8):2849–2860. doi:10.1109/TIM.2018.2871353
41. Harb J, Rébéna N, Chosidow R, Roblin G, Potarusov R, Hajri H (2020) FRSign Dataset. System X, Institute of Research and Technology. https://frsign.irt-systemx.fr/. Zugegriffen: 27. Mai 2025
42. Harb J, Rébéna N, Chosidow R, Roblin G, Potarusov R, Hajri H (2020) FRSign: A Large-Scale Traffic Light Dataset for Autonomous Trains
43. Leibner P, Hampel F, Schindler C (2023) The GERALD dataset. RWTH Aachen. https://github.com/ifs-rwth-aachen/GERALD?tab=readme-ov-file. Zugegriffen: 27. Mai 2025
44. Leibner P, Hampel F, Schindler C (2023) GERALD: A novel dataset for the detection of German mainline railway signals. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 237(10):1332–1342. doi:10.1177/09544097231166472
45. Song Q, Wu Y, Xin X, Yang L, Yang M, Chen H, Liu C, Hu M, Chai X, Li J (2019) Real-Time Tunnel Crack Analysis System via Deep Learning. IEEE Access 7:64186–64197. doi:10.1109/ACCESS.2019.2916330

46. Lamas D, Soilán M, Grandío J, Riveiro B (2021) Automatic Point Cloud Semantic Segmentation of Complex Railway Environments. Remote Sensing 13(12):2332. doi:10.3390/rs13122332
47. Dekker B, Ton B, Meijer J, Bouali N, Linssen J, Ahmed F (2023) Point Cloud Analysis of Railway Infrastructure: A Systematic Literature Review IEEE Access, Bd 11, S 134355–134373
48. Grandio J, Riveiro B, Soilán M, Arias P (2022) Point cloud semantic segmentation of complex railway environments using deep learning. Automation in Construction 141:104425. doi:10.1016/j.autcon.2022.104425
49. Ton B, Strukton Rail (2024) Annotated mobile laser scans of the Dutch railway environment. https://data.4tu.nl/datasets/fa259c52-a585-420c-8a0c-af5e91518e29/1
50. Kellner M, König T, Jäkel J-I, Klemt-Albert K, Reiterer A (2025) 3D bridge segmentation using semi-supervised domain adaptation. Automation in Construction 172:106021. doi:10.1016/j.autcon.2025.106021
51. Wu X, DeTone D, Frost D, Shen T, Xie C, Yang N, Engel J, Newcombe R, Zhao H, Straub J (2025) Sonata: Self-Supervised Learning of Reliable Point Representations
52. Abid M, Teixeira M, Mahtani A, Laurent T (2024) RailCloud-HdF: A Large-Scale Point Cloud Dataset for Railway Scene Semantic Segmentation Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. SCITEPRESS - Science and Technology Publications, S 159–170
53. Kharroubi A, Ballouch Z, Hajji R, Yarroudh A, Billen R (2024) Multi-Context Point Cloud Dataset and Machine Learning for Railway Semantic Segmentation. Infrastructures 9(4):71. doi:10.3390/infrastructures9040071
54. Kharroubi A (2025) Rail3D dataset. https://github.com/akharroubi/Rail3D. Zugegriffen: 28. Mai 2025
55. Jiang T, Li S, Zhang Q, Wang G, Zhang Z, Zeng F, An P, Jin X, Liu S, Wang Y (2024) RailPC Dataset. https://github.com/NNU-GISA/GISA-RailPC. Zugegriffen: 27. Mai 2025
56. Jiang T, Li S, Zhang Q, Wang G, Zhang Z, Zeng F, An P, Jin X, Liu S, Wang Y (2024) RailPC: A large-scale railway point cloud semantic segmentation dataset. CAAI Trans on Intel Tech 9(6):1548–1560. doi:10.1049/cit2.12349
57. Lin W, Sheil B, Zhang P, Zhou B, Wang C, Xie X (2024) Seg2Tunnel: A hierarchical point cloud dataset and benchmarks for segmentation of segmental tunnel linings. Tunnelling and Underground Space Technology 147:105735. doi:10.1016/j.tust.2024.105735

58. Ton B (2023) Labelled high resolution point cloud dataset of 15 catenary arches in the Netherlands (dataset). https://data.4tu.nl/articles/dataset/Labelled_high_resolution_point_cloud_dataset_of_15_catenary_arches_in_the_Netherlands/17048816/1. Zugegriffen: 02. Mai 2025
59. Ton B, Ahmed F, Linssen J (2022) Semantic Segmentation of Terrestrial Laser Scans of Railway Catenary Arches: A Use Case Perspective. Sensors 23(1):222. doi:10.3390/s23010222
60. Lin S, Xu C, Chen L, Li S, Tu X (2020) LiDAR Point Cloud Recognition of Overhead Catenary System with Deep Learning. Sensors 20(8):2212. doi:10.3390/s20082212
61. Riquelme JLA, Ruffo M, Tomás R, Riquelme A, Pagán JI, Cano M, Pastor JL MOMIT (Multi-scale observation and monitoring of railway infrastructure threats) EU project - H2020-EU.3.4.8.3. - Grant agreement ID: 777630. 3D Point Cloud of a railway slope. EU project - H2020-EU.3.4.8.3. Zenodo
62. Soilán M, Lamas-Novoa D, Martínez-Sánchez J, Arias P, Lorenzo H, Riveiro B (2024) SemanticRail3D: A 3D Point Cloud dataset with semantic annotations of railway environments. https://zenodo.org/records/11143767. Zugegriffen: 12. Juni 2025
63. Lauer M (2025) Rail-Bench | Daten+. Karsruher Institut für Technologie, KIT. https://daten.plus/projekte/rail-bench. Zugegriffen: 17. Juni 2025
64. van Hees O (2020) Finding railway fasteners in image data - ProRail. https://www.kaggle.com/datasets/oscarvanhees/finding-railway-fasteners-in-image-data-prorail. Zugegriffen: 26. Mai 2025
65. van Hees O (2019) Image data of insulation joints - ProRail. https://www.kaggle.com/datasets/oscarvanhees/insulation-joint-training-set-prorail/data. Zugegriffen: 26. Mai 2025
66. van Hees O (2020) Image data of spark erosion - ProRail. https://www.kaggle.com/datasets/oscarvanhees/image-data-of-spark-erosion-prorail/data?select=traingset_small. Zugegriffen: 26. Mai 2025
67. Wallner M, Steininger D, Widhalm V, Schörghuber M, Beleznai C RGB-D Railway Platform Monitoring and Scene Understanding for Enhanced Passenger Safety. Pattern Recognition. ICPR International Workshops and Challenges. ICPR 12667:656–671. doi:10.1007/978-3-030-68787-8_47
68. Wallner M, Steininger D, Widhalm V, Schörghuber M (2021) RailEye3D Dataset. Pattern Recognition. ICPR International Workshops and Challenges. https://github.com/raileye3d/raileye3d_dataset. Zugegriffen: 11. Juni 2025

69. Yuan H, Mei Z, Chen Y, Nui W, Wu C (Hrsg) (2022) RailVID: A Dataset for Rail Environment Semantic. The seventeenth International Conference on Systems: April 24-28, 2022, Barcelona, Spain. IARIA, Wilmington, DE, USA
70. Yuan H (2022) RailVID: A Dataset for Rail Environment Semantic [Dataset]. https://doi.org/10.5281/zenodo.7063921. Zugegriffen: 11. Juni 2025
71. Rail (2022) Rail Computer Vision Dataset. rail-hp8ij dataset: Rail Computer Vision Project. https://universe.roboflow.com/rail-psseq/rail-hp8ij. Zugegriffen: 10. Juni 2025
72. Aydin I, Akin E, Karakose M (2021) RailDefectDetection. https://github.com/ilhanaydintr/RailDefectDetection. Zugegriffen: 01. Juli 2025
73. Arastounia M (2015) Automated Recognition of Railroad Infrastructure in Rural Areas from LIDAR Data. Remote Sensing 7(11):14916–14938. doi:10.3390/rs71114916
74. Qiu B, Zhou Y, Dai L, Wang B, Li J, Dong Z, Wen C, Ma Z, Yang B (2024) WHU-Railway3D: A Diverse Dataset and Benchmark for Railway Point Cloud Semantic Segmentation. IEEE Trans. Intell. Transport. Syst. 25(12):20900–20916. doi:10.1109/TITS.2024.3469546
75. Chen Z (2023) RailFOD23 [Dataset]. https://figshare.com/articles/figure/RailFOD23_zip/24180738?file=43616139. Zugegriffen: Juni 2025
76. Neri M, Battisti F (2022 - 2022) 3D Object Detection on Synthetic Point Clouds for Railway Applications 2022 10th European Workshop on Visual Information Processing (EUVIP). IEEE, S 1–6
77. SNCF Réseau (2017) Nuage de points 3D des infrastructures ferroviaires. https://ressources.data.sncf.com/explore/dataset/nuage-points-3d/information/. Zugegriffen: 30. Mai 2025

Session 13
Kira Zschiesche, Johannes Pelz, Alexander Reiterer

Autoren



Zschiesche, Kira

Dr. Kira Zschiesche leitet die Gruppe „Mobile Bahnmesstechnik“ am Fraunhofer-Institut für Physikalische Messtechnik IPM in Freiburg. Sie hat Geoinformation und Kommunaltechnik an der Frankfurt University of Applied Sciences und Geoinformatik und Vermessung an der Mainz University of Applied Sciences studiert. 2023 promovierte sie an der Universität Freiburg. Der Schwerpunkt ihrer Arbeit liegt auf der 3D-Erfassung mittels Laserscanning und Photogrammetrie im Bereich der Infrastrukturen und komplexen Umgebungen.



Pelz, Johannes

Johannes Pelz (M. Sc.) arbeitet als wissenschaftlicher Mitarbeiter in der Gruppe „Mobile Bahnmesstechnik“ am Fraunhofer-Institut für Physikalische Messtechnik IPM in Freiburg. Er hat Wirtschaftsingenieurwesen für Elektro- und Informationstechnik an der HTWG Konstanz und Embedded Systems Engineering an der Universität Freiburg studiert. Schwerpunkt seiner Arbeit ist die hardwarenahe Programmierung von Messtechnik im Bahnbereich.



Reiterer, Alexander

Prof. Dr. Alexander Reiterer leitet die Abteilung „Objekt- und Formerfassung“ am Fraunhofer-Institut für Physikalische Messtechnik IPM in Freiburg. Er hat Geodäsie und Geophysik an der TU Wien studiert und 2001 an der TU Wien promoviert. 2010 bis 2012 war er mit einem Alexander-von-Humboldt-Stipendium an der TU München. Weitere Forschungsaufenthalte absolvierte er u. a. an der University of Calgary in Kanada und an der ETH Zürich. Parallel zu seiner Tätigkeit beim Fraunhofer-Institut für Physikalische Messtechnik IPM ist er Inhaber des Lehrstuhls „Monitoring von Großstrukturen“ an der Universität Freiburg.

Session 13
Felix Lampe, Felix Doose, Christian Schaper, Björn Dickenbrok

Quantifizierung der Wechselwirkungen von Instandhaltungs- und Erneuerungsmaßnahmen auf den Zustand des Schienennetzes

Lampe, Felix[1]; Doose, Felix[1]; Schaper, Christian[1]; Prof. Dr.-Ing. Dickenbrok, Björn[2]

[1]DB InfraGO AG, Infrastrukturzustandsbewertung (I.IAP 13) Frankfurt (Main);
[2]Technische Hochschule Mittelhessen, Gießen

Zusammenfassung

Die nachhaltige Bewirtschaftung des Schienennetzes erfordert ein fundiertes Verständnis der Wirkung von Instandhaltungs- und Erneuerungsmaßnahmen auf den Netzzustand, um die Effizienz der eingesetzten Mittel und die Verfügbarkeit der Infrastruktur langfristig zu sichern. Mithilfe der Netzzustandsnotenlogik, Verschleißmodellen und heuristischer Ansätze für die Maßnahmenwirkung werden die Wechselwirkungen zwischen verschiedenen Erneuerungs- und Instandhaltungsstrategien analysiert und quantifiziert, um daraus gezielte Handlungsempfehlungen für eine strategische Maßnahmenplanung ableiten zu können.

Keywords: Infrastrukturmanagement; Infrastrukturzustand; Infrastrukturqualität; Zustandsnote; Netzzustandsnote; Prognose

1 Einleitung und Zielsetzung

Um den Zustand der Eisenbahninfrastruktur in Deutschland zu verbessern und langfristig in einem guten Zustand zu halten, sind kontinuierliche Investitionen und eine nachhaltige Bewirtschaftung des Bestandsnetzes erforderlich. Eine nachhaltige Bewirtschaftung erfordert sowohl reaktive als auch präventive Instandhaltungs- sowie langfristige Erneuerungsmaßnahmen, welche in direkter Wechselwirkung zueinanderstehen. Über den Lebenszyklus einer technischen Anlage werden reaktive und präventive Instandhaltungsmaßnahmen durchgeführt, um die Anlage in einem betriebsbereiten Soll-

Zustand zu erhalten. Sobald eine Instandhaltung nicht mehr wirtschaftlich sinnvoll ist, sollte die Anlage erneuert werden. Sofern es nicht möglich ist, eine Anlage zum vorgesehenen Zeitpunkt zu erneuern, sind zusätzliche Instandhaltungsmaßnahmen erforderlich, um diese betriebsbereit zu halten.

Für die Entwicklung einer effizienten und ressourcenschonenden Strategie ist nicht nur das Wissen über den Zustand der Infrastruktur im Status quo ausschlaggebend, sondern auch ein fundiertes Verständnis des Verschlechterungsverhaltens der Infrastruktur sowie der Wirkung von Erneuerungs- und Instandhaltungsmaßnahmen auf den künftigen Zustand. Bei Instandhaltungsmaßnahmen ist es bis dato nur eingeschränkt gelungen, die Wirkung auf den Zustand aus den verfügbaren Daten zu quantifizieren.

Ziel ist die Entwicklung eines Ansatzes, der die Auswirkungen von Instandhaltungsmaßnahmen auf den künftigen Netzzustand ganzheitlich quantifiziert. Zentral ist dabei die modelltechnische Abbildung ihrer Wirkung auf die Zustandsbewertung, insbesondere da diese oft zeitlich versetzt wirksam wird und sich mit anderen Maßnahmen sowie Verschleißprozessen überlagern kann. (vgl. Abbildung 1). Der Netzzustand wird dabei durch die Netzzustandsnote beschrieben.



Abbildung 1: Abhängigkeit des Zustands von Maßnahmen und Alterung

Der zeitliche Versatz der Wirkung ergibt sich aus der aktuellen Methodik der Netzzustandsnotenlogik: So zielen bestimmte Instandhaltungsmaßnahmen beispielhaft auf die Reduktion von Störmeldungen ab und tragen zur Senkung der Störanfälligkeit bei. Aufgrund der Jahresscheibenbetrachtung werden die bis zum Zeitpunkt der Maßnahme im Betrachtungsjahr angefallenen Störmeldungen jedoch vollständig in die Berechnung der Netzzustandsnote einbezogen. Die Wirkung der Maßnahme – ein Rückgang der Störmeldungen – wird daher erst zeitversetzt in den Netzzustandsnoten der darauffolgenden Jahresscheiben sichtbar. In diesem Sinne äußert sich der Effekt primär in einer Vermeidung zukünftiger Störereignisse, was die rückwirkende Erfassung und Bewertung der Maßnahme erschwert. Zusätzlich setzt sich auch nach der Instandhaltungsmaßnahme der natürliche Verschleiß der Anlage weiter fort, was die Störanfälligkeit der Anlage wieder negativ beeinflussen und die Wirkung der Maßnahme partiell kompensieren kann.

Durch die Verknüpfung von Verschleißansätzen mit der Wirkung von Instandhaltungs- und Erneuerungsmaßnahmen wird ein Werkzeug entwickelt, das zukünftig als datenbasiertes Entscheidungsinstrument eine zielgerichtete Planung unterstützen soll. Dabei sollen die Wechselwirkungen zwischen Erneuerungs- und Instandhaltungsmaßnahmen, analysiert und bewertet werden. Dies ermöglicht ein fundiertes Verständnis für die gegenseitigen Einflüsse auf den Netzzustand. Darüber hinaus können kurzfristige Instandhaltungsmaßnahmen identifiziert werden, die wahrscheinlich zur Erreichung eines definierten Netzzustands beitragen können. In weiteren Schritten soll der Ansatz verfeinert und somit für mikroskopische Analysen, Prognosen und eine zielgerichtete Planung passfähig werden.

Hierfür wird zunächst in Abschnitt 2 die Methodik der Netzzustandsnote als Grundlage für die Abbildung des Anlagenzustands beschrieben. In Abschnitt 3 wird auf Basis der Netzzustandsnote und des Anlagenalters eine Verschleißrate für die Anlagen ermittelt. Anschließend erfolgt in Abschnitt 4 die Herleitung der Instandhaltungswirkung auf die Netzzustandsnote mittels verschiedener Ansätze. Der Verschleiß und die Maßnahmenwirkung werden in ein Softwaretool implementiert. Mit dem Softwaretool werden verschiedene Szenarien berechnet und Aussagen zu den Wechselwirkungen zwischen Instandhaltungs- und Erneuerungsmaßnahmen auf den Zustand des Schienennetzes getroffen (vgl. Abschnitt 5). Abschließend erfolgt in Abschnitt 6 eine Zusammenfassung der Ergebnisse und ein Ausblick in weitere Handlungsstränge.

2 Netzzustandsnote

Der Zustand von Fahrweganlagen kann mithilfe der Netzzustandsnotenlogik präzise beschrieben und quantifiziert werden. Die Logik wurde im Rahmen eines Gemeinschaftsprojekts zwischen der DB InfraGO AG, der AFRY Schweiz AG und dem Verkehrswissenschaftlichen Institut der RWTH Aachen entwickelt und wird im Rahmen eines Weiterentwicklungsprozesses kontinuierlich geschärft. Seit der Entwicklung wird der Anlagenzustand einmal jährlich mittels der Netzzustandsnotenlogik im Netzzustandsbericht veröffentlicht. [1] Dabei wird der Zustand einer Einzelanlage auf einer Notenskala veranschaulicht, die von Note 1 für *neuwertig* bis hin zu Note 5 für *mangelhaft* reicht (vgl. Abbildung 2).

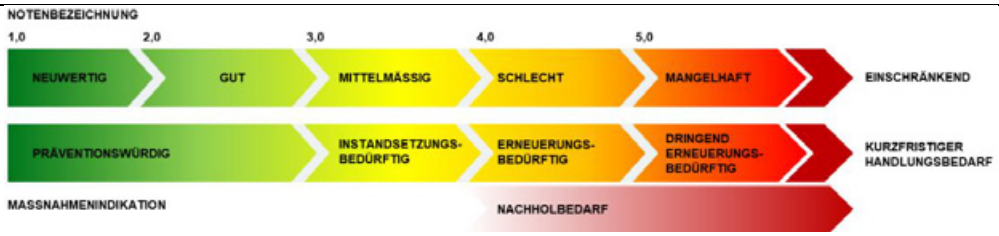


Abbildung 2: Bewertungsskala der Netzzustandsnote [1]

Die Zustandsnote einer Infrastrukturanlage ergibt sich aus der gewichteten Zusammenführung der vier Qualitätsdimensionen Inspektion, Verfügbarkeit, Lebenszyklus und besondere Sachverhalte. Jede Dimension wird durch spezifische Merkmale bewertet. Die Merkmale und deren Gewichtungen sind für jeden Anlagentyp festgelegt, um anlagenspezifische Eigenschaften zu berücksichtigen. Bei den Gleisen werden exemplarisch die Gleislage in Form der Längshöhenabweichung (Inspektion), die Anzahl der Störmeldungen (Verfügbarkeit) und die Restlebensdauer des Gleises (Lebenszyklus) als Merkmal betrachtet. In den Qualitätsdimensionen Inspektion und Verfügbarkeit kann grundsätzlich zwischen messtechnisch erhobenen Merkmalen (z.B. Gleislage, Zustand Schiene) und manuell angelegten Meldungen (z.B. Befundmeldungen, Störmeldungen, Vormerkmeldungen) differenziert werden.

Für eine Einzelanlage kann die Zustandsnotenlogik mittels eines Notenbaums dargestellt werden, in dem neben den Merkmalen selbst auch die Aggregation der Merkmale und der Qualitätsdimensionen enthalten sind. In Abbildung 3 ist eine vereinfachte Version des Notenbaums für Gleise dargestellt.

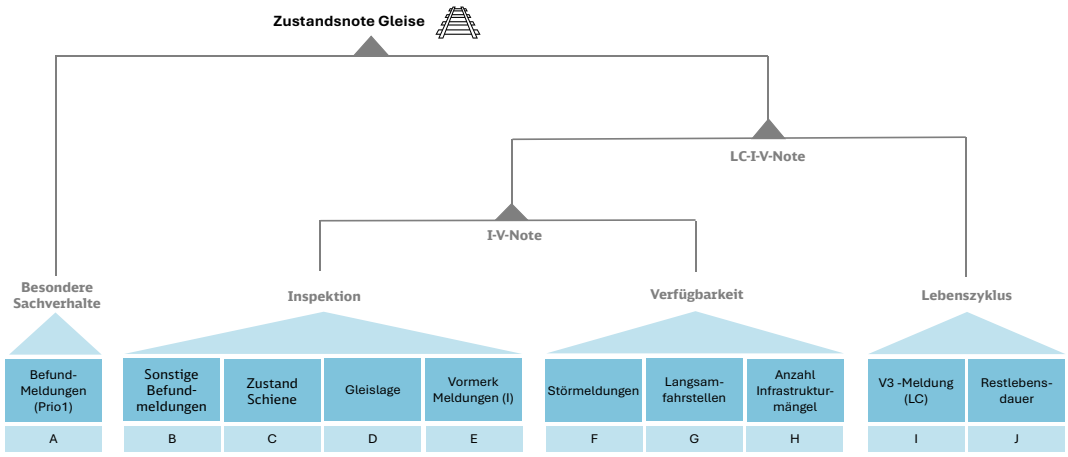


Abbildung 3: Notenbaum der Gleise

Neben der Aggregation innerhalb einer Anlage kann die Zustandsnote zudem anlagenübergreifend aggregiert werden. Damit ist es möglich, den Anlagenzustand für Netzabschnitte wie Strecken, Korridore, Regionen oder das Gesamtnetz mittels einer

mittleren Netzzustandsnote zu beschreiben. Für die Aggregation wird ein gewichteter Mittelwert der Zustandsnoten aller betrachteten Anlagen berechnet, wobei das Gewicht jeder einzelnen Zustandsnote durch den Quotienten aus dem Wiederbeschaffungswert und der technischen Nutzungsdauer der jeweiligen Anlage definiert ist. [2]

3 Anlagenverschleiß auf Basis der Netzzustandsnote

Für die Prognose einer gesamtheitlichen Entwicklung des Anlagenzustands ist die Kenntnis über den Verschleiß der Anlagen elementar. Der Verschleiß auf Ebene der Zustandsnote wird dabei mit einer linearen Regression über das Alter der Anlagen ermittelt. Dafür wird für jede Alterskohorte die durchschnittliche Zustandsnote berechnet und anschließend eine univariate lineare Regression durch die gebildeten durchschnittlichen Zustandsnoten durchgeführt. Dabei werden jeweils nur die Anlagen berücksichtigt, die sich innerhalb ihrer technischen Nutzungsdauern befinden, um den „Überlebens“-Bias, also die Berücksichtigung von überrepräsentierter Anzahl von alten Anlagen im guten Zustand zu vermeiden: Nach dem Erreichen der technischen Nutzungsdauern werden überwiegend Anlagen im schlechten Zustand erneuert, sodass vermehrt Anlagen in guten Zuständen überleben. [3] Zudem wird der Intercept der Regressionsfunktion durch die Auswertung der durchschnittlichen Zustände neuer Anlagen (Alter 0-2) festgelegt.

Aus der Steigung der linearen Regressionsfunktion lässt sich eine Verschleißrate in Abhängigkeit des Alters für die Zustandsnote ableiten. Diese Auswertung wird für jede Anlagenklasse durchgeführt und ist in Abbildung 4 beispielhaft für Gleise dargestellt.

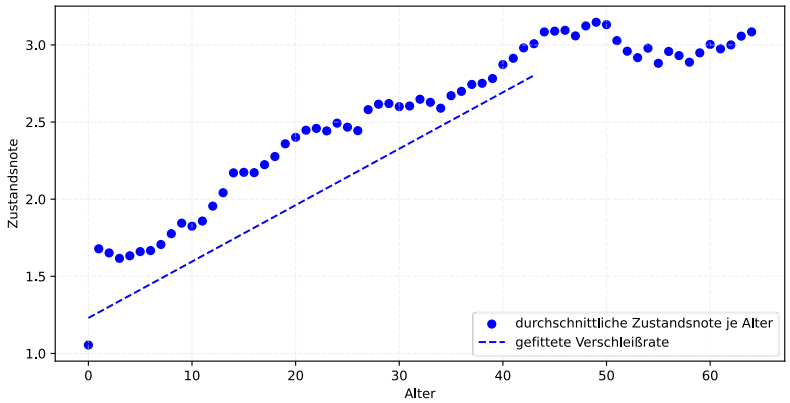


Abbildung 4: Lineare Regression über die Zustandsentwicklung für Gleise in Abhängigkeit vom Alter

Die Regression wird dabei durch eine lineare Funktion gebildet und ermöglicht damit eine konstante Verschleißrate über das Anlagenalter anzusetzen.

4 Herleitung des Wirkzusammenhangs zwischen Maßnahmen und Anlagenzustand

Für die Abbildung der Zusammenhänge zwischen Anlagenzustand und Instandhaltungsmaßnahmen sind neben dem Verschleiß einer Anlage auch die Wirkung spezifischer Instandhaltungs- und Erneuerungsmaßnahmen zu modellieren. Für die verschiedenen Anlagenklassen kann zwischen verschiedenen Maßnahmenbündeln differenziert werden. Diese werden als Aktivität zum Erhalt bzw. zur Verbesserung des Anlagenzustands durchgeführt. Bei Gleisen wird im Rahmen der Analysen zwischen folgenden vier Maßnahmenbündeln unterschieden:

- Gleiserneuerung
- Schienenerneuerung
- Verbesserung der Gleislage
- Beseitigung von Schienenfehler

Die Gleiserneuerung entspricht einer Erneuerungsmaßnahme mit vollständigem Ersatz des Oberbaus einschließlich Verbesserung der Gleisanlage. Bei der Schienenerneuerung hingegen werden nur die Schienen vollständig getauscht. Die anderen Maßnahmenbündel stellen Instandhaltungsmaßnahmen dar und können verschiedene Maßnahmen umfassen. So können beispielsweise unter dem Maßnahmenbündel „Verbesserung der Gleislage“ auch Schwellenwechsel und die Bettungsreinigung gefasst werden, da diese Maßnahmen Aktivitäten enthalten, die auf die Verbesserung der Gleislage durch Stopfen des Gleises abzielen. Neben den einzelnen Maßnahmenbündeln treten auch Kombinationen aus den Maßnahmenbündeln auf, sodass beispielsweise sowohl die Schienenerneuerung als auch eine Schwellenerneuerung an einer Gleisanlage stattfinden.

Weiterhin ist zu beachten, dass nicht nur unterschiedliche Maßnahmen verschiedene Wirkungen auf den Zustand der Anlage, sondern dieselbe Maßnahme an verschiedenen Anlagen eine unterschiedliche Wirkung auf den Zustand haben kann. Unterschiedliche Wirkungen treten beispielsweise aufgrund unterschiedlicher Ausgangszustände und verschiedener Einflussfaktoren, wie örtlicher Gegebenheiten, oder Ausführungsqualitäten der Maßnahmen auf. Daher wird eine Spannweite der Maßnahmenwirkung je Maßnahmenbündel angenommen. Dies erfolgt durch die Definition von zwei Extremszenarien – „Best-case“ und „Worst-case“ – für jedes Maßnahmenbündel. Diese Szenarien repräsentieren die Bandbreite möglicher Auswirkungen auf die jeweiligen Zustandsmerkmale und ermöglichen eine Einschätzung der Maßnahmenwirkung unter Berücksichtigung unterschiedlicher Rahmenbedingungen.

Dafür wird für jedes Maßnahmenbündel (1-4) eine Wirkung auf die verschiedenen Merkmale der Zustandsnote (A-O vgl. Abbildung 3) in den beiden Szenarien „Best-case“ und „Worst-case“ definiert. Mittels der Veränderung der einzelnen Merkmale kann die Wirkung auf die Zustandsnote der Anlage berechnet und anschließend eine Differenz zwischen dem neuen (ZN neu) und dem alten Zustand (ZN alt) ermittelt werden (vgl. Abbildung 5).

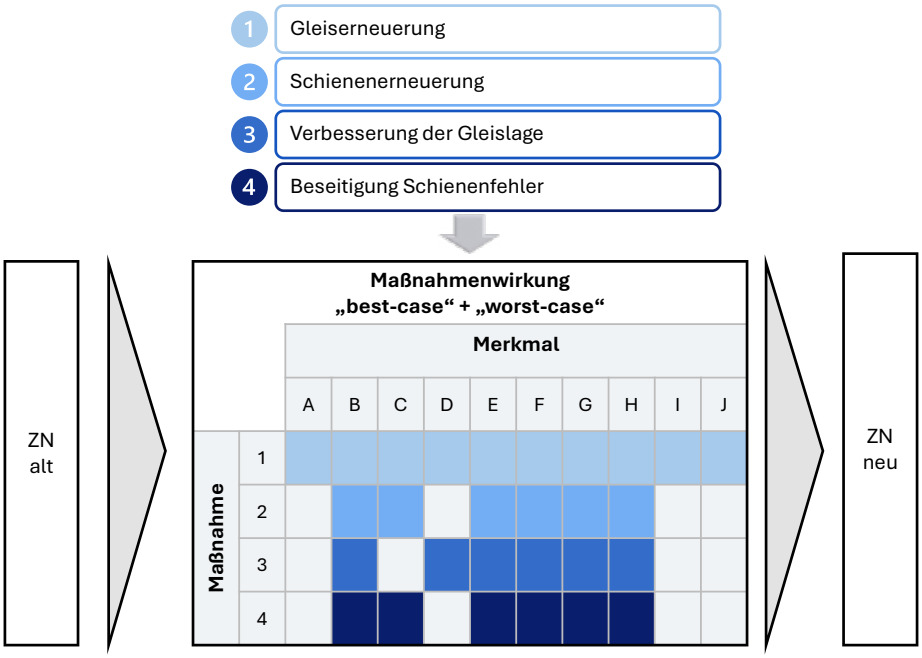


Abbildung 5: Konzept zur Maßnahmenwirkung

Während die Gleiserneuerung aufgrund ihrer vollumfänglichen Sanierung Einfluss auf alle Zustandsmerkmale hat, unterscheiden sich die Instandhaltungsaktivitäten hinsichtlich der Wirkung auf Merkmalsebene. Zur Quantifizierung der Wirkung der Maßnahmen werden zwei heuristische Ansätze herangezogen: Ein richtlinienbasierter Ansatz sowie ein statistischer Ansatz.

Der richtlinienbasierte Ansatz findet Anwendung bei denjenigen Merkmalen der Gleisanlage, die messtechnisch erfasst werden und für die normierte Abnahmewerte gemäß den Richtlinien der DB InfraGO AG vorliegen. [4] Diese Abnahmewerte definieren den erforderlichen Mindestzustand nach Durchführung einer Maßnahme. Wenn eine Maßnahme ein entsprechendes Merkmal beeinflusst, wird im ungünstigsten Fall (Worst-case) davon ausgegangen, dass der Zustand nach Umsetzung der Maßnahme exakt dem festgelegten Abnahmewert entspricht. Im günstigsten Fall (Best-case) wird ein idealer Zustand (z.B. keine Längshöhenabweichung) angenommen. Dieser Ansatz ist jedoch nur für messbare Merkmale mit definierten Abnahmekriterien anwendbar.

In allen anderen Fällen wird ein statistischer Ansatz verwendet. Für Merkmale wie Befundmeldungen oder Störmeldungen existieren keine entsprechenden Grenz- oder Abnahmewerte. Befundmeldungen dokumentieren im Rahmen von Inspektionen identifizierte Schäden oder Unregelmäßigkeiten an der Anlage. Störmeldungen hingegen erfassen Funktionseinschränkungen oder Ausfälle von Anlagen.

Die grundlegende Annahme hierbei ist, dass eine Maßnahme, die auf ein bestimmtes Schadensbild oder Merkmal wirkt, auch sämtliche damit verknüpfte Meldungen beseitigt. Die Zuordnung der Meldungen zu den Schadensbildern erfolgt anhand von Ursachencodes, die systematisch erfasst werden und der Meldung zugehörig sind. Darüber lässt sich abgrenzen, ob eine Meldung beispielsweise auf einen Schienenfehler oder die Gleislage zurückzuführen ist.

Zur Quantifizierung des Worst-Case-Szenarios wird in historischen Daten für jedes Maßnahmenbündel der prozentuale Anteil der Meldungen ermittelt, die einem spezifischen Schadensbild zugeordnet werden können. Dieser Anteil wird in Relation zur Gesamtzahl aller Meldungen gesetzt. Für die Best-Case-Betrachtung erfolgt eine differenzierte Analyse auf Ebene der einzelnen Anlagen. Dabei wird der durchschnittliche prozentuale Anteil der Meldungen berechnet, die einem bestimmten Schadensbild zugewiesen sind – bezogen auf die Gesamtheit der Meldungen pro Anlage. In die Analyse einbezogen werden ausschließlich Anlagen, bei denen mindestens eine Meldung dem jeweiligen Schadensbild zugeordnet wurde.

Die auf diese Weise ermittelten prozentualen Anteile zeigen, wie viele Meldungen durch die jeweils angewandte Maßnahme potenziell vermieden werden können. Die Wirksamkeit der Maßnahmen wird demnach über eine modellhafte Reduktion der Meldungsanzahl in Höhe der aus historischen Daten ermittelten prozentualen Werte abgebildet.

Die Wirkung einzelner Maßnahmen auf Anlagenebene kann direkt beschrieben werden. Um darüber hinaus eine generalisierte Wirkung auf die Zustandsnote abzuleiten, wird jede Maßnahme exemplarisch auf eine Stichprobe geeigneter Anlagen angewendet. Die resultierenden Wirkungen auf die einzelnen Anlagen werden anschließend durch Bildung des gegenüber Ausreißern stabilen und für ordinalskalierte Daten geeigneten Medians aggregiert. Zur Abbildung möglicher Schwankungsbreiten werden sowohl der Best-Case- als auch der Worst-Case-Median ermittelt. Dadurch lässt sich für jede Maßnahme ein potenzieller Wirkungskorridor definieren. Die Auswahl der Stichprobe erfolgt maßnahmenspezifisch anhand festgelegter Kriterien, beispielsweise dem Vorhandensein von Störmeldungen. (vgl. Abbildung 6)

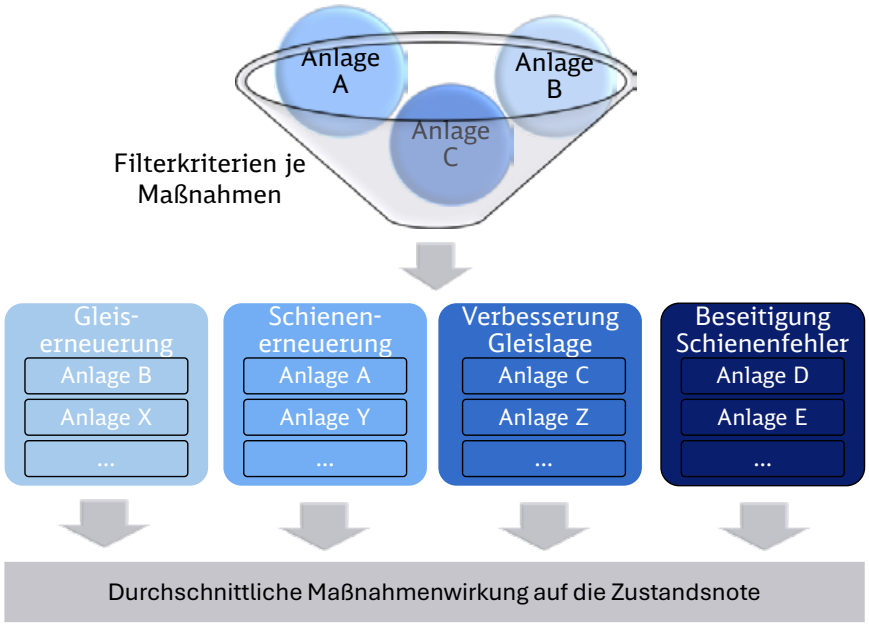


Abbildung 6: Konzept zur Bildung der Stichproben

Auf Grundlage dieser Vorgehensweise ergeben sich die in Tabelle 1 dargestellten mittleren Effekte der Maßnahmen auf die Zustandsnote der Gleise. Zu beachten ist hier, dass keinerlei gegenläufige Verschleißeffekte betrachtet werden. Es handelt sich hier nur um die Positivwirkung der Maßnahme.

Tabelle 1: Verbesserung der Zustandsnote bei den Gleisen durch Maßnahmen

Maßnahme	Best-case	Worst-case
Gleiserneuerung	3,74	3,54
Schienenerneuerung	0,86	0,02
Verbesserung Gleislage	0,70	0,03
Beseitigung Schienenfehler	0,39	0,01

Dabei zeigt sich, dass die Erneuerung der Gleisanlagen den größten positiven Einfluss auf die Zustandsnote hat. Diese hohe Wirksamkeit ist insbesondere auf die Reduktion des Anlagenalters zurückzuführen und wirkt sich damit deutlich auf die Qualitätsdimension „Lebenszyklus“ aus. Im Vergleich dazu zeigen Instandhaltungsmaßnahmen eine geringere Wirkung. Insbesondere im Worst-Case-Szenario ist ihre Effektivität kaum sichtbar. Dieses Ergebnis lässt sich unter anderem durch die Methodik der Netzzustandsnote erklären, bei der verschiedene Merkmale miteinander verrechnet werden und somit punktuelle Verbesserungen einzelner Merkmale nur begrenzt zum Tragen kommen.

5 Ergebnisse

Sowohl der Verschleiß als auch die Maßnahmenwirkung werden in einem selbstprogrammierten Softwaretool implementiert, mit dem der Anlagenzustand unter Vorgabe der Erneuerungs- und Instandhaltungsmengen prognostiziert werden kann. Mit diesem Softwaretool können verschiedene Szenarien verglichen und Rückschlüsse auf die Wirksamkeit und die Wechselwirkungen zwischen den Maßnahmenarten gezogen werden.

Dafür werden fiktive Szenarien am Beispiel der Gleise erarbeitet, um Effekte auf die Netzzustandsnote darzustellen. Zunächst wird eine Zufallsstichprobe an Anlagen, die nur aus Gleisen besteht, gebildet. Anschließend werden verschiedene Mengenkombinationen aus jährlich durchgeführten Erneuerungs- und Instandhaltungsmaßnahmen angesetzt und ein Rechendurchlauf bis zum Jahr 2035 für jede Mengenkombination durchgeführt. Dabei werden die Maßnahmenmengen in jedem Simulationsjahr durchgeführt und bei den Instandhaltungsmengen gleichmäßig auf die verschiedenen Instandhaltungsaktivitäten verteilt. Es werden verschiedene Szenarien mit anteiligen Erneuerungs- bzw. Instandhaltungsmengen und einer Kombination aus beidem gebildet und als Input für das Tool genutzt. Insgesamt werden 25 verschiedene Szenarien betrachtet. Dabei ist zu berücksichtigen, dass die Auswahl der Anlagen, denen eine Maßnahme zugewiesen wird, einen erheblichen Einfluss auf die Ergebnisse hat. In den durchgeführten Szenarien erfolgt die Verteilung der Maßnahmen auf die Gleise derart, dass vorrangig jene Anlagen berücksichtigt werden, die sich in einem besonders schlechten Zustand befinden. Dies stellt eine effektive Form der Steuerung dar, die jedoch unter den gegebenen Rahmenbedingungen lediglich theoretisch realisierbar ist.

Für jedes dieser Szenarien wird die Netzzustandsnote der Gleise sowohl im Best-case als auch im Worst-case für die gebildete Stichprobe je Simulationsjahr ermittelt. Die Netzzustandsnote im Jahr 2035 wird als Funktion von jährlicher Instandhaltungs- und Erneuerungsmenge dargestellt. Für das Jahr 2035 ergibt sich im Best-case die Abbildung 7. Im Ausgangsjahr 2024 hat die Stichprobe eine aggregierte Netzzustandsnote von 2,89, die als Ebene in der Abbildung dargestellt ist.

Die Analyse zeigt erwartungsgemäß, dass Erneuerungsmaßnahmen grundsätzlich eine höhere Wirksamkeit auf die Verbesserung des aggregierten Zustandsniveaus von Infrastrukturanlagen haben als reine Instandhaltungsmaßnahmen.

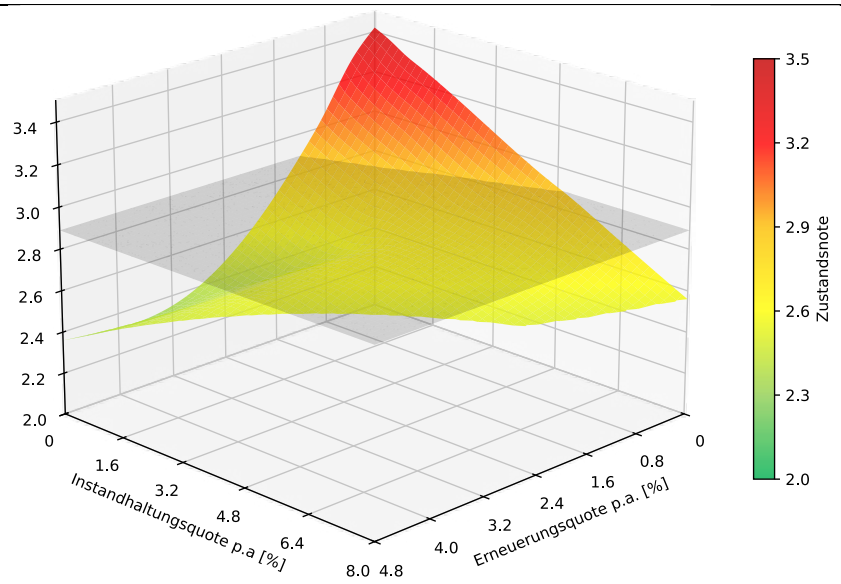


Abbildung 7: Netzzustandsnote der Gleise in Abhängigkeit der jährlichen Erneuerungs- und Instandhaltungsmengen im Jahr 2035 bei der Best-case Wirkung

Bei ausschließlicher Durchführung von Erneuerungsmaßnahmen lässt sich im Best-case ab einer jährlichen Erneuerungsquote von ca. 1,2 % der Gesamtanlagemenge eine Verbesserung der aggregierten Netzzustandsnote feststellen. Im Gegensatz dazu ist bei ausschließlicher Durchführung von Instandhaltungsmaßnahmen eine deutlich höhere jährliche Maßnahmenmenge von etwa 4,8 % erforderlich, um eine Verbesserung der Netzzustandsnote zu erzielen. Dabei wird unterstellt, dass die Instandhaltungsmaßnahmen gleichmäßig über die unterschiedlichen Maßnahmen verteilt sind. Ein weiterer Effekt zeigt sich ab einer jährlichen Erneuerungsquote von etwa 2,4 % (ohne Instandhaltung), von der die Verbesserung der Netzzustandsnote merklich abflacht. Diese Entwicklung lässt sich darauf zurückführen, dass mit steigender Erneuerungsmenge weniger Anlagenteile mit schlechtem Zustand zur Verfügung stehen. Hier kann der maximale Verbesserungseffekt von 3,74 nicht mehr vollständig ausgeschöpft werden, da die Notenverbesserung durch die beste Note von 1,0 begrenzt ist.

Aus ökonomischer und technischer Sicht ist die alleinige Anwendung einer einzelnen Maßnahmenart nicht zielführend. Eine rein instandhaltungsorientierte Bewirtschaftung bspw. würde langfristig zu einer starken Überalterung der Infrastruktur führen, da Instandhaltungsmaßnahmen keinen Einfluss auf das Anlagenalter nehmen. Die Überalterung von Anlagen führt zu steigenden Instandhaltungsbedarfen [5] und nicht abzuschätzenden Folgen für die Zustandsentwicklung der Infrastrukturanlagen. Nur durch Erneuerung kann das Durchschnittsalter des Anlagenportfolios gesenkt werden.

Zudem können auch die langfristigen Effekte des Ausbleibens von Maßnahmen im Modell nur begrenzt abgebildet werden. Daher erweist sich eine kombinierte Strategie aus Erneuerung und Instandhaltung als die effektivste und wirtschaftlich sinnvollste Lösung.

Zur Bewertung der Wechselwirkungen zwischen beiden Maßnahmenarten werden Szenarien mit ausschließlich Erneuerungs- bzw. ausschließlich Instandhaltungsmaßnahmen analysiert. Hierzu wurde die Veränderung der Netzzustandsnote in Abhängigkeit von den jeweiligen Maßnahmenmengen und über die Jahre mittels linearer Regression ausgewertet. Die resultierenden Steigungskoeffizienten der Regressionsgeraden wurden miteinander ins Verhältnis gesetzt. Im günstigsten Fall ergibt sich ein Wirkungsverhältnis von 3,9. Das bedeutet, dass etwa die vierfache Menge an Instandhaltungsmaßnahmen notwendig wäre, um im hier betrachteten Spektrum der Netzzustandsnote den gleichen Effekt auf die Netzzustandsnote zu erzielen wie mit Erneuerungsmaßnahmen.

Für Instandhaltungsmaßnahmen wurde eine Spannbreite möglicher Wirkungen ermittelt, da deren Effektivität stark von örtlichen Gegebenheiten und den Zuständen der betroffenen Anlagen abhängt. Um dieser Unsicherheit Rechnung zu tragen, wurde die zuvor beschriebene Auswertung nicht nur für den Best-Case, sondern auch für den Worst-Case durchgeführt. Die Ergebnisse dieser Analyse sind für den Worst-Case in Abbildung 8 dargestellt. Diese zeigen deutlich, wie stark die Effektivität von Instandhaltungsmaßnahmen im Vergleich zum Best-Case (vgl. Abbildung 7) variieren kann.

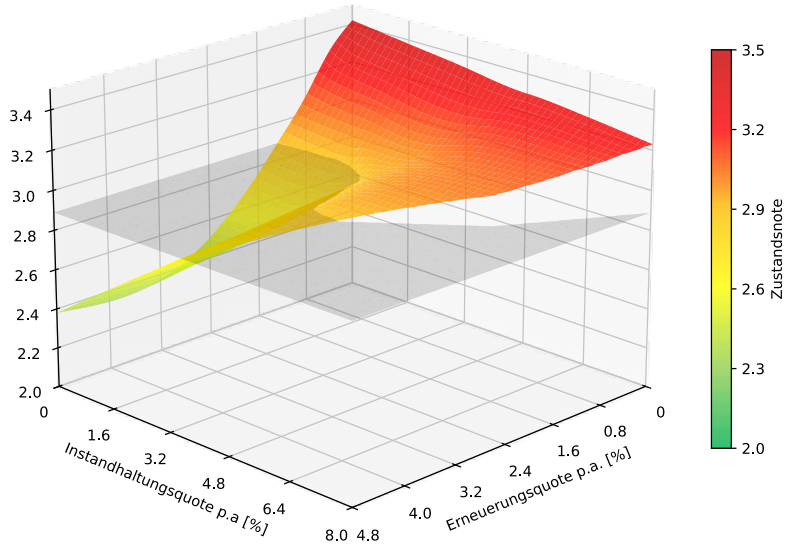


Abbildung 8: Netzzustandsnote der Gleise in Abhängigkeit der Erneuerungs- und Instandhaltungsmengen im Jahr 2035 bei der Worst-case Wirkung

Im Worst-Case-Szenario zeigt sich bei den Instandhaltungsmaßnahmen eine deutlich geringere Wirkung auf die Netzzustandsnote. In den durchgeführten Berechnungen konnte in diesem Fall keine Verbesserung der aggregierten Netzzustandsnote allein durch Instandhaltungsmaßnahmen erzielt werden. Dies liegt an der geringen Maßnahmenwirkung im Worst-Case. Im Gegensatz dazu fällt die Spannweite der Wirkungen bei Erneuerungsmaßnahmen deutlich geringer aus, was auf eine konsistentere Wirksamkeit dieser Maßnahmenart hinweist. Entsprechend steigt auch der Wechselwirkungsfaktor, der das Verhältnis der erforderlichen Maßnahmenmengen beschreibt, im Worst-Case auf 20,1 an. Somit wäre im Worst-case die 20-fache Menge an Instandhaltungsmaßnahmen notwendig, um im betrachteten Spektrum der Netzzustandsnote denselben Effekt auf die Netzzustandsnote zu erzielen wie mit Erneuerungsmaßnahmen.

Der resultierende Korridor der Wechselwirkungen, also die Spannweite zwischen Best-Case (Faktor 3,9) und Worst-Case (Faktor 20,1), ist damit als sehr groß zu bewerten. Diese hohe Spannweite ist primär auf die Unsicherheit in der Wirkungsabschätzung von Instandhaltungsmaßnahmen zurückzuführen. Die breite Streuung der Wirkung verdeutlicht, dass eine exakte Vorhersage der Zustandsveränderung auf Einzelanlagenebene derzeit nur eingeschränkt möglich ist. Der gezeigte Ansatz eignet sich für eine erste Abschätzung größerer Stichproben.

Eine vielversprechende Möglichkeit zur Reduktion dieser Unsicherheit besteht in der detaillierteren Modellierung auf Merkmalsebene. Durch eine differenzierte Betrachtung einzelner Anlagen und deren Merkmalsausprägung können präzisere Aussagen zur Maßnahmenwirkung auf Ebene der Anlagen getroffen und somit die Schwankungsbreite reduziert werden.

6 Zusammenfassung und Ausblick

Die zukünftige Entwicklung des Infrastrukturzustands ist maßgeblich von Instandhaltungs- und Erneuerungsaktivitäten abhängig. Durch solche Maßnahmen ist es möglich, dem Verschleiß der Anlage entgegenzuwirken und den Zustand der Anlage zu halten oder zu verbessern. Mit der Netzzustandsnote besteht bereits ein Instrument zur Beschreibung des Infrastrukturzustands. Mit den gezeigten Auswertungen ist es möglich, die Maßnahmenwirkung auf Ebene der Netzzustandsnote abzubilden und Vorschauen für verschiedene Planungsszenarien zu erstellen. Dabei können verschiedene Instandhaltungsmaßnahmen berücksichtigt werden, was es ermöglicht,

Wechselwirkungen zwischen verschiedenen Maßnahmenarten zu analysieren und zu quantifizieren.

In einer ersten Analyse auf Zustandsnotenebene zeigt sich, dass die Wirkung von Instandhaltungsmaßnahmen in Abhängigkeit von Ausgangszustand und örtlichen Gegebenheiten sehr verschieden ausfallen kann. Erste Aussagen über Wechselwirkungen zwischen den Maßnahmenarten lassen sich bereits mit dem derzeitigen Softwaretool treffen. Obwohl diese Wechselwirkungen mit den gezeigten Ansätzen quantifiziert werden können, zeigt sich noch eine große Spannweite möglicher Wechselwirkungsfaktoren.

Bisher wurden in diesem Tool sowohl der Verschleiß als auch die Maßnahmenwirkung auf Ebene der Zustandsnote dargestellt. Dies ist ein vereinfachter Ansatz, der auf heuristischen Ansätzen beruht. Sofern Aussagen über kleinere Netzausschnitte oder Einzelanlagen getroffen werden sollen, ist ein mikroskopischer Ansatz zielführender. Daher gilt es in einem nächsten Schritt, die Wirkung und den Verschleiß in einem Tool zu implementieren, das die Prognose des Zustands auf Merkmalsebene durchführt und damit eine detailliertere und mikroskopische Betrachtung ermöglicht.

Im Rahmen des Projekts „Ursache-Wirkung-Zusammenhang“ wurden bereits Verschleißmodelle für verschiedene Zustandsmerkmale entwickelt und in ein Prognosetool integriert [6]. Damit können Alterungs- und Verschleißprozesse für einige Merkmale jetzt schon auf Merkmalsebene vorhergesagt werden. Dieses Prognosetool ist um alle Wirkungen von Maßnahmen auf Merkmalsebene zu erweitern, um umfangreiche Prognosen durchführen zu können und Auswertungen auf Anlagenebene zu ermöglichen.

Literatur

[1] DB InfraGO AG, „Netzzustandsbericht DB Netz AG 2021 - Fokus Netzzustandsnote,“ Frankfurt a.M., 2022.

[2] DB InfraGO AG, „InfraGO-Zustandsbericht Netz und Personenbahnhöfe 2024,“ Frankfurt a.M., 2025.

[3] M. Heckmann, L. Elfert, S. Höhn, C. Schubert, H. Aktas, F. Dittberner, F. Grob, F. Lampe und N. Nießen, „Ursache-Wirkungs-Zusammenhang zwischen Mitteleinsatz und erzielbarer Anlagenqualität am Beispiel von Verkehrsstationen,“ in *IRSA 2021: Tagungsband, Proceedings: 3rd International Railway Symposium Aachen*, Aachen, 2022.

[4] DB InfraGO AG, „Richtlinie 8240550A01 - Oberbauarbeiten durchführen," 2018.

[5] A. Blazey, F. Gonguet und P. Stokoe, „Chapter 14 Maintaining and Managing Public Infrastructure Assets," in *Well Spent: How Strong Infrastructure Governance Can End Waste in Public Investment*, International Monetary Fund, 2020, p. 344.

[6] S. Bluhm, C. Wink, B. Dickenbrok, A. Helfenstein und N. Nießen, „Qualität der Eisenbahninfrastruktur - Eine methodische Herangehensweise zur umfassenden Bestimmung des Infrastrukturzustands," in *IRSA 2021: Tagungsband, Proceedings: 3rd International Railway Symposium Aachen*, Aachen, 2022.

Autoren



Lampe, Felix

Felix Lampe schloss 2018 sein Masterstudium des Bauingenieurwesens an der RWTH Aachen ab und arbeitete anschließend als Wissenschaftlicher Mitarbeiter am Verkehrswissenschaftlichen Institut der RWTH Aachen im Bereich der Eisenbahnbetriebswissenschaft. Seit 2024 ist er als Experte Netzzustandsnote bei der DB InfraGO AG tätig.



Doose, Felix

Nach seinem Masterabschluss in Mathematik an der Rheinischen Friedrich-Wilhelms-Universität Bonn im Jahr 2019, arbeitete Felix Doose zuerst bei der KPMG AG. 2021 stieg er bei der DB AG ein. Erst war er dort als Trainee bei der Konzernleitung tätig, bevor er 2023 als Data Scientist für Infrastrukturbewertung zur DB InfraGO AG wechselte.



Schaper, Christian

Christian Schaper schloss 2015 sein Masterstudium des Wirtschaftsingenieurwesens an der Leibniz Universität Hannover ab. Im Jahr 2016 nahm er seine Tätigkeit bei der Planung und Steuerung der DB Netz AG (heute: DB InfraGO AG) in München auf. Seit 2022 ist er als Experte Netzzustandsnote bei der DB InfraGO AG tätig.

Session 13

Felix Lampe, Felix Doose, Christian Schaper, Björn Dickenbrok



Dickenbrok, Björn

Prof. Dr.-Ing. Björn Dickenbrok studierte bis 2004 Bauingenieurwesen an der RWTH Aachen und promovierte anschließend am Verkehrswissenschaftlichen Institut der RWTH Aachen. Von 2010 bis 2025 arbeitete er bei der DB InfraGO AG und war dort zuletzt als Leiter „Infrastrukturzustandsbewertung“ in der Planung und Segmentsteuerung für die Netzzustandsnotenlogik verantwortlich. Seit Oktober 2025 ist er Professor für Bahningenieurwesen an der TH Mittelhessen in Gießen.

Session 13
Sebastian Ude, Alexander Rhein, Tobias Jacke,
Sascha Willgardt-Levin

Bedarfsermittlung bei der DB InfraGO AG – Chancen und Herausforderungen der Integration des Anlagenzustands

Eisenbahninfrastrukturmanagement

Ude, Sebastian[1], Dr. Rhein, Alexander[1], Dr. Jacke, Tobias[1],
Willgardt-Levin, Sascha[1]

[1]DB InfraGO AG

Zusammenfassung

Mit der Integration der Netzzustandsnote in die bisher altersbasierte Asset Simulation des 3-i Modells der DB InfraGO AG eröffnen sich neue Möglichkeiten bei der Berechnung des technischen Bedarfs. Dieser Beitrag beleuchtet die Herausforderungen bei der Erarbeitung eines Zusammenhangs zwischen Alter und Zustand des Anlagenkollektivs sowie die Chancen, die sich daraus ergeben.

Keywords: Bedarfsermittlung; 3-i Modell; Anlagenmanagement; Netzzustandsnote; Technischer Bedarf

1 Einleitung

Die DB InfraGO AG betreibt rund 33.000 km Schienennetz in Deutschland. Für die Bewirtschaftung eines Anlagenportfolios dieser Größe benötigt es eine Asset Management-Strategie. Das grundsätzliche Ziel einer solchen Strategie ist es, das Anlagenportfolio hinsichtlich des gesamten Lebenszyklus zu optimieren. Dieser beginnt mit der Planung und Beschaffung, dem anschließenden Betrieb sowie der begleitenden Instandsetzung und endet mit der Nutzungsdauer der Anlage [1].

Im Geschäftsbereich Fahrweg der DB InfraGO AG wird die 3-i Strategie als Asset Management-Strategie genutzt, welche für die integrierte Betrachtung der Instandhaltungs- und Investitionsstrategie steht. Diese wurde zwischen 2006 und 2008 im Projekt „3-i“ der

DB Netz AG erarbeitet, wobei erste gewerkespezifische Ansätze für Lebenszyklusmodelle entwickelt wurden. Im Projekt „Nachhaltigkeit 3-i“ erfolgte die Ausweitung und Verstetigung der Ansätze. Diese verwendet die DB InfraGO AG seit dem Jahr 2012 im 3-i Modell, welches als Asset Simulation die technischen Bedarfe für die Instandhaltung und die Ersatzinvestition des gesamten Netzes berechnet [2]. Aktuell beruhen die Lebenszyklusmodelle des 3-i Modells auf deren Interaktion mit dem Alter der Einzelanlagen. Im Rahmen des Projekts „Weiterentwicklung des 3-i Modells 2.0“ (WiiiM 2.0) wird der Anlagenzustand in Form der Netzzustandsnote (NZN) integriert, wodurch diese zukünftig die Steuerung der Simulation über das Alter ersetzen kann.

2 Motivation zur Integration der Netzzustandsnote

Mit der im Jahre 2021 entwickelten NZN besteht erstmals die Möglichkeit, die Einzelanlagen über deren Zustand statt deren Alter im Lebenszyklusmodell zu verorten. Die NZN entwickelte die DB InfraGO AG gemeinsam mit der RWTH Aachen sowie der Beratungsgesellschaft AFRY und veröffentlichte diese erstmals im Rahmen des Netzzustandsberichts des Jahres 2021 [3]. ÖBB, SBB und BaneNor verwenden vergleichbare Ansätze, welche mit derselben Beratungsgesellschaft entwickelt wurden. ÖBB [4] und SBB [5] veröffentlichen ihren jeweiligen Infrastrukturzustand ebenfalls im Rahmen von Netzzustandsberichten.

2.1 Altersbasiertes 3-i Modell

Die Grundlogik des 3-i Modells basiert auf Lebenszyklusmodellen und der Einordnung von Einzelanlagen über deren Alter in ebendiese. Ein Lebenszyklusmodell wird strategisch für ein Kollektiv aus Anlagen mit vergleichbaren Eigenschaften entworfen. Diese Kollektive werden im 3-i Modell Anlagencluster genannt. Als Merkmale für die Differenzierung der Cluster werden beispielweise für Weichen die zulässige Geschwindigkeit, die Belastung, der Weichenhalbmesser sowie die Schwellenart herangezogen. [6]

Session 13
Sebastian Ude, Alexander Rhein, Tobias Jacke,
Sascha Willgardt-Levin



Inhalte eines Lebenszyklusmodells

ID	Tätigkeit	Kompo-nente	Alter der Anlage									tND
			0	1	2	3	4	5	6	7	...	28
Einzelfehlerbeseitigung												
100	Stopfen	EFB	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	...	0,02
300	Herzstückwechsel	EFB	0,00	0,00	0,00	0,01	0,02	0,02	0,03	0,04	...	0,06
400	Radlenkerwechsel	EFB	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	...	0,06
...	Weitere Tätigkeiten	EFB									...	
Entstörung												
9210	Entstörung	EN	0,00	0,01	0,03	0,06	0,09	0,13	0,17	0,22	...	2
Inspektion und Wartung												
9110	Inspektion & Wartung	IW	2,50	2,50	2,50	2,50	2,50	2,50	2,50	2,50	...	2,50
Prävention												
100	Stopfen	PR						1,00			...	
...	Weitere Tätigkeiten	PR									...	
Investition												
9653	Vollerneuerung	VE	1,00								...	

Lebenszyklusmodell

Das Lebenszyklusmodell gilt für ein Kollektiv von Anlagen. Die Anlagen eines Gewerks mit vergleichbaren betrieblichen sowie bautechnischen Eigenschaften bilden ein Anlagencluster.

Tätigkeiten

Die Tätigkeiten bilden das „Scheckheft" der Anlage ab. Je Tätigkeit und Anlagenalter ist der Anteil der zu behandelnden Anlagen hinterlegt.

Technische Nutzungsdauer (tND)

Die tND beschreibt den mittleren, optimalen Investitionszeitpunkt. Sobald das Alter einer Anlage die tND überschreitet, ist die Anlage im Investitionsrückstau.

Komponenten

Die Komponenten bündeln die Tätigkeiten der Lebenszyklusmodelle nach der Art der Instandhaltung bzw. Investition.

Abbildung 1: Beispiel für ein Lebenszyklusmodell einer Weiche mit fiktiven Zahlen inkl. Erläuterungen [6]

Das strategische Lebenszyklusmodell eines Anlagenclusters besteht aus Jahresscheiben, in welchen das durchschnittliche Soll an notwendigen Instandhaltungstätigkeiten je Alter für das betroffene Anlagenkollektiv beschrieben wird. Zudem werden die Instandhaltungstätigkeiten nach Instandhaltungskomponenten differenziert (siehe Abbildung 1). Die investive Vollerneuerung ist der Startpunkt des Lebenszyklusmodells und die durchschnittliche technische Nutzungsdauer (tND) begrenzt es. Nach der initialen Einordnung

aller Anlagen aus der sogenannten Mengendatenbank in die Lebenszyklusmodelle durchläuft das 3-i Modell den Simulationszeitraum. Dabei wandern die Anlagen über die Fortschreibung des Alters durch die Lebenszyklusmodelle. Mithilfe der hinterlegten Preisdatenbank werden aus den technischen Bedarfen je Simulationsjahr Mittelbedarfe für die Instandhaltung sowie für die Ersatzinvestition errechnet. Der Bedarf an Ersatzinvestitionen ergibt sich aus den überalterten Anlagen. Eine Anlage gilt als überaltert, wenn diese die tND überschritten hat. Damit befindet sich die Anlage im Investitionsrückstau. [6]

Die Inhalte der Lebenszyklusmodelle basieren auf Datenanalysen, Erfahrungswerten, Herstellerangaben, Expertenwissen sowie dem Austausch mit anderen Infrastrukturunternehmen [2]. Im jährlichen Überarbeitungsprozess können Rückschlüsse aus der Umsetzung der Instandhaltung bzw. der Investition, technische Innovationen oder weiterer Verbesserungsbedarf in die Lebenszyklusmodelle eingearbeitet werden. Dazu werden kontinuierlich Themen gesammelt und priorisiert in die bestehenden Modelle eingepflegt. [6]

2.2 Notwendigkeit der Zustandsintegration

In der Gerontologie finden sich die Konzepte des funktionellen und des biologischen Alter, um den Alterungsprozess eines Menschen mithilfe von Biomarkern zu bewerten. Das chronologische Alter einer Person ist dabei nicht per se ein guter Indikator für den Alterungsprozess, insbesondere in den späteren Phasen der Lebensspanne [7]. Analog verhält es sich bei der Betrachtung von Eisenbahnanlagen. Das Alter der Einzelanlage beschreibt mitunter nicht die korrekte Position im Lebenszyklus. Begünstigende wie nachteilige Faktoren können diese Position einer Anlage beeinflussen.

Das altersbasierte 3-i Modell berücksichtigt den Zustand der Anlage in Form einer NZN nicht. Lediglich über das Auslaufregime wird ein mangelhafter Zustand durch eine Verfehlung von Instandhaltungsleistungen an den Anlagen pauschal abgebildet und diesen Anlagen eine kürzere tND zugewiesen. Dieser generalistische Ansatz, der ebenfalls an das Anlagenalter geknüpft ist, wird nun weiterentwickelt, da sich auf Ebene der Einzelanlagen unterschiedliche Extreme zeigen. Hier gibt es zum einen alte Anlagen, die aufgrund guter Instandhaltung und weiterer Faktoren in gutem Zustand sind. Zum anderen finden sich junge Anlagen, die aufgrund von Anfangsmängeln oder besonderen Sachverhalten bereits in so schlechtem Zustand sind, dass Investitionen zur Wiederherstellung der Verfügbarkeit unabdingbar werden. Daraus folgt die Erkenntnis, dass der tatsächlich vorliegende Zustand der Anlagen als Resultat der Instandhaltung und Belastung zur Einordnung in das Lebenszyklusmodell eingeführt werden sollte.

Mit der Einführung der NZN besteht erstmals die Möglichkeit, das altersbasierte in ein zustandsbasiertes Simulationsmodell zu überführen. Mithilfe der in der NZN abgebildeten Qualitätsdimensionen ist ein Schlüssel für die Implementierung des Zustands in das 3-i Modell entwickelt, der simpel, aber signifikant Indikatoren für die Verortung der Anlagen innerhalb des Lebenszyklus liefert. Somit können die individuellen Anlagenverhalten antizipiert und auf die weiterhin altersbasierten Lebenszyklusmodelle projiziert werden. Eine Entwicklung des Anlagenportfolios kann unter Voraussetzung der strategisch erwarteten und tatsächlich umgesetzten Maßnahmen abgeschätzt werden.

3 Integration der Netzzustandsnote in das 3-i Modell

Um die bestehenden Lebenszyklusmodelle der DB InfraGO AG nach Integration der NZN weiterzuverwenden, muss ein Soll-Zusammenhang zwischen Alter und NZN geschaffen werden. Über diesen Zusammenhang kann die erwartete Soll-NZN mit dem erfassten Ist-NZN verglichen und die verbleibende Restnutzungsdauer der Anlagen angepasst werden. Mithilfe der zustandsbasierten Restnutzungsdauer erfolgt die Einordnung in die Lebenszyklusmodelle (siehe Abbildung 2).

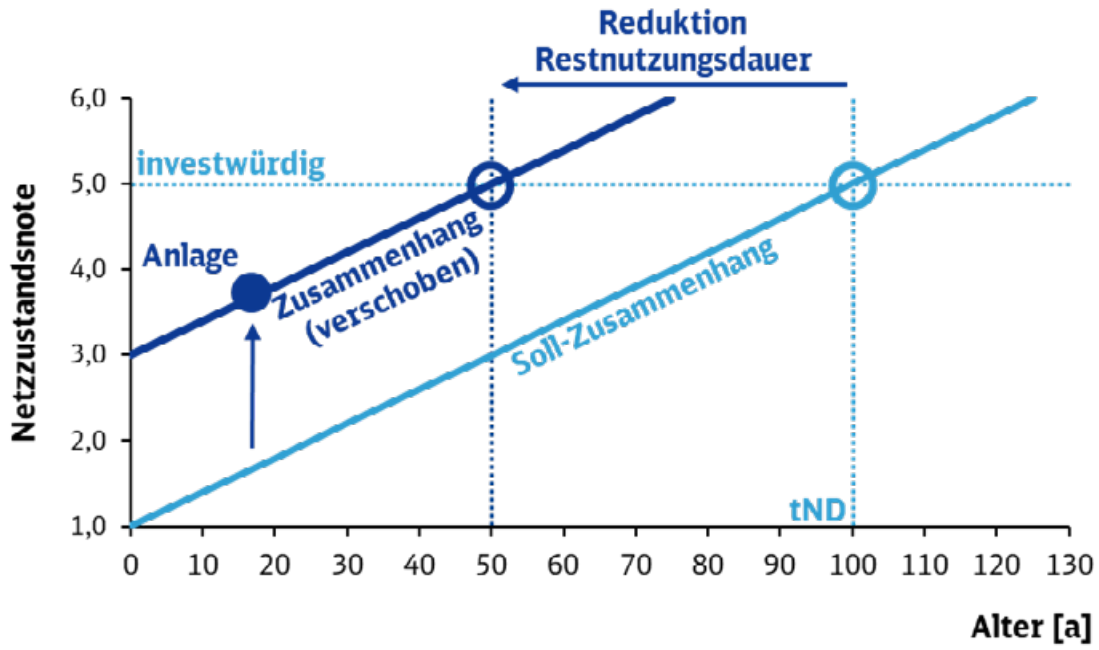


Abbildung 2: Prinzipskizze für die Ableitung der Restnutzungsdauer über den Zusammenhang zwischen Alter und NZN [eigene Darstellung].

Die Grundlage für die nachfolgenden Analysen bilden die NZN sowie deren vier zugrundeliegenden Qualitätsdimensionen (siehe Abbildung 3).

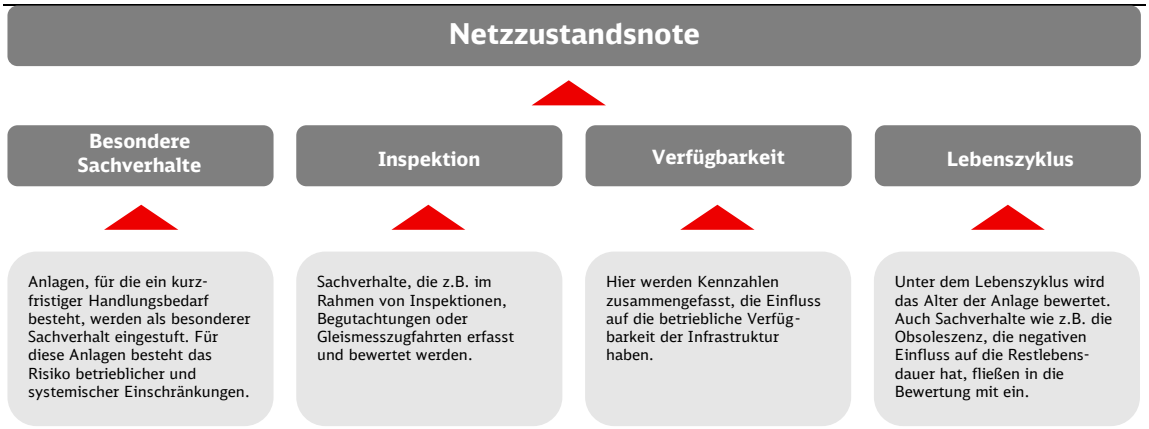


Abbildung 3: Die vier Qualitätsdimensionen der NZN [3]

3.1 Alter und Zustand in Zusammenhang bringen

Mit der zur Verfügung stehenden NZN und den Inbetriebnahmedaten je Einzelanlage sind Untersuchungen zu etwaigen Zusammenhängen zwischen beiden Größen möglich. Die Zielstellung ist dabei, Annahmen aus den Ist-Daten abzuleiten, mit welchen die gemessene NZN der Einzelanlage strategisch fortgeschrieben werden kann. Das Änderungsverhalten des Zustands wird dazu über das Anlagenkollektiv untersucht. Je Objektgruppe erfolgen Analysen zwischen Alter und Zustand bzw. den Qualitätsdimensionen der Inspektionsnote (I-Note), der Verfügbarkeitsnote (V-Note) und der Verrechnung der beiden Dimensionen zur Inspektions- und Verfügbarkeitsnote (IV-Note).

Abbildung 4 liegen die NZN-Jahrgänge 2021 bis 2023 zugrunde. Für die Analyse wird das relative Anlagenalter verwendet, um verschiedene tNDen innerhalb der Objektgruppe vergleichbar zu machen. Das relative Alter von eins entspricht der tND. Im Mittel und im Median ist für die Objektgruppe Weichen ein linearer Zusammenhang zum Zustand bis zum Erreichen der tND erkennbar. Ab diesem Punkt wird der Zusammenhang vergleichbar zu den Beobachtungen aus der Gerontologie zunehmend diffuser, da der Effekt der Ersatzinvestition von schlechten Anlagen wirkt und sich die Menge somit auf die Gutläufer reduziert (vgl. Absatz 2.2).

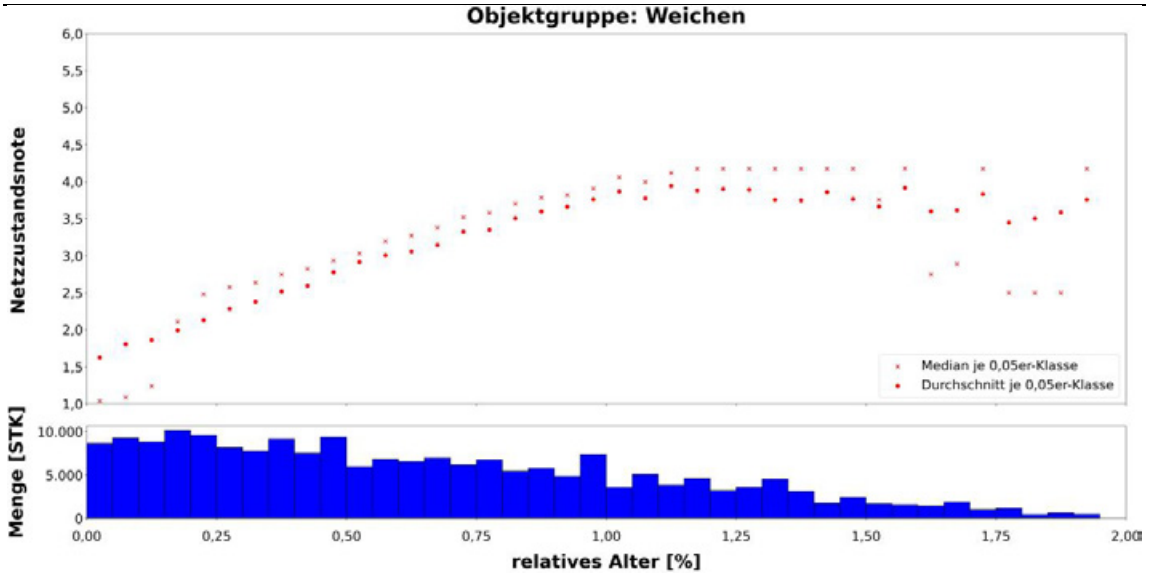


Abbildung 4: Darstellung des relativen Alters von Weichen im Verhältnis zur NZN [eigene Darstellung].

3.2 Annahmen für die Integration der Netzzustandsnote

Die wesentliche Herausforderung der Integration der NZN ist es, die richtige Auswahl an Annahmen zu treffen. Dabei werden die in bestimmten Objektgruppen ungleichmäßig besetzten Notenbereiche sowie die Berechnungslogik der NZN beachtet. Folgende vier Annahmen für den strategischen Zusammenhang zwischen Alter und NZN werden mithilfe der Erkenntnisse aus der Analyse der Ist-Daten extrahiert:

- Bei einem Alter von null Jahren (Abnahme der Anlage) ist die NZN eins und bei Erreichen der tND fünf. Diese Stützpunkte korrespondieren mit der Definition der Note. Dazwischen ist die Änderungsrate der IV-Note konstant.
- Die IV-Änderungsrate setzt sich nach Erreichen der tND linear fort.
- Es werden keine besonderen Sachverhalte und keine weiteren Merkmale neben der Lebenszyklusnote (LC-Note) in der Lebenszyklus-Dimension modelliert.
- Die Berechnung der Restnutzungsdauer ergibt sich aus der Verrechnung der Änderungsraten der IV-Note und LC-Note mit der Ist-NZN der Einzelanlage.

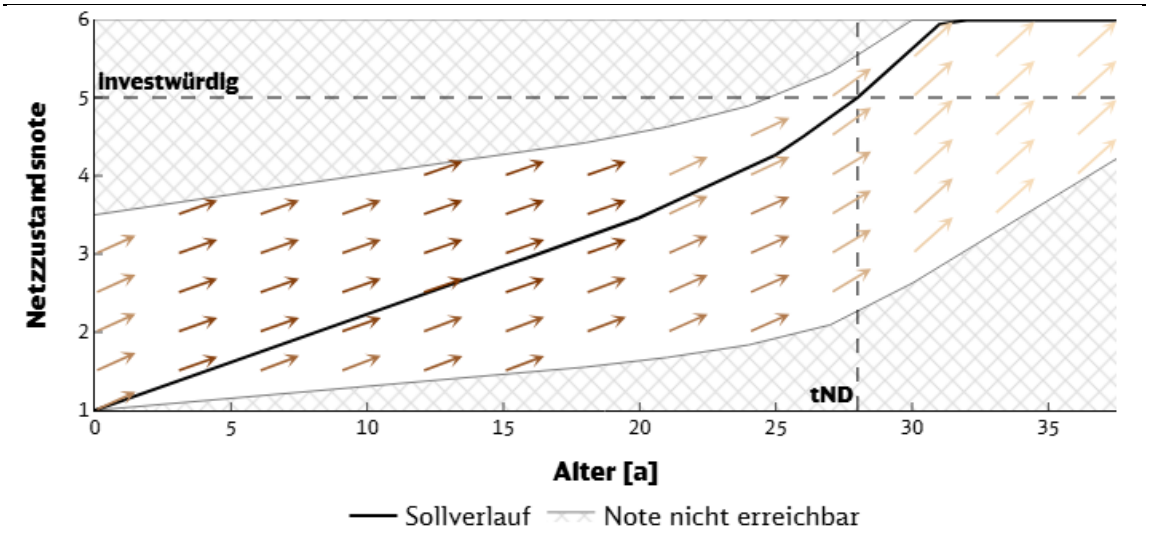


Abbildung 5: Darstellung der strategischen Änderungsrate der NZN im Vektorfeld für ein ausgewähltes Weichencluster [eigene Darstellung]

Dieses Set an Grundannahmen kann in einem Vektorfeld abgebildet werden (siehe Abbildung 5). Auf Ebene der Einzelanlagen kann mithilfe des Vektorfelds die Fortschreibung der NZN aus jeder erreichbaren Kombination von Alter und NZN erfolgen, ohne komplexe analytische Ansätze anzuwenden. Für jede Kombination ist, auf Basis der Berechnungslogik der NZN, eine strategische Änderungsrate hinterlegt, welche in Abbildung 5 vereinfacht durch Pfeile dargestellt ist. Die Einzelanlage wird in das Vektorfeld mit ihrer Kombination aus NZN und Alter eingesetzt. Über die hinterlegten Änderungsraten ergibt sich die Notenentwicklung über das Alter der Anlage. Bei Erreichen der Note fünf wird das Alter von der tND subtrahiert und das Ergebnis als Restnutzungsdauer für die neue Positionierung im Lebenszyklusmodell genutzt (vgl. Abbildung 2). Damit ergibt sich die Chance, zustandsabhängige Bedarfe für die Erneuerung und die Instandhaltung der Anlagen zu bestimmen. Der Sollverlauf ist der Pfad durch das Vektorfeld mit den Startwerten Alter gleich null Jahre und NZN gleich eins hin zur NZN fünf bei Erreichung der tND. Dieser Pfad bildet die strategische Entwicklung des Soll-NZN über das Anlagenkollektiv ab.

3.3 Zusammenhang zwischen Sollverlauf und Instandhaltung

Die Interaktion zwischen dem Sollverlauf und der Instandhaltung kann in zwei Varianten unterschieden werden. Durch die Anpassung des Anlagenalters über den Abgleich zwischen Ist- und Soll-NZN ergeben sich neue Instandhaltungsbedarfe, da sich im Lebenszyklusmodell die Instandhaltungsmengen je Alter unterscheiden (vgl. *Abbildung* 1). Bei

zustandsabhängigen Instandhaltungstätigkeiten erfährt die Anlage damit den Tätigkeitsumfang von Anlagen in einer schlechteren NZN. Der Zusammenhang dieser Variante ist eng an die Gestaltung der Instandhaltung im altersbasierten Lebenszyklusmodell geknüpft. Der sich ergebende Instandhaltungsbedarf beschreibt die Soll-Instandhaltung, die ab der ermittelten NZN im Durchschnitt zu erwarten ist.

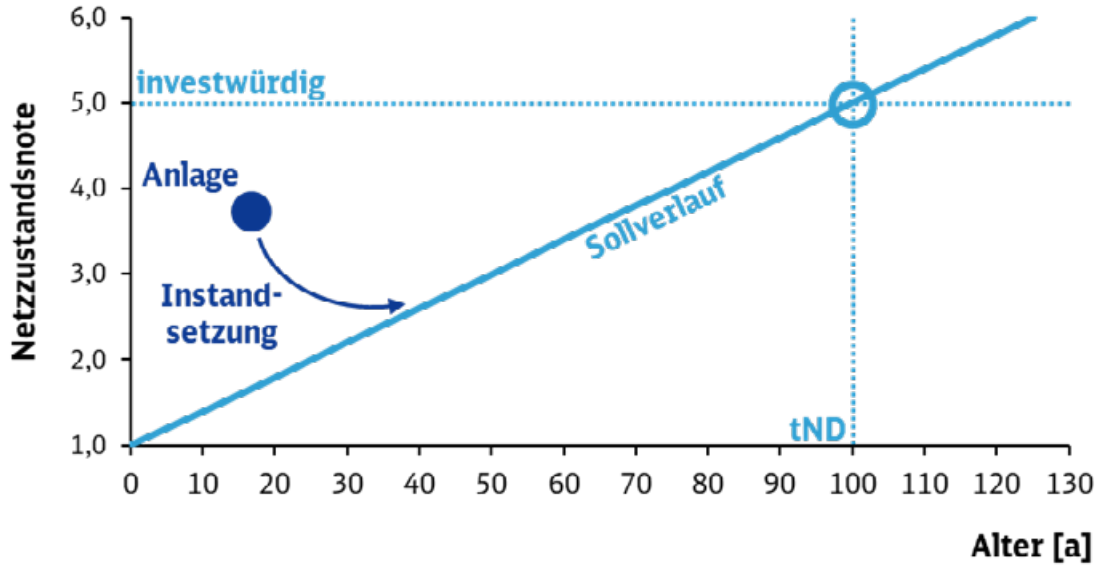


Abbildung 6: Prinzipskizze für den Mechanismus der Zustandsverbesserung durch Instandsetzung (zweite Variante) [eigene Darstellung]

Ein wesentlicher Faktor für die Weiterentwicklung des zustandsbasierten Ansatzes im 3-i Modell ist die Implementierung der zweiten Variante, bei welcher die Abweichung zwischen Ist- und Soll-NZN über die Instandsetzung der Anlage korrigiert wird (siehe *Abbildung* 6). Wenn die Anlagen deutlich von der erwarteten Soll-NZN abweichen, liegen meist technische Sachverhalte vor, die einer Migrationsstrategie der Instandhaltung und Teilerneuerung bedürfen. Hierfür existieren bereits erste Ansätze, wie die Migration der NZN vom Ist- in das Soll-Portfolio über beispielsweise gesteigerte Instandhaltungsmittel mehrjährig modelliert werden kann. Diese Ansätze sollen im Anschluss an das Projekt WiiiM 2.0 vertieft werden.

3.4 Chancen und Herausforderungen der Integration des Zustands

Das 3-i Modell ist und bleibt ein lebenszyklusbasiertes Verteilungsmodell von Instandhaltungs- und Investitionsbedarfen für das Anlagenkollektiv. Zukünftig wird es durch die Implementierung der NZN allerdings möglich sein, die für die 3-i Anlagenstrategie erforderlichen Mittel zustandsbasiert und damit bedarfsgerechter in die Regionen und Netze

je Objektgruppe zu verteilen. Zudem können die Entwicklungen der NZN und die daraus entstehenden Mittelbedarfe verfolgt werden. Aus dem Vergleich der Bedarfe mit angefallenen Instandhaltungs- und Investitionsmengen können Handlungsempfehlungen für die Regionen abgeleitet oder Verbesserungsbedarfe am 3-i Modell aufgezeigt werden.

Im Kontext der Finanzmittelverhandlungen mit dem Eigentümer kann in Zukunft auf ein zustandsbasiertes 3-i Modell zurückgegriffen werden, das die notwendigen Mittel für Investitionen zur Erreichung eines Soll-Infrastrukturzustands auf die Regionen verteilt, sowie die aus diesem Investitionsportfolio abzuleitenden Instandhaltungsbedarfe ebenfalls je Region berechnet. Aus diesem Wechselspiel können Infrastrukturfinanzierungsdiskussionen begleitet werden, die auch auf alternative Finanzierungslösungen wie der Rahmenplanung in Österreich [8] oder des Bahninfrastrukturfonds (BIF) in der Schweiz abzielen [9]. Die sich aus diesen Diskussionen ergebenden Leitlinien können dann Grundlage für die Planungsrunden der Investition sowie der Instandhaltung bei der DB InfraGO AG werden.

4 Ausblick für ein Strecken- und Knotenmanagement

Die DB InfraGO AG sowie der Bund zielen mit der Entwicklung der nächsten Leistungsvereinbarung sowie des InfraPlans [10] auf die Implementierung eines Strecken- und Knotenmanagements nach Vorbild der ÖBB und SBB ab. Das Ziel soll es hier sein, je Korridor Ziele und Maßnahmen zu definieren, die anhand von transparenten Kenngrößen nachvollzogen und dadurch finanziert werden können.

Die DB InfraGO AG hat im Kontext verschiedener Konzernprojekte bereits einen ersten Prototypen eines Strecken- und Knotenentwicklungsplans (SKP) erstellt, der unter anderem die auf Basis von NZN, Verspätungsschwere [11] und ETCS-Strategie notwendigen Investitionsbedarfe auf baubetriebliche Korridore (Bautaktkorridore) verortet und in Sperrzeiten übersetzt [12]. Diese Sperrzeiten (intern Container genannt) lassen sich durch die Werkzeuge des Kapazitätsmanagements vom mittelfristigen bis in den langfristigen Horizont koordinieren, sodass ein erstes Sperrzeitgerüst lange vor Umsetzung individueller Maßnahmen entsteht. Durch den interdisziplinären Ansatz und die gewerkeübergreifende Planung werden die einzelnen Sperrzeiten sehr effizient genutzt. Das ermöglicht im Kontext Fahren und Bauen mehr Platz für das Fahren, ohne Abstriche beim Bauen hinnehmen zu müssen. Auf das interdisziplinäre sowie regionsübergreifend berechnete und abgestimmte Sperrzeitgerüst kann langfristig hingearbeitet werden, sodass Ineffizienzen durch ausfallende Maßnahmen und erhöhter ungeplanter Instandhaltungsbedarf re-

duziert werden. Nicht zuletzt können mit dem SKP die notwendigen Instandhaltungsmittel in Bezug zum abgestimmten Investitionsportfolio mittels zustandsbasiertem 3-i Modell bestimmt, sodass sich der Kreis im Sinne der 3-i Strategie schließt. Dies trägt auch wesentlich zur Pünktlichkeitsstrategie bei, da über klare Investitionszeitpunkte langfristige Baustellenstabilität gewährleistet werden kann, welche wiederum zu fristgerechter Netzfahrplanung und Baustellenfahrplanung führt. Die Instandhaltung wird zukünftig besser auf die klaren Investitionszeitpunkte ausgerichtet werden können, sodass die Infrastrukturverfügbarkeit zielgerichtet gesteuert werden kann.

5 Nächste Schritte

Mit den beschriebenen Annahmen konnte für jede Objektgruppe und deren Anlagencluster ein Sollverlauf konstruiert werden. Ein selbstentwickelter Prototyp enthält die vorgestellten Mechanismen und ist in der Lage, erste Ergebnisse zu errechnen. Mit zunehmender Verfügbarkeit von weiteren ermittelten NZN sollten die vier Annahmen hinterfragt und weiterentwickelt werden. Mit dem Vektorfeld existiert die Möglichkeit, beliebige Sollverläufe abzubilden und neue Erkenntnisse zu integrieren. Darüber hinaus kann das Vektorfeld auch mehrdimensional ausgelegt werden. Bei den Sollverläufen ist es denkbar, diese aus den auf Ist-Daten basierenden internen wie externen Prognosemodellen abzuleiten, den Wissensaustausch bezüglich der Zustandsentwicklung von Anlagenklassen mit anderen Infrastrukturunternehmen zu vertiefen oder Forschungsarbeiten auf dem Themengebiet zu initiieren. Unabhängig von der Art der Sollverläufe müssen die bestehenden Lebenszyklusmodelle auf die Integration der NZN angepasst werden. Dazu laufen im Projekt WiiiM 2.0 Entwicklungen für eine pauschale Überarbeitung der Lebenszyklusmodelle, welche in den kommenden Jahren detailliert ausgearbeitet werden sollen.

Literatur

[1] P. Veit, „Instandhaltung und Anlagenmanagement des Fahrwegs," in *Handbuch Eisenbahninfrastruktur*, Bd. 3. überarb. u. aktualisierte Auflage, L. Fendrich und W. Fengler, Hrsg., Berlin, Heidelberg, Springer Berlin Heidelberg, 2019, pp. 1003 - 1049.

[2] M. Gramer, „Die integrierte Investitions- und Instandhaltungsstrategie (3-i) bei der DB Netz AG," *Eisenbahntechnische Rundschau,* pp. 12-15, 2013.

[3] DB Netz AG, „Netzzustandsbericht DB Netz AG 2021 - Fokus Netzzustandsnote," DB Netz AG, Frankfurt am Main, 2022.

[4] ÖBB Infrastruktur AG, „Netzzustandsbericht 2023," ÖBB-Infrastruktur AG, GB Asset Management und Strategische Planung, Wien, 2024.

[5] SBB Infrastruktur, „Netzzustandsbericht 2023," SBB Infrastruktur, Bern, 2024.

[6] A. Rhein und S. Ude, „DB InfraGO AG: Bedarfsermittlung mit dem 3-i Modell," *Österreichische Zeitschrift für Verkehrswissenschaft,* Nr. 1, pp. 2-4, 2025.

[7] G. T. Baker und R. L. Sprott, „Biomarkers of aging," *Experimental Gerontology,* Bd. 23, Nr. 4-5, pp. 223-239, 1988.

[8] ÖBB Infrastruktur AG, Bundesministerium Innovation, Mobilität und Infrastruktur, „Rahmenplan 2025–2030: Investitionen mit Weitblick - Für eine starke Bahninfrastruktur," 2025. [Online]. Available: https://infrastruktur.oebb.at/de/dam/jcr:e09950c5-7298-44a0-82b5-29becf5a9856/%C3%96BB_Rahmenplan25-30_Pr%C3%A4sentation-final.pdf. [Zugriff am 30 Juni 2025].

[9] Bundesamt für Verkehr BAV, „Der Bahninfrastrukturfonds BIF," 2025. [Online]. Available: https://www.bav.admin.ch/bav/de/home/verkehrsmittel/eisenbahn/bahninfrastruktur/bahninfrastrukturfonds.html. [Zugriff am 30 Juni 2025].

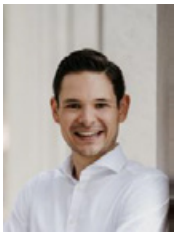
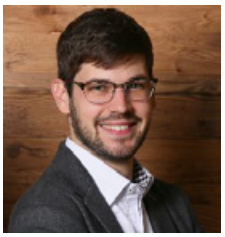
[10] Bundesministerium für Verkehr, „Umsetzungsbericht Beschleunigungskommission Schiene (20. Legislaturperiode)," 04 April 2025. [Online]. Available: https://www.bmv.de/SharedDocs/DE/Publikationen/E/umsetzungsbericht-beschleunigungskommission-schiene.html. [Zugriff am 01 Juli 2025].

[11] P. Scherer, B. Kogel, W. Lenze, M. Maus, N. Nießen und M. Fietze, „Identifikation von Kapazitätsengpässen," in *IRSA 2023 : Tagungsband/proceedings: Aachen, Germany 22-23 November 2023*, N. Nießen, C. Schindler und DVV Media Group GmbH, Hrsg., Aachen, Aachen : Universitätsbibliothek der RWTH Aachen, 2024, pp. 493-509.

[12] A. Rhein, „Alters- und zustandsbasierte Bedarfsermittlung mit dem 3-i Modell - Die Asset Simulation der DB InfraGO AG," in *23. Wiener Eisenbahnkolloquium, Unveröffentlicht*, Wien, 2025.

Session 13
Sebastian Ude, Alexander Rhein, Tobias Jacke,
Sascha Willgardt-Levin

Autoren

	**Ude, Sebastian** Sebastian Ude ist als Spezialist für Lebenszyklusmodelle und Anlagenstrategie bei der DB InfraGO AG tätig. Dort verantwortet er die Integration der Netzzustandsnote in das 3-i Modell als Arbeitspaketleiter im Projekt WiiiM 2.0. Herr Ude bringt umfassende Erfahrung aus der Entwicklung von Lebenszyklusmodellen sowie der Erstentwicklung der Netzzustandsnote mit. Vor seiner beruflichen Laufbahn studierte er Mobilität & Verkehr an der RWTH Aachen, war Hilfswissenschaftler am Verkehrswissenschaftlichen Institut und konnte im Rahmen einer Masterarbeit bei der DB Netz AG seine Kenntnisse im Bereich der eisenbahnbetriebswissenschaftlichen Untersuchungen vertiefen.
	**Dr. Rhein, Alexander** Dr. Alexander Rhein ist seit 2020 Leiter der OE Modelle 3-i Strategie Segmente bei der DB InfraGO AG. Dabei verantwortet er die Ermittlung des technischen Bedarfs für die Instandhaltung und die Erneuerung der Anlagen im Geschäftsbereich Fahrweg. Darüber hinaus bringt er die Entwicklung des strecken- und zustandsbasierten Asset Managements voran. Zuvor war Herr Dr. Rhein von 2019 bis 2020 Spezialist für IH-Strategie und Analysen bei der DB Netz AG. Zwischen 2013 und 2018 war er wissenschaftlicher Mitarbeiter an der Technischen Universität Darmstadt. Seine Dissertation hat er über das Asset Management in elektrischen Energieversorgungsnetzen verfasst. Herr Dr. Rhein bringt umfassende Erfahrungen zu strategischen und analytischen Fragestellungen in der Infrastrukturplanung mit.
	**Dr. Jacke, Tobias** Dr. Tobias Jacke ist seit 2021 im Bereich Infrastrukturlebenszyklusmodellierung bei der DB InfraGO AG tätig und verantwortet hier den Bereich der Weiterentwicklung. Zudem ist er bei verschiedenen Konzernprojekten tätig und konnte dabei wesentlich bei der Entwicklung der Netzzustandsnote sowie des Hochleistungsnetzes und den Hochleistungskorridoren unterstützen. Aktuell liegt der Fokus auf dem Aufbau eines international abgestimmten zustands- und kapazitätsgetriebenen Strecken- und Knotenmanagements,

	wofür er auf seine Erfahrungen aus der Promotion im Asset Management sowie der Eisenbahnbetriebswissenschaft zurückgreifen kann.
	**Willgardt-Levin, Sascha**
	Sascha Willgardt-Levin ist als Spezialist 3-i Modell bei der DB InfraGO AG tätig. Zuvor hat er Energietechnik und Wirtschaftswissenschaften an der RWTH Aachen studiert. Sein Aufgabengebiet umfasst die Analyse und Modellierung von Daten im Kontext des strategischen Asset Managements von Eisenbahninfrastruktur.

IRSA 2025

SESSION 14

INFRASTRUKTURZUSTAND II

INFRASTRUCTURE CONDITION II

Session 14
Thorsten Neumann, Martin Rabel, Oana-Iuliana Popescu, Andreas Gerhardus

Causal Discovery for Railway Health Condition Monitoring – A Case Study

Neumann, Thorsten[1], Rabel, Martin[2],
Popescu, Oana-Iuliana[2], Gerhardus, Andreas[3]

[1]German Aerospace Center (DLR),
Institute of Transportation Systems, Berlin
[2]University of Potsdam,
Institute of Computer Science, Potsdam
[3]German Aerospace Center (DLR),
Institute of Data Science, Jena

Abstract

Integrating knowledge of causal relationships into machine-learning models might help to overcome the problem of limited interpretability of these models. Causal methods can thus be expected to enable innovative data-driven approaches for transparent and reliable diagnostics in safety-critical domains such as railway health condition monitoring. Based on available interlocking data, this paper exemplarily demonstrates how causal discovery can be applied to learn the causal influences of certain weather parameters on the insulation resistance of the electric installation of electronic interlockings.

Keywords: Causal Discovery; Explainable AI; PHM; Interlocking

1 Introduction

Capacities and reliability of the railway infrastructure are key success factors for the rail sector in competition with other modes of transport. Innovative data-driven approaches using machine learning are expected to essentially support asset health monitoring including predictive maintenance. In this context, common machine learning relies on finding and utilizing *statistical associations* (e.g., correlations) based on large scale data.

But, typically, it is not capable of actually deciding about *causality* which however is crucial for transparent and holistic diagnostics. This is one of the main reasons why machine learning is still facing big challenges with regard to validation and verification in safety-critical domains such as rail, in particular.

Recent research in the field of *"Explainable AI"* – still without explicitly addressing causality – therefore attempts to make (black-box) machine learning models better accessible for human interpretation (cf. [1]). Besides that, the *"Causal Revolution"* – as it is called in [2] – has brought a profound new mathematical understanding of causality during the last decades with promising algorithmic tools for causal learning and reasoning (cf. [3], [4], [5], [6]) that are now ready to be widely applied and tested in practice. The remainder of this paper focusses on this second approach.

2 Causal graphs and causal discovery

It is an important and well-proven fact that "correlation is not causation" (see [2]). Indeed, there are numerous academic and real-world examples where correlation must not be misinterpreted in terms of causality. Though, given knowledge about the data-generating process or in case that certain assumptions such as *faithfulness* or *causal sufficiency* (cf. [7]) hold, associations in the data nevertheless carry important qualitative and quantitative information about the causal relations between the variables under consideration. Formalizing and utilizing this connection between statistical and causal associations in order to reason about causality in the given data is the topic of *causal inference* (cf. [4]).

Here, the task of *causal discovery* refers to learning cause-and-effect relationships between variables. Basically, constraint-based causal discovery algorithms retrieve the causal information from marginal and conditional independencies in the data distribution (see [5] and, for example, the review article [8]). In this paper, the PCMCI algorithm [9] (see Section 2.2) together with a partial correlation test (see Section 2.1) as implemented in the Python package Tigramite (see [10]) is employed. As a result, the found qualitative causal relations are described by *causal graphs* in which nodes stand for the variables under consideration and directed edges represent their causal interaction, respectively.

2.1 Partial correlation test

By definition, two random variables A and B are marginally independent if and only if their joint distribution $p(A, B)$ equals the product $p(A) \cdot p(B)$ of their marginal distributions. In this case, observing the value of B does not contain any information about

the value of A. Similarly, A and B are conditionally independent given a third variable C if and only if the factorization $p(A,B|C) = p(A|C) \cdot p(B|C)$ holds, where $p(X|Y) = p(X,Y)/p(Y)$ denotes the conditional distribution of X given Y. In this case, if the value of C is already observed, then additionally observing the value of B does not yield any information about the value of A beyond the information already obtained by observing C. For more details see textbooks on probability theory or statistics (e.g., [11]).

In the case study presented here, the simplifying assumption is made that the data distribution is multivariate Gaussian. Then, A and B are marginally independent if and only if their correlation is zero. Similarly, A and B are conditionally independent given C under this assumption if and only if the so-called partial correlation of A and B given C is zero (cf. [12]). To calculate this partial correlation, one can proceed in three steps: First, linearly regress A on C and calculate the residuals r_A as the difference $A - \hat{A}$ where $\hat{A}$ is the value of A as predicted by the linear regression on C. Second, repeat the first step with B instead of A. Third, calculate the correlation between the residuals r_A and r_B (cf. [12]).

In practice, the correlation and partial correlation values are not given but rather need to be estimated from finite data. To build a statistical hypothesis test, with marginal respectively conditional independence being the null hypothesis, one then utilizes the fact that a certain transformation of the estimated correlation respectively estimated partial correlation follows a Student's t-distribution with a certain number of degrees of freedom (cf. [12]). The Python package Tigramite referenced above (see [10]) provides an implementation of this kind of partial correlation test.

2.2 The PCMCI algorithm

The constraint-based approach to causal discovery builds on the observation that the particular form of a causal graph typically imprints several marginal and conditional independencies into the data (cf. [13], [14]). For example, if there are three variables A, B, C and the causal graph is $A \to B \to C$, then A and C are conditionally independent given B. Intuitively speaking, this independence holds because all information about A that is contained in C must have "passed through" B and, hence, is already contained in B. Thus, if one has already observed B, then one does not learn more about C by additionally observing A.

Now suppose that, in addition to $A \to B \to C$, the causal graph also contains the directed edge $A \to C$. Then, A and C are no longer conditionally independent given B because there is a "pathway of information" from A to C, namely $A \to C$, that does not pass

through B. The idea of constraint-based causal discovery is to find these marginal and conditional independencies based on data by statistical tests of independence, for instance by the partial correlation test described above, and to then reverse-engineer the causal graph. For example, if one finds that A and C are conditionally independent given B, then one can already conclude that the edge $A \to C$ is not part of the causal graph.

While such types of arguments appear intuitively plausible, they nevertheless do rely on certain assumptions that one has to be aware of. These are the causal Markov assumption and causal faithfulness assumptions (or variations thereof) as well as, depending on the particular algorithm that one decides to use, the causal sufficiency assumption and the assumption that there are no cyclic causal relationships (cf. [5]).

A prominent example for constraint-based causal discovery algorithms is the PC algorithm (cf. [5]). This algorithm utilizes the four assumptions listed in the previous paragraph and has originally been developed for non-time series data. The PCMCI algorithm (see [15]) is an adaptation of PC to time series data that is specifically designed to handle the statistical challenge of applying causal discovery to time series with a large number of variables. In addition to the assumptions of the PC algorithm, PCMCI assumes that the causal relationships remain constant in time.

The result of PCMCI is an estimated causal graph (see Figure 2 further below for an example), in which the curved directed edges signify causal relationships between variables. These edges are supplemented by their so-called “lag”, written as white digits on top of the edges in the plot, which specifies how many time steps the respective causal relationships take to manifest themselves. For example, if A causally influences B at lag 2, then A_{t-2} has a causal influence on B_t. Here, the time index t is arbitrary due to the assumption of constant-in-time causal relationships. The same edge can have multiple lags. Causal relationships that act on time scales below the sampling interval, which thus have a time lag of zero, are drawn as straight-line edges with circles at their ends. This notation signifies that the algorithm does not decide about their direction; that is, if $A_t \circ\!\!-\!\!\circ B_t$ is in the estimated causal graph, then either A influences B or B influences A at time lag zero. An extension of PCMCI that attempts to orient this type of edge, if possible, is the PCMCI+ algorithm (see [9]).

3 Case Study: Interlocking

Properly monitoring the health condition of the control, communication and safety systems of the railway infrastructure is crucial for its maintenance as well as reliable train operations. Besides monitoring other components such as switches and crossings or the

track itself, electronic interlockings are an example where continuously measuring meaningful health related parameters helps to detect failures and thus to avoid safety-critical situations.

Typically, the electrical systems of electronic interlockings in Germany are isolated against earth. The insulation resistance (IR) thus is a good measure for defects in the electric installation. While a single defect is not that problematic in general, parallel defects at more than one location may result in uncontrollable and safety-critical behavior of the interlocking system. Events of temporarily or suddenly reduced IR therefore require special attention. At the same time, it is known that weather has a major influence on the overall level of the IR (see [16]).

The goal of this case study is to analyze how and to what extent these known causal relations between relevant weather parameters (e.g., temperature or precipitation) and available IR measurements can be detected automatically from data by means of time-series based causal discovery (cf., for example, [17]). In practical applications, knowing such causal relations usually helps to better interpret the data which again can lead to better and faster maintenance decisions.

Note that the presentation in this paper is mostly to demonstrate the general potential of innovative causal methods for railway condition monitoring purposes rather than a deep analysis of the specific use case itself. Nonetheless, the selection of the use case is motivated by the fact that diagnostics of insulation faults as in context of interlockings, for instance, are an urgent and important topic related to railway infrastructure operation and safety (cf. [18]).

3.1 Data and Preprocessing

The IR data used in this case study comprise measurements for about 1.5 years between 2015 and 2016 collected from an electronic interlocking in southern Germany. Details on the data collection can be found in [19]. For the purpose of causal discovery, the data have then been aggregated to one sample per 10 minutes (see Figure 1a) with the option of further down-sampling, and synchronized with freely available weather data (see [16]) provided by Deutscher Wetterdienst (DWD). Missing values were interpolated linearly. Regarding the IR measurements, time intervals with missing (i.e., interpolated) data are indicated by the orange line in Figure 1a and were left out as input during the causal learning process (see Section 3.2).

Furthermore, in the middle (around time index 40,000) and at the end (around time index 80,000) it seems that there is a systematic drop in the overall level of the IR measurements where the true reason is not clear so far. One cause, of course, could be a drift or other sensor artefacts in the measurement system. But, also the connection to a temporary (untypical) configuration of the interlocking system depending on the combined operational states of all connected switches could be a conceivable explanation.

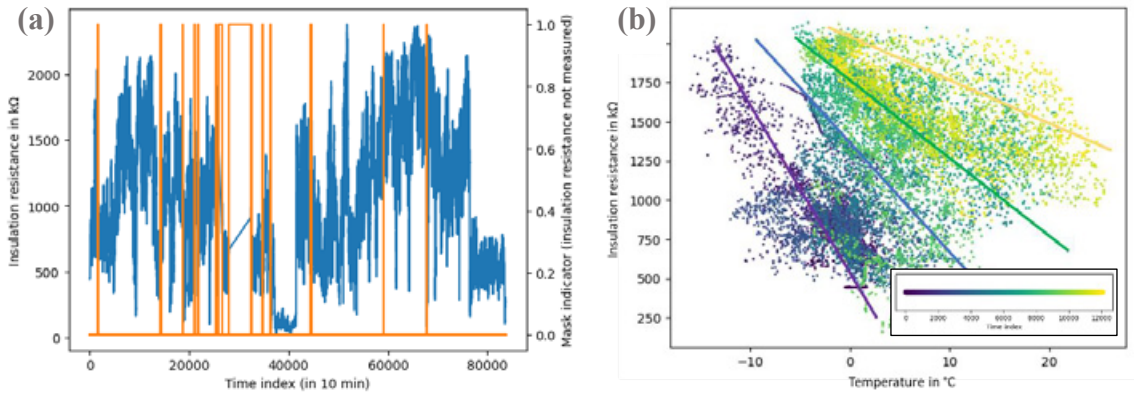


Figure 1: a) IR with a time-based resolution of 10 min (regions with missing data marked by the orange line) and b) negative correlation between temperature and IR based on a data subset covering the first three months of the available data.

Note that some relations between weather and IR variables have already been observed in previous studies and can thus perfectly be used for evaluating the plausibility of the causal discovery results in the remainder of this paper. For a more detailed discussion about known dependencies between IR and certain weather parameter (such as temperature and precipitation) see [16]. In addition, Figure 1b shows a (time-varying) negative correlation of IR and temperature which was found in this way during conducting the case study presented here.

3.2 Results and Discussion

Due to multiple periodicities on different time scales (e.g., daily, seasonal) overlapping with sudden effects, ensuring *(causal) stationarity* (cf. [7]) of the input time series for the learning process can be challenging which, by the way, is also supported by the previously mentioned finding of time-varying correlations as shown in Figure 1b. In the following, long-term effects are therefore addressed by learning separate causal graphs per season (e.g., summer or winter). At the same time, *hourly* aggregated samples are used to tolerably capture the short-to-medium-term effects while sufficiently limiting the number

of backward time slices (i.e., four) as one important hyperparameter in the causal discovery process using PCMCI (cf. Section 2.2). Other relevant hyperparameters of the PCMCI algorithm were specified based on some common default values.

Figure 2 shows the derived causal graph based on data for summer seasons (June to August). Following this graph, both precipitation and ground temperature appear to have a negative causal association with the IR at small time lags which is in good agreement with previous findings (cf. Figure 1b and/or [16]). Interestingly, the (air) temperature is detected to only have an *indirect effect* on the IR via the ground temperature, while both temperature measures strongly correlate. Also, humidity seems to negatively affect the IR, but with a larger delay. Sunshine duration, on the other hand, has a positive impact on the IR, probably because it facilitates evaporation and thus a drier environment for the cables and field elements of the interlocking system opposed to the effect of humidity and precipitation. Lastly, temperature and sunshine duration are found to positively correlate while humidity and precipitation (at least in the short term) have a reverse trend as compared to temperature.

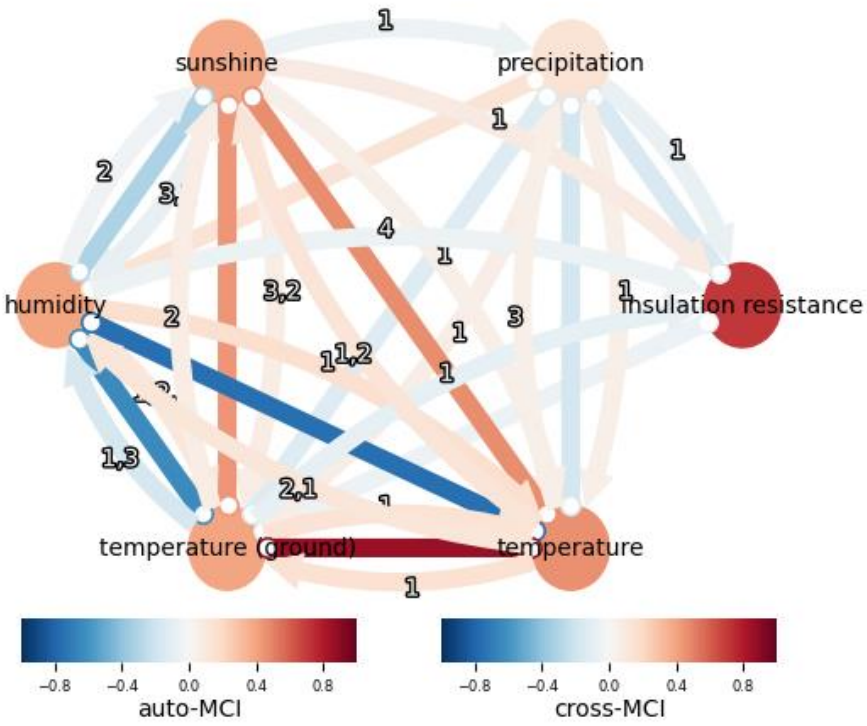


Figure 2: Causal graph exemplarily showing the results of causal discovery for summer based on hourly data. Undirected edges refer to causal relations whose direction is non-decidable via PCMCI (cf., last paragraph in Section 2.2), and numbers denote the time lag(s) per edge (in hours) related to the respective causal interactions. Further, red and blue colors respectively represent positive and negative correlations (see [10]).

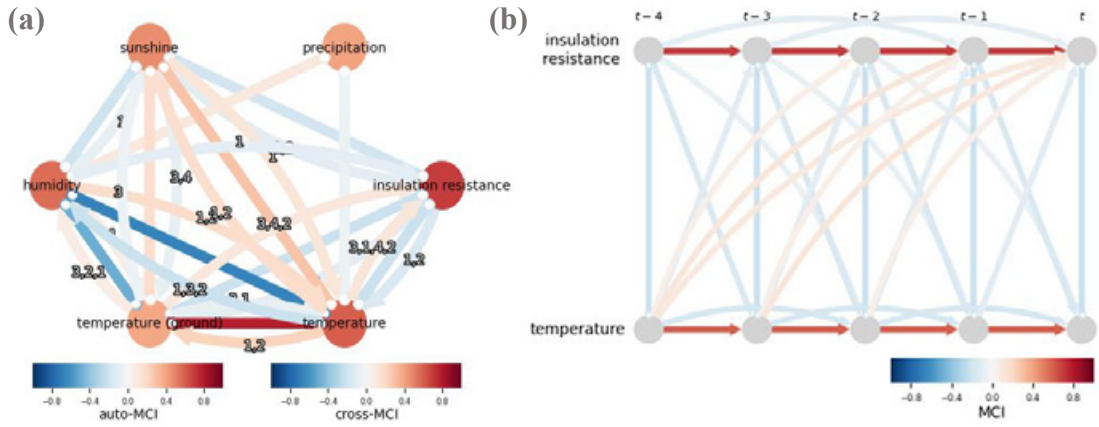


Figure 3: a) Causal graph exemplarily showing the results of causal discovery for winter based on hourly data and b) Unrolled graph for the relation between temperature and IR with explicit time slices. (Colors and labels as in Figure 2.)

In comparison to the summer graph, Figure 3a shows the respective causal discovery results for the winter season (December to February). As can be seen, there are at least three significant differences between both graphs.

First, the precipitation node now seems a bit more isolated from the IR measurements meaning that the former direct effect on IR vanishes. Whether or to what extent the latter is replaced by *indirect effects* (e.g., via the humidity node) however remains unclear due to the undirectedness of the zero-lag-edges that connect precipitation to the rest of the causal graph in Figure 3a. Yet unproven, a possible explanation for the divergent findings concerning the precipitation effect on IR could be the different forms of precipitation in summer and winter together with their distinct electric implications. That is, fluid water can be expected to penetrate the electric installations in the field much easier than snow which simply accumulates above with much less electrically relevant contact. In addition, water typically is a good conductor while snow is mostly non-conducting. Consequently, snow is indeed likely not to significantly affect the IR, but only when it is being reduced into its fluid state again later and thus not directly during the precipitation event as represented by the precipitation node in the causal graph.

Interestingly, if true, this might also explain a second observation, namely the inverted algebraic sign of the correlation between sunshine duration and IR in Figure 3a. That is, while summer sunshine facilitates evaporation and drying ground with the likely result of *higher* IR values as discussed earlier, in winter it may cause snow and ice getting melted which contrarily makes the direct environment of the cables and field elements of the interlocking system more wet and thus potentially *lowers* the IR level.

Finally, a third difference between Figure 2 and Figure 3a is the existence of a quite complex interaction of temperature and IR in the winter graph (cf. also the unrolled subgraph in Figure 3b). In particular, the causal relation from IR towards temperature looks very counterintuitive and has very likely to be considered as wrongly detected.

In this context, it has to be clearly stressed that the above results and discussions are subject to substantial uncertainty because of several reasons. Namely, as detailed above, the employed causal discovery method relies on various assumptions about the data-generating process, and some of these assumptions might be violated for the data at hand. For example, in potential violation of the causal sufficiency assumption, there might be unobserved variables that act as confounders which introduce additional dependencies that are not modelled by causal relationships between the observed variables. Moreover, in potential violation of the assumption behind the partial correlation test, the data distribution might be more complex than multivariate Gaussian.

In addition to that, the presented analysis has an explorative nature with an admittedly rather limited scope. To improve, one could repeat the causal discovery steps in which one systematically varies the hyperparameters of the analysis. This, for example, means modifying the hyperparameters of the PCMCI algorithm itself such as the maximal considered time lag, but also the choice of different conditional independence tests, the selection of data, or the details of the data preprocessing. Needless to say, relatively stable results across a proper range of hyperparameter settings in context of such an extended analysis would clearly improve the trustiness of the final results.

4 Conclusion and outlook

While further analyses are necessary, the presented case study nevertheless hints at the applicability and reasonableness of causal discovery in the field of railway health condition monitoring. Note that the general concept is not limited to the specific use case of monitoring electronic interlockings of course, but can also be applied to other critical components of the rail infrastructure. With regard to refining the results from Section 3.2, one could take an additional causal effect analysis based on the estimated causal graphs (see [6]) into account, and/or support the causal discovery step by integrating expert knowledge on the existence or non-existence of edges in the causal graphs (incl. their causal direction), for instance. Also, one might explore further causal discovery methods for time series and dynamical systems and analyze the consistency among the results.

Acknowledgement

This work was financed via institutional funding at the German Aerospace Center (DLR). All authors contributed in their role as scientific staff members of the DLR in the project "CausalAnomalies".

Literature

[1] A. Sajid, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez and F. Herrera, "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," *Information Fusion,* vol. 99, no. 11, p. 101805, 2023.

[2] J. Pearl and D. Mackenzie, The Book of Why, First Trade Paperback ed., New York: Basic Books, 2020.

[3] A. R. Nogeira, A. Pugnana, S. Ruggieri, D. Pedreschi and J. Gama, "Methods and tools for causal discovery and causal inference," *WIREs Data Mining and Knowledge Discovery,* vol. 12, no. 2, 2022.

[4] J. Peters, D. Janzing and B. Schölkopf, Elements of causal inference, Cambridge, London: MIT Press, 2017.

[5] P. Spirtes, C. Glymour and R. Scheines, Causation, Prediction, and Search, 2nd ed., Cambridge, Massachussetts: MIT Press, 2001.

[6] J. Pearl, Causality: Models, Reasoning, and Inference, 2nd ed., Cambridge: Cambridge University Press, 2009.

[7] J. Runge, "Causal network reconstruction from time series: From theoretical assumptions to practical estimation," *Chaos,* vol. 28, no. 7, 2018.

[8] C. Glymour, K. Zhang and P. Spirtes, "Review of Causal Discovery Methods Based on Graphical Models," *Front. Genet.,* vol. 10, 2019.

[9] J. Runge, "Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets," in *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence*, Toronto, Canada, 2020.

[10] J. Runge, "Tigramite - Causal inference for time series datasets," [Online]. Available: https://github.com/jakobrunge/tigramite. [Accessed 26/03/2025].

[11] L. Wasserman, All of Statistics: A Concise Course in Statistical Inference, New York: Springer, 2004.

[12] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, 3rd ed., Hoboken, New Jersey: John Wiley & Sons, 2003.

[13] T. Verma and J. Pearl, "Causal Networks: Semantics and Expressiveness," *Machine Intelligence and Pattern Recognition,* vol. 9, pp. 69-76, 1990.

[14] D. Geiger, T. Verma and J. Pearl, "Identifying independence in Bayesian networks," *Networks,* vol. 20, no. 5, pp. 507-534, 1990.

[15] J. Runge, M. Kretschmer, S. Flaxman and D. Sejdinovic, "Detecting and quantifying causal associations in large nonlinear time series datasets," *Sci. Adv.,* vol. 5, no. 11, 2019.

[16] J. Heusel and J. C. Groos, "Analysis of Systematic Influences on the Insulation Resistance of Electronic Railway Interlocking Systems," in *IAI 2021. Lecture Notes in Mechanical Engineering*, Cham, 2022.

[17] G. Camps-Valls, A. Gerhardus, U. Ninad, G. Varando, G. Martius, E. Balaguer-Ballester, R. Vinuesa, E. Diaz, L. Zanna and J. Runge, "Discovering causal relations and equations from data," *Physics Reports,* vol. 1044, pp. 1-68, 2023.

[18] C. Ziegler, "Condition-based maintenance and planting an apple tree," in *The Rise of IoT & Big Data in Rail*, Cologne, Germany, 2025.

[19] C. Linder and R. Schenkendorf, "Datengetriebene Diagnoseansätze für ESTW-Kabelanlagen," *SIGNAL + DRAHT,* vol. 107, no. 10, pp. 16-21, 2015.

Session 14

Thorsten Neumann, Martin Rabel, Oana-Iuliana Popescu, Andreas Gerhardus

Authors

	**Neumann, Thorsten (Dr.-Ing.)**
	Dipl.-Math. (Mathematics) in Münster, PhD (Engineering) in Braunschweig. Researcher at the DLR-Institute of Transportation Systems in Berlin as member of the working group on "Asset Monitoring and Management". Research interests: Transparent and holistic diagnostics for railway infrastructures.
	**Rabel, Martin (Dr. rer. nat.)**
	M.Sc. (Physics), PhD (Mathematics) in Heidelberg. Formerly staff member at the DLR-Institute of Data Science in Jena; now Postdoc with focus on causal methods with Jakob Runge's group, currently affiliated with the University of Potsdam. Research interests: Causal methods.
	**Popescu, Oana-Iuliana (M.Sc.)**
	M.Sc. (Computer Science). Formerly staff member at the DLR-Institute of Data Science in Jena; now PhD student affiliated with the Institute of Computer Science at the University of Potsdam. Research interests: Causal discovery for data with anomalies and distribution shifts.
	**Gerhardus, Andreas (Dr. rer. nat.)**
	M.Sc. in Physics (University of Bonn), Dr. rer. nat. (Faculty of Mathematics and Natural Sciences at the University of Bonn), now group leader at the DLR-Institute of Data Science in Jena. Research interests: Causal inference methods and their application

Session 14
Laura T. Rodríguez-Bayona, Héctor A. Fernández-Bobadilla, Ullrich Martin

Monitoring Data Mapping for Infrastructure-oriented Visualization of Railway Track Condition

Rodríguez-Bayona, Laura T.[1], Fernández-Bobadilla, Héctor A. [1] and Martin, Ullrich[1]

[1]Institute of Railway and Transportation Engineering, University of Stuttgart, Pfaffenwaldring 7, 70569 Stuttgart, Germany

Abstract

The railway system, as a complex network of interdependent components, is facing increasing transportation demand and aged infrastructure, leading to significant challenges for an efficient infrastructure maintenance planing. In this context, infrastructure digitalization and data-driven maintenance strategies, supported by condition monitoring (CM), have emerged to alleviate this issue. For the railway track, even though a large amount of monitoring data can be collected, this information is often indexed in time, meaning that a direct and sufficient association with the infrastructure is not provided and must be established. A possible approach to overcome this situation is by leveraging infrastructure-oriented data visualization. By mapping monitoring data to the infrastructure, a first approximation for the track condition is obtained. This work explores the development of a tool for monitoring data visualization, which integrates a three-dimensional (3D) infrastructure digital model with a database. On the basis of this representation, relevant physical, operational and condition features are depicted, providing a baseline for a comprehensive analysis and contributing to a condition-based maintenance decision-making for infrastructure owners and operators.

Keywords: Infrastructure Digitalization; Data Mapping and Visualization; Condition Monitoring; Track Maintenance Planning

Session 14
Laura T. Rodríguez-Bayona, Héctor A. Fernández-Bobadilla, Ullrich Martin

1 Introduction

Digitalization plays nowadays a remarkable role in almost all areas of knowledge. Recently, the implementation of digital tools for infrastructure planning, construction, monitoring and maintenance has gained enormous relevance. These tools not only allow a smooth schematization of infrastructure elements for the creation of drawings and blueprints, but also facilitate data representation, offering valuable insights for monitoring.

In the framework of digitalization, a concept that arises frequently is digital twin (DT). This is a virtual representation of a physical system constantly updated by the automatic and bidirectional exchange of information between the virtual and the physical part [8]. DT supports simulation, control and monitoring of physical systems over time, under diverse operating regimes, facilitating applications in the infrastructure sector such as automation, fault detection and predictive maintenance [1]. However, in cases where emulating the dynamic behavior of the physical system is not required, a static virtual representation, also called a digital model, could be adequate to support monitoring and maintenance activities.

Recently, advanced proactive approaches have emerged for railway infrastructure maintenance, in order to carry out intervention actions before a fault causes significant disruptions in the operation. These strategies rely on condition monitoring (CM), which can be defined as the continuous and automatic tracking of the operative state (condition) of a dynamical system to detect early signs of deterioration and developing faults [3]. In the context of vehicle-based infrastructure monitoring, a large amount of data can be collected via either specialized supervision vehicles or regular in-service vehicles equipped with data acquisition devices. However, since most of the monitoring data are indexed over time, a direct and sufficient association between data and infrastructure is still missing to some extent. An infrastructure-oriented monitoring data mapping serves the purpose of associating the measurements to the asset, obtaining thus an initial meaningful condition assessment.

This work focuses on the development of a visualization tool for infrastructure-oriented monitoring data depiction in a railway short line, the so-called "Tälesbahn", located in the southeast of Stuttgart, Germany. This visualization tool is based on a three-dimensional (3D) digital model of the track and a relational database containing infrastructure and monitoring information. In addition to infrastructure digitalization and data storage, the

digital model and database act as the core elements for visualizing the railway track condition, providing an insightful infrastructure-oriented data representation.

2 Background

2.1 Test Railway Line

The Tälesbahn is a single-track, standard gauge non-electrified regional passenger railway line, connecting the cities of Nürtingen and Neuffen, both located in the district of Esslingen, at the southeast of Stuttgart, Baden-Württemberg, Germany. This line has a length of approximately 8.9 km, exhibits a height difference of 111 m and is officially identified by DB InfraGO AG with the line number 9465. The infrastructure operator is the Württembergische Eisenbahn-Gesellschaft mbH (WEG), which belongs to Transdev GmbH Germany, a subsidiary company of the Transdev Group. The lines serves seven stations, namely Nürtingen, Nürtingen-Vorstadt, Nürtingen-Roßdorf, Frickenhausen, Frickenhausen-Kelterstraße, Linsenhofen and Neuffen. The maximum operating speed is 80 km/h and the travel time between terminals is 12 minutes, according to the timetable. The line is served by the route Regional Bahn RB65.

2.2 Track Monitoring System

The monitoring data used in this work were obtained from the *Reallabor 2.0*, which is an efficient and cost-effective vehicle-based monitoring system for the detection and localization of track structural and geometrical faults using inertial measurements. This monitoring system was realized by means of regular in-service vehicles to allow for a continuous track monitoring without affecting the normal railway operation [4]. The system does not require any special authorization from the German Federal Railway Authority (EBA, by its German acronym).

The measurement technology chosen for this application was inertial sensors, which consist of accelerometers and rate gyros, measuring, respectively, acceleration and angular rates along and around the vehicle's longitudinal, transversal and vertical axes. Inertial measurement systems stand up for their superior dynamical characteristics, the suitability for in-service vehicle and their low cost. Specifically for this application, two inertial

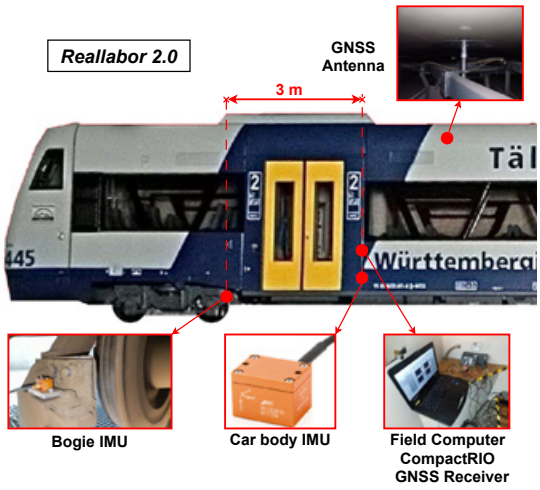


Figure 1: Component arrangement for the Reallabor 2.0.

measurement units (IMU) were mounted on the bogie and car bogie of the railway vehicle. The overall components arrangement is depicted in Figure 1.

The monitoring system was designed to detect middle-length track defects, i.e., those within the wavelength interval between 3 m and 70 m [5]. Most of the structural and geometrical vertical track defects are included within this range, except for the short-wavelength corrugation. The long-wavelength design geometry lies also out of the scope of this system. The sampling rate chosen for this application was 300 Hz, corresponding to 300 samples of inertial data per second.

To spatially locate the collected data, the system was provided with a Global Navigation Satellite System (GNSS) receiver and an antenna, mounted inside the vehicle car body. The sampling rate for geographical coordinates was set to 1 Hz. This means that the receiver delivers one sample of coordinate data every second.

3 3D Digital Model

To develop the infrastructure-oriented visualization tool, the creation of a georeferenced 3D digital model of the Tälesbahn was necessary as a foundation for monitoring data depiction. This digital representation was developed using Computer-Aided Design (CAD) software. In this model, only the fundamental elements were considered, including rails, ballast and frost protection layer. Other components such as sleepers, fasteners and

drainage systems were excluded for the sake of simplicity and efficiency in the development process. This model was georeferenced to align with the existing infrastructure based on available track drawings.

To define the 3D model, additionally to the horizontal and vertical alignment, functional cross sections capturing the variations of the track's elements were required. Here, the infrastructure had to be modeled with sufficient spatial resolution to allow a meaningful representation of the collected data. Therefore, a discretization process was carried out to obtain independent and representative solid elements. The creation of these elements is described in the next section.

3.1 Infrastructure Model Discretization

In the context of this work, discretization refers to the spatial segmentation of the infrastructure into pieces of a certain length in order to assign specific data to them. These segments are associated to the track geometry and further physical characteristics. Discretization is important to accurately represent the data variations (corresponding in this application to monitoring information) that will be visualized later in the digital model.

To perform the infrastructure model discretization, the most important factor to consider is the wavelength of the different track features, including track irregularities and defects. Similarly to the Nyquist-Shannon sampling theorem [6] (which establishes that the minimum rate at which a continuous signal must be sampled to correctly and unequivocally reconstruct it must be twice the maximum frequency contained in the signal), it is possible to establish the minimum spatial resolution to represent a track characteristic as half its minimum wavelength. Considering the minimum wavelength (3 m) within the scope of the Reallabor 2.0, a spatial discretization of 1.5 m could be proposed. However, a more conservative approach to the Nyquist-Shannon sampling theorem is to take a sampling rate of 10 times the maximum signal frequency; this approach is often adopted in practice [3]. Analogously, the minimum spatial resolution for the representation of track features can be defined as one tenth of the smallest wavelength; therefore, a segment length of 0.3 m for the infrastructure discretization was selected. This segment length ensures that track features with the minimum wavelength are captured by the model. Considering this length, 30200 segments were required to represent the complete Tälesbahn line.

4 Relational Database

Through monitoring, the condition of the railway track can be assessed. However, this process also produces large volumes of data. In order to store and manage these data, as well as to integrate them with other available relevant information, a well-structured database is required. In this research, a relational database was developed to centralize and organize information. This database is characterized by its entities and the relationships between them. An entity models a real-world object, represented as a table, and a relationship defines how two or more entities are associated. Specifically, the goal was to integrate infrastructure and monitoring data of the Tälesbahn line for visualization purposes. To achieve this, entities and their mutual relationships were defined to represent both static infrastructure data and collected monitoring data. The entities were divided into two categories:

1. **Monitoring-related data**: They consist of inertial and positioning measurements obtained from Reallabor 2.0, as well as environmental information of the day and time when the measurements were carried out, including precipitation, pressure, temperature and relative humidity. This information was obtained from [2].
2. **Infrastructure-related data**: They represent geometrical, physical and operational characteristics of the Talësbahn line. These data focus not only on those elements that define the line route alignment, but also on factors that could influence the inertial measurements collected during monitoring, which are crucial for assessing the track condition, e.g., turnouts, bridges and level crossings.

For the association of the georeferenced digital model with the monitoring data, an integration method was required, considering that the collected data were not precisely aligned with the digital representation. Typically, data integration is carried out via map-matching. This method merges positioning data with spatial network data to determine the precise location on which a vehicle is traveling [7]. Map-matching can incorporate different techniques such as geometric, topological and probabilistic analysis, as well as advanced algorithms, e.g., Kalman filter, extended Kalman filter, fuzzy logic, etc. [7]. However, this study adopted a simplified analysis in order to obtain an initial approximation to the actual position of the measurements along the track. This process is described in the following section.

4.1 Monitoring Data Aggregation and Association with the Infrastructure

In the Reallabor 2.0, the collection of inertial measurements and positioning data occurs simultaneously (i.e., the processes are synchronous), but with different sampling rates. Inertial measurements were recorded at 300 Hz, while positioning data was collected at 1 Hz. This means that, for every 300 inertial measurements done in one second, only one positioning point was collected. However, when these two measurements points are simultaneously available, they both share the same timestamp.

Thanks to the synchronicity of the data collection, it was possible to align both inertial and positioning data according to their common timestamps. However, in principle only 1 of 300 inertial measurements every second could be merged with positioning data. Since there was a large number of inertial measurements lacking a corresponding position, the latter had to be upsampled via linear interpolation. Although this approach involves several assumptions that could affect the precise localization of the measurements along the infrastructure, it provided an initial baseline for monitoring data visualization.

As previously mentioned, the vehicle position obtained via GNSS did not exactly match the georeferencing of the digital model for the Tälesbahn. This means that each position had to be adjusted; hence, a proximity analysis was carried out. This approximation was made between the positions corresponding to the inertial measurements and the fixed georeferenced segments defined for the digital model (Section 3.1).

Finally, given the discretization of the infrastructure, each segment was longer than the spatial distance between inertial data points, as consequence of the high sampling frequency. Thus, multiple measurements had to be aggregated and associated to each segment. Consequently, a statistically representative value summarizing the data associated to each segment was required. For this purpose, the root mean square (RMS) value of the measurements associated with each segment was chosen. This metric is defined as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1}$$

where x_i represents an individual inertial data point and N is the number of data points per segment. The RMS value is not affected by the sign of the measurements, which makes

it an adequate metric for analyzing signal amplitudes, which in turn is useful to detect potential anomalies and defects. However, this metric is not suitable to represent longitudinal acceleration values along the track, which are associated to vehicle acceleration and braking processes. For this reason, when aggregating longitudinal acceleration values, the positive or negative sign of the average acceleration in the segment was retained and assigned to the RMS value. In this work, this metric is defined as signed RMS.

5 Monitoring Data Visualization Tool

Data visualization has a particular relevance in infrastructure management, as it contributes to a quick and clear understanding of large data volumes. The visual representation developed in this work supports infrastructure-oriented depiction of monitoring data, useful for trend identification in the track condition as well as operation analysis. It also allows an initial approximation to fault detection, which ultimately enhances a condition-based maintenance decision-making process.

5.1 Data Categorization

To achieve a meaningful visualization of the monitoring data, the aggregation via RMS values described in Section 4.1 was employed. The analysis builds upon the determination of the data distribution via the shape of the probability density functions inferred from the data histograms. Two methods for RMS data categorization are used in this work:

1. Categorization via the interquartile range (IQR) criterion.
2. Categorization via pre-established percentiles thresholds.

The first approach allows anomaly detection in the signal amplitudes via outliers identification, which was carried out on the basis of the boxplot representation of the data and the IQR, defined as:

$$IQR = q_3 - q_1 \tag{2}$$

In Eq. (2), q_1 and q_3 correspond, respectively, to the first and third quartiles of the data distribution. The outlier identification works as follows: let X represent the data set of

(signed) RMS values, as discussed in Section 4.1; according to the IQR criterion, the outliers set $\mathscr{O}$ can be defined as:

$$\mathscr{O} = \{x \in X \mid x < q_1 - k_1 \cdot IQR \text{ or } x > q_3 + k_2 \cdot IQR\} \tag{3}$$

Where k_1 and k_2 are positive constants that help to define the thresholds for outlier detection. Based on this criterion, the RMS values can be categorized as:

- Normal values: $\{x \in X \mid q_1 - 1.5 \cdot IQR \leq x \leq q_3 + 1.5 \cdot IQR\}$
- Outliers values:
  - Lower outliers: $\{x \in X \mid q_1 - 3 \cdot IQR \leq x < q_1 - 1.5 \cdot IQR\}$
  - Upper outliers: $\{x \in X \mid q_3 + 1.5 \cdot IQR < x \leq q_3 + 3 \cdot IQR\}$
- Extreme outliers:
  - Lower extreme outliers: $\{x \in X \mid x < q_1 - 3 \cdot IQR\}$
  - Upper extreme outliers: $\{x \in X \mid x > q_3 + 3 \cdot IQR\}$

The second RMS values categorization approach relies on thresholds definition based on specific distribution percentiles. These percentiles are useful to establish lower or upper bounds of accelerations and angular velocities, in order to identify events outside the normal operation, and geometric characteristics of the infrastructure itself.

Considering both data categorization approaches, two types of indicators were established: (1) potential fault indicators, and (2) operation or infrastructure geometrical indicators. The IQR criterion was utilized to define the categorization of the former indicators and percentiles thresholds were applied to the latter indicators.

The inertial signals selected to build potential fault indicators were the vertical acceleration and the pitch angular rate. Since the IQR criterion is applied to these indicators, only upper outliers and upper extreme outliers are considered, as the RMS values for these signals are positive. The remaining signals, including longitudinal acceleration, transverse acceleration, roll and yaw angular rate, were used to form operation or geometrical indicators.

In the rest of this section, three illustrative case studies for data representation (vertical acceleration, longitudinal acceleration and yaw angular rate) are presented, based on data

from a particular run within the Reallabor 2.0 data set. Its behavior is equivalent and representative of all other runs.

5.2 Case Study I: Detection of Potential Track Faults via Visualization of the Vertical Acceleration

Vertical acceleration is particularly significant for potential detection of medium- wavelength track defects. This signal also provides information regarding structural singularities of the track such as level crossings, bridges, turnouts, etc. This means, several discontinuities in the track could be detected by analyzing the vertical acceleration.

For this case study, since there was a significant amount of upper outliers (7.04% for bogie signals and 5.44% for car body signals) and upper extreme outliers (2.65% for bogie signals and 1.35% for car body signals), the IQR method from Section 5.1 was chosen to define the categorization.

Vertical acceleration measurements are influenced by the railway vehicle suspension dynamics; the further the sensor is located from the track, the lower the amplitude and frequency content of the signals. This explains the lower proportion of extreme outliers present in the car body vertical acceleration. Both bogie and car body signals were relevant for fault detection. However, extreme outliers identified in car body measurements are of particular relevance, since these points indicate high signal amplitudes, even after the filtering effect of the primary and secondary suspensions. In the case that these outliers values do not correspond to known structural singularities, they could be related to potential track irregularities and defects. Three categories were defined for potential fault detection, according to the IQR method: (1) normal values (depicted in green color in the visualization over the infrastructure), (2) upper outliers (yellow) and (3) upper extreme outliers (red).

Figure 2 presents an extract of the representation of the RMS values characterization for vertical acceleration measured in the car body. Most of the observations related to outliers and extreme outliers had a clear justification linked to the infrastructure, such as level crossings, turnouts and bridges. However, there was a specific section classified mainly as extreme outlier, between Im Roßdorf and Feldweg crossings, which could not be explained or supported by the characteristics of the track. This result is particularly

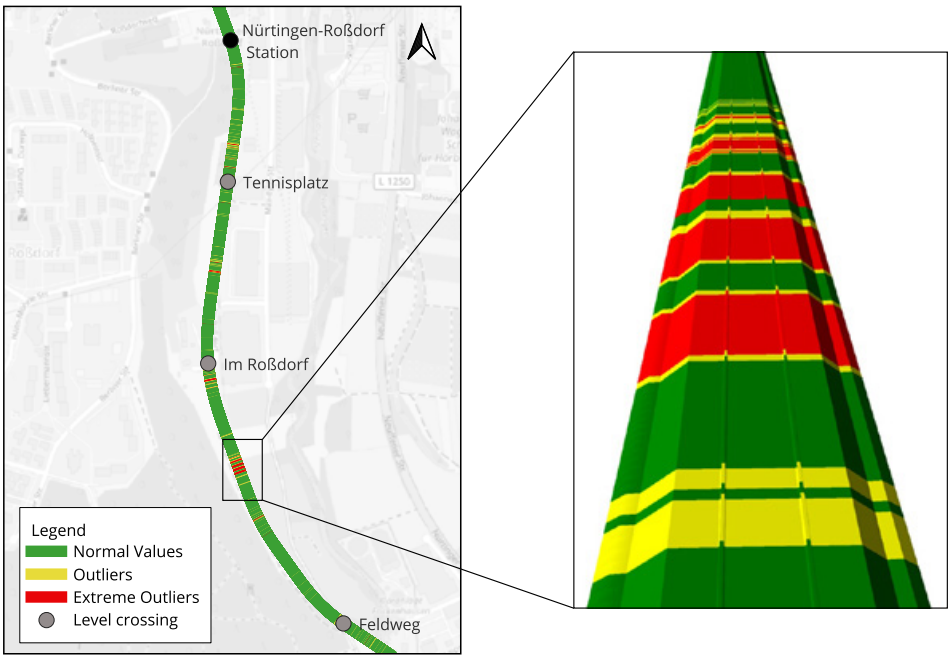


Figure 2: RMS categorization for vertical acceleration on the car body.

relevant, because it can be related to a potential fault. Figure 2 also presents a close-up view of the potential fault section in the 3D digital model. Figure 3 shows the RMS values relative to the segments of the digital model over a short extract of the track. The section highlighted in red corresponds to the potential fault. This section was declared as a temporary speed restriction section (TSR) due to track condition by the infrastructure operator, supporting the results obtained through this method.

5.3 Case Study II: Vehicle Speed Changes via Visualization of the Longitudinal Acceleration

The longitudinal acceleration collected on the vehicle does not directly provide information about the condition of the railway track. However, it offers information about the operation of the railway vehicle, specifically regarding speed changes. Here, the visualization was intended to reflect the significant increases and decreases in speed observed over the infrastructure.

For this case study, as there exists a low percentage of outliers for both sensor locations (bogie and car body), the IQR-based categorization method described in Section 5.1 was

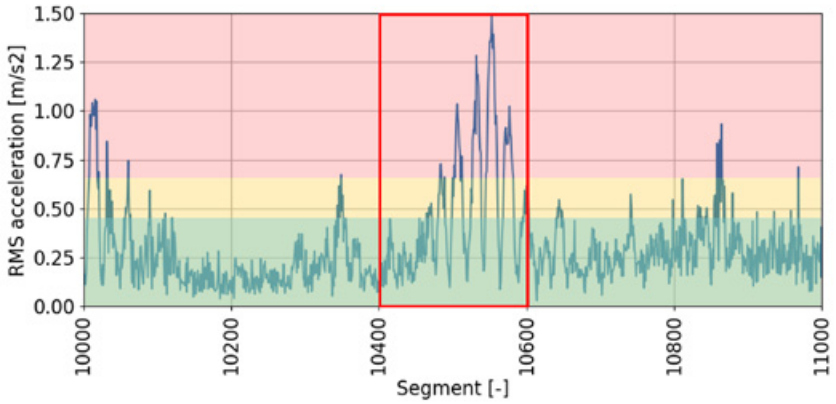


Figure 3: RMS values of vertical acceleration in the potential fault section.

not particularly effective for visualization of speed variations. As the analysis focused on the extreme events of acceleration and braking, the second categorization approach in Section 5.1 resulted more suitable for this application. Here, the 10th and 90th percentiles were chosen to determine the variation thresholds. The values between the 10th and 90th percentiles represented nearly constant speeds. On the other hand, values below the 10th percentile and above 90th percentile corresponded to the most pronounced speed decrements and increments, respectively. As result, the categories were defined as: (1) significant decrease in speed (depicted in red color), (2) quasi-constant speed (green), and (3) significant increase in speed (yellow).

Figure 4 shows an extract of the representation of the signed RMS values for longitudinal acceleration in the digital infrastructure model, measured on the car body. As expected, the most pronounced speed decrements are associated to the train approaching stations and curve entries, while the largest speed increments are observed during train departures from stations. In other cases, substantial speed reductions could be related to emergency braking due to the presence of obstacles in the infrastructure. Furthermore, a potential application of longitudinal acceleration measurements could be the investigation of compressive and tensile forces in wagons, with potential impact on coupling performance. Another possible application is the modeling of travel time distribution and the representation of driver behavior, as well as the repercussion of this behavior on travel time.

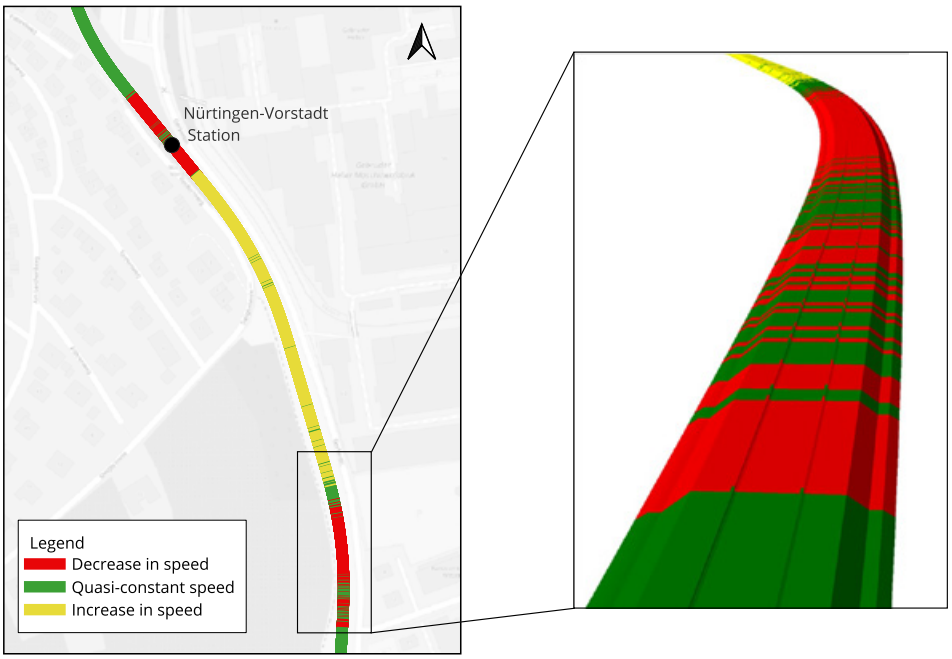


Figure 4: Signed RMS categorization for longitudinal acceleration on the car body.

5.4 Case Study III: Detection of Curves via Visualization of the Yaw Angular Rate

This signal represents the rotation velocity of the railway vehicle along its vertical axis, which reflects the left and right turns when the vehicle moves through a curve. The 80th percentile was selected as threshold for the RMS value of yaw angular rate to visualize the highest angular rates, which are related to the most pronounced curves in the alignment. Two categories were defined: (1) tangent section (depicted in green color) and (2) curve section (red).

Figure 5 presents an extract of the representation of the RMS values characterization for yaw angular rate. The visualized colors are derived from the inertial data measured from the car body, matching the infrastructure geometry. In this way, the depiction provides an effective representation of the vehicle movement, and its validation with the railway layout. This means that the identification of curve elements through yaw angular rate measurements, facilitates track mapping and vehicle localization applications.

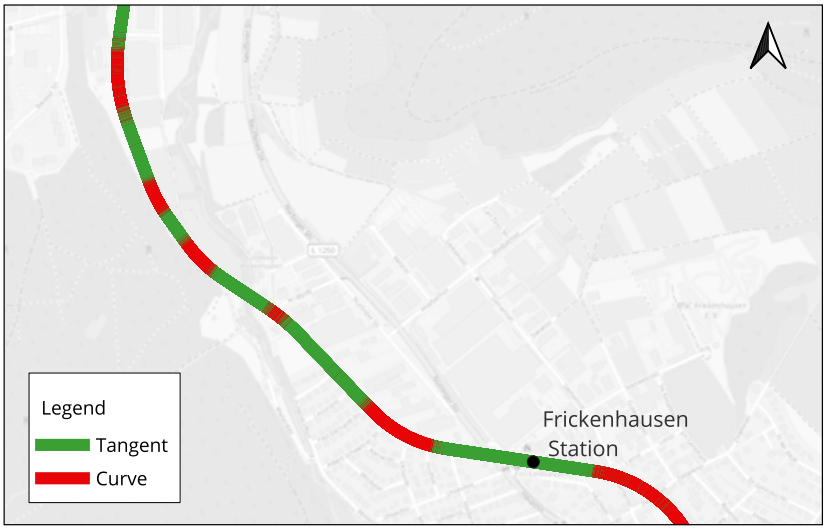


Figure 5: RMS categorization for yaw angular rate on the car body.

6 Conclusions

In this work, an infrastructure-oriented data visualization tool was developed to map monitoring-related signals into the railway track and depict them in a meaningful way. This tool consisted of a 3D digital infrastructure model and a relational database for data modeling. The measurements were collected via the vehicle-based track monitoring system Reallabor 2.0.

The georeferenced digital model of the railway line developed in this work has the capacity to represent the geometrical and physical characteristics of the track, offering a comprehensive overview of its elements in a 3D environment. A spatial discretization of the model was performed to assure that even the track characteristics with short wavelengths within the scope of the monitoring system could be properly represented. This model not only provides a detailed depiction of the infrastructure, but also constitutes the foundation for further integration and visualization of monitoring data.

The developed relational database provides a well-structured solution for storing and integrating the monitoring data with the available infrastructure-related information. Data integration is effectively achieved using spatial and statistical analysis, providing a meaningful categorization of the inertial measurements via the RMS value.

The visualization tool consolidates the 3D digital model with the integrated data from

the database. By the means of a color-coding representation of the defined categories, indicators regarding the infrastructure geometry, operation and potential track faults are achieved. Some of these indicators were analyzed in the case studies in this article. An opportunity for future research is the assessment of the temporal evolution of the track condition based on monitoring data, contributing to predictive maintenance strategies.

Acknowledgment

This work was supported by the German Research Foundation (DFG) under the cooperation project "Efficient Sensor-Based Condition Monitoring Methodology for the Detection and Localization of Faults on the Railway Track (ConMoRAIL)" with the Institute of Engineering Geodesy (IIGS) of the University of Stuttgart and the Württembergische Eisenbahn-Gesellschaft mbH, grant no. 515687155.

Literature

[1] Bado, Mattia F. ; Tonelli, Daniel ; Poli, Francesca ; Zonta, Daniele ; Casas, Joan R.: Digital Twin for Civil Engineering Systems: An Exploratory Review for Distributed Sensing Updating. In: *Sensors* 22 (2022), Nr. 9. – ISSN 1424–8220

[2] BrightSky: *Bright Sky Demo*. https://brightsky.dev/, 2024

[3] Fernández-Bobadilla, Héctor A. ; Martin, Ullrich: Modern Tendencies in Vehicle-Based Condition Monitoring of the Railway Track. In: *IEEE Transactions on Instrumentation and Measurement* 72 (2023), S. 1–44

[4] Lerke, Otto ; Bahamon-Blanco, Sebastian ; Metzner, Martin ; Martin, Ullrich ; Schwieger, Volker: Vorarbeiten zur Entwicklung eines Gleisfehlerdetektionssystems mit Regelzügen und Low-Cost-Sensorik. In: *Zeitschrift für Geodäsie, Geoinformation und Landmanagement – zfv 146* (2021), S. 207–217. http://dx.doi.org/10.12902/zfv-0339-2021. – DOI 10.12902/zfv–0339–2021

[5] Podwórna, M.: Modelling Of Random Vertical Irregularities Of Railway Tracks. In: *International Journal of Applied Mechanics and Engineering* 20 (2015), Nr. 3, 647–655. http://dx.doi.org/10.1515/ijame-2015-0043. – DOI 10.1515/ijame–2015–0043. – ISSN 1734–4492

[6] Prandoni, P. ; Vetterli, M.: *Signal Processing for Communica-*

tions. PRESSES POLYTECHNIQUES ET, 2008 (Communication Sciences). https://books.google.de/books?id=25XmDAAAQBAJ. – ISBN 9782940222209

[7] QUDDUS, Mohammed A. ; OCHIENG, Washington Y. ; NOLAND, Robert B.: Current map-matching algorithms for transport applications: State-of-the art and future research directions. In: *Transportation Research Part C: Emerging Technologies* 15 (2007), Nr. 5, 312-328. http://dx.doi.org/https://doi.org/10.1016/j.trc.2007.05.002. – DOI https://doi.org/10.1016/j.trc.2007.05.002. – ISSN 0968–090X

[8] VANDERHORN, Eric ; MAHADEVAN, Sankaran: Digital Twin: Generalization, characterization and implementation. In: *Decision Support Systems* 145 (2021), S. 113524. – ISSN 0167–9236

Authors



Laura T. Rodríguez-Bayona, M.Sc.
Academic staff member at the Institute of Railway and Transportation Engineering at the University of Stuttgart. laura.rodriguez@ievvwi.uni-stuttgart.de



Héctor A. Fernández-Bobadilla, M.Eng.
Academic staff member at the Institute of Railway and Transportation Engineering at the University of Stuttgart. hector.fernandez@ievvwi.uni-stuttgart.de



Prof. Dr.-Ing. Ullrich Martin
Professor and Director of the Institute of Railway and Transportation Engineering at the University of Stuttgart and Director of the Institute of Transportation Research Stuttgart. ullrich.martin@ievvwi.uni-stuttgart.de

Session 14
Ute Gläser, Joseph Feng, Mirjana Jovanovic

Decision support for tamping activities

A predictive maintenance approach

Ute, Gläser[1], Joseph, Feng[2], Mirjana, Jovanovic[2]

[1] Fraunhofer Institute for Transportation and Infrastructure Systems IVI
[2] DB InfraGO AG

Summary

The central task of every railroad infrastructure operator is to keep the track properly maintained and available for operation. To achieve this, maintenance management must meet both safety and economic challenges. In detail, the threshold values considered for the superstructure are defined at European level by the EN13848 [1] and at the national level by DB Group Guideline (KoRil) 821.2001 [2]. The condition assessment of the track geometry for example is determined by the measured values obtained during prescribed regular inspections. Due to the individual track condition development of each segment, the variety of machine types and the wide range of possible possession times, an exact solution of the resulting linear mixed-integer program would require a high numerical effort that results in an unfeasible computation time. Therefore, we present a heuristic solution approach that solves the described optimization problem with sufficient accuracy in a significantly shorter runtime.

Keywords: predictive maintenance, operational planning, decision support

1 Problem description

Predictive and condition-based maintenance planning are integral to decision support for railway tamping activities, transforming maintenance from a fixed-interval schedule to an optimized, data-driven process, see [3] and [4]. The primary goals of using prediction of track geometry for tamping are to enhance efficiency, safety, and cost-effectiveness [5].

In more detail, effective maintenance planning involves not just maximizing the number of interventions during a limited shift but ensuring the right actions are taken at the right time. Therefore, a planning approach is proposed that enables asset managers to create predictive maintenance plans for various scenarios and evaluate them based on costs and quality of track geometry. These scenarios include different ownership conditions and planning parameters. A sound database that not only describes the network itself but also reflects the development of its condition with sufficient accuracy forms the basis for reliable operational planning.

The first section therefore explains the necessary database for the planning approach before moving on to explaining and analyzing the algorithm itself in the second section.

1.1 Prediction of the degradation

A valid prediction of the degradation of track geometry is one of the most important foundations for reliable planning and scheduling of maintenance. Only a reliable prediction paves the way from reactive or preventive to predictive maintenance planning to extend the lifetime of components and improve the reliability of the infrastructure. In [6], it is analyzed that the amplitude of single failures can reach the threshold that triggers maintenance before the standard deviation reaches this value, resulting in the need to maintain the single failure before planned maintenance. At the same time, this approach implies two requirements, namely a high quality of historical data in terms of reliability and reproducibility and a high positional accuracy, which are essential for a reliable prediction. The database used from DB InfraGO meets these necessary criteria, allowing the model to predict track geometry deterioration for individual defects or standard deviations. To facilitate the prediction of track degradation, the railway track is segmented into sections of 25 meters in length. Measurement data for each section is subjected to pre-processing to eliminate measurement errors and anomalies, ensuring data integrity. The degradation model developed in this study is grounded in historical inspection data. Accordingly, the planning model presented herein accounts for the distinct conditions of each 25-meter section, enabling the consolidation of maintenance activities and associated tasks into organized work shifts.

1.2 Data base used

A comprehensive database is essential for predictive maintenance planning. The database required for the development of the decision support tool can be subdivided as follows:

Network data. The network under consideration must be defined in terms of its structural characteristics (i.e., type of asset, type of structure, and buildings etc.) and geographic location to enable effective and cost-efficient planning. Based on this data, a graph-theoretical model is created, where operational points are interpreted as nodes and routes as edges. This model forms the basis for path optimization and requires information on the localization, the type, the maximum permitted speeds, the maximum load and the available capacity of the individual route sections.

Condition data. For each 25 m segment an accurate and reliable prediction of the track geometry parameter with respect to the maintenance threshold is provided. The measurement data is pre-processed for each segment to remove measurement errors and anomalies. The results of the used degradation model which is based on historical inspection data are shown as an example in Figure 1, where the prediction for a 20 km long section is given. Each 25 m segment is represented by a thin stripe whose color reflects the state of its track geometry. In detail the red area of the color scale, i.e. the range from 90 % to 100 %, implies that the track geometry parameter reaches the threshold value. Furthermore, the historical values are illustrated by a time less than or equal to zero whereas the predicted track geometry is represented by a time greater than zero.

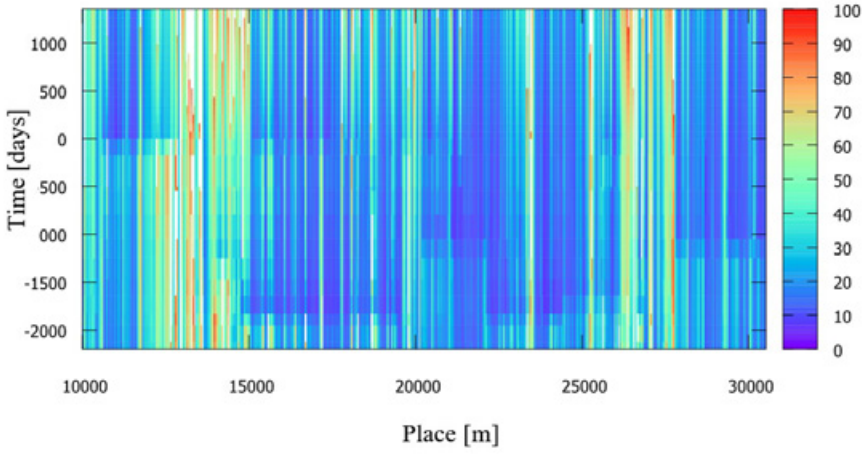


Figure 1: Historical and predicted values of the normalized longitudinal level with maintenance thresholds (color code)

Specification of machines. The tamping machines differ in their working speed, their working area (switches and/or tracks) and the depot in which they are currently located. The user can set these parameters and also vary the number of machines available.

Planning parameters. The maintenance planners can customize the threshold values and time horizon of planning. Among other parameters, the threshold value from which segments are to be assessed as critical, the time horizon of the planning, the maximum shift length and maintenance corridors can be selected.

Maintenance container. So-called maintenance containers are defined in order to allow the construction or maintenance work to take place in stable, pre-planned and announced, temporally and spatially defined cycles. Containers are standardized time windows for the scheduled implementation of maintenance activities (maintenance containers) as well as for larger modernization projects, see [7]. The use of maintenance containers therefore also supports the bundling or combination of maintenance activities.

2 Solution approach

This planning process has two objectives: to improve the quality and reliability of the network and at the same time to minimize the costs of maintenance. These costs include both the construction costs, and the costs associated with setting up the track closure. To ensure the latter, interventions are planned in three different types of possession times, maintenance container times, shadow times and extra times. Maintenance containers are pre-planned and have relatively long traffic free times, see [7] for more information. Shadow times were created initially for other interventions, and it must be checked whether a combination of construction works is possible. And finally extra times are booked, if inevitable, for they are planned for dedicated maintenance works and imply higher costs. The planning process considering these different types of possession times is shown in Figure 2. Once the input data has been read and processed, all work shifts are scheduled on the assumption that all possible time slots and planning containers can be fully booked. The planning step therefore results in shifts that are either in suggested maintenance containers or in suggested shorter time slots. It is then checked whether the planned tamping activities are compatible with maintenance activities already planned in the containers. If this is the case, the shifts are allocated to containers and added to the maintenance plan; otherwise, a new planning step is started with the critical segments that were in containers that could not be planned. It is then examined whether the calculated shifts fit into shadow times or whether new time slots need to be booked for this purpose. The planning process shown in Figure 2. should be understood as a continuous process, i.e. segments that cannot be planned in the current planning horizon are moved to the next one and given higher priority there.

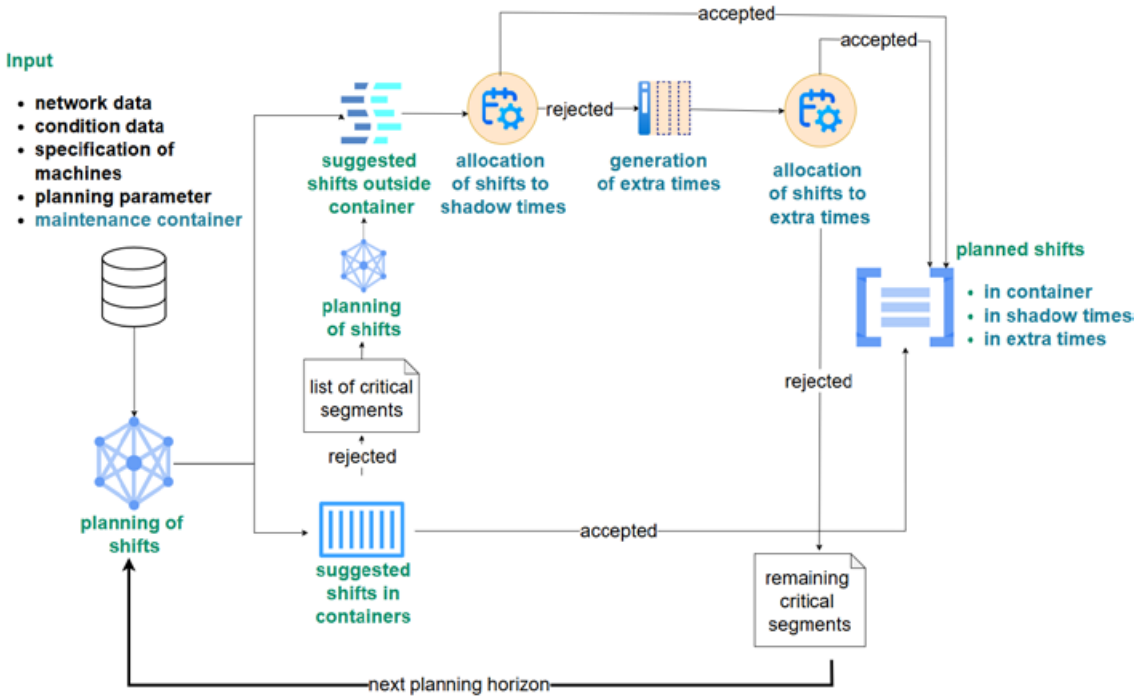


Figure 2: Planning process

The size of the infrastructure network and the the multi-layered planning criteria to be considered lead to a very high complexity of the optimization problem, whereby the exact solution of the resulting linear mixed-integer program would require a high effort and accordingly a long computing time. To solve the described optimization problem with an acceptable runtime and sufficient accuracy, a heuristic solution approach is proposed in the following.

2.1 Heuristic solution approach for the planning of shifts

The flow chart of the planning step for the tamping shifts is shown in Figure 3. As already mentioned, a valid prediction of the deterioration of track geometry is most important for reliable planning. This and a graph-theoretical representation of the infrastructure network form the starting point for planning the tamping shifts. Based on the prediction data a list of 25 m segments, that will exceed the predefined maintenance threshold in the considered time horizon, is created. We distinguish between segments of the ‘switch’ or ‘track’ type, as not all tamping machines are able to tamp switches, but they can certainly tamp track. Furthermore, switches are considered particularly critical elements in the

network, meaning that they should be planned first with the greatest possible degree of freedom, i.e. the highest possible machine and time window availability.

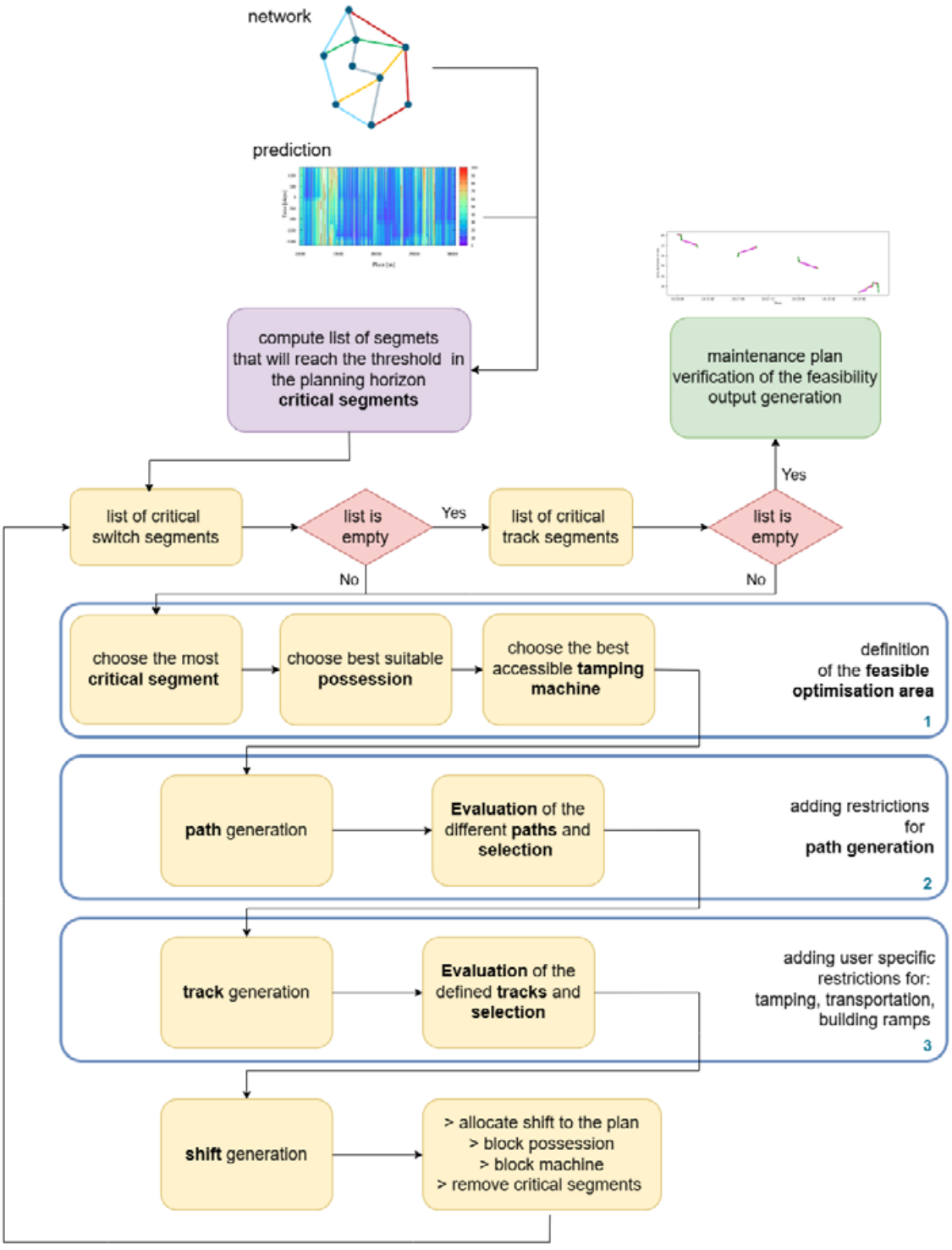


Figure 3 Flow chart for shift planning

As shown in Figure 3 Figure, the planning process is divided into three steps, namely:

1. definition of the feasible optimization area,
2. path selection,
3. tamping track generation and shift definition.

In the first step initially the most critical segment is determined. For this segment a check is performed to determine whether a predefined maintenance container is available within the planning horizon or whether a time window must be allocated. Once the container/time window has been defined, the fastest and most suitable (available and appropriate type) tamping machine is selected and assigned to the tamping shift to be planned. Thus the local optimization area is defined based on the critical segment, the selected maintenance container/time window and the speed of the selected tamping machine. In the second step all paths, i.e., a sequence of nodes and edges that can be reached by the considered tamping machine within the specified time window, are defined and subsequently evaluated, whereby those paths that contain many critical segments are given high priority. In order to plan the exact tamping sections for a path, it is important to note that during a possession time, not only the working time but also the preparation and transition time must be considered to develop a reliable shift schedule. In Figure 4, a schematic representation of a shift, including transport and working mode as well as the construction of ramps at both ends of the tamping section, is illustrated. The transport mode of the tamping machine is shown by the green line and the working mode by the red line. The speeds of the two modes differ significantly. Consequently, the algorithm must distinguish between two areas to be tamped whether it is better to tamp through or to convert the machine. The blue lines indicate that a ramp has to be tamped at the beginning and end of each tamping area, which takes additional 5 minutes.

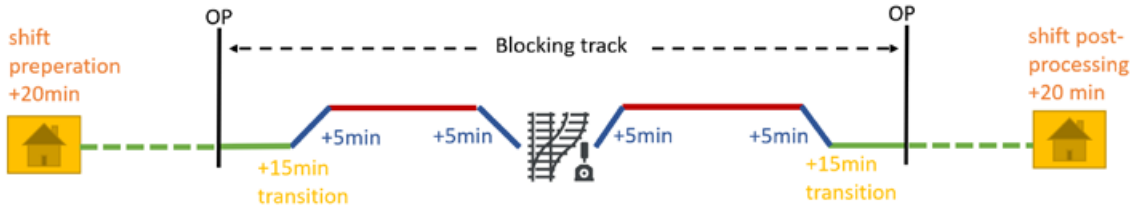


Figure 4: Schematic representation of a shift

In the third step all possible shifts are generated for each path of high priority, taking into account technical and company-specific planning conditions. Examples of these conditions are:

- The construction of a ramp (obligatory at the beginning and end of the tamping area, see Figure 3) is only carried out on sections in good condition (track geometry parameters are below a certain threshold).
- Ramps must not be built in transition curves or at level crossings.
- If the time required to set up and dismantle the tamping machine is longer than the time required to tamp the section between two critical areas to be tamped, the entire area should be tamped.
- Switches can only be tamped by a universal tamping machine. Special treatment of switches: Switches are always maintained as a whole.
- If the path includes a change of direction or a change of route during the shift, additional time needs to be added
- The tamping machine is only allowed to enter or leave the track at so called operating points
- The tamping machine can drive in both directions but to tamp only in one direction.

The shifts are generated considering these conditions, the machine specification and the length of possession time. To ensure that the critical segment is tamped even with slight variations in the execution, care is taken to ensure that the critical segment is located in the middle of the shift. Finally, each computed shift is evaluated by the priority value. This prioritization value is defined as the sum of the track geometry parameter across all segments of a shift to be tamped that exceed a user-specified threshold. This definition ensures that shifts are favored which contain many segments with high criticality. According to Figure 2, the best possible time window is now selected to schedule the shift, considering maintenance container, shadow times, and extra times. If several time slots are available, the costs and the earliest possible availability are considered. After planning this shift, all segments to be processed in the shift are deleted from the list of segments to be planned (corresponds to the list of critical segments) and the planning procedure starts again with the next critical segment until all critical segments are planned, or a capacity limit has been reached in terms of available machines or available time windows.

This planning results in a schedule of tamping activities, with other words in a list of work shifts, where each shift is assigned to a specific possession time, a specific machine, and a list of sections to be tamped during that shift. In addition, a list of critical sections that cannot be processed during the planning horizon due to capacity constraints is provided. These segments are shifted to the next planning horizon where they have a higher priority.

This proposed planning approach will lead to more consistent quality development across the network by maintaining contiguous areas rather than isolated patches.

3 Developed planning tool

The planning tool is accessed via a user-friendly graphical user interface (GUI), the start page of which is shown in Figure 5.

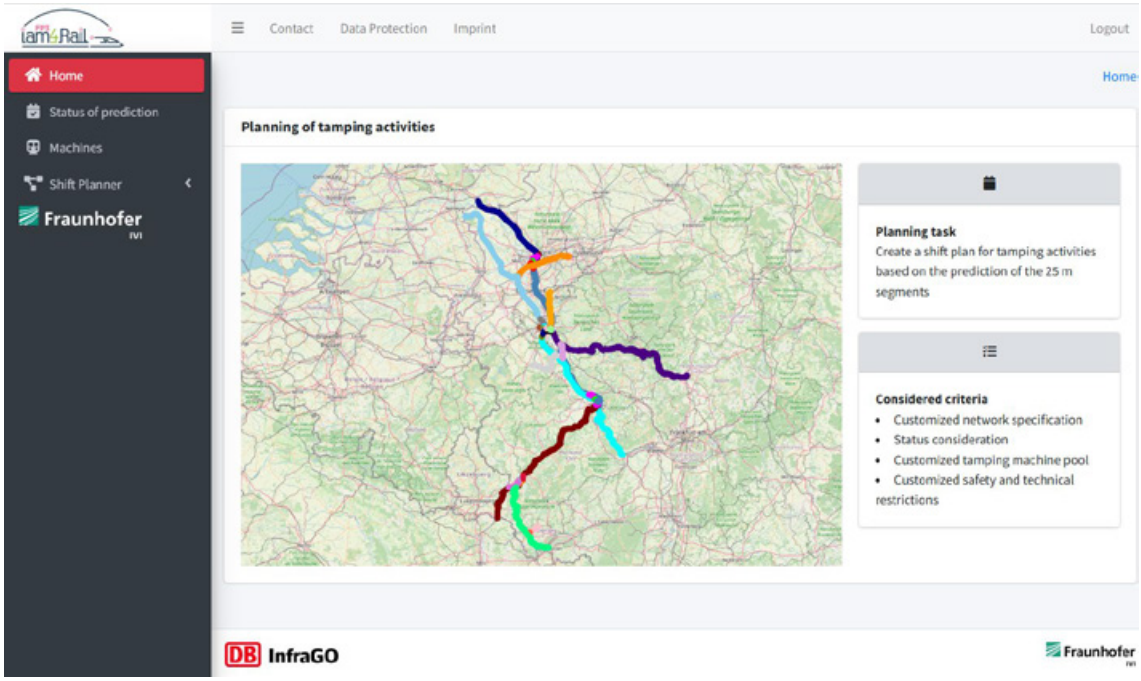


Figure 5: GUI of the planning tool

After user authentication, the user can check the current status of the predictions used, define tamping machine types to be used in the planning and set planning parameters such as, among others start date, planning horizon, ramp size and the criticality threshold of the geometric parameter.

The computed maintenance schedules are made readily accessible not only for download but also through GUI. Within the GUI, users can have a comprehensive overview of the generated shifts, which includes detailed information for each shift, such as the start time, duration, utilized machine, and possession ID. Additionally, to enhance the understanding of shift utilization, the interface provides a breakdown of the total duration, including specific time allocations for tamping, transport, transition, and ramp construction. This

feature ensures that users can easily assess and manage maintenance activities while gaining insights into resource allocation and operational efficiency.

4 Summary and Outlook

The process of creating maintenance task plans typically requires significant human resources. This tool provides the capability to rapidly develop a plan while adhering to existing capacity, safety and technical constraints. Additionally, it allows for an effective comparison of various planning configurations and their respective impacts on network quality and the efficiency of maintenance activities.

Furthermore, the tool attempts to formalize and standardize the current planning process for tamping activities. In particular, it consolidates the expert knowledge of maintenance planners, which is essential for preserving expert knowledge in light of demographic change.

However, it is important to note that our analysis has focused solely on tamping activities, which means certain simplifying assumptions have been made.

Looking ahead, the next stages of development will involve mapping out a more intricate planning process that takes into account a broader range of maintenance activities and variables. This refinement aims to enhance the tool's accuracy and applicability in real-world scenarios, ensuring that it can better support maintenance planning efforts in a comprehensive manner.

Acknowledgements

Funded by the European Union. Views and opinion expressed are however those of the authors only and do not necessarily reflect those of the European Union or Europe’s Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them. The project is supported by the Europe's Rail Joint Undertaking and its members. This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101101966.

5 Bibliography

[1] EN13848, „Railway Applications - Track - Track geometry quality -Part 1," 2019.

[2] D. Bahn, „Richtlinien Oberbau 821: Oberbau inspizieren".

[3] M. Wen, R. Li und K. B. Salling, „Optimization of preventive condition-based tamping for railway tracks," Optimization of preventive condition-based tamping for railway tracks, Bd. 252, Nr. 2, pp. 455-465, 2016.

[4] A. Jamshidi, S. Hajizadeh, Z. Su, M. Naeimi, A. Núñez, R. Dollevoet, B. Schutter und Z. Li, „A decision support approach for condition-based maintenance of rails based on big data analysis," Transportation Research Part C: Emerging Technologies, 2018.

[5] M. Sedghi, E. V. B. Bergquist und A. Migdalas, „Data-driven maintenance planning and scheduling based on predicted railway track condition," Quality and Reliability Engineering International, 2022.

[6] IN2SMART2, „Deliverable 10.2 Data Analytics," 2022.

[7] DB InfraGO, „Paradigmenwechsel beim Bauen," 2024. [Online]. Available:https://www.deutschebahn.com/resource/blob/13191934/e60b1f1086d25b033fd0ed99da41a5b6/241120_ZIB-Forum_Vortrag_Bauen-im-Takt-data.pdf.

Session 14

Ute Gläser, Joseph Feng, Mirjana Jovanovic

Autoren



Ute Gläser is a postdoctoral researcher at the Fraunhofer Institute for Transportation and Infrastructure Systems (IVI). Her research focusses on the development of decision support systems for linear infrastructures. She holds a doctorate in mathematics and previously worked as a research assistant at the Technical University of Berlin.



Joseph Feng, expert in railway dynamics and assessment in DB InfraGO AG (I.IAI 24), Munich, is the task leader of WP 9.4 in the EU Project IAM4RAIL. He is working on measurement systems and innovative maintenances for track geometry. He holds a doctorate in Engineering Sciences of Technical University Berlin.



Mirjana Jovanovic holds a doctorate in civil engineering (Technical University of Karlsruhe) and is a member of the department for the measurement engineering and the wheel-rail-system of the DB InfaGO (I.IAI 24). For many years she has been working on development and calculation of prediction of track geometry for tamping maintenance planning.

IRSA 2025

SESSION 15

DATEN UND PLANUNG
DATA AND PLANNING

Session 15
Mario Trappein, Dennis Götz

LeiDa-D – Versionierte, zentralisierte Datenhaltung für Basis- und Infrastrukturdaten

Trappein, Mario[1], Götz, Dennis[2]

[1]quattron GmbH
[2]DB InfraGO AG

Zusammenfassung

Im Rahmen des Programms PRISMA („Projekt Redesign der Informationssysteme Betrieb und Modernisierung der Architektur“), einem der größten IT-Projekte der DB InfraGO AG, wurde das zentrale System LeiDa-D („Leitsystem der Betriebsführung Zentrale Datenhaltung – Datenhaltung“) der DB InfraGO AG neu entwickelt. LeiDa-D übernimmt die Konsolidierung der Daten aus den regionalen Bestandssystemen und gewährleistet über ein Versionierungskonzept die netzweite Bereitstellung eines aktuellen, einheitlichen Infrastruktur- und Basisdatenbestandes für Dispositions-, Kunden- und Statistiksysteme der DB InfraGO AG.

Keywords: Infrastrukturdaten; Zentralisierung; Basisdaten; versionierte Datenhaltung; netzweiter Datenbestand; Modernisierung der IT-Infrastruktur;

Session 15
Mario Trappein, Dennis Götz

1 Digitalisierung und Zentralisierung als Grundlage für den modernen Bahnbetrieb

Die fortschreitende Digitalisierung und Zentralisierung prägen nicht nur die moderne IT-Welt, sondern auch das System Bahn: unverzichtbare Treiber für Effizienz, Innovation und Vernetzung. Vor diesem Hintergrund adressiert das Programm PRISMA („Projekt Redesign der Informationssysteme Betrieb und Modernisierung der Architektur“), eines der bisher größten IT-Projekte der DB InfraGO, genau diese Herausforderung: die Modernisierung und Zentralisierung der IT-Infrastruktur in den Bereichen Fahrplanbearbeitung, Disposition, Zugmeldung und statistische Auswertungen.

Ein zentralisierter, netzweit einheitlicher Datenbestand bildet dabei die Grundlage für die Harmonisierung und Zentralisierung der genannten Komponenten des Eisenbahnbetriebs. Im Rahmen von PRISMA hat das Projekt BaSIs-BZ („Basis-, Stamm- und Infrastrukturdaten Betriebszentrale“) diese Kernaufgabe übernommen. In enger Kooperation zwischen der DB InfraGO AG, der DB Systel GmbH und der quattron GmbH entstand das Produkt LeiDa-D („Leitsystem der Betriebsführung Zentrale Datenhaltung – Datenhaltung“), das konzipiert, entwickelt und erfolgreich in den neuen Systemverbund integriert wurde. Die DB InfraGO AG war dabei die treibende Kraft, die die Entwicklung initiiert, beauftragt und mit ihrem Engagement den Fortschritt maßgeblich vorangetrieben hat. Die quattron GmbH konnte im Auftrag der DB InfraGO AG einen wesentlichen Beitrag im Bereich Architektur sowie bei den Entwicklungsleistungen für das Produkt leisten. Dieser Erfolg wurde durch die fachkundige Unterstützung und langjährige Expertise der DB Systel GmbH zusätzlich bereichert.

2 Einheitliche und zuverlässige Datenbasis für den Bahnbetrieb

Die Kernfunktion von LeiDa-D besteht darin, regionale Daten aus den Betriebszentralen aufzubereiten und als einheitlichen, netzweiten Datenbestand bereitzustellen. Dieser bildet die Grundlage für alle angebundenen Dispositions-, Statistik- und Kundensysteme. Der Datenbestand umfasst primär Basis-, Stamm- und Infrastrukturdaten, die derzeit noch im etablierten regionalen Bestandssystem LeiDa-S („Leitsystem Datenhaltung – Systemdaten“) projektiert und anschließend an LeiDa-D übermittelt werden.

Ein zentrales Ziel von LeiDa-D ist die Sicherstellung der regionalen Autonomie der Betriebszentralen. Dementsprechend wurde das System so gestaltet, dass regionale Daten unabhängig verarbeitet und erst anschließend zu einem zentralisierten, netzweiten Datenbestand konsolidiert werden. Hierbei kommt ein speziell entwickeltes Versionierungskonzept zum Einsatz, das die Zusammenführung und Konsolidierung dieser regionalen Daten zu einer netzweit gültigen Grundlage unterstützt.

LeiDa-D ermöglicht die Abbildung unterschiedlicher Datenstände, wodurch jederzeit ein konsistenter und vollständiger netzweiter Datenbestand erzeugt werden kann. Zu festgelegten Aktivierungszeitpunkten – wie dem europaweiten Fahrplanwechsel im Dezember oder bei unterjährigen Ereignissen – werden die Daten für alle angebundenen Systeme freigeschaltet. Dieses Verfahren gewährleistet, dass sämtliche Systeme synchronisierte Zugriffe auf denselben zentralisierten Datenbestand haben und somit eine einheitliche Systemsicht entsteht.

Neben den Basis- und Infrastrukturdaten aus regionalen Bestandssystemen integriert LeiDa-D auch zusätzliche Datenquellen, teils mit eigener Datenpflegemöglichkeit. Beispielsweise können Ersatzfahrplanhefte oder spezifische Geschäftspartner-Daten in den zentralen Datenbestand aufgenommen werden. Um den unterschiedlichen Anforderungen der angeschlossenen Systemkomponenten gerecht zu werden, stellt LeiDa-D eine webbasierte, tabellarische Projektierungsplattform bereit, über die beispielsweise die Stammdaten für Verspätungsbegründungen und die Störfall- und Statistikbearbeitung erfasst werden. Diese Funktionalitäten werden nicht nur von betrieblichen Administratoren genutzt, sondern auch von der zentralen fachlichen Betriebsführung und der Netzleitzentrale.

Die Qualität der bereitgestellten Daten ist entscheidend für eine fehlerfreie und effiziente Disposition. LeiDa-D führt daher umfangreiche Konsistenzprüfungen durch, um die Datenbestände zu validieren und den Anwendern übersichtlich aufzubereiten. Mithilfe hunderter Prüfmechanismen wird der netzweite Datenbestand regelmäßig analysiert – und das innerhalb kürzester Zeit. Ein besonderer Fokus liegt dabei auf den Übergangsbereichen der jeweiligen Regionaldaten, um ein einheitliches Verständnis der Infrastruktur sicherzustellen. Regionale Besonderheiten in der Infrastruktur-Darstellung werden dabei individuell berücksichtigt.

Zur weiteren Unterstützung der Anwender bietet LeiDa-D eine webbasierte grafische Visualisierung der Infrastruktur. Diese umfasst unter anderem Darstellungen von Spurplantopologien, Betriebsstellenfahrwegen und ZN-Fahrstraßen („Zugnummer-Fahrstraße“). Neben der grafischen Darstellung stehen zusätzlich umfangreiche

Analysewerkzeuge zur Verfügung, mit denen die Infrastruktur detailliert untersucht und potenzielle Problembereiche identifiziert und bewertet werden können.

3 Zentralisierung von Daten zu einem netzweiten Datenbestand

Das System LeiDa-D hat die zentrale Aufgabe, alle regionalen Basis- und Infrastrukturdaten zu zentralisieren. Diese Daten werden zunächst in den regionalen Bestandssystemen („LeiDa-S“) erfasst und anschließend an LeiDa-D übertragen. Dabei arbeiten die Betriebszentralen autonom und bestehen aus den sieben Fernbahnregionen sowie der zentralen fachlichen Betriebsführung (ZFBF). Die Autonomie dieser acht Einheiten muss weiterhin gewährleistet werden. Die Aufbereitung der regionalen Daten zu einem netzweiten Datenbestand erfolgt dabei ohne Beeinträchtigung der Bestandsprozesse.

LeiDa-D verarbeitet die regionalen Datenpakete umgehend und analysiert sie mithilfe umfangreicher Prüfmechanismen und liefert gezielte Rückmeldungen an die Anwender. Der durch LeiDa-D aufbereitete netzweite Datenbestand umfasst vorrangig zwei Kategorien: **Basisdaten** und **Infrastrukturdaten**. Diese beiden Kategorien werden unter dem Aspekt der Zentralisierung differenziert behandelt und in den folgenden Abschnitten detailliert beschrieben. Im Kontext von LeiDa-D wird außerdem von sogenannten **Fachobjekten** gesprochen. Diese bezeichnen fachlich zusammengehörige Datenobjekte, bei denen sich die Datensätze inhaltlich, aber nicht strukturell unterscheiden. Der Begriff wird häufig auch synonym für Dialoge und Tabellen verwendet.

3.1 Basisdaten

Basisdaten – oder auch Stammdaten – sind primär durch die zentrale fachliche Betriebsführung (ZFBF) erfasste Informationen, die in den bestehenden regionalen Bestandssystemen (LeiDa-S) größtenteils identisch vorliegen sollten. Daher wird die Zentralisierung der Basisdaten im LeiDa-D als reine Zusammenführung behandelt und unabhängig von der Verarbeitung der Infrastrukturdaten betrachtet. LeiDa-D übernimmt aus sämtlichen Datenpaketen aller LeiDa-S-Systeme die Basisdaten und konsolidiert sie intern.

Bei der Überführung der Basisdaten aus LeiDa-S nach LeiDa-D wird jeder Datensatz einzeln geprüft, ob er anhand fachlicher Merkmale – dem sogenannten fachlichen Schlüssel oder Fachschlüssel – identifiziert werden kann. Der Fachschlüssel ist für jedes

Fachobjekt fest definiert und dient der eindeutigen fachlichen Zuordnung eines Datensatzes. Hierbei gliedern sich die Bestandteile eines Datensatzes in **Objektidentität**, welche die fachlichen Merkmale beschreibt, und **Objektausprägung**, die die weiteren Attribute umfasst.

Ein Beispiel dafür ist das Fachobjekt „Eisenbahnunternehmen“. Ein Datensatz dieses Fachobjekts umfasst die folgenden vier Attribute:

- EBU-UIC/NE-Kz
- EBU-Nr
- Kurzname
- Langname

Die Attribute „EBU-UIC/NE-Kz“ und „EBU-Nr“ bilden den fachlichen Schlüssel und sind somit die Objektidentität des Fachobjekts. Die Felder „Kurzname“ und „Langname“ gehören zur Objektausprägung. Ein Eisenbahnunternehmen mit dem Fachschlüssel „U 8000“ ist in LeiDa-D beispielsweise das Unternehmen mit dem Kurznamen „DB“ und dem Langnamen „Deutsche Bahn AG“.

Der große Vorteil dieses Konzepts liegt darin, dass Basisdaten anhand ihrer fachlichen Schlüssel eindeutig identifiziert und in LeiDa-D zusammengeführt werden können. Das zugrunde liegende architektonische Konzept wurde entwickelt, um die weitgehend identischen Basisdaten, die aus den acht LeiDa-S-Systemen stammen, effektiv zu konsolidieren. Wenn alle LeiDa-S-Systeme beispielsweise einen Datensatz mit dem Fachschlüssel „U 8000“ projektieren, erstellt LeiDa-D eine Objektidentität für das Fachobjekt „Eisenbahnunternehmen“ mit mehreren Objektausprägungen. Im Normalfall existieren pro Fachobjekt und pro Objektidentität acht Objektausprägungen – sieben aus den Fernbahnregionen und eine aus der ZFBF.

Da jedoch die fachlichen Anwender in den Betriebszentralen der Fernbahnregionen sowie die ZFBF mit ihren jeweiligen LeiDa-S-Systeme eigenständig agieren und nicht verpflichtet sind, die Datenerfassung vollständig aufeinander abzustimmen, kommt es zwangsläufig zu Abweichungen. Diese Abweichungen treten in zwei verschiedenen Formen auf:

1. **Unterschiedliche Objektidentitäten:** Nicht alle Betriebszentralen der Fernbahnregionen sowie die ZFBF projektieren innerhalb ihres LeiDa-S-Systemes dieselben Objektidentitäten eines Fachobjekts. Somit kann es vorkommen, dass für bestimmte Objektidentitäten weniger als acht Objektausprägungen existieren.

2. **Unterschiedliche Objektausprägungen:** Auch wenn die Objektidentität übereinstimmt, können die Inhalte der Objektausprägungen zwischen den autonomen LeiDa-S-Systemen variieren.

Damit LeiDa-D jedoch stets einen konsistenten, fachlich korrekten und netzweiten Datenbestand bereitstellen kann, wurden spezifische Mechanismen entwickelt. Zu diesem Zweck implementiert LeiDa-D die sogenannte **Vorrangregel**. Die Vorrangregel ist ein Algorithmus, der für jedes Fachobjekt und jede Objektidentität entscheidet, welche Objektausprägung als fachlich korrekt gilt. Diese wird anschließend den fachlichen Anwendern und den Abnehmersystemen bereitgestellt.

Die Vorrangregel wendet zur Bestimmung der fachlich korrekten Objektausprägung folgende Kriterien an:

1. **Herkunft der Objektausprägung:** Daten, die aus dem LeiDa-S-System der ZFBF übermittelt wurden, haben grundsätzlich Vorrang vor allen anderen.
2. **Aktualität der Objektausprägungen:** Falls keine ZFBF-Daten vorliegen – zum Beispiel, weil Fernbahnregionen aufgrund lokaler Anforderungen einen Datensatz vorzeitig erfassen mussten – wird die zeitlich neuste Objektausprägung aus den Fernbahnregionen ausgewählt.

Dieser Mechanismus wird konsistent für sämtliche Basisdaten angewandt, die LeiDa-D aus den autonomen, regionalen LeiDa-S-Systemen erhält. Dadurch wird sichergestellt, dass LeiDa-D immer eine zentrale und korrekte Sicht auf die netzweiten Basisdaten bietet.

Zusätzlich zu den Basisdaten aus den LeiDa-S-Systemen gibt es Fachobjekte in der Kategorie der Basisdaten, die direkt von LeiDa-D verwaltet und dort projektiert werden können. Für diese Fachobjekte entfällt die Notwendigkeit der Vorrangregel, da LeiDa-D selbst als zentrales System fungiert und die Daten bereits in konsolidierter Form vorliegen. Die Unterscheidung zwischen Objektidentität und Objektausprägung bleibt jedoch weiterhin gültig. Beispiele solcher Fachobjekte, die von der zentralen fachlichen Betriebsführung und der Netzleitzentrale in LeiDa-D direkt gepflegt werden, umfassen unter anderem:

- Ersatzfahrplanhefte
- Geschäftspartner-Daten
- Verspätungsbegründungen
- Daten für die Störfall- und Statistikbearbeitung

3.2 Infrastrukturdaten

Infrastrukturdaten sind, im Gegensatz zu Basisdaten, in LeiDa-S nur regional vorhanden. Eine Ausnahme bildet das LeiDa-S-System der ZFBF, da dort keine Infrastrukturdaten vorliegen. Die sieben Betriebszentralen der Fernbahnen projektieren ihre regionale Infrastruktur innerhalb ihrer jeweiligen Region. Darüber hinaus werden sogenannte Vorschaubereiche angrenzender Regionen projektiert, um regionale Anforderungen an die Disposition zu erfüllen. Diese Vorschaubereiche spiegeln die Autonomie der einzelnen LeiDa-S-Systeme wider und verdeutlichen zugleich die Herausforderung der Zentralisierung durch LeiDa-D. Die Projektierung der Vorschaubereiche ist notwendig, da die Disposition regional organisiert ist. Die Übernahme der Infrastrukturdaten aus den regionalen LeiDa-S-Systemen und deren Aufbereitung zu einem konsistenten, netzweiten Datenbestand bringt jedoch systematisch andere Herausforderungen mit sich als bei den Basisdaten.

Infrastrukturdaten werden in LeiDa-D, wie auch die Basisdaten, in **Objektidentität** und **Objektausprägung** gegliedert, allerdings mit einem anderen Schwerpunkt. Die Objektidentität von Infrastrukturdaten basiert nicht auf einem fachlichen Schlüssel, sondern auf einem technischen Wiedererkennungsmerkmal: dem **technischen Schlüssel** eines Datensatzes aus dem entsprechenden LeiDa-S-System. Zusätzlich enthält die Objektidentität die Information, aus welcher der sieben Regionen der Datensatz stammt. Der technische Schlüssel bleibt über die gesamte Lebensdauer eines Datenobjekts in LeiDa-S stabil. Dies löst die Herausforderung, dass der fachliche Schlüssel bei Infrastrukturelementen geändert werden kann. Da die Fachobjekte der Infrastruktur, anders als die Basisdaten, aus verschiedenen Regionen stammen, erweist sich diese technische Identifizierung als eine elegante Lösung, um die Daten zwischen den Systemen konsistent zu verknüpfen.

Eine exakte Übernahme der Infrastrukturdaten inklusive der Vorschaubereiche in LeiDa-D würde jedoch dazu führen, dass Übergangsbereiche zwischen Regionen doppelt vorkämen. Dies würde den Aufbau eines eindeutigen netzweiten Datenbestands verhindern. Um dies zu vermeiden, muss LeiDa-D die gelieferte Infrastruktur bereits während der Datenübernahme zurechtschneiden. Dabei werden die Vorschaubereiche entfernt, und es werden nur die Datensätze übernommen, die von der jeweiligen Region entsprechend ihrer Pflegehoheit geliefert wurden. Grundlage dieses Verfahrens ist die **Betriebsstellen-Übernahmeverantwortung**, die durch die zentrale fachliche Betriebsführung definiert wird. Sie legt fest, welche Region für welche Betriebsstelle die Infrastruktur liefern darf. LeiDa-D gleicht algorithmisch ab, welche Infrastrukturdaten

innerhalb einer Betriebsstelle liegen und verwirft alle anderen Datensätze. Dadurch wird sichergestellt, dass jedes Infrastrukturelement nur einmal zu einem bestimmten Zeitpunkt im LeiDa-D-Datenbestand vorhanden ist.

Der algorithmische Prozess, um zu bestimmen, welche Infrastrukturelemente zu einer Betriebsstelle gehören, erfolgt in mehreren Schritten. Dieser Prozess beginnt bei den Fachobjekten, die direkt auf Betriebsstellen referenzieren, und überprüft anschließend, ob die identifizierten Datensätze von anderen Infrastrukturelementen referenziert werden. So wird eine Menge von Datensätzen ermittelt, die bei der Übernahme aus einer LeiDa-S-Region entfernt wird. Für viele Fachobjekte können diese Schritte ohne Rückwirkungen durchgeführt werden. Ein Beispiel hierfür sind Betriebsstellenfahrwege, die vollständig in einer Betriebsstelle liegen. Wenn eine Betriebsstelle nicht der Region zugeordnet ist, die die Daten geliefert hat, können solche Datensätze direkt verworfen werden. Eigenschaften dieser Betriebsstellenfahrwege, wie Prioritäten oder zusätzliche Attribute, die in eigenständigen Tabellen geführt werden, werden anschließend ebenfalls entfernt. Für die meisten Infrastrukturfachobjekte funktioniert dieser Ansatz analog zu den Betriebsstellenfahrwegen.

Es gibt jedoch auch Ausnahmen, bei denen dieser Mechanismus nicht uneingeschränkt anwendbar ist. Zwei prominente Beispiele hierfür sind **Betriebsstellengrenzen** und **ZN-Fahrstraßen**.

Betriebsstellengrenzen besitzen die Besonderheit, dass sie im Rahmen der dispositiven Daten Informationen für ihr jeweiliges Gegenstück enthalten. Ein Betriebsstellengrenzeanfang verweist technisch auf das zugehörige Betriebsstellengrenzeende und umgekehrt. Diese Datensätze befinden sich daher in zwei verschiedenen Betriebsstellen. Liegen beide Betriebsstellen in einer Region und haben damit dieselbe Betriebszentrale als Verantwortliche für die Datenübernahme festgelegt, kann LeiDa-D die Daten problemlos zusammenführen. Liegt jedoch die Übernahmeverantwortung einer der Betriebsstellen in einer benachbarten Region, entstehen an der Regionsgrenze Vorschaubereiche, die entfernt werden müssten. Ein reines Entfernen der Betriebsstelle und damit der Betriebsstellengrenze würde jedoch dazu führen, dass LeiDa-D Informationen verliert, die nicht rekonstruiert werden können. Um dies zu verhindern, wurde in LeiDa-D das Konzept der sogenannten **Konnektoren** eingeführt.

Ein Konnektor ist eine fachlich definierte Position im Gesamtnetz, die als Datensatz in LeiDa-D existiert und verwendet wird, wenn es notwendig ist, Übergänge von einer Region in eine andere zu definieren. Durch die Konnektoren kann LeiDa-D die

Referenzen von Betriebsstellengrenzen an Regionsübergängen entfernen, ohne dass dabei entscheidende Informationen verloren gehen. Damit ein konsistenter netzweiter Datenbestand hergestellt werden kann, ist es jedoch erforderlich, dass beide benachbarten Regionen identische Informationen zu jedem Konnektor liefern. Sind die Daten unvollständig oder widersprüchlich, erkennt LeiDa-D diese Abweichungen mithilfe diverser Prüfalgorithmen und informiert die fachlichen Anwender. Diese müssen – gegebenenfalls in Abstimmung mit der benachbarten Region – die Modellierung der Infrastruktur an den Regionsübergängen anpassen und die Daten erneut bereitstellen.

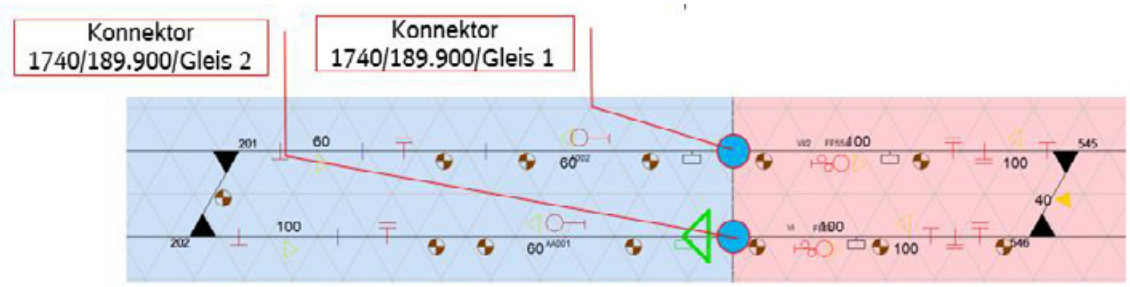


Abbildung 1: Konnektoren Betriebsstellengrenze

Eine ZN-Fahrstraße beschreibt einen sicherungstechnischen befahrbaren Weg zwischen zwei Signalen. Die Ausgangsdaten der ZN-Fahrstraßen gelangen über das System LeiDa-Z (Leitsystem Datenhaltung – Zugstandorte) in die regionalen LeiDa-S-Systeme und werden von dort an LeiDa-D übermittelt. Der Start und das Ende einer ZN-Fahrstraße wird durch Referenzen auf Spurplanknoten beschrieben, die wiederum in Betriebsstellen liegen. Ähnlich wie bei den Betriebsstellengrenzen muss LeiDa-D auch bei ZN-Fahrstraßen Informationen an Regionsübergängen durch Konnektoren festhalten. Im Gegensatz zu Betriebsstellengrenzen, bei denen die Basisinformationen der Konnektoren ausreichen, ist es bei ZN-Fahrstraßen notwendig, dass LeiDa-D zusätzliche Details speichert und abgleicht.

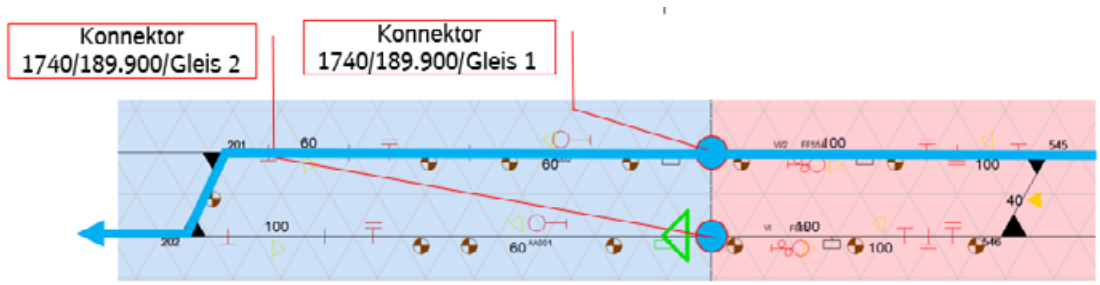


Abbildung 2: Konnektor ZN-Fahrtsraße

Die Abbildung zwei zeigt einen möglichen Verlauf einer ZN-Fahrstraße, die in einer Betriebsstelle beginnt und über eine Regionsgrenze hinweg verläuft. LeiDa-D identifiziert die ZN-Fahrstraße ähnlich wie Betriebsstellengrenzen über den Konnektor, muss jedoch beim Erstellen und Verwalten des Konnektors zusätzliche Informationen speichern und abgleichen.

Eine ZN-Fahrstraße kann nur dann als konsistent betrachtet und in den netzweiten Datenbestand von LeiDa-D übernommen werden, wenn die Daten aus beiden Regionen exakt übereinstimmen. Der algorithmische Abgleich erfolgt dabei anhand folgender Kriterien:

1. **Übereinstimmung des Startsignals:** Das Signal, an dem die ZN-Fahrstraße beginnt, muss in den Daten beider Regionen übereinstimmen.
2. **Übereinstimmung des Zielsignals:** Das Signal, an dem die ZN-Fahrstraße endet, muss ebenfalls eindeutig übereinstimmen.
3. **Identität der Teilverläufe:** Der Verlauf der ZN-Fahrstraße zwischen Start und Ziel muss aus beiden Regionen identisch geliefert werden.
4. **Unterschiedliche Richtung:** Die Richtung der ZN-Fahrstraße muss für Hin- und Rückfahrt unterschiedlich modelliert sein, damit technische und fachliche Korrektheit gewährleistet werden.

Nur wenn diese Kriterien erfüllt sind, wird eine ZN-Fahrstraße, die über eine Regionsgrenze hinweg verläuft, als konsistent betrachtet und in den netzweiten Datenbestand von LeiDa-D übernommen. Abweichungen werden durch LeiDa-D erkannt und den fachlichen Anwendern gemeldet, um frühzeitig Anpassungen vorzunehmen.

4 Versionierte Datenhaltung

Neben der Zentralisierung von Daten ist die **versionierte Datenhaltung** die zweite Hauptaufgabe von LeiDa-D. LeiDa-D speichert nicht nur die aktuell gültigen Basis- und Infrastrukturdaten, sondern stellt zudem historische sowie zukünftige Daten in abrufbereiter Form bereit. Die Architektur der versionierten Datenhaltung basiert im Kern auf den in Kapitel 3 erläuterten Konzepten der Objektidentitäten und Objektausprägungen. Dabei spielen diese Konzepte auch eine zeitliche Rolle: Jede Objektidentität verweist auf eine interne Version in LeiDa-D, die mit einem Zeitpunkt verknüpft ist, an dem diese Version aktiviert wird.

4.1 Der Prozess der Gültigschaltung

Die **Gültigschaltung** beschreibt den zentralen Prozess, mit dem ein netzweiter Datenbestand zu einem definierten Zeitpunkt aktiviert wird. Dabei werden die projektierten Daten für LeiDa-D selbst sowie für sämtliche angeschlossenen Abnehmersysteme freigegeben.

Ein zentraler Bestandteil des Prozesses ist der sogenannte **Aktivierungszeitpunkt**, der den genauen Zeitpunkt festlegt, zu dem der projektierte, netzweite Datenbestand aktiv wird. Ab diesem Zeitpunkt wird sichergestellt, dass alle Systeme, die Daten von LeiDa-D beziehen, ausschließlich mit diesem aktiven Datenbestand arbeiten. Der Aktivierungszeitpunkt dient gleichzeitig als Ziel für die fachlichen Anwender, die auf diesen Zeitpunkt hinarbeiten, indem sie Basis- und Infrastrukturdaten entsprechend anpassen und veredeln.

LeiDa-D ermöglicht es, mehrere zukünftige Aktivierungszeitpunkte parallel zu halten, um den fachlichen Anwendern maximale Flexibilität bei der Datenprojektierung zu bieten. Daten, die für einen Aktivierungszeitpunkt projektiert wurden, werden erst mit Eintreten dieses Zeitpunkts gültig. Bereits gültige Daten bleiben unverändert gültig, sofern sich an diesen Fachobjekten oder zugehörigen Datenquellen keine Änderungen ergeben haben.

4.2 Datenquellen und Versionierung

Im Kontext von LeiDa-D bezeichnet eine **Datenquelle** einen funktionalen Baustein, aus dem ein netzweiter Datenbestand generiert wird. Eine Datenquelle enthält mindestens ein Fachobjekt, kann jedoch auch mehrere Fachobjekte umfassen. Jede der acht LeiDa-S-Systeme stellt eine umfassende Datenquelle dar.

Für jede Datenquelle existieren verschiedene Versionen, die jeweils einem Aktivierungszeitpunkt zugeordnet sind. Wird zu einem bestehenden Aktivierungszeitpunkt eine neue Version einer Datenquelle erstellt, ersetzt diese die vorherige Version. Solche neuen Versionen entstehen beispielsweise, wenn innerhalb eines Fachobjekts einer Datenquelle neue Daten projektiert werden.

Um zu verstehen, wie LeiDa-D einen netzweiten Datenbestand zu einem beliebigen Zeitpunkt erzeugt, hilft die Vorstellung, dass ein solcher Datenbestand aus der Menge aller in LeiDa-D bekannten Datenquellen besteht. Für jede Datenquelle muss es zu dem jeweils gewünschten Aktivierungszeitpunkt eine gültige Version geben.

Dieser Prozess der Bereitstellung und Aktivierung eines netzweiten Datenbestands wird durch die **Gültigschaltung** ermöglicht, die sich in mehrere Phasen unterteilen lässt.

4.3 Das Aktivierungszeitpunkt-Phasenmodell

Der Ablauf der Gültigschaltung ist in folgende Phasen unterteilt:

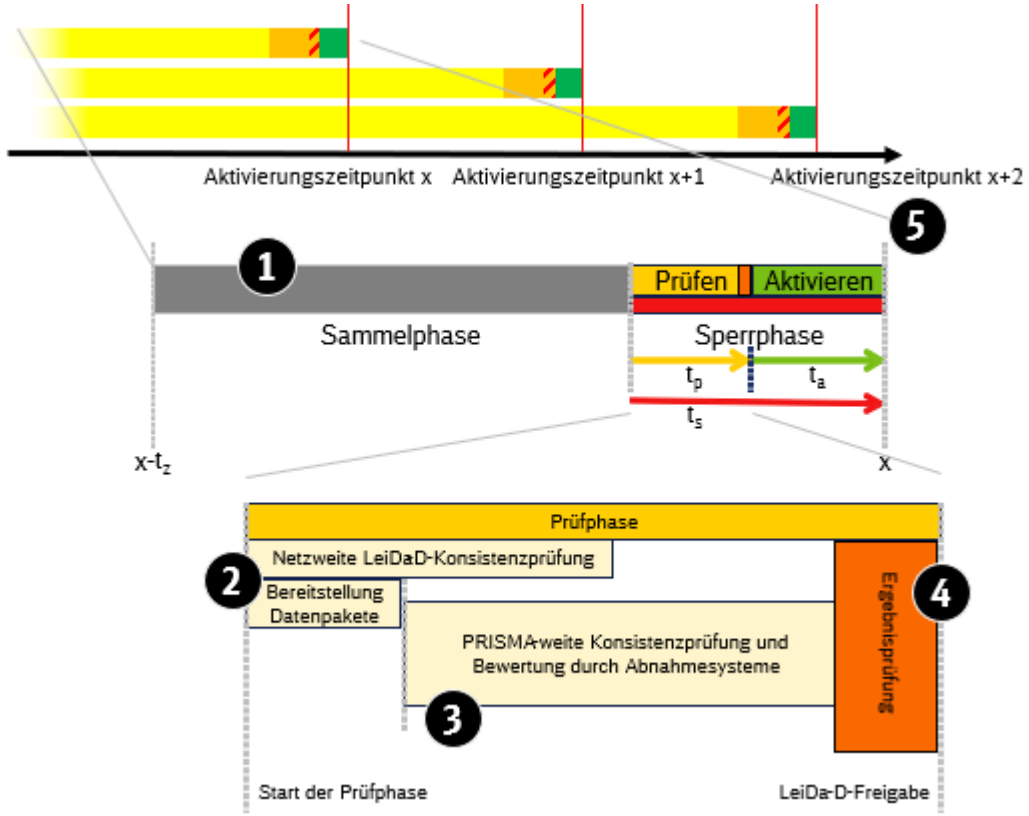


Abbildung 3: Detaillierte Darstellung des Aktivierungszeitpunkt-Phasenmodells

Der Prozess der Gültigschaltung folgt einem mehrstufigen Modell, das den gesamten Ablauf der Bereitstellung eines netzweiten Datenbestands systematisch gliedert. Mit der Definition eines Aktivierungszeitpunkts durch die zuständigen fachlichen Anwender beginnt die **Sammelphase** (Punkt 1, Abbildung 3). Während dieser Phase können fachliche Anwender die benötigten Daten einliefern oder in LeiDa-D projektieren. LeiDa-D prüft jede einzelne Datenlieferung kontinuierlich und stellt Prüfergebnisse frühzeitig zur Verfügung, damit potenzielle Probleme direkt adressiert werden können. In dieser Phase wird nicht vorgegeben, wie oft und in welchem Umfang Daten geliefert werden

dürfen – die Frequenz und Menge der Daten ist für LeiDa-D irrelevant. Fachliche Anwender können jederzeit den netzweiten Datenbestand auf Plausibilität prüfen lassen, um sicherzustellen, dass er für den Aktivierungszeitpunkt vollständig und konsistent vorbereitet ist. Zusätzlich können Abnehmersysteme während der Sammelphase vorläufige Versionen des projektierten Datenbestands abrufen, um eigene Prüfungen durchzuführen. Diese Zwischenstände für die Vorabprüfung werden von LeiDa-D eigenständig einmal täglich generiert und an die Abnehmersysteme weitergegeben. Die Sammelphase endet, wenn keine weiteren Änderungen mehr vorgenommen werden und der Termin für den Übergang in die nächste Phase erreicht ist.

Mit Abschluss der Sammelphase beginnt die **Prüfphase** (Punkt 2, Abbildung 3). In diesem Abschnitt werden alle Daten für den Aktivierungszeitpunkt eingefroren, wodurch Änderungen an den projektieren Daten nicht mehr möglich sind. LeiDa-D definiert den finalen, netzweiten Datenbestand und informiert alle Abnehmersysteme über die Bereitstellung dieser Daten. Gleichzeitig führt LeiDa-D eine Reihe interner Prüfroutinen durch, um potenzielle Datenfehler oder Inkonsistenzen zu identifizieren. Parallel dazu fordern die Abnehmersysteme die aufbereiteten Daten an, um ebenfalls eigene Prüfungen durchzuführen. Die Prüfergebnisse der Abnehmersysteme werden detailliert dokumentiert und zurück an LeiDa-D übermittelt (Punkt 3, Abbildung 3). LeiDa-D sammelt diese Rückmeldungen und stellt sie dem zentralen Koordinator der Gültigschaltung zur Verfügung.

Der zentrale Koordinator analysiert die gesammelten Ergebnisse und trifft auf dieser Grundlage die Entscheidung, ob der Aktivierungszeitpunkt freigegeben werden kann (Punkt 4, Abbildung 3). Diese Freigabe markiert das Ende der Prüfphase und leitet die **Aktivierungsphase** ein. Falls die Freigabe nicht erteilt wird, bleiben die neuen Daten in LeiDa-D gespeichert und werden auf einen neuen Aktivierungszeitpunkt verschoben. Datenverluste treten in diesem Fall nicht auf.

Die **Aktivierungsphase** (Punkt 5, Abbildung 3) beschreibt den Zeitraum, in dem die Abnehmersysteme die finalen Schritte zur Implementierung des neuen Datenbestands durchführen. Der bereinigte und validierte netzweite Datenbestand wird exakt zum festgelegten Zeitpunkt gültig und steht allen angeschlossenen Systemen synchronisiert zur Verfügung. Die Datenlieferungen werden dann produktiv genutzt und bilden die Grundlage für alle von LeiDa-D abhängigen Anwendungen und Prozesse.

Ein Beispiel für einen Aktivierungszeitpunkt ist der **jährliche europaweite Fahrplanwechsel**, der typischerweise im Dezember stattfindet. Bei diesem Ereignis stellt LeiDa-D sicher, dass ab einem definierten Zeitpunkt die aktualisierten Fahrpläne und

Infrastrukturdaten gültig sind. Dies umfasst Änderungen an Strecken, Fahrwegen, und Betriebsstellen, die von Abnehmersystemen wie dem neuen Dispositionssystem genutzt werden. Neben dem Fahrplanwechsel sind auch **unterjährige Ereignisse**, wie etwa die Inbetriebnahme eines neuen elektronischen Stellwerks, typische Aktivierungszeitpunkte. Ebenfalls können größere Streckenerweiterungen oder umfangreiche Baumaßnahmen durch die Mechanismen der Gültigschaltung und Aktivierungszeitpunkte effektiv abgebildet werden. Dadurch ist LeiDa-D in der Lage, sowohl langfristige Planungen als auch kurzfristige Anpassungen verlässlich und konsistent für einen netzweiten Datenbestand umzusetzen.

Um als fachliche Anwender längere Projektierungen bereits in den Abnehmersystemen frühzeitig zu erkennen, gibt es die sogenannte Vorabprüfung, welche täglich erfolgt. Hierbei stellt LeiDa-D einen vorläufigen, netzweiten Datenbestand bereit. Dieser kann sich jedoch bis zur Gültigschaltung ändern und spiegelt den aktuellen Stand der Projektierung innerhalb von LeiDa-D wider.

5 LeiDa-D als System

LeiDa-D besteht aus drei Hauptanwendungen: zwei Webapplikationen sowie einem zentralen Server. Die Webapplikationen stehen den fachlichen Anwendern zur Verfügung und bieten eine benutzerfreundliche Oberfläche für verschiedene Aufgaben. Sie ermöglichen das Erfassen und Bearbeiten von Daten, die in LeiDa-D verwaltet werden, die Sichtung von Prüfergebnissen, eine grafische Visualisierung der Infrastrukturdaten sowie die Steuerung und fachliche Bewertung der Gültigschaltungen. Ergänzt werden diese Anwendungen durch den LeiDa-D Server, der speziell für die Durchführung rechenintensiver Prozesse und die Verarbeitung großer Datenmengen entwickelt wurde. Der Server empfängt die Rohdaten aus den LeiDa-S-Systemen, führt die Prüfalgorithmen aus, koordiniert die Gültigschaltungen und stellt den netzweiten Datenbestand auf Abruf für sämtliche Abnehmersysteme bereit. Diese Infrastruktur gewährleistet, dass LeiDa-D effizient mit der Vielzahl von Datenquellen sowie den komplexen Prozessen im Zusammenhang mit der Datenverarbeitung umgehen kann.

Neben den drei Anwendungen umfasst LeiDa-D eine zentrale Datenbank, die sämtliche Informationen speichert, die für die Bereitstellung eines netzweiten Datenbestands erforderlich sind. Dazu gehören nicht nur die aktuellen Daten, sondern auch historische Informationen, die auf frühere Aktivierungszeitpunkte zurückzuführen sind. Diese Architektur ermöglicht es LeiDa-D, Datenprozesse sowohl rückverfolgen als auch zukünftige Aktivierungszeitpunkte flexibel planen zu können.

LeiDa-D wurde mit modernsten Technologien entwickelt und zukunftssicher konzipiert. Ein zentraler Fokus liegt auf der Hochverfügbarkeit des Systems. Ausfallsicherheit wird unter anderem dadurch gewährleistet, dass jede der drei Anwendungen mehrfach bereitgestellt wird. Sollte eine Instanz ausfallen, leitet eine interne Lastverteilung sowohl die Anfragen der fachlichen Anwender als auch die Anforderungen der Abnehmersysteme zuverlässig an verfügbare Instanzen weiter. Auf diese Weise bleiben die Anwendungen jederzeit erreichbar, und LeiDa-D kann Anfragen schnell und effizient bearbeiten.

Ein weiterer Vorteil liegt in der Skalierbarkeit des Systems. Wenn zusätzliche Kapazitäten für neue fachliche Anwender oder Abnehmersysteme benötigt werden, kann LeiDa-D problemlos erweitert werden. Die Systemarchitektur ermöglicht es, weitere Anwendungskomponenten in den bestehenden Gesamtverbund zu integrieren, ohne dass dies die Stabilität oder Leistung des Systems beeinträchtigt. Diese Flexibilität stellt sicher, dass LeiDa-D den Anforderungen eines wachsenden Netzwerks von Nutzern und Abnehmersystemen auch langfristig gerecht wird.

Darüber hinaus erfüllt LeiDa-D sämtliche gängigen sicherheitsrelevanten Anforderungen, um die Vertraulichkeit und Integrität der Daten zu gewährleisten. Dazu gehören beispielsweise eine durchgängig verschlüsselte Kommunikation zwischen den Systemen und ein umfassendes Rechte- und Rollenkonzept für die fachlichen Anwender. Letzteres stellt sicher, dass Nutzer nur auf die für ihre Aufgaben relevanten Daten und Funktionen zugreifen können. Somit vereint LeiDa-D sowohl effiziente Performance als auch ein hohes Maß an Sicherheit sowie Benutzerfreundlichkeit.

6 Zukunftsorientierte Weiterentwicklung und Integration neuer Technologien

Die zukünftige Weiterentwicklung von LeiDa-D fokussiert sich auf zwei zentrale Ziele: die Ablösung der derzeitigen Bestandssysteme zur Erfassung von Basis- und Infrastrukturdaten (LeiDa-S) sowie die Ertüchtigung von LeiDa-D zur Unterstützung von ETCS Level 2.

Zukünftig ist geplant, die Bestandssysteme LeiDa-S vollständig durch die Funktionen von LeiDa-D zu ersetzen. Dies erfordert eine umfassende Migration der aktuell in LeiDa-S gepflegten Daten und Prozesse, einschließlich der automatisierten Verarbeitung der Datenpakete, die LeiDa-D derzeit aus LeiDa-S erhält. Diese Daten müssen vollständig in den Mechanismus der LeiDa-D-Versionierung eingebunden werden, insbesondere in

Bezug auf die Verwaltung und Weiterentwicklung der Datenquellen. Dazu gehört nicht nur die tabellarische Bearbeitung von Fachobjekten, sondern auch die Integration und Weiterentwicklung eines **grafischen Infrastruktureditors**, um die Pflege der Infrastrukturdaten effizient und anwenderfreundlich zu gestalten.

Darüber hinaus müssen die bisher von LeiDa-S bereitgestellten Schnittstellen, insbesondere zu den Systemen DaViT und LeiDa-Z (zuständig für zustandsorientierte Daten), nahtlos in LeiDa-D integriert werden. Dies gewährleistet, dass bestehende Datenströme sowie zentrale als auch regionale Prozesse auch nach der Ablösung von LeiDa-S weiterhin vollständig abgebildet werden können. Dieser Umstieg stellt einen weiteren Schritt in der fortlaufenden Modernisierung und Zentralisierung dar und ist eine logische Konsequenz aus den übergeordneten Zielen von LeiDa-D.

Eine zweite große Weiterentwicklung ergibt sich aus den Anforderungen des Projekts **Digitaler Knoten Stuttgart (DKS)**. Unter anderem hierfür soll die Disposition um das europäische Zugsicherungssystem **ETCS (European Train Control System) Level 2** ertüchtigt werden. Um dieser Anforderung gerecht zu werden, muss LeiDa-D so erweitert werden, dass es ETCS Level 2 vollständig unterstützt und die entsprechenden Daten an das neue Kerndispositionssystem liefern kann. Dies stellt eine erhebliche fachliche und technische Neuerung dar, da eine derartige Funktionalität bislang weder in LeiDa-D noch in den LeiDa-S-Systemen umgesetzt wurde.

Die Umsetzung dieser Erweiterung in LeiDa-D ist ein zentraler Baustein der Abbildung von ETCS Level 2 im Dispositionssystem und somit ein wesentlicher Beitrag zur Digitalisierung und Zukunftsfähigkeit des Bahnbetriebs in Deutschland.

Die geplanten Weiterentwicklungen spiegeln die langfristige Vision von LeiDa-D wider: die vollständige Modernisierung und Zentralisierung der Datenlandschaft sowie die Anpassung an innovative Technologien wie ETCS Level 2. Mit der Ablösung der LeiDa-S-Systeme und der Unterstützung von ETCS übernimmt LeiDa-D eine zentrale Rolle in der Datenversorgung der Dispositionssysteme. Beide Entwicklungsschritte stellen sicher, dass LeiDa-D optimal für die Anforderungen einer zukunftsorientierten und effizienten Nutzung im Schienenverkehr vorbereitet ist.

Danksagung

Ich möchte an dieser Stelle meinen besonderen Dank der „DB InfraGO AG“ aussprechen. Sie hat als treibende Kraft die Entwicklung des Produkts „LeiDa-D“ beauftragt, vorangetrieben und durch ihr Engagement den Fortschritt maßgeblich ermöglicht. Die „quattron GmbH“ hat in diesem Rahmen die Entwicklung übernommen und durch fachkundige Beratung und professionelle Umsetzung einen wertvollen Beitrag geleistet. Ebenso gilt mein Dank der „DB Systel GmbH“, deren langjährige Unterstützung und Expertise einen wichtigen Bestandteil dieser erfolgreichen Zusammenarbeit darstellte. Diese partnerschaftliche und vertrauensvolle Zusammenarbeit, initiiert und gefördert durch die „DB InfraGO AG“, war die Grundlage für den gemeinsamen Erfolg, der nun in einem gemeinsamen Vortrag präsentiert wird.

Autoren



Trappein, Mario

Mario Trappein begann nach seinem Abschluss an der FH Aachen bei der quattron GmbH (ehemals VIA Consulting & Development GmbH) seine Karriere als IT-Ingenieur. Seine Expertise umfasst die Softwarearchtitekur, Entwicklung und Qualitätssicherung.

quattron GmbH

Römerstr. 50, 52064 Aachen

mario.trappein@quattron.com



Götz, Dennis

Dennis Götz arbeitet bei der DB InfraGO AG für das Anforderungsmanagement und die Informationssicherheit der dispositiven Leittechnik in den Betriebszentralen.

Aktuell betreut er in der Rolle als Product Manager die Anwendungen für Infrastrukturdaten in einem Agile Release Train.

DB InfraGO AG

dennis.d.goetz@deutschebahn.com

Session 15
Ernst Zollinger, Gabrio Caimi, Vanessa Hänni

Traffic Management System (TMS) in der Schweiz – Neue Herausforderungen bez. Sicherheit und Zuverlässigkeit

Zollinger, Ernst[1], Caimi, Gabrio[2], Hänni, Vanessa[1]

[1]Enotrac AG
[2]Schweizerische Bundesbahnen AG

Zusammenfassung

Das Traffic Management System (TMS) der Schweiz unterstützt die Trassenproduktion von der Planung bis zur Durchführung. Es wird von mehreren Infrastrukturbetreibern (ISB) unter dem Lead der SBB gemeinsam entwickelt und genutzt. Die Weiterentwicklung fokussiert auf Automatisierungen und Behebung von Systembrüchen mit den Zielen Effizienzsteigerung, Robustheit, Kapazität und Sicherheit. Anhand der neuen Funktionalität «Warnfunktion Rangieren» wird aufgezeigt, wie damit auch Sicherheit und Zuverlässigkeit des Systems selbst mehr in den Fokus rücken und wie mit den diesbezüglichen Herausforderungen inkl. Zulassung umgegangen wird.

Keywords: Traffic Management System, Fahrassistenz, Rangieren, RAMSS, Zulassung.

1 Einleitung

Das Traffic Management System (TMS) ist eine Familie von IT-Systemen zur optimalen Bereitstellung von Trassenkapazitäten durch die Infrastrukturbetreiber (ISB), welche die Trassenproduktion von der Erstellung des Fahrplans bis zur Betriebsführung unterstützt.

Das TMS umfasst folgende Funktionen:

Planung (dispositive Ebene)

- Die Bestellung von Kapazitäten (Trassen, Abstellungen, …) durch die Eisenbahnverkehrsunternehmen (EVU).
- Die Ausplanung der Kapazitäten (Einlegen von Trassen) durch ISB.
- Die Auswertung sowie Abrechnung genutzter Trassen.

Durchführung (operative Ebene)

- Die Bedieneroberfläche der Betriebszentralen für Überwachung und Eingriffe in den Verkehr.
- Die Ansteuerung der Stellwerke via Leittechnik für die automatisierte Durchführung des Fahrplans.
- Die teil-automatisierte lokale real-time Optimierung des Verkehrs bei Konflikten.
- Real-time Informationen (Fahrempfehlungen) für Fahrassistenzsysteme.
- Die Informationsdrehscheibe für alle Beteiligten (ISB und EVU) im Abweichungsmanagement (Störungsmanagement).
- Grundlage für die Kundeninformation (Information zu Verspätungen und anderen Abweichungen).

TMS wird als Familie von IT-Anwendungen unter der Leitung der SBB mit den schweizerischen ISB BLS Netz AG, SOB und TPF partnerschaftlich entwickelt und betrieben. Die Zusammenarbeit im gleichen System erlaubt die Optimierung von Zugläufen und Anschlüssen über mehrere Gebiete, die Vermeidung zahlreicher Schnittstellen, die Nutzung gemeinsamer Prozesse der ISB und die Reduktion der IT-Kosten in Entwicklung und Betrieb. TMS basiert auf vorhergehenden Systemgenerationen und wird schrittweise weiterentwickelt. Die Entwicklung erfolgt in-house bei der SBB IT.

Die wichtigsten Wirkungsachsen für die Weiterentwicklung von TMS sind Effizienz (Kosteneinsparungen), Robustheit des Betriebs, Flexibilität, Kapazität und Sicherheit (im Sinne von „Safety“).

Mit ersten TMS-Anwendungen in den Betriebszentralen konnte die Durchgängigkeit des Datenaustausches und die Effizienz im Störungsfall verbessert werden. Mit real-time

Fahrassistenzinformationen leistet TMS einen Beitrag zur Robustheit des Fahrplans und zu namhaften Energieeinsparungen. Zentrale Stoßrichtungen der aktuellen Weiterentwicklung von TMS umfassen:

- Ablösung der aktuellen Planungssysteme: Anpassung an europäische Normen (TSI TAF/TAP; Technische Spezifikation für die Interoperabilität TSI Teilsystem Telematikanwendungen für den Güterverkehr TAF und den Personenverkehr TAP) und (Teil-) Automatisierung der Fahrplanerstellung.
- Rangieren: Digitalisierung von Bestellung, Planung, Fahren und Abrechnung.
- Weiterführung der Automatisierung in der Betriebsführung (automatisierte Dispositionsentscheide im Konfliktfall auch in komplexen Gebieten wie z.B. Flankenschutz).

Anhand der Entwicklung der „Warnfunktion Rangieren“ WaRa werden nachstehend beispielhaft einige der angetroffenen Herausforderungen und Erkenntnisse aufgezeigt.

2 Herausforderungen und gewonnene Erkenntnisse

2.1 Erkennen und Einschätzen der RAMSS-Relevanz

Für die Funktionen der Vorgängersysteme von TMS wurde keine besondere RAMSS-Relevanz (Reliability, Availability, Maintainability, Safety and Security) ermittelt. Sie verfügten auch nicht über einen steuernden Durchgriff auf die Leittechnik- und Stellwerkebene. Die Systeme wurden daher nicht normenbasiert und ohne Zulassung entwickelt und betrieben. Mit der zunehmenden Automatisierung von TMS erhöht sich die Verantwortung des Systems. Gewisse Funktionen von TMS sind damit als RAMSS-relevant einzustufen. Da die für TMS zuständige Organisation sich bis zu diesem Zeitpunkt nicht mit Themen der Zulassung befasst hatte, war die initiale Bereitschaft, sich mit normenbasierter IT-Entwicklung zu befassen, gering. Klare Botschaften seitens des Auftraggebers „SBB Infrastruktur“ und der Einbezug von respektierten Experten haben geholfen, diese Ablehnung zu überwinden. Ergänzend wurden den Teams konkrete Methoden und Werkzeuge für normenbasierte Entwicklung bereitgestellt.

2.2 Vereinbarkeit von normenbasierter und agiler Entwicklung

In der SBB werden in der IT-Entwicklung mehrheitlich agile Verfahren angewendet, welche auf kurzen Entwicklungszyklen (10 Wochen) und raschem Feedback beruhen. Die

starke Phasenorientierung und Dokumentationslastigkeit der Normen stehen auf den ersten Blick in Widerspruch zu agilen Prinzipien. Der initiale Release von WaRa war (auch aufgrund der notwendigen Lernkurve) mit einem substanziellen Aufwand verbunden. Durch Fokus auf eine Delta-Betrachtung für Folgereleases und Reduktion der Redundanz in der Dokumentation konnte dieser Aufwand reduziert werden. Weitere Verbesserungen sind geplant.

3 Anwendungsbeispiel „Warnfunktion Rangieren“ (WaRa)

3.1 Übersicht

Rangierfahrten sind technisch nicht abgesichert: die Beachtung der Rangiersignale, die Geschwindigkeit des Fahrens und das Bremsen können heutzutage technisch nicht beeinflusst werden. Diese Verantwortung liegt daher allein beim Rangiermitarbeiter. Die Einführung der Anwendung „Warnfunktion Rangieren“ (WaRa) ändert daran nichts.

WaRa ist ein Assistenzsystem zur Unterstützung der Rangiermitarbeiter mit dem Ziel, die Safety im Rangierumfeld zu erhöhen. WaRa überwacht anhand der Belegung der Gleisfreimelde-(GFM)-Abschnitte die Rangierfahrt und warnt optisch und akustisch den Rangiermitarbeiter bei Annäherung bzw. Vorbeifahrt an einem „Halt“ zeigenden Rangiersignal. Dazu bezieht WaRa Informationen aus der Leittechnik. Es besteht infrastrukturseitig aus einer zuverlässigen, hochverfügbaren Backend-Applikation (Betrieb in der Cloud) und fahrzeugseitig aus einer Frontend-Applikation auf kommerziell erhältlichen Mobiltelefonen bzw. Tablets mit Android- oder iOS-Betriebssystem als Endgerät.

3.2 Wie funktioniert WaRa

WaRa kennt die vorhandenen GFM-Abschnitte des jeweiligen Bahnhofs, die Registrierungen der Rangierfahrzeuge auf den GFM-Abschnitten sowie die aktuelle Belegung der GFM-Abschnitte durch Fahrzeuge. Ist eine Rangierfahrstraße eingestellt und werden die GFM-Abschnitte in korrekter Reihenfolge belegt und wieder freigegeben, so erkennt WaRa dies als „korrekte Rangierfahrt“.

WaRa hat keine Kenntnisse, ob und was für Fahrzeuge sich bewegen. WaRa kennt auch keine Fahrdynamik oder erkennt nicht, ob Fahrzeuge stillstehen. WaRa kennt nur die aktuell belegten GFM-Abschnitte und die aktuellen eingelaufenen Fahrstraßen aus der Leittechnik.

WaRa erfordert initial eine Anmeldung durch einen dazu berechtigten Benutzer. Anschließend registriert der Rangiermitarbeiter sein Fahrzeug auf dem Endgerät mit einer eindeutigen Kennung auf dem Startgleis im jeweiligen Bahnhof.



Quelle: Schweizerische Bundesbahn (SBB) – © 2025

Abbildung 1: Darstellung des registrierten Fahrzeugs auf Gleis 5 mit der Kennung RD1 auf dem mobilen Endgerät

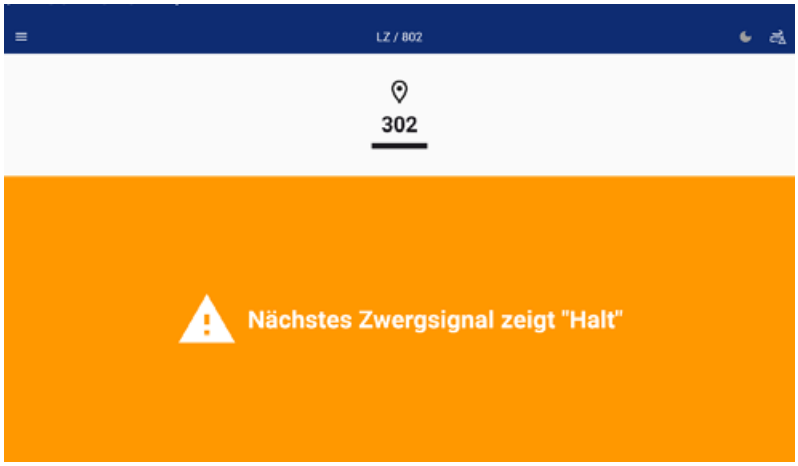
In einem zweiten Schritt wird der gewünschte Fahrweg beantragt. Nach der Freigabe des Fahrwegs durch das Stellwerk erscheint auf dem Endgerät die Abfolge der GFM-Abschnitte vom Start- bis zum Zielgleis als Perlschnur.



Quelle: Schweizerische Bundesbahn (SBB) – © 2025

Abbildung 2: Darstellung der Rangierfahrstraße auf dem mobilen Endgerät in Form einer „Perlschnur“

Belegt die Rangierkomposition den GFM-Abschnitt am Ende ihrer Fahrstraße, gibt WaRa auf dem Endgerät akustisch und optisch eine Warnung aus.



Quelle: Schweizerische Bundesbahn (SBB) – © 2025

Abbildung 3: Optische Warnung auf dem mobilen Endgerät am Ende einer Fahrstraße

Wird das „Halt zeigende Rangiersignal“ überfahren, wird optisch und akustisch ein Alarm ausgelöst.



Quelle: Schweizerische Bundesbahn (SBB) – © 2025

Abbildung 4: Optischer Alarm bei Vorbeifahrt an einem „Halt“ zeigenden Rangiersignal

3.3 Bewertung von Assistenzsystemen

Die Funktion WaRa ist eine Assistenzfunktion, da ihr Einsatz keine Voraussetzung für den Rangierbetrieb darstellt. Für das Rangierpersonal sind unverändert wie bisher die Schweizerische Fahrdienstvorschriften (FDV) und die Ausführungsbestimmungen der jeweiligen ISB / EVU maßgebend. Für den Rangiermitarbeiter gilt weiterhin ausschließlich das Signalbild der Rangiersignale. WaRa stellt auf dem mobilen Endgerät zusätzliche Informationen zur Verfügung. Bei korrekter Funktionsweise informiert WaRa den Rangiermitarbeiter auf seinem Gerät über

- den Status des angeforderten Fahrweges,
- die Codepunkte der eingestellten Rangierfahrstraße sowie die aktuelle Belegung der GFM-Abschnitte der Spitze der Rangierfahrt entlang der Rangierfahrstraße, und
- die potenziell kritischen Situationen in Form von einer Meldung „Rangierwarnung/Rangieralarm“ (akustisch und visuell).

Damit wird WaRa einen positiven Beitrag zur Vermeidung von Rangierunfällen leisten.

3.4 Benutzerverhalten und HOF-Analysen

Dem Nutzen von WaRa stehen potenzielle Gefahren entgegen, wie bspw.

- Ablenkungen des Rangiermitarbeiters durch WaRa,
- Ablenkung des Rangiermitarbeiters durch Fremdanwendungen auf dem mobilen Endgerät,
- ein ungerechtfertigtes Abstützen auf die Warnungen durch WaRa und
- Vernachlässigen der Sicht ins Gleis.

Um diese Mechanismen zu verstehen und zu adressieren, wurden umfangreiche Untersuchungen zum Benutzerverhalten und HOF (human and organizational factors) inklusive Tests von frühen Versionen im Feld durchgeführt.

3.5 Ermittlung der Gefährdungen

Sechs (6) Gefährdungen, hervorgerufen durch WaRa, wurden im Rahmen des Gefährdungs-Workshops von Experten:innen identifiziert. Es sind dies im Einzelnen:

1. **Vernachlässigung der Rangiersignale:** Der Lokführer verlässt sich unerlaubterweise auf WaRa, da er erwartet, dass WaRa rechtzeitig warnt bzw. alarmiert sowie die Fahrweginformationen richtig anzeigt.
2. **Ablenkung des Lokführers:** Der Einsatz von WaRa kann im Führerstand zu besonderen Belastungssituationen des Lokführers führen, wodurch dieser durch die optischen und akustischen Warn- oder Alarmsignale abgelenkt wird.
3. **Erweckung falscher Erwartungshaltungen beim Lokführer (Verwirrung):** Alarme und/oder die Darstellung der Fahrstraßen-Information auf der Anzeige des mobilen Endgeräts ruft eine Erwartungshaltung bzgl. der Rangiersituation beim Lokführer hervor, die nicht der Realität entspricht.
4. **Hörschaden durch zu lautes Alarm- oder Warnsignal:** Wenn ein Nutzer das mobile Endgerät für ein Telefonat am Ohr hat, könnte ein ausgegebener Alarm bei maximaler Lautstärke zu einem einseitigen Gehörschaden führen.
5. **Störende Blendung des Lokführers:** Wenn die Anzeige des mobilen Endgeräts nachts oder in einem Tunnel zu hell leuchtet, könnte der Lokführer geblendet werden, wodurch er das Signalbild eines Rangiersignals nicht korrekt erkennt bzw. übersieht.
6. **Überlagerung von anderen Kontroll- oder Warnsignalen im Führerstand:** WaRa warnt bzw. alarmiert den Lokführer so laut, dass andere Warn- oder Kontrollsignale im Führerstand überlagert werden und deshalb überhört werden könnten.

3.6 Festlegung des Risikoakzeptanzprinzips

Alle Gefährdungen von WaRa werden nach dem Risikoakzeptanzprinzip „explizite Risikoabschätzung“ gemäß der „gemeinsamen Sicherheitsmethode für die Evaluierung und Bewertung von Risiken“ ([1], [2]) beurteilt.

3.7 Einstufung der ermittelten Gefährdungen

Die Einstufung der sechs (6) ermittelten Gefährdungen erfolgte bezüglich deren Häufigkeit und Schadensausmaß.

Die Kollision einer Rangierfahrt mit einer Zugfahrt darf im Hinblick auf das zu erwartende Schadensausmaß als kritischster Fall angenommen werden. Für eine solche Kollision wird die Schadensausmaßklasse „*mehrere Schwerverletzte, 1 Todesopfer*“ angenommen. In den sieben (7) durch die SUST (Schweizerische Sicherheitsuntersuchungsstelle) zwischen dem 01.01.2018 und 31.12.2020 untersuchten Ereignissen wurden insgesamt eine Person schwer und 6 Personen leicht verletzt. Darüber hinaus kam es nur zu Sachschaden. Es kann davon ausgegangen werden, dass in den restlichen Fällen, die in der Ereignisdatenbank der SBB aufgenommen, aber nicht durch die SUST untersucht wurden, keine Personen verletzt wurden. Für das Überfahren eines „Halt“ zeigenden Rangiersignals soll daher im Folgenden die Schadensausmaßklasse „*mehrere leicht verletzte Personen, eine mittelschwer verletzte Person*“ angenommen werden.

Die folgenden Gefährdungen wurden analysiert:

- Vernachlässigung der Rangiersignale
- Ablenkung des Lokführers
- Erweckung falscher Erwartungshaltungen beim Lokführer (Verwirrung)
- Hörschaden durch zu lautes Alarm- oder Warnsignal
- Störende Blendung des Lokführers
- Überlagerung von anderen Kontroll- oder Warnsignalen im Führerstand

Für jede Gefährdung wurden folgende Merkmale ermittelt:

- Häufigkeit
- Schadensausmaßklasse
- Herleitung der Häufigkeit oder Beschreibung der Sicherheitsbarrieren
- Risiko

Gemäß der Risikobeurteilungsmethode der SBB liegen die Risiken (Häufigkeiten und Schadensausmaße) im akzeptablen Bereich.

3.8 SIL-Einstufung der Funktion WaRa

Die mit den sechs (6) Gefährdungen verbundenen Risiken sind alle akzeptabel. Ihre Beherrschung zur Gewährleistung ausreichender Safety auf Stufe „Basisintegrität“ erfordert daher keine weiteren normativen ([3], [4], [5]) Maßnahmen.

4 Die Rolle der Security

Während der Projektlaufzeit wurde in der Schweiz eine neue Vorgabe an die Cybersicherheit im Schienenverkehr eingeführt. Neue Technologien (in WaRa zum Beispiel Cloud, öffentliches Netzwerk, mobile Geräte) und die Bedrohungslage führen zu einer höheren Risikoexposition durch Schwachstellen, deren Ausnutzung wiederum Safety oder Verfügbarkeit gefährden kann.

„*If it's not secure, it's probably not safe, not reliable, not available and not maintainable!*"

Die Gewichtung der Security wurde darum im Verlauf des Vorhabens erhöht und enger mit der Gefährdungsanalyse aus RAM- und Safety-Sicht verknüpft.

5 Zusammenfassung

Das in der Schweiz gemeinsam genutzte TMS liefert messbare Wirkungen in Effizienz, Robustheit, Flexibilität, Kapazität und Sicherheit für die Trassenproduktion und das Gesamtsystem Eisenbahn. Die architektonische Verortung dieser Funktionen in Informationssystemen (statt in der Sicherungstechnik strecken- oder fahrzeugseitig) bietet günstigere Entwicklungs- und Betriebskosten und schnellere Entwicklungszyklen. Gleichzeitig führen die umgesetzten Automatisierungen dazu, dass Sicherheitsverantwortungen verschoben werden und für gewisse Funktionen auch im TMS eine Zulassung notwendig wird.

Das Beispiel der Warnfunktion Rangieren (WaRa) zeigt, welche Herausforderungen damit verbunden sind und wie diese gelöst wurden. WaRa hat die Zulassung erhalten und ist erfolgreich in Betrieb. Dank diesen Erfahrungen kann der Zulassungsprozess agilisiert und TMS effizient weiterentwickelt werden.

Literatur

[1] „Durchführungsverordnung (EU) Nr. 402/2013 der Kommission vom 30. April 2013 über die gemeinsame Sicherheitsmethode für die Evaluierung und Bewertung von Risiken und zur Aufhebung der Verordnung (EG) Nr. 352/2009“, Amtsblatt der Europäischen Union, L 121/8 – 121/25., 3.5.2013.

[2] „Durchführungsverordnung (EU) 2015/1136 der Kommission vom 13. Juli 2915 zur Änderung der Durchführungsverordnung (EU) Nr. 402/2013 über die gemeinsamen Sicherheitsmethoden für die Evaluierung und Bewertung von Risiken“, Amtsblatt der Europäischen Union, L 185/ 6 – L 185/10, 14.7.2015.

[3] EN 50126-1: 2017, „Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) – Teil 1: Generischer RAMS-Prozess“, Europäisches Komitee für Elektrotechnische Normung (CENELEC), Oktober 2017.

[4] EN 50126-2: 2017, „Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) – Teil 2: Systembezogene Sicherheitsmethodik“, Europäisches Komitee für Elektrotechnische Normung (CENELEC), Oktober 2017.

[5] EN 50716: 2023, „Railway Applications – Requirements for software development”, European Committee for Electrotechnical Standardization (CENELEC), November 2023.

Session 15
Ernst Zollinger, Gabrio Caimi, Vanessa Hänni

Autoren



Zollinger, Ernst

Prof. Dr. sc. techn. Dipl. El.-Ing ETH, executive MBA HSG, ist spezialisiert in RAMSS und Senior Projektingenieur bei der Enotrac AG. Er hat im Programm TMS der SBB die Rolle des Validierers gemäß EN 50126 inne.



Caimi, Gabrio

Dr. sc. ETH, Dipl. Math. ETH, ist Solution Architekt der DSO (Digital Solution) Trassen, der Organisation bei SBB, welche TMS betreibt und weiterentwickelt. Er verantwortet in dieser Rolle die Gesamtarchitektur von TMS und das Zusammenspiel der verschiedenen Pakete.



Hänni, Vanessa

MAS in Eisenbahntechnik und nachhaltiger Mobilität, ist spezialisiert in Safety und Projektleiterin bei der Enotrac AG. Sie hat im Programm TMS der SBB die Leitung der Validierungsstelle und ist stellvertretende Validiererin gemäß EN 50126.

Session 15
Tobias Körner, Danny Rohrberg, Tobias Mohn, Heiko Güthenke

Optimierung der betrieblichen Planungsprozesse bei der DB Fernverkehr AG zur Umsetzung der Anforderungen des AnnexVII

Darstellung über die Auswirkungen und Vorteile auf die Prozesse und Kunden der DB Fernverkehr AG

Körner, Tobias; Rohrberg, Danny; Mohn, Tobias; Güthenke, Heiko

DB Fernverkehr AG

Zusammenfassung

Die Einführung des AnnexVII durch die DB InfraGO AG bedeutet einen großen Schritt für die Vereinheitlichung und Verbesserung der Planung und Kommunikation von Baumaßnahmen im europäischen Schienenverkehr. Um auf die Auswirkungen zu reagieren und die Vorteile nutzen zu können, hat die DB Fernverkehr AG das Projekt „Planung neu denken“ aufgesetzt. Im Zuge der Optimierung der betrieblichen Planungsprozesse wird ebenfalls der Einsatz eines IT gestützten Umlaufoptimierers angestrebt, der den Planungsprozess deutlich beschleunigen soll. In diesem Paper wird die methodische Vorgehensweise bei der Umsetzung sowie die Vorteile aus Sicht der Kunden beschrieben.

Keywords: Bauplanung, EVU, Prozessentwicklung, IT-Prozessautomatisierung

1 Einleitung und Problemstellung

Mit dem Annex VII der Richtlinie 2012/34/EU gibt die Europäische Union (EU) den Eisenbahninfrastrukturbetreibern (EIU) in der EU einheitliche Vorgaben zu Fristen in der Kommunikation und Abstimmung von Baumaßnahmen und legt verbindlich fest, wann welche Art von Baumaßnahmen in der Fahrplanerstellung zu berücksichtigen sind. Der

Infrastrukturbetreiber DB InfraGO AG plant, die Richtlinie der EU in zwei Stufen umzusetzen. Die erste Stufe, der Übergangszeitraum, beginnt ab Erarbeitung des Fahrplanjahrs 2027 (ab 12/2026 | Vorbereitungen ab Sommer 2026) und soll ab Fahrplanjahr 2028 (ab 12/2027 | Vorbereitungen ab Sommer 2027) von der zweiten Stufe, dem Zielzustand, abgelöst werden.

Mit der Umsetzung des Annex VII ändern sich die Planungsstrukturen und insbesondere die Bauplanungsprozesse sowie Baudokumente, welche die DB InfraGO AG an die Eisenbahnverkehrsunternehmen (EVU) liefert, grundlegend. Das heutige baustellenbezogene Planen mit stetiger Weiterentwicklung des Fahrplans wandelt sich zu einem sich wöchentlich wiederholenden und standardisierten Planungsprozess mit gesammelten Baustellengrundlagen. Das Planungsobjekt hat künftig Kalenderwochenbezug statt Baustellenbezug. Ein Grobplanungsergebnis (GPE) benennt 24 Wochen vor Betriebswoche (grobe) Bauauswirkungen für die jeweilige Kalenderwoche. Da dieses Dokument Schätzzeiten enthält und Ortsbrüche nicht auszuschließen sind, kann dieses Planungsobjekt lediglich im Sinne einer ersten groben Planungsiteration verwendet werden. In der 1. Umsetzungsstufe (ab Fahrplan 2027) liefert DB InfraGO acht Wochen vor der Betriebswoche mit dem Feinplanungsergebnis (FPE) ein Dokument, das als Grundlage für eine konkrete Fahr- und Umlaufplanung sowie verbindliche Kundeninformation genutzt werden kann.

Neben der Einführung des Annex VII wirken noch weitere externe sowie interne Faktoren auf die Planungsprozesse bei DB Fernverkehr und verstärken den Handlungsdruck die Planungsprozesse neu zu denken sowie effektiver und zugleich effizienter zu gestalten. Ein geringes Maß an Standardisierung und Automatisierung führt heute zeitweise zu Überlastung von Mitarbeitern und einer Beeinträchtigung der Planungsgüte. Exemplarisch kann dieses Spannungsfeld anhand der von Baumaßnahmen betroffenen Züge im DB Fernverkehr verdeutlicht werden. Im Jahr 2019 lag die Baubetroffenheit aller Züge im DB Fernverkehr bei ca. 50%. Im Jahr 2024 waren es ca. 61%[1]. Gleichzeitig haben sich die unterjährig angekündigten Baumaßnahmen um knapp 40% gegenüber 2019 erhöht. Die zu erwartende Intensivierung von Bautätigkeiten in den nächsten Jahren sowie nicht auszuschließende Außerfristigkeiten von Baudokumenten werden die Problematik sehr wahrscheinlich verschärfen.

[1] Quelle: Interne Auswertung DB Fernverkehr AG

Die Auswirkungen zeigen sich für die Kunden des DB Fernverkehrs vor allem in schlechter Pünktlichkeit sowie Unzuverlässigkeit und ggf. mangelhafter Kundeninformation. Die besondere Herausforderung in der Kundeninformation besteht vor allem darin, den Kunden entsprechende Informationen über etwaige Änderungen der gebuchten Reiseverbindung präzise und zeitgerecht zur Verfügung zu stellen und ihnen gleichzeitig mit der Vielzahl an Änderungen, die ab dem Buchungszeitraum auf den Fahrplan wirken, nicht zu „überhäufen“.

Wie in Abbildung 1 dargestellt, beeinflussen die Änderungen des Annex VII, inkl. der benannten neuen Inputgrößen, die gesamte Prozesskette von der Jahresplanung über die unterjährige Bauplanung bis zur Disposition und lassen sich nicht einfach in die bestehenden Prozesse integrieren.

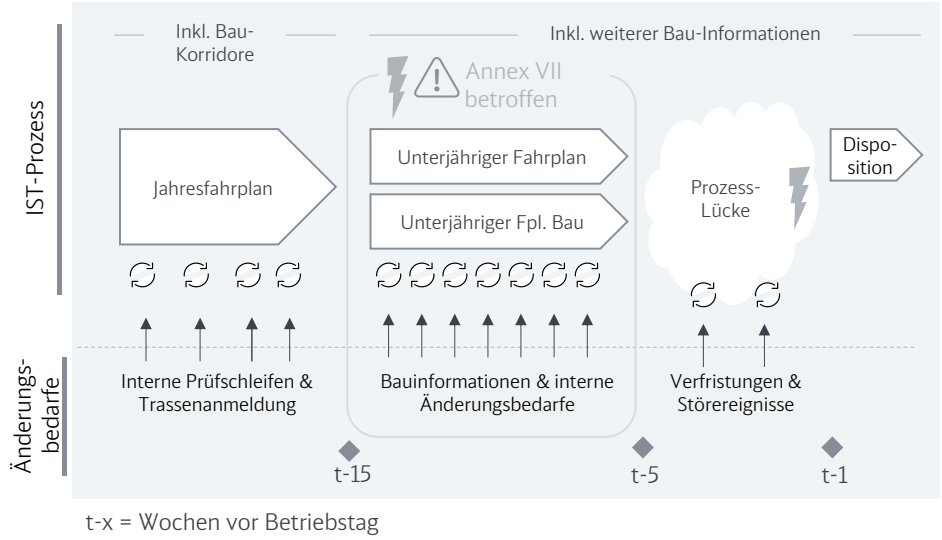


Abbildung 1: Auswirkungen des AnnexVII auf die Prozesskette in der Planung

Die größten Änderungen ergeben sich in der unterjährigen Planung und in direkter Abhängigkeit auch in den Prozessen der angrenzenden Schnittstellen, wie der Personalplanung von Bordservice und Triebfahrzeugführern, in den Werken und beim Fahrzeugmanagement sowie in den Bereichen Angebotsmanagement und den Informationsmedien für Kunden und Vertrieb (z.B. Fahrplanauskunft, Fahrscheinverkauf, Reservierungen u. dgl.).

In der 2. Umsetzungsstufe soll das FPE mit einem Vorlauf von 18 Wochen vor Betriebstag deutlich früher kommen. Dadurch werden sich viele der genannten Problemstellungen, die sich durch die erste Umsetzungsstufe ergeben, deutlich reduzieren.

2 Die Neugestaltung des Planungsprozesses durch „Planung neu denken“

Aufgrund der zahlreichen Einflussfaktoren, insbesondere durch die Änderungen von Annex VII, hat der Bereich „Fahrplan und Verkehrsleitung“ der DB Fernverkehr AG das Projekt „Planung neu denken“ gestartet. Ziel ist die Neugestaltung der Planungsprozesse für die Jahresplanung, die unterjährige Planung und die Schnittstelle zur Disposition. Darauf aufbauend werden spezifische Detailprozesse und ein Organisationsmodell ausgearbeitet, das Strukturen und Verantwortlichkeiten klar festlegt. Ein zentraler Schritt ist die umfassende Pilotierung, bei der die neuen Prozesse und Strukturen in einem zeitlich begrenzten Rahmen im Regelbetrieb getestet werden, um evtl. Schwachstellen zu erkennen und Verbesserungen vorzunehmen. Gleichzeitig werden IT-Anpassungen definiert, um die neuen Prozesse optimal zu unterstützen. Schließlich erfolgt die (Neu-) Dimensionierung der Planungsorganisation, gefolgt von der Einführung der neuen Organisation, begleitet durch Change-Management und ein Monitoring für eine nachhaltige Umsetzung.

Um diese Vielzahl an Handlungsfeldern projektseitig umzusetzen, gliedert sich „Planung neu denken“, wie in Abbildung 2 dargestellt, in fünf Projektmodule, die nachfolgend näher beschrieben werden.



Abbildung 2: Projektmodule von „Planung neu denken“

Der Prozess der Jahresfahrplanentwicklung beginnt mit der Erstellung des Jahresfahrplans als stabile Grundlage. Anschließend startet die unterjährige Planung, bei der die weitere Ausarbeitung in Kalenderwochen erfolgt. Dabei werden zunächst EVU-seitige Fahrplanänderungen gesammelt, die als Basis für die Grobplanung dienen. Etwa 24 Wochen vor dem Betriebstag erfolgt die erste Anpassung der Fahrlagen- & Umlaufplanung, bei der die gesammelten EVU-seitigen Änderungen sowie grobe Bauinformationen aus dem Grobplanungsergebnis (GPE) berücksichtigt werden. Die Umlaufplanung erfolgt dabei für eine sogenannte Planperiode von 8 Wochen inkl. der Prüfung von Abstellmöglichkeiten für die Fahrzeuge in diesem Zeitraum. Die

Informationen aus Fahrlage und Umlauf werden am Ende des Prozessabschnitts für die Einsatz- und Schichtplanung bereitgestellt.

Etwa acht Wochen vor dem Betriebstag erfolgt die Feinplanung auf Grundlage des Feinplanungsergebnisses (FPE). In dieser Phase werden die Planungen, unter Berücksichtigung konkreterer Bauinformationen und neuer EVU-seitiger Anpassungsbedarfe sowie durch Anwendung eines IT-gestützten Umlaufoptimierers, nochmals optimiert, um den Zielzustand fünf Wochen vor dem Betriebstag zu erreichen. Dabei wird auch die Umlaufplanung für eine Kalenderwoche aktualisiert. Somit steigt die Taktung der Planung von einer periodenweisen Planung zu einer Kalenderwochenplanung. Die in der Feinplanung vorgenommenen Anpassungen werden in die Kundeninformationen eingepflegt, um sicherzustellen, dass Fahrgäste zeitnah über alle relevanten Änderungen informiert werden. Dieser beschriebene Ablauf der Grob- & Feinplanung wird im Zielprozess einen rollierenden, systematischen Prozess darstellen, der entlang standardisierter Inputs funktioniert. Durch die Orientierung an standardisierten Eingaben wird gewährleistet, dass alle Prozessschritte effizient, einheitlich und fehlerfrei durchgeführt werden können, während gleichzeitig eine flexible Anpassung an neue Anforderungen möglich bleibt. Innerhalb dieses Systems können kleinere Sonderfälle oder minimale Verfristungen gezielt behandelt werden, indem definierte Handlungsoptionen oder alternative Prozesswege integriert werden. Diese Ansätze sorgen dafür, dass Abweichungen in einem kontrollierten Rahmen bleiben und den Gesamtprozess nicht stören. Dadurch soll ein störungsfreier „Fließband-Prozess“, wie in Abbildung 3 dargestellt, gewährleistet werden.

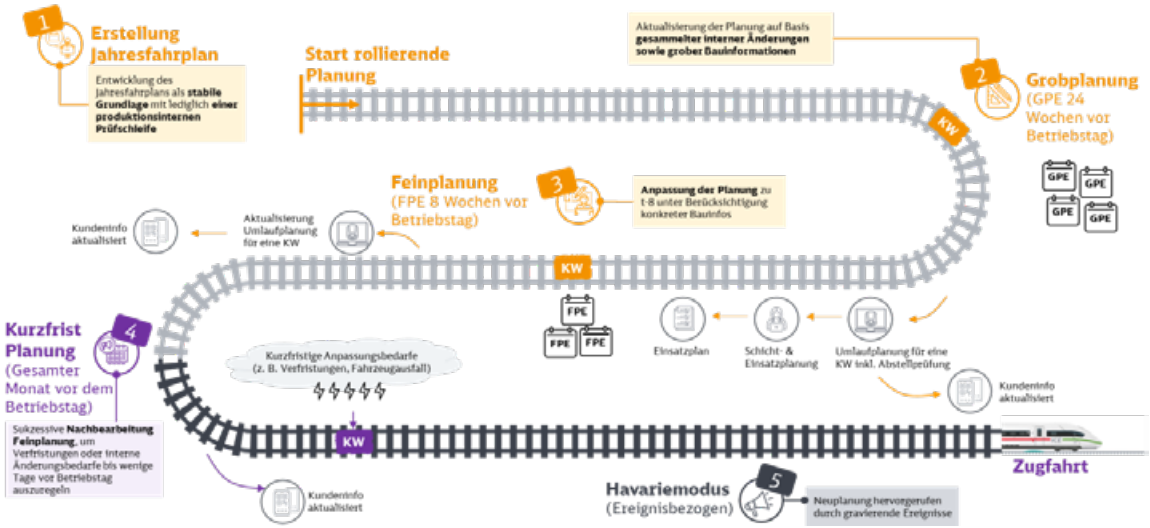


Abbildung 3: Vereinfachtes Prozess-Zielbild von "Planung neu denken"

Im letzten Monat vor dem Betriebstag erfolgt eine intensive Kurzfristplanung, bei der die Feinplanung sukzessive nachbearbeitet wird. Ziel dabei ist es, Verzögerungen, Fahrzeugausfälle oder kurzfristige andere Änderungsbedarfe bis wenige Tage vor dem Betriebstag auszugleichen, ohne den Regelprozess negativ zu beeinträchtigen. Herausfordernd ist in diesem Zeitraum, dass alle an der Produktion beteiligten Ressourcen bereits ihre Planung begonnen haben und Änderungen konsistent in allen Planungssystemen umgesetzt werden müssen. Dazu wird ein crossfunktionales Planungsteam, das die Kompetenzen der Bereiche Fahrplan, Umlauf, (Zug-) Personal sowie Abstellung, Behandlung, Instandhaltung und Disposition abbildet, gemeinsam an der Produktionsplanung arbeiten. Durch den kontinuierlichen / täglichen Austausch und die koordinierte Zusammenarbeit zwischen den Gewerken sowie die klare Zuordnung von Verantwortlichkeiten kann auf Störungen schnellstmöglich und konsistent reagiert werden. Die beschriebene Vorgehensweise des crossfunktionalen Planungsteams wird vereinfacht in Abbildung 4 dargestellt. Eine abschließende Konsistenzprüfung stellt sicher, dass die Fahrplan- und Produktionsplanungsdatenstände an allen relevanten Schnittstellen im Sinne der “Vollständigkeit und Synchronität” übereinstimmen. Den Abschluss dieser Kurzfristplanung bildet die konsolidierte Übergabe der Fahrplan- und Produktionsplanungsdaten an die disponierenden Abteilungen im Fernverkehr.



Abbildung 4: Vereinfachter Ablauf des der Kurzfristplanung

Sollte es zu gravierenden Ereignissen kommen, die eine grundlegende Neuplanung erforderlich machen, tritt der sogenannte „Havariemodus“ in Kraft. In diesem Modus wird der Fahrplan ereignisbezogen angepasst, wobei klare fachliche und kommunikative Verantwortlichkeiten gelten, um die Auswirkungen so schnell und effektiv wie möglich zu managen.

3 Auswirkungen des Einsatzes eines Umlaufoptimierers auf die Prozesseffizienz

Eine wöchentlich rollierende Planung – zukünftig auch auf Einzelzugbasis – steigert die Anzahl der notwendigen Planungsvorgänge. Verstärkend ist davon auszugehen, dass auch in Zukunft ein hohes und ggf. noch weiter steigendes Baugeschehen zusätzliche Umplanungsvorgänge generiert. In diesem Szenario hat eine IT-gestützte Umlaufoptimierung mehrere Zielrichtungen: Einerseits eine deutliche Effizienzsteigerung durch eine messbare Verkürzung der Erstellungszeiten für eine Neu- oder Umplanung, andererseits einen Qualitätshub im Sinne der Fahrzeugeinsatzquote bzw. Laufleistung je eingesetztem Fahrzeug. Ausgehend von den individuellen Änderungen der zu planenden Woche, erstellt der Optimierer defektfreie Umläufe, die die Fahrzeuge und Ressourcen mathematisch optimiert nutzen.

Ein kompletter Umlauf besteht dabei aus einer Sammlung von Leistungsketten je verwendetem logischen Fahrzeug, die sich über einen definierten Zeitraum (hier im Optimierungsszenario die Kalenderwoche (KW)) erstreckt. In der Leistungskette befinden sich neben der reinen Abfolge von Fahrzeugbewegungen (Ort, Zeit, Strecken, Gleise, Identifizierungsmerkmale einer Zugfahrt wie Zugnummer, Last-, Leer-, Rangierfahrt etc.) auch Informationen über Wartung und Instandhaltungen, zugehörige Fristen, Reihungsinformationen etc. Die Leistungsketten sind dabei stets so zu halten, dass diese an beiden Grenzen eines Optimierungszeitfensters als fixe Randbedingungen direkt anschlussfähig sind.

Mathematisch kann die Erstellung eines optimalen Umlaufs als ein Problem der ganzzahligen Optimierung, gesehen werden, da natürlich nur „ganze“ Fahrzeuge Teil eines Umlaufs sein können.

Die Herausforderungen in der Umsetzung bestehen in verschiedenen Dimensionen:

1. Grundsätzliche Formulierung der Umlaufproblematik als ein ganzzahliges lineares Optimierungsproblem

2. Ergänzung um spezifische Rahmenbedingungen und Integration in bestehende IT- und Prozesswelt
3. Finden einer gemäß Kostenfaktoren optimalen Lösung in angemessener Zeit

Im Markt sind zahlreiche kommerzielle (Gurobi, IBM ILOG CPLEx, …) oder auch Open Source Optimierer vorhanden, die Umlaufprobleme erfolgreich lösen können. Die Produkte sind im praktischen Einsatz bewährt und werden kontinuierlich weiterentwickelt. Die Hürde bei der Anwendung eines generellen Optimierers besteht in der Formulierung und Adaptierung an bahnspezifische Randbedingungen wie Kopplungs- und Ausrichtungsrestriktionen, Unterscheidung in Zug- und Rangierfahrten, Streckentauglichkeiten sowie der Vielzahl an sonstigen Randbedingungen wie Instandhaltungsfristen, Werkeanbindungen und Abstellkapazitäten etc. Zudem kommt es erfahrungsgemäß im Sektor Bahn zu anlassbezogenen oder prozessual vorgesehenen mehrfachen Neu- oder Umplanungen, bspw. in Folge von Baustellen, Änderungen bei den verfügbaren Fahrzeugen oder langanhaltenden Betriebsstörungen. Für das Planungswerkzeug und auch den eingesetzten Optimierer bedeutet dies, dass diese mehrfach im Prozess eingesetzt werden müssen, zyklisch und teilweise parallel. Diese Anforderungen muss auch der eingesetzte Optimierer erfüllen. Der Gesamtprozess im zeitlichen Ablauf ist in Abb. 5 schematisch dargestellt. Im Sinne des „Fließband-Prozesses“ aus Abb. 3 in Abschnitt 2 wird jeweils eine Kalenderwoche im jeweiligen Schritt vollständig bearbeitet.

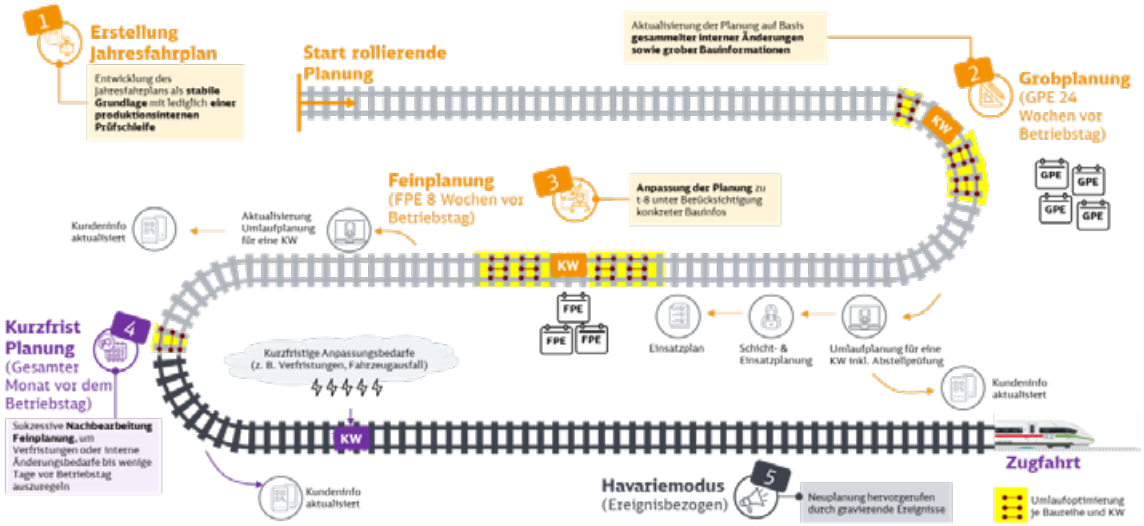


Abbildung 5: Gesamtablauf Planungs- und Optimierungsprozess

Vor diesem Hintergrund empfiehlt sich eine Lösung, die den Bahnbetrieb im Sinne einer vollintegrierten Enterprise Ressource Planning-Plattform (ERP) bereits abbildet und dann um Optimierungskompetenzen erweitert wird. So wird automatisch eine Konsistenz der Rahmenbedingungen über Betrieb und Optimierung hergestellt sowie Schnittstellen, unterschiedliche Taxonomie und Synchronisationsabläufe vermieden.

Im Rahmen der Umstellung auf den neuen Planungsprozess ist eine vollständige und exakte Ausformulierung und Hinterlegung von Rahmenbedingungen in den IT-Systemen notwendig. Dabei ist die Verschriftlichung von Anwendungsregeln, die einem Umlaufplaner intuitiv oder durch lange Ausübung seiner Tätigkeit bekannt sind, eine große Herausforderung und wesentlicher Erfolgsfaktor.

Jede der Randbedingungen könnte mit einem individuellen Kostenbeitrag versehen werden. In der Praxis liegt dieser Detaillierungsgrad aber nicht vor, so dass verallgemeinernde Annahmen zu treffen sind, die dann wiederum im Rahmen einer Evaluation auf Gesamtplausibilität zu überprüfen sind. Bei der Einführung erweist sich dies als eine Thematik, die sich durch iteratives Wiederholen lösen lässt. Die Formulierung der einzelnen Kostenfunktion in Form von einfachen linearen, exponentiellen oder Stufenfunktionen ist dabei ausreichend.

Die Dauer eines Optimierungslaufs bleibt in dem oben beschriebenen Szenario bis auf weiteres mit mehreren Stunden ein limitierender Faktor. Trotz sehr performanter Hardware werden mehrere Stunden für die Berechnung einer vollen Umlaufwoche (7 Tage) für eine Baureihe mit ca. 50 Fahrzeugen benötigt. Eine Ausweitung auf längere Zeiträume oder baureihenübergreifend für noch bessere Ergebnisse ist damit bei aktueller Performance noch nicht in Reichweite.

4 Ergebnisse und Nutzen der Prozessumstellungen für den Kunden

Aus Sicht unserer Kunden ergeben sich durch den Abschluss der ersten Projektphase von „Planung neu denken“ einige Vorteile. Der neue Planungsprozess bietet Stabilität und Flexibilität. Kunden können sich vom länger aktuell gehaltenen Angebot auf Basis des Jahresfahrplans inspirieren lassen. Die zwei Planungsiterationen (Jahresfahrplan & Grobplanung) lösen das Spannungsfeld aus kontinuierlicher Fahrplanaktualisierung zu einem (zu) frühen Zeitpunkt und rechtzeitiger Informationsbereitstellung auf. Die effektive Kurzfristplanung gewährleistet eine produktionsintern abgestimmte und zügige Einarbeitung aller kurzfristigen Änderungen in die jeweiligen IT-Systeme und schafft

dadurch planerische Handlungsfähigkeit, um auf die volatilen (Fahrplan-)Einflüsse wie beispielsweise verfristete Baudokumente oder kurzfristige Instandhaltung reagieren zu können. Dies resultiert in einer aktuelleren Fahrplanauskunft und weniger dispositiven Eingriffen kurz vor dem Betriebstag.

5 Ausblick

Die besondere Herausforderung bei allen Umsetzungsvorhaben im Rahmen von Planung neu denken ist, dass ein reibungsloser Übergang der alten Prozesse auf die neuen Prozesse gewährleistet und zusätzlich immer eine Rückfallebene sichergestellt sein muss. Daher werden die Ergebnisse der einzelnen Projektmodule schrittweise umgesetzt und vorab im Rahmen von Pilotierungen unter Realbedingungen verprobt. Erst wenn sich der neue Prozess bezüglich der Leistungsparameter Qualität, Effizienz und Kundenwirkung bewährt hat, wird das betreffende Projektmodul als neuer Regelplanungsprozess eingeführt. Für das Jahr 2025 stehen dabei das Modul 5 (Havariemodus) und das Modul 4 (Kurzfristplanung) im Fokus. Ein Ziel ist es, hierbei schnellstmöglich an Reaktionsfähigkeit auf die zunehmenden kurzfristigen (Prozess-) Störfaktoren zu gewinnen sowie die Ergebnisse zu stabilisieren und weitere Anpassungen im Rahmen des kontinuierlichen Verbesserungsprozesses vorzunehmen. Ein weiteres Ziel ist es, Vollständigkeit und höchste Qualität bei der Formulierung und Hinterlegung der Rahmenbedingungen im verwendeten IT-Verfahren zu erreichen. Damit wäre eine breite Einsetzbarkeit für Neuplanungen, Umplanungen oder Szenarienbewertungen in der Umlaufplanung jederzeit im laufenden Betrieb möglich. Mit Einführung des Annex VII werden die Ergebnisse aller Projektmodule in einen neuen Regelprozess überführt.

Session 15
Tobias Körner, Danny Rohrberg, Tobias Mohn, Heiko Güthenke

Autoren



Körner, Tobias
Leiter Entwicklung Betrieb; DB Fernverkehr AG, Frankfurt am Main



Rohrberg, Danny
Projektleiter „Planung neu denken“ im Bereich Entwicklung Betrieb, DB Fernverkehr AG, Frankfurt am Main



Mohn, Tobias
Senior Referent Strategie und Konzepte Betrieb, DB Fernverkehr AG, Frankfurt am Main



Güthenke, Heiko
Leiter Digitalisierung Betrieb; DB Fernverkehr AG, Frankfurt am Main

IRSA 2025

SESSION 16

FAHRZEUGTECHNIK
VEHICLE TECHNOLOGY

Session 16
Dieter Knabl, Alexander Plotho, Stefan Marschnig

Load-dependent influences on rail damage mechanisms

Methodology for wear assessment based on empirical data

Knabl, Dieter [1], Plotho, Alexander[1], Stefan Marschnig[1]

[1] Institute of Railway Engineering and Transport Economy at Graz University of Technology, Rechbauerstrasse 12/II, 8010 Graz, Austria

Summary

This paper presents a methodology for evaluating rail wear in order to support predictive maintenance. Although predictive models are widely used for aspects such as track geometry, their application to rail wear is still in its infancy. This study uses processed measurement data from ÖBB's track recording car, combined with infrastructure and loading information, to enable large-scale, data-driven analysis.

Track sections are divided into specific elements, and the development of wear is analysed over time using robust linear regression. Deterioration periods are identified between maintenance events or operational changes. The first step in assessing load-dependent wear is to normalise wear rates by cumulative gross tonnage.

The analysis confirms well-known correlations, such as increased wear in tighter curves and variations in rail steel grades. This scalable methodology provides a foundation for forecasting models that can improve maintenance planning and efficiency.

Future work will focus on integrating vehicle-specific data and exploring stress-based indicators to improve the ability to predict rolling contact fatigue as well as wear.

Keywords: Rail wear; Predictive maintenance; Wheel–rail interaction; RCF; Rail infrastructure

1 Introduction

Rising traffic volumes put increasing demands on railway infrastructure, leading to a corresponding rise in maintenance requirements. Meanwhile, the time available for maintenance activities is becoming more limited due to tighter operational schedules and the need to minimise disruption or delays to train services. Consequently, infrastructure operators face the dual challenge of maintaining system reliability within increasingly narrow timeframes.

The efficiency of maintenance processes, in terms of both planning and execution, has therefore become critically important. Maintenance activities must be timed precisely, coordinated effectively and carried out with minimal impact on railway operations and to affordable costs. In this context, the role of forecasting is paramount. Reliable forecasting methods are essential for meeting these challenges as they enable infrastructure managers to anticipate accurately when and where maintenance will be required and allocate resources accordingly.

This paper explores the development and application of such forecasting approaches in the context of modern railway maintenance management, paying particular attention to rails. As one of the most heavily loaded components of the track system, the rail is subject to continuous wear and degradation. The ability to predict its condition in advance is crucial for ensuring safe, cost-efficient and timely maintenance interventions. The focus of this study is therefore on evaluation methods enabling for sophisticated forecasting methods that support the proactive and data-driven planning of rail maintenance activities.

2 Background and state of the art

In recent years, predictive maintenance strategies have become an integral part of railway infrastructure management, particularly with regard to track geometry maintenance [1–5]. These strategies use historical data, current asset conditions and predictive modelling techniques to forecast maintenance needs. This makes planning more efficient, reduces service disruption and optimises the allocation of resources. While such methodologies are now well established for several track components, their application to rail-specific issues, especially rail wear, remains comparatively underdeveloped.

This disparity highlights the need for a more comprehensive understanding of the mechanisms and boundary conditions that cause rail degradation. Rail wear is not solely a function of cumulative loading; it is also significantly influenced by various vehicle-

and track-related parameters. Vehicle characteristics such as axle spacing, pivot pitch and bogie steering behaviour affect dynamic wheel–rail interactions and substantially influence wear development. Similarly, track attributes such as curvature, cant, inclination and superstructure configuration contribute to the transmission and distribution of mechanical forces at the contact interface. Identifying and quantifying the relevance of these parameters is essential for advancing robust prediction models.

A key prerequisite for a predictive maintenance approach at the network level is the implementation of automated, large-scale rail wear assessment. This would enable consistent, high-resolution data to be collected across extensive sections of track, providing a reliable basis for empirical analysis and data-driven decision-making. However, there are several practical limitations that complicate this effort. One of the primary challenges is undocumented or only partially documented maintenance interventions, such as rail surface treatments or rail replacements, which are often not systematically recorded. These interventions may be carried out over entire track segments or only in isolated portions, which introduces discontinuities into the wear progression data. Consequently, these inconsistencies can compromise the reliability of statistical evaluations and forecasting algorithms. The methodology presented in Section 3 addresses approaches for managing and mitigating these uncertainties in the data.

From a mechanical perspective, rail wear results from complex interactions at the wheel–rail interface, where rolling and sliding contacts produce elevated stresses and temperatures. These conditions cause material loss through a combination of oxidative and fatigue-related wear mechanisms [6,7]. The rate and extent of wear progression are influenced by various factors, such as axle loads, operational speeds, lubrication states, track geometry and contact dynamics [8–13]. In particular, increased axle loads and train speeds are associated with elevated contact pressures and frictional stress, exacerbating rail degradation and leading to changes in contact geometry over time [8].

In summary, although predictive maintenance for rails, especially in the context of wear, lags behind other infrastructure domains, it has considerable potential. Realising this potential requires an improved empirical understanding of wear mechanisms and systematic collection of reliable input data. Ongoing research aims to bridge this gap by investigating the correlation between wear progression and influencing parameters using real-world operational data.

3 Methodology

The following sections outline the development of evaluation methods for automated wear data analysis.

3.1 Data basis

This study's wear assessment is based on data provided by ÖBB's track recording car. While the recording car takes detailed measurements of the rail profile across the network, it processes the raw data internally to calculate wear values, which are then made available for further analysis. These processed wear metrics form the primary dataset for assessing rail condition and wear progression.

In addition to wear data, comprehensive asset information is required to accurately describe track superstructure parameters. This includes details on rail steel and profile, sleeper type, and other structural elements that may influence rail behaviour and wear characteristics. Furthermore, routing information like curve radius, inclination and cant is provided.

To evaluate the impact of operational conditions, we use track loading data to determine section-specific loadings. Currently, gross tonnage is available to express traffic volume, which is quite common for track loading and thus utilised as first step in this study. However, other relevant parameters, such as axle load distributions and traffic respectively train types, are not yet accessible, and must be evaluated in further research to enhance the accuracy of load-related wear assessments.

Furthermore, we integrate records of rail surface treatments such as grinding and milling into the analysis. These maintenance activities modify the rail surface and can significantly influence wear patterns. Accurate knowledge of the timing and range (length) of such interventions is therefore crucial for properly interpreting and forecasting rail wear.

3.2 Selection of suitable routes and partitioning in elements

Data from a representative cross-section of the Austrian Federal Railways' network is used for the development and validation of the proposed methodology. We select suitable routes to ensure that the approach can be applied to a wide range of infrastructural and operational conditions, and can be extended in future research to investigate the influence of various parameters on rail wear.

To this end, the selection of routes is based on the inclusion of different line types that vary in key characteristics such as traffic intensity, operational speed, routing complexity and topographical context. We consider both, heavily used main lines and lower-traffic secondary lines, as well as routes involving both high-speed and conventional operations. Geometric diversity is addressed by including sections with straight and curved alignments, and topographical variation is covered by integrating routes in mountainous and flat regions. This broad coverage is essential for ensuring the general applicability of the developed methodology and enabling future studies to investigate the influence of specific factors, such as curvature, gradient, load spectra or traffic type, on rail wear behaviour.

Once we have selected the routes, they are divided into discrete analysis units, known as 'elements'. Each element is defined by consistent geometric and structural characteristics. A new element is introduced whenever there is a change in routing (e.g. a change in radius) or in superstructure configuration (e.g. changes in rail type, rail steel grade or sleeper design). This segmentation enables wear behaviour to be directly linked to identifiable physical and operational conditions, providing a foundation for targeted analysis and potential refinement of wear prediction models in future research.

3.3 Evaluation method

In this study, the evaluation of rail wear is based on processed measurement data collected by ÖBB's track recording car, which systematically surveys the rail network. The wear signal is derived from rail profile measurements and is presented as two key indicators: vertical wear (W1) and lateral or side wear (W3). W1 represents material loss at the rail head due to vertical loading and rolling contact, and W3 reflects lateral wear on the gauge face. These metrics allow rail degradation to be assessed quantitatively over time and across different track segments. Combining W1 and W3 values provides a comprehensive overview of rail wear, supporting condition monitoring and the development of predictive maintenance models. Figure 1 provides a schematic illustration clarifying the definition and location of W1 and W3 on the rail cross-section.

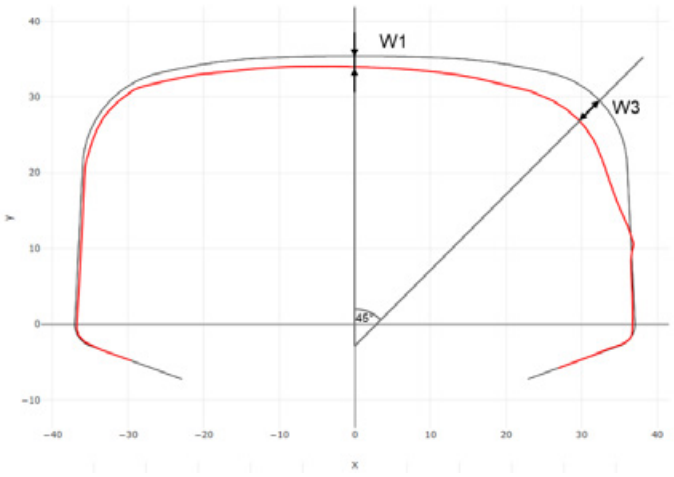


Figure 1: Scheme of different types of rail wear between design (black) and worn (red) profile

To derive a consistent wear indicator for each curve, we compute a representative value considering localised effects and data uncertainties. The biggest challenges are the curve-in effect, whereby wear typically increases with the length of the travelled curve until saturation is reached, and maintenance measures that are either undocumented or only partially recorded, such as rail surface treatments or rail replacement. These can lead to irregularities in the wear signal.

To address the former, each curve (including transition curves) is divided into two equal halves. For each half, we calculate the trimmed mean wear value by excluding the lowest and highest 5% of valid measurements to reduce the influence of outliers and local anomalies. If at least 50 metres of consecutive valid data are available in each half, the higher of the two trimmed mean values is selected as the representative wear value for the entire curve, if not, we exclude the specific element from further analysis. This approach ensures robustness against local deviations while identifying the maximum wear tendency within the curve that requires maintenance. The methodology is designed to detect dominant wear patterns regardless of operational flow, so traffic direction is not explicitly considered.

The methodology for dealing with the latter requires two input datasets. Firstly, the wear data for vertical and side wear over time must be obtained for each route, track and rail of interest. The second dataset required is track information, including identification data such as the track and curve number, and routing data such as radii and kilometre markings.

Prior to executing the calculations, the data must be prepared. Firstly, we sum up the wear data for each rail (vertical and side wear are added) to make any fluctuations in the data

more substantial. In the next stage, some of the data will be excluded: a) If the total wear is less than a certain amount or if the radii are particularly large (a combination which is often true), we exclude the element. b) We calculate the standard deviation of the wear data for each measurement run. If the standard deviation is zero, this indicates an error in the measurement, the data-set will be excluded from further analysis. After erroneous data has been eliminated, the mean sum wear is calculated for each measurement run and element. As illustrated in Figure 2, it is evident that a mean wear development over time has been achieved.

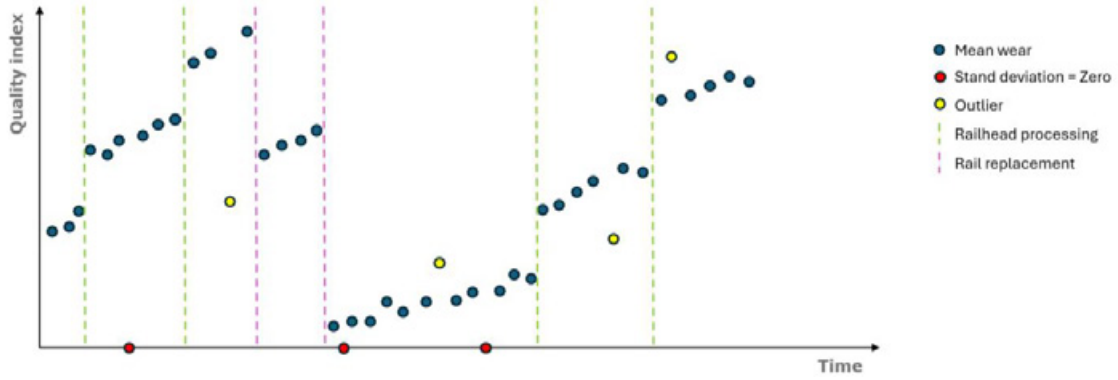


Figure 2: Exemplary time series of wear data with detected maintenance actions,

Given the potential for measurement errors and inaccuracies, it is essential to use a robust method to detect outliers. In this instance, we use a combination of three distinct methods. If two of these methods are applicable, the mean wear is identified as a true outlier. Common methods employed in the literature include k-Nearest Neighbours (k-NN) [14], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [15], and Local Outlier Factor (LOF) [16]. These methodologies are effective because they can identify outliers within local densities. Figure 2 shows the identified outliers.

The jump detection method, which actually detects significant changes in the signal, can be divided into two sections. The first section uses kernel change-point detection (KernelCPD), while the second uses the initial step of the cross-section and RANSAC-based algorithm (CRAB).

KernelCPD can be defined as a methodology for the identification of structural breaks (change points) in time series or sequential data. We divide time series into segments based on the kernel method to ensure the statistical properties of the data are as homogeneous as possible whilst allowing for significant changes between the segments.

The fundamental concept underpinning this approach is the utilisation of a kernel function. The function $k(x, x')$ is employed to map the original data into a higher-

dimensional space, thereby enabling linear separations between different segments. In such cases, the radial basis function (RBF) kernel is typically employed [17,18]. This preliminary step is designed to organise the data in a manner that optimises the efficacy of subsequent analysis.

The CRAB algorithm is used once the data has been divided into smaller segments. The algorithm's methodology involves identifying the optimal pair of data. This is achieved by establishing a margin around each combination of data pairs. The pair retaining the most data points defines a newly found structural break [19].

Following the acknowledgement of all pertinent factors, the subsequent data is presented in graphical form (see Figure 2) and tabular format.

3.4 Evaluation of time series and determination of valid deterioration branches

We plot previously calculated mean wear values chronologically for each track segment in order to analyse the development of rail wear over time. Based on the delivered time series, we can now evaluate the progress of wear and identify consistent deterioration periods.

A critical step in the analysis is to segment the time series temporally into valid deterioration branches – time intervals during which wear evolves under stable conditions. The boundaries of these branches are defined by significant interventions or changes in the operational context. These include maintenance actions, such as rail surface treatments or rail replacements, which are identified through either documented maintenance records or detection of abrupt changes in the wear signal as described in the previous section. Additionally, we consider operational changes such as the introduction of new vehicle types or alterations in traffic direction as segmentation points since they can lead to shifts in loading conditions and consequently in wear behaviour.

Within each identified deterioration branch, we model wear progression using robust linear regression. This approach enables wear rates to be estimated while minimising the influence of outliers and measurement noise, ensuring reliable trend extraction even in the presence of limited or imperfect data.

The spatial segmentation, based on the previously defined track elements, remains fixed throughout the analysis. This ensures that the temporal evaluation is always linked to clearly defined physical and operational conditions, enabling consistent interpretation and comparison of wear behaviour in space and over time.

Figure 3 shows all the available wear data for one rail – the left rail, which corresponds to the high rail in this example – within a selected track element, covering the whole period for which measurement data is available. The beginning and end of a clearly defined deterioration branch are marked in the figure, corresponding to two maintenance actions: a rail exchange at the beginning, and a rail surface treatment at the end. These maintenance events define a stable phase that is suitable for wear rate analysis.

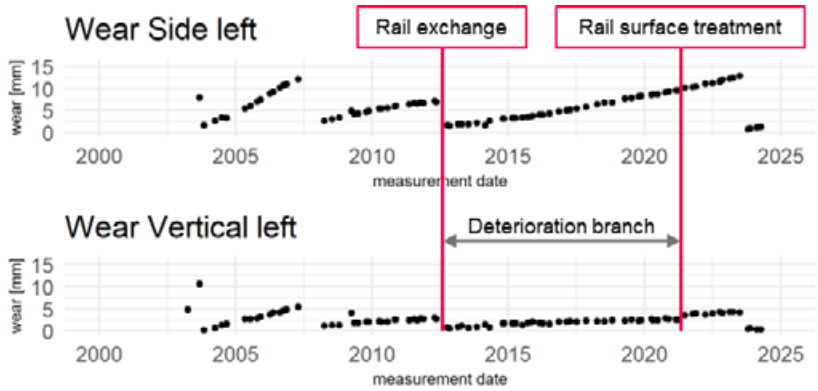


Figure 3: Wear data plot for high rail (left) of an exemplary right curve

Figure 4 zooms in on this deterioration branch, showing the side wear values recorded during the interval. A robust linear regression line is overlaid in red to represent the average wear rate for this period. This visualisation highlights the consistency of wear progression within the branch and demonstrates the applicability of the regression approach for estimating wear trends under stable conditions.

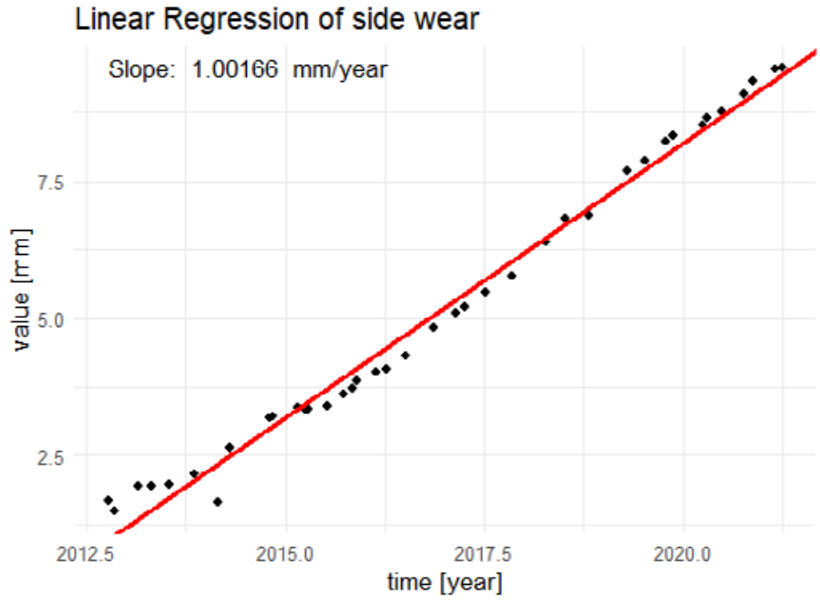


Figure 4: Linear regression of side wear of high rail over one deterioration branch

This systematic approach to temporal and spatial segmentation provides a reliable basis for determining wear rates, forming the foundation for data-driven forecasting and targeted maintenance planning.

3.5 Load-dependent wear

As a first step in quantifying the relationship between track usage and rail wear we calculate the cumulative loading, expressed in gross tons, for each identified deterioration branch. To this end, we collect track loading data for all relevant subsections and aggregate them over the duration of each branch. This yields a total load per time interval for each specific element, forming the basis for evaluating wear progression under defined loading conditions.

Then, we convert the wear rates, originally expressed in millimetres per year, into load-dependent wear by normalising them against the cumulative load. This yields a wear rate expressed in millimetres per million gross tons $\left[\frac{mm}{10^6\ t}\right]$, which provides a more meaningful indicator of wear relative to usage.

While gross tonnage is used in the current analysis due to data availability, it is recognised as a relatively coarse indicator of actual mechanical loading. Therefore, replacing gross tons with more suitable, specific loading measures is a key objective of future research aimed at improving the accuracy and explanatory power of wear modelling.

3.6 Identification of dominant influencing factors

In order to deepen our understanding of wear progression mechanisms, the next step in the methodology involves identifying correlations between rail wear and relevant influencing parameters from the track and vehicle domains. Linking empirical wear data with available infrastructure and operational characteristics enables key dependencies to be identified and quantified.

One of the most apparent relationships is the influence of track geometry, particularly curvature. As expected, sections of track with a greater curvature – equalling smaller radius – generally exhibit higher wear rates than sections with a lesser curvature (or even straight sections), due to the increased lateral forces at the wheel–rail interface. The analysis therefore begins by examining the correlation between curvature (resp. track radius) and calculated wear rates across all deterioration branches.

Figure 5 illustrates the relationship between load-dependent side wear of high rails in curved track sections and curve radius. This visualisation highlights the tendency for increasing wear with decreasing track radius, thus confirming the importance of geometric alignment as a key wear driver.

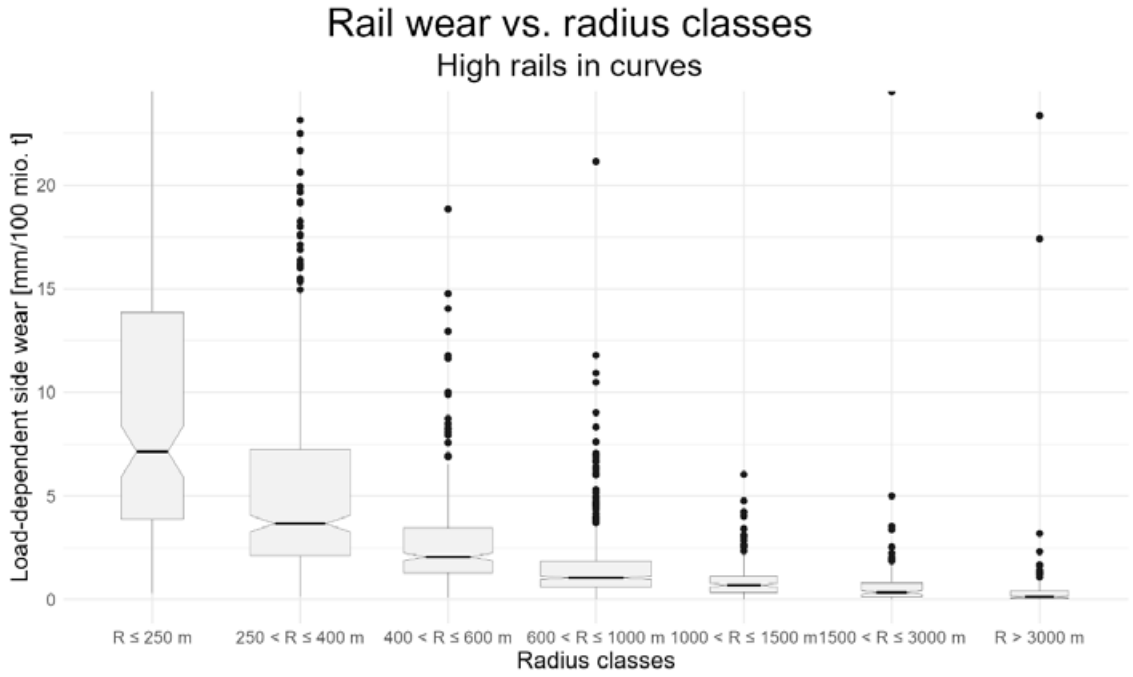


Figure 5: Influence of curve radius on load-dependent side wear of high rails

In addition to geometric effects, we also consider material properties. Figure 6 illustrates the impact of rail steel grade on load-dependent side wear in curves with different radii. The comparison reveals significant differences in wear behaviour between various steel grades, suggesting that material selection is a key factor in wear resistance, particularly under high-stress conditions, such as those encountered in tight curves.

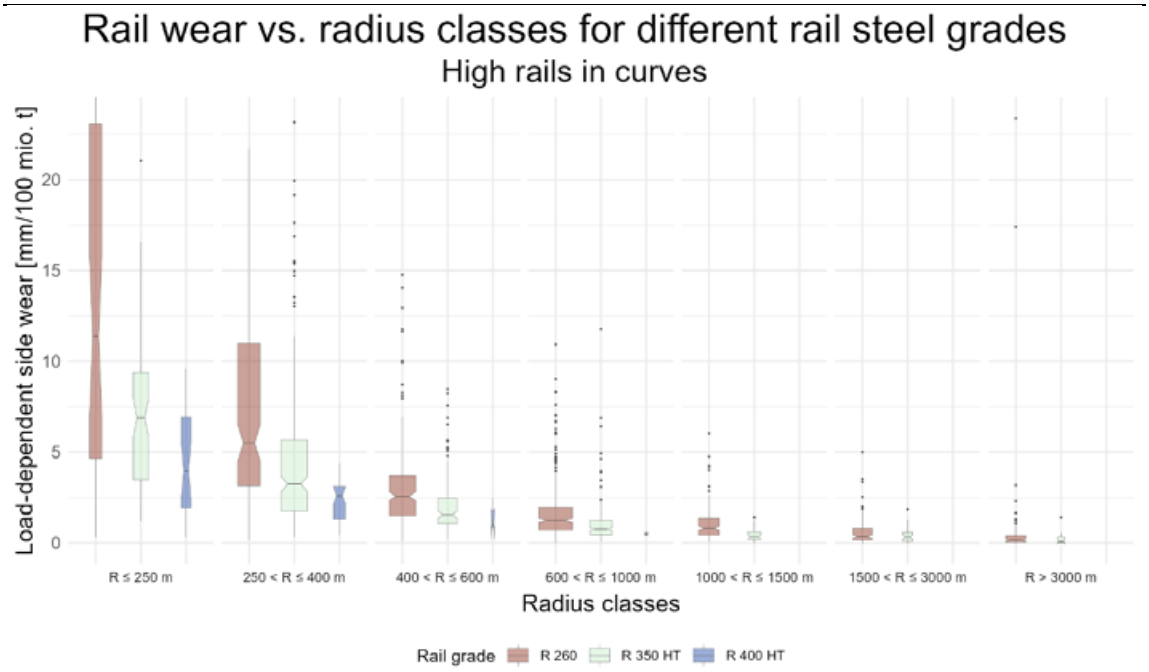


Figure 6: Influence of rail steel grade on load-dependent side wear of high rails in curves

As the data basis expands, we will progressively include additional factors such as track construction features and operational characteristics in the analysis. This stepwise approach enables a systematic evaluation of potential dependencies and supports the long-term objective of creating more comprehensive, data-driven models for wear prediction and infrastructure planning.

4 Preliminary results

Preliminary evaluations confirm well-known correlations between rail wear and key influencing factors, such as track curvature radius and rail steel grade [20]. These findings validate the effectiveness of the proposed methodology in identifying important wear patterns and dependencies in the data.

By way of example, we investigate the influence of transition curve length on wear propagation by ongoing and future work. Figure 7 shows preliminary findings relating to this parameter and demonstrates its potential impact on wear progression. It also highlights the importance of considering detailed track geometry when forecasting wear.

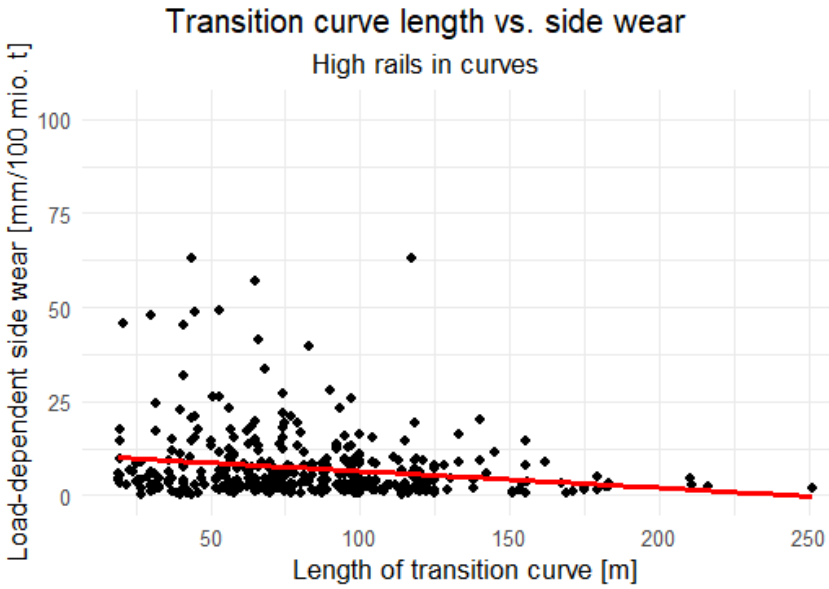


Figure 7: Influence of transition curve length on load-dependent side wear of high rails in curves with radii ≤ 400 m

As can be seen, there is a clear dependency between load-dependent side wear on high rails in curves and the length of the preceding transition curve in the direction of traffic flow. This dependency is depicted by the red regression line.

5 Discussion & further research

In this study we use data that reflects the specific boundary conditions of the Austrian railway network, including track construction standards, routing characteristics and operational parameters. Notably, the dataset contains the tightest curves with radii of around 170 metres and maximum operating speeds of up to 230 km/h, which reflect Austria's high-speed corridors. Additionally, the vehicle fleet represented in the data corresponds exclusively to rolling stock in service in Austria. Consequently, the developed methodology is tailored to these particular conditions and may require adaptation before being applied to networks with different track geometries, speed profiles or vehicle characteristics.

The restrictions we had to impose during development relates to low wear values or rates. These usually occur in conjunction with large radii or straight sections of track. This means that the described methodology, particularly the maintenance detection aspect, can only be applied reliably if there are significant changes to the wear signal over time.

The methodology presented in this study provides a structured basis for data-driven analysis of rail wear. However, several key areas of further research are identified in order to advance towards more comprehensive and predictive modelling of rail degradation.

One important area for investigation is the evaluation of crack propagation rates, particularly in relation to head checks. Therefor we need the development of an appropriate methodology, as well as a better understanding of eddy current testing data, on which crack detection and assessment are based.

In parallel, a more detailed elaboration of the section-specific vehicle collective is essential. Currently, information on the types and characteristics of vehicles operating on individual sections is limited, yet it is critical for refining the interpretation of wear behaviour and understanding localised loading conditions.

A major focus of future work will be the transition from load-based to stress-based indicators. Although we use cumulative gross tonnage as a practical loading measure in the current study, it only provides a rough approximation of the actual mechanical stress conditions acting on the rail. In modern track deterioration models (TDMs), the Tγ model is the standard for estimating damage based on combined vertical and lateral forces [21,22]. However, this research aims to identify alternative parameters that are potentially more suitable for wear and RCF evaluation. One such candidate is the lateral guiding force of the leading wheelset in the bogie, commonly denoted as Y_R, which may provide a more direct link to lateral contact conditions and material stress.

Additionally, we plan clustering track sections or curves according to boundary conditions to support identifying wear-relevant groupings and behavioural patterns. This will also help to identify and validate the most effective parameters for evaluating the contributions of both the track and the vehicle to wear and RCF initiation.

The ultimate goal is to derive reliable, condition-dependent forecast models for wear progression and RCF crack development. These models will form the basis of a forward-looking assessment of maintenance needs, including the derivation of grinding requirements aligned with actual rail degradation.

Acknowledgment

The authors would like to thank the Austrian Federal Railways (ÖBB) for supplying the operational data on which this research is based.

Literature

[1] Wen M, Li R and Salling K B 2016 Optimization of preventive condition-based tamping for railway tracks *European Journal of Operational Research* **252** 455–65

[2] Binder M, Mezhuyev V and Tschandl M 2023 Predictive Maintenance for Railway Domain: A Systematic Literature Review *IEEE Eng. Manag. Rev.* **51** 120–40

[3] prof. Stjepan Lakušić and Lakusic S 2015 Track degradation analysis in the scope of railway infrastructure maintenance management systems *JCE* **67**

[4] Neuhold J, Vidovic I and Marschnig S 2020 Preparing Track Geometry Data for Automated Maintenance Planning *J. Transp. Eng., Part A: Systems* **146** 04020032

[5] Sedghi M, Bergquist B, Vanhatalo E and Migdalas A 2022 Data-driven maintenance planning and scheduling based on predicted railway track condition *Quality & Reliability Eng* **38** 3689–709

[6] Kalker J J 1991 Wheel-rail rolling contact theory *Wear* **144** 243–61

[7] Yılmaz Sönmez H and Öztürk Z 2020 Effects of Traffic Loads and Track Parameters on Rail Wear: A Case Study for Yenikapi–Ataturk Airport Light Rail Transit Line *Urban Rail Transit* **6** 244–64

[8] Jin X, Wen Z, Xiao X and Zhou Z 2007 A numerical method for prediction of curved rail wear *Multibody Syst Dyn* **18** 531–57

[9] Dirks B, Enblom R and Berg M 2016 Prediction of wheel profile wear and crack growth – comparisons with measurements *Wear* **366–367** 84–94

[10] Meghoe A, Loendersloot R and Tinga T 2020 Rail wear and remaining life prediction using meta-models *International Journal of Rail Transportation* **8** 1–26

[11] Sadeghi J and Akbari B 2006 Field investigation on effects of railway track geometric parameters on rail wear *J. Zhejiang Univ. - Sci. A* **7** 1846–55

[12] Krause H and Poll G 1986 Wear of wheel-rail surfaces *Wear* **113** 103–22

[13] Yang Y, Wang J, An W, Cai J and Bai H 2025 Advances in prediction methods of wear of rails *Wear* **564–565** 205676

[14] Bansal M, Goyal A and Choudhary A 2022 A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning *Decision Analytics Journal* **3** 100071

[15] Wang L, Li M, Han X and Zheng K 2018 An improved density-based spatial clustering of application with noise *International Journal of Computers and Applications* **40** 1–7

[16] Breunig M M, Kriegel H-P, Ng R T and Sander J LOF: Identifying Density-Based Local Outliers

[17] Harchaoui Z and Cappe O 2007 Retrospective Mutiple Change-Point Estimation with Kernels *2007 IEEE/SP 14th Workshop on Statistical Signal Processing* 2007 IEEE/SP 14th Workshop on Statistical Signal Processing (Madison, WI, USA: IEEE) pp 768–72

[18] Aminikhanghahi S and Cook D J 2017 A survey of methods for time series change point detection *Knowl Inf Syst* **51** 339–67

[19] Schatzl J, Gerhold F, Loidolt M and Marschnig S 2024 Enhancing Predictive Maintenance Through Detection of Unrecorded Track Work *Infrastructures* **9** 204

[20] Egger J 2023 *Schienenverschleiß und Schienenstahlgüte* Master's Thesis (Graz: Technische Universität Graz)

[21] Ehrhart U, Knabl D and Marschnig S 2024 Track Deterioration Model—State of the Art and Research Potentials *Infrastructures* **9** 86

[22] Burstow M C 2004 *Whole Life Rail Model Application and Development for RSSB – Continued Development of an RCF Damage Parameter* (Derby: AEA Technology Rail)

Authors



Knabl, Dieter

He is a PhD student at the Institute of Railway Engineering and Transport Economics at Graz University of Technology. His research focuses on rail damage mechanisms in the context of vehicle–track resp. wheel–rail interaction.



Plotho, Alexander

He is currently pursuing a master's degree at Graz University of Technology and working at the Institute of Railway Engineering and Transport Economics, where he is writing his thesis on rail wear.



Marschnig, Stefan

He is the head of the Institute of Railway Engineering and Transport Economics.

Session 16
Kensuke Nagai, Haw-Shyang Chang, Wataru Ohnishi, Takafumi Koseki, Shota Miyoshi, Tomonori Hasegawa

Proposal of Reliable Onboard Train Localization Method by Integration of RTK-GNSS, IMU, and Multiple One-dimensional LiDAR Sensors

Nagai, Kensuke[1], Chang, Haw-Shyang[1], Ohnishi, Wataru[1], Koseki, Takafumi[1], Miyoshi, Shota[2] and Hasegawa, Tomonori[2]

[1]The University of Tokyo, Japan
[2]National Traffic Safety and Environment Laboratory, Japan

Abstract

In railway systems, train localization is an important technological factor; however, maintaining a large amount of ground equipment for localization is costly. The objective of this research is to propose and validate a reliable train localization method that utilizes onboard sensors. The key concept is that compensating for RTK-GNSS latency, which has been less focused on, can be achieved through structure identification using 1D LiDAR sensors. The result of the experiment in a real railway environment shows our proposed method improves the localization accuracy up to 70 % in RMSE.

Keywords: Onboard train localization, RTK-GNSS, IMU, 1D LiDAR sensor

1 Introduction

Accurate train localization is important for preventing collisions and enabling advanced railway technologies. Currently, many railway operators rely on ground equipment such

as track circuits and balises to determine train positions. While these ground-based systems have been used for decades, maintaining such ground equipment has become increasingly challenging due to the high labor requirements. As a result, there is increasing interest in localization methods that rely solely on onboard sensors, reducing dependency on ground equipment [7].

The Global Navigation Satellite System (GNSS) is a widely used technology in vehicle localization research. This system can detect the receiver's position without requiring any additional ground equipment. However, GNSS is affected by its surrounding environment [11]. For example, urban canyons and forests reflecting the satellite signals result in poor receiving conditions for satellite signals. Therefore, many researchers combine GNSS and other sensors, such as inertial measurement unit (IMU) [2, 14] and Light Detection And Ranging (LiDAR) sensor [4, 8, 15]. When combining GNSS and inertial navigation systems, there are two major approaches: loosely-coupled (LC) and tightly-coupled (TC) [3]. LC uses the GNSS receiver's output, while TC uses raw GNSS measurement data for localization. LC is easy to implement and can be used with inexpensive GNSS receivers. TC can estimate position even if there are fewer than four observable satellites; however, localization accuracy depends on the performance of the GNSS receivers. In this paper, we focus on the LC approach because our research centers on developing inexpensive and reliable localization systems.

Recent studies focus on augmentation systems for GNSS. Satellite-Based Augmentation System (SBAS) was first developed for aviation; however, it is widely used for localization, such as railway and car [9]. Real-Time Kinematic (RTK)-GNSS is another augmentation-based approach [11]. RTK-GNSS requires a base station with a fixed and known position, as well as data transmission from the base station to the receiver [10]. Although RTK-GNSS can achieve centimeter-level localization accuracy under open-sky conditions, there is a time delay derived from the data transmission and calculation in the receiver. This may degrade the localization accuracy. Some researchers focuses on the latency of GNSS. The latency is likely to be 100 ms order [5, 6] however, the latency of RTK-GNSS has been less investigated.

For these reasons, we focus on the LC approach and RTK-GNSS. The RTK-GNSS time delay is compensated using multiple 1D LiDAR sensors and the identification of surrounding structures. The basic concept of structure identification using a 1D LiDAR sensor is described in our previous research [12, 13]. The main contribution of this research

is as follows:

- RTK-GNSS time delay is estimated and compensated using multiple 1D LiDAR sensors.
- The usability and robustness of the train localization are improved by RTK-GNSS time delay compensation.
- Experiments were conducted in a real railway environment in several weather conditions.

The outline is as follows: In Section 2, our proposed method is described. In Section 3, the results of experimental validation are presented. In Section 4, conclusions and future research directions are presented.

2 Methodology

2.1 System overview

The sensor placement is shown in Figure 1, and the system structure is shown in Figure 2. We placed a GNSS antenna and receiver, an IMU, and three 1D LiDAR sensors on the railway vehicle. GNSS receiver calculates the position and velocity using messages from satellites and correction data. Correction data is transferred via the Internet. The time delay of internet communication and filtering in the receiver can affect localization accuracy, which is difficult to measure. The key point of this research is to estimate and

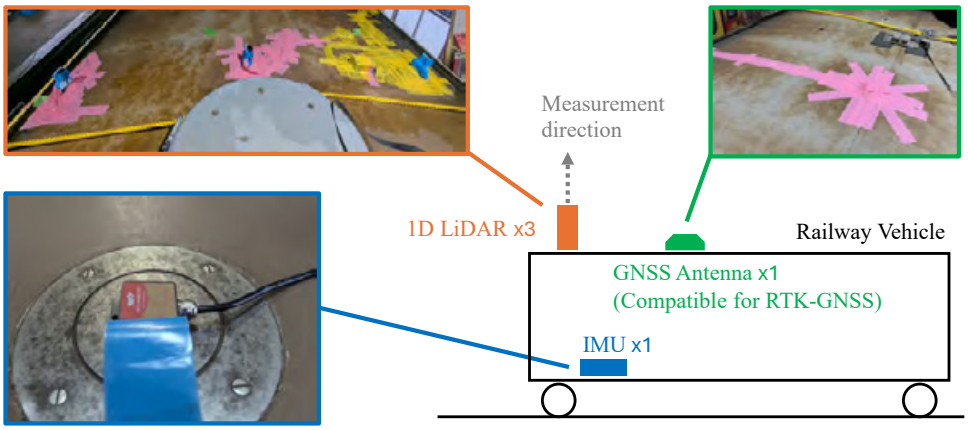


Figure 1: The sensor placement on a railway vehicle.

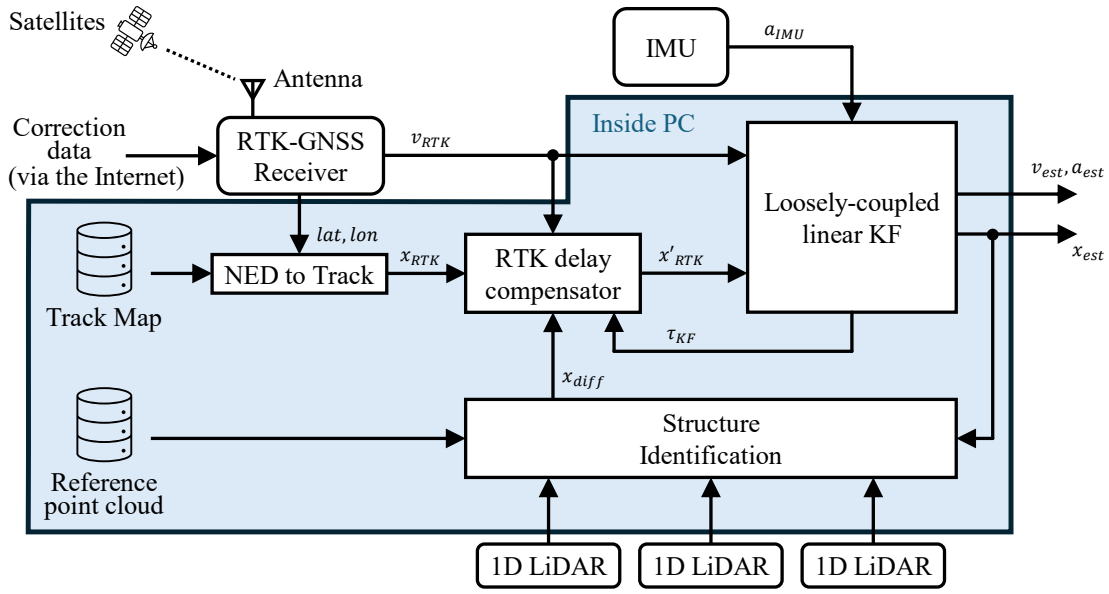


Figure 2: The system structure of the proposed method.

compensate for this time delay by using 1D LiDAR sensors and structure identification. We describe the detail of the proposed method in two subsections. Firstly, in Subsection 2.2, the loosely-coupled linear Kalman Filter (LCKF) is described. This LCKF estimates the position, velocity, and acceleration of the railway vehicle using delay-compensated RTK-GNSS position x'_{RTK}, velocity v_{RTK}, and acceleration a_{IMU} by IMU. Next, the delay compensating method using 1D LiDAR sensors is presented in Subsection 2.3.

2.2 Loosely-coupled Kalman Filter using RTK-GNSS and IMU

In this subsection, an LCKF for continuous train localization is introduced. Before using LCKF, it is necessary to convert the RTK-GNSS measurement data because it is provided in a North-East-Down (NED) frame. The method for converting GNSS measurements to track positions is described in our previous research [12].

The state-space model is shown in Equation 1. The size of the state vector is 3×1. It contains the estimated train position x_{est}, velocity v_{est}, and acceleration a_{est} as shown in

Figure 2. There is no control input u.

$$\begin{bmatrix} \dot{x}_{est} \\ v_{est} \\ a_{est} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{est} \\ v_{est} \\ a_{est} \end{bmatrix}, \begin{bmatrix} x'_{RTK} \\ v_{RTK} \\ a_{IMU} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{est} \\ v_{est} \\ a_{est} \end{bmatrix} \tag{1}$$

$$a_{IMU} = a_{est} \tag{2}$$

x'_{RTK} is the measured train position with time delay compensation. The detail of time delay compensation is described in Subsection 2.3. v_{RTK} is the measured velocity by RTK-GNSS, and a_{IMU} is measured acceleration by IMU. Note that if there is no GNSS measurement data, the observation equation is changed to Equation 2. We convert this continuous system to a discrete system. The time step is denoted as T_s. This value is decided by the measurement frequency of sensors.

The estimation of LCKF also has a time delay because LCKF is a kind of low-pass filter. The approximated time constant is calculated using the eigenvalues of the system. It is denoted as τ_{KF} in this paper.

2.3 Time delay estimation by structure identification using multi 1D LiDAR sensors

In this subsection, the concept of time delay estimation using structure identification is described. The flowchart in Figure 3 shows the concept of time delay estimation and compensation of RTK-GNSS. This algorithm calculates the time delay compensated train position x'_{RTK} using $x_{est}, x_{RTK}, v_{RTK}, \tau_{KF}$, reference point cloud, and the measurement data of three 1D LiDAR sensors. This procedure corresponds to "Structure Identification" and "RTK delay compensator" in Figure 2. 1D LiDAR sensors are placed on the roof of the railway vehicle as shown in Figure 1. Each sensor measures the separation distances between the roof and structures above, such as bridges, tunnels, stations, poles, wires. These separation distances are associated with the estimated train position x_{est} after the train has passed through the structures. Created three point clouds are matched to reference point clouds using Iterative Closest Point (ICP) [1]. By Structure Identification, we can get the train position x_{id}. Using Equation 3, the difference between the estimated position x_{est} and x_{id} is obtained as x_{diff}. The detail of the scan matching process is also described in

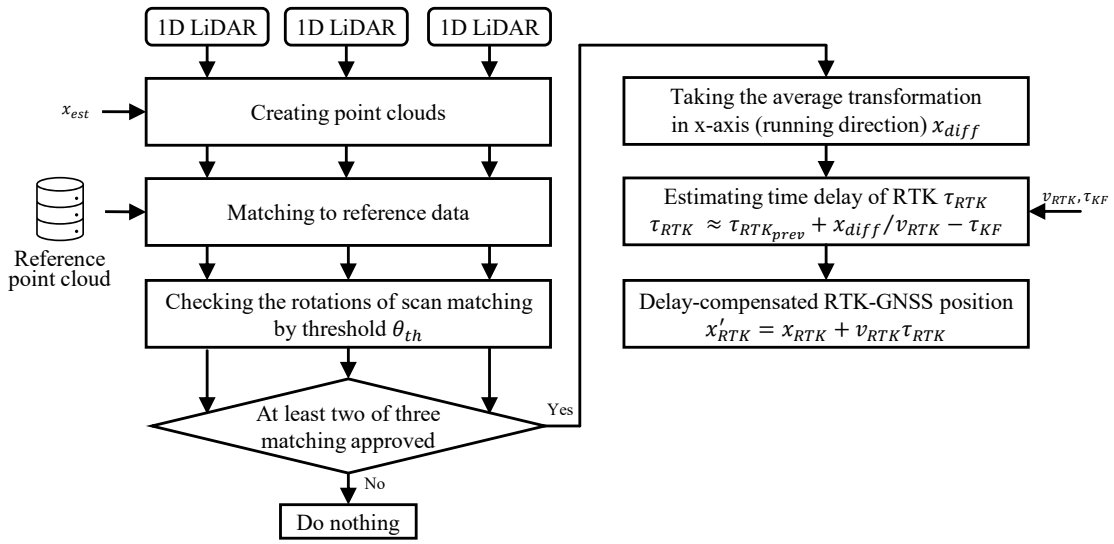


Figure 3: The flowchart of RTK-GNSS time delay estimation and compensation.

our previous research [12].

$$x_{diff} = x_{id} - x_{est} \tag{3}$$

The key point of this paper is estimating and compensating for the time delay of RTK-GNSS τ_{RTK}. We describe the details of the method below. First of all, we assume that the localization error by LCKF is mainly derived from the unknown time delay of RTK-GNSS τ_{RTK} and the known time delay of LCKF τ_{KF}. Then, τ_{RTK} is approximated using previous estimated time delay $\tau_{RTK_{prev}}$ by Equation 4.

$$\tau_{RTK} \approx \tau_{RTK_{prev}} + \frac{x_{diff}}{v_{RTK}} - \tau_{KF} \tag{4}$$

Using τ_{RTK}, the time delay is compensated like Equation 5.

$$x'_{RTK} = x_{RTK} + v_{RTK}\tau_{RTK} \tag{5}$$

In general, the maximum acceleration and deceleration of the railway vehicle are relatively smaller than that of cars and robots, so we ignore the time delay of velocity measurement by RTK-GNSS v_{RTK}.



Figure 4: The test train vehicle of the Yamagata Railway Co., Ltd.

3 Experimental validation

3.1 Experimental setup

The experiment was conducted on a suburban railway line in Japan. Figure 4 shows the test railway vehicle. The test runs 1, 2, 3, and 4 were conducted on the same day, while test runs 5, 6, 7, and 8 were conducted on another day. An RTK-GNSS antenna and receiver, an IMU, and three 1D LiDAR sensors were placed on the vehicle. Figure 1 shows the exact placement of sensors. We also installed a Post-Processed Kinematic (PPK)-GNSS antenna and receiver, as well as a laser Doppler speedometer, for acquiring ground truth. The measurement frequencies of RTK-GNSS and PPK-GNSS were set to 5 Hz, while those of IMU and laser Doppler speedometer were 100 Hz. The measurement frequency of 1D LiDAR sensors was 1 kHz. The length of the test track is approximately 12 km. We conducted eight test runs, resulting in a total running distance of approximately 96 km for this experiment. All of the sensor data were collected using several laptops. We conducted an offline validation using this dataset on a computer equipped with an Intel Core i7-11700 and 32GB RAM.

Out of eight runs, data from two runs are used as reference data. The remaining six runs are for test data. The detailed configuration of eight runs is shown in Table 1. Because of the rainy weather during the first to fourth runs, reference data is created by the seventh and eighth runs. All runs are conducted during the day.

There are ten structures on the test track. Eight of them are thin electric wires, which

Table 1: The detailed configuration of eight runs.

Day	Run number	Direction	Weather	Used for
1	1	inbound	Rainy	Test data
1	2	outbound	Rainy	Test data
1	3	inbound	Rainy	Test data
1	4	outbound	Rainy	Test data
2	5	inbound	Sunny	Test data
2	6	outbound	Sunny	Test data
2	7	inbound	Sunny	Reference data
2	8	outbound	Sunny	Reference data

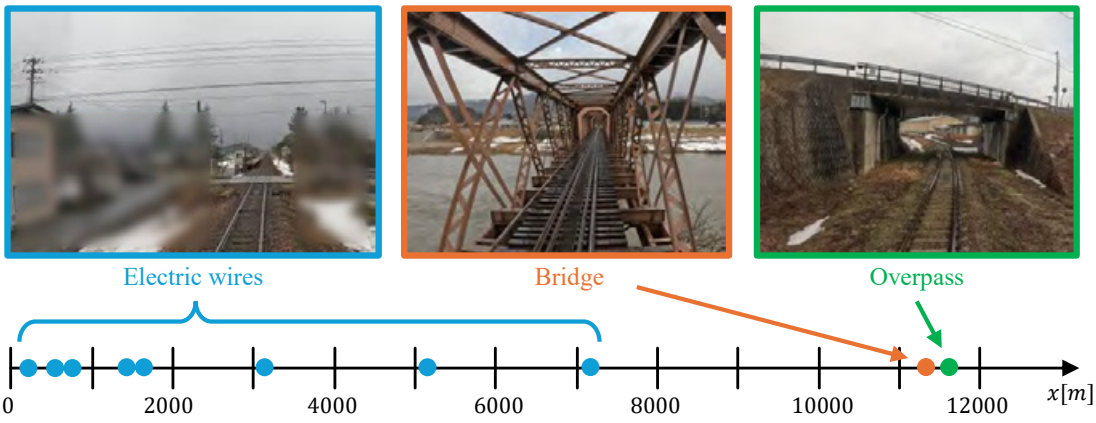


Figure 5: The positions and photographs of structures on the test track.

are under 10 cm in diameter and seem to be hard to detect. Another one is a bridge, and the other one is an overpass. The position of these structures and the appearance of the structures are shown in Figure 5.

3.2 Results and discussion

The entire set of experimental results is summarized in Table 2. "Final τ_{RTK}" means τ_{RTK} at the finishing point of an experiment. Apparently, the RMSEs of runs 1, 3, and 5 are much smaller than those of runs 2, 4, and 6. When running in the inbound direction, there is an overpass and a bridge near the starting point of the experiment. This means RTK-GNSS delay compensation tends to begin from the early timing of the experiment. On the other hand, when running in the outbound direction, there are only tiny electric wires near the starting point of the experiment, as shown in Figure 5. In general, the

scan matching for these tiny structures often fails. In other words, the performance of the proposed method depends on the presence of the structures. We might solve this problem by changing the initial value of τ_{RTK}. For this experiment, we set the initial value of τ_{RTK} to zero. If we set it to a predetermined value based on past runs, the localization performance should improve, even if there is no structure around the starting point of the experiment.

Figure 6 and Figure 7 display the velocity, position error, and velocity error for test runs 1 and 4. The top figure shows the velocity of the ground truth, as well as the velocity estimated with and without RTK-GNSS time delay compensation. The middle figure shows the error of the estimated position. The bottom figure shows the error of the estimated velocity. The vertical dotted lines indicate the timings when RTK-GNSS time delay estimation is executed. In the test run 1, the delay estimation is executed at an early timing. This improves the positioning accuracy, as shown in the middle of Figure 6. In contrast, the delay estimation is executed at a late timing in test run 4. In this situation, the positioning accuracy does not change until the delay estimation is executed. For the same reason, the maximum position errors do not differ whether delay compensation is applied or not.

Table 3 shows the experimental results only after τ_{RTK} is estimated. The differences of the

Table 2: The localization accuracy with/without the proposed method for the whole test tracks.

Run num.	Method	RMSE [m]	Max. error (abs) [m]	Final τ_{RTK} [s]
1	w/o delay comp.	11.5	16.5	-
1	w/ delay comp.	5.09	15.2	0.648
2	w/o delay comp.	11.5	18.5	-
2	w/ delay comp.	11.3	18.5	1.29
3	w/o delay comp.	12.2	18.4	-
3	w/ delay comp.	5.11	18.4	0.736
4	w/o delay comp.	11.3	17.8	-
4	w/ delay comp.	11.1	17.8	0.550
5	w/o delay comp.	10.2	14.9	-
5	w/ delay comp.	3.78	8.33	0.860
6	w/o delay comp.	9.84	14.5	-
6	w/ delay comp.	9.57	14.5	1.31

*num.: number, comp.: compensation

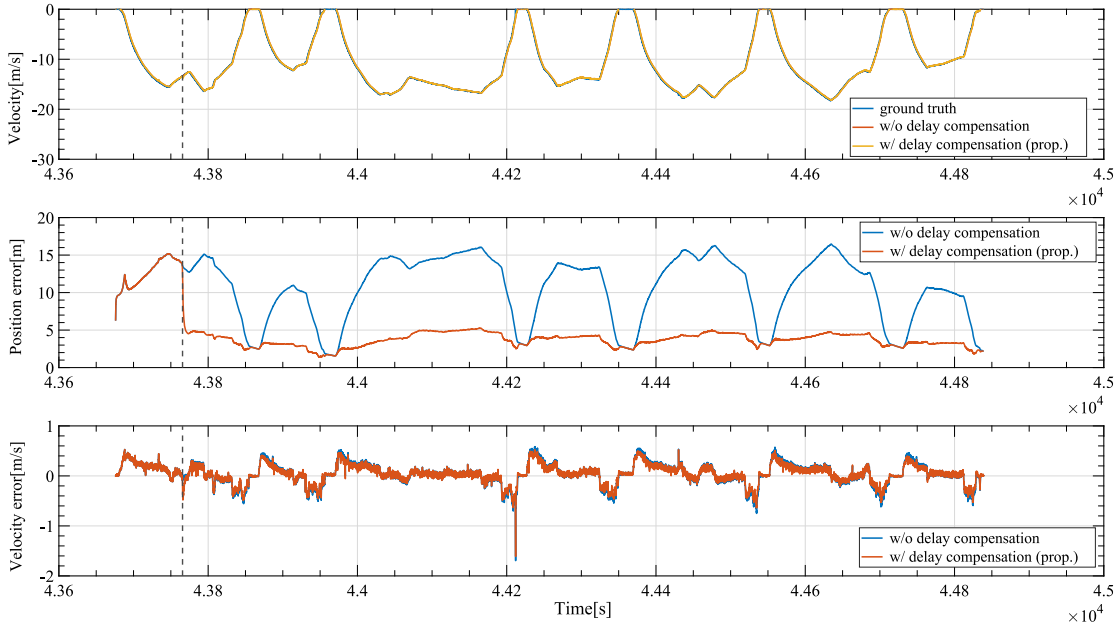


Figure 6: The position and velocity errors during the test run 1 (inbound direction).

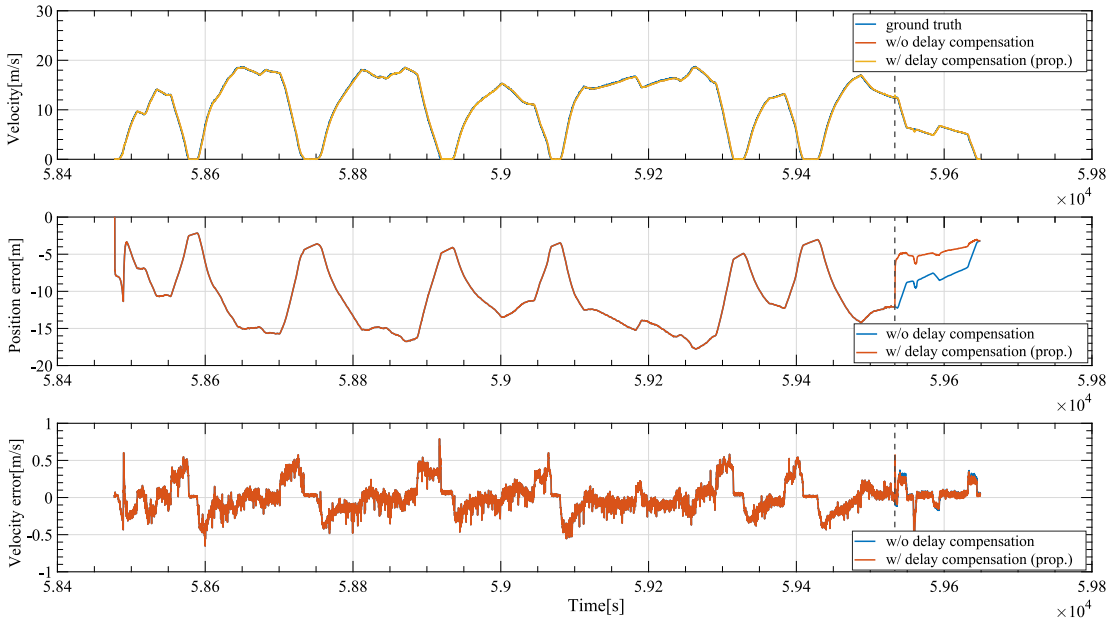


Figure 7: The position and velocity errors during the test run 4 (outbound direction).

RMSEs and maximum errors are directly related to the proposed method. After time delay compensation, the proposed method improves localization performance approximately 40 % to 70 % in RMSE. These results indicate the higher localization accuracy of the proposed method.

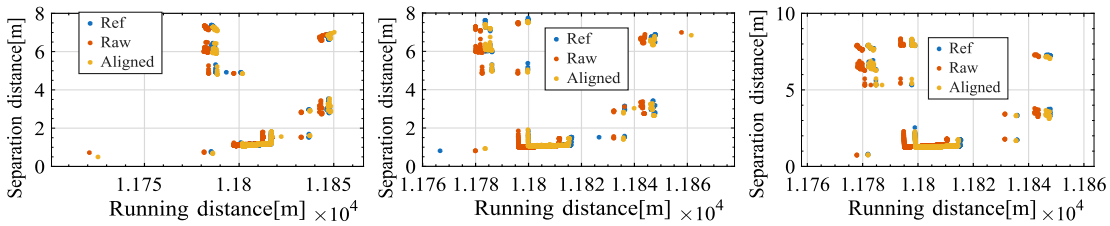


Figure 8: The scan matching result of the overpass using 1D LiDAR sensors.

However, it should be noted that the RMSE includes several factors of error. First, the measured latitude and longitude by GNSS are transformed onto the center line of the track. This procedure requires the detailed data of the center line of the track, and we approximate it by a sequence of points created on Google Maps. The ambiguity of aerial photographic maps degrades the quality of points. A similar problem is also reported in the previous research [11]. Another problem is that we use a PVT solution that GNSS receiver outputs. In general, the positioning calculation in the GNSS receiver contains smoothing. This filtering may cause the latency.

Test runs 1 to 4 were conducted in rainy conditions. According to the experimental results, using multiple 1D LiDAR sensors enhances the robustness of our proposed method. There is no misidentification case and wrong estimation of RTK-GNSS time delay. If we use

Table 3: The localization accuracy with/without the proposed method only after time delay estimation.

Run num.	Method	RMSE [m]	Max. error (abs) [m]	Final τ_{RTK} [s]
1	w/o delay comp.	11.4	16.5	-
1	w/ delay comp.	3.80	13.5	0.648
2	w/o delay comp.	6.61	11.4	-
2	w/ delay comp.	2.23	11.4	1.29
3	w/o delay comp.	12.0	18.0	-
3	w/ delay comp.	3.28	17.4	0.736
4	w/o delay comp.	8.12	12.2	-
4	w/ delay comp.	4.71	12.1	0.550
5	w/o delay comp.	10.5	14.9	-
5	w/ delay comp.	3.69	8.33	0.860
6	w/o delay comp.	8.11	11.3	-
6	w/ delay comp.	3.06	11.2	1.31

*num.: number, comp.: compensation

only one LiDAR sensor, we can hardly get an exact scan matching result [13]. A 1D LiDAR sensor is so cost-effective that it seems to be easy to install multiple 1D LiDAR sensors even in rural railways. Overall, our proposed method improves the usability and robustness of localization.

The estimated time delay τ_{RTK} is similar to test runs 1, 3, 4, and 5, while those of test runs 2 and 6 differ from the others. The time delay depends on the latency of internet communication, calculation in the receiver, communication between the sensor and the laptop, and the accuracy of structural identification. In these cases, this difference seems to be related to the overpass shown in Figure 5. Although the scan matching finished correctly, as shown in Figure 8, the estimated time delay τ_{RTK} increased by approximately 0.6 seconds after identifying this overpass in test runs 2 and 6. Based on the above observations, it is likely that τ_{RTK} increases only near the specific structure. This effect is probably due to signal obstruction or reflection caused by surrounding objects. A straightforward approach would be to exclude time delay estimation in such areas. However, to fully validate the proposed method, further test runs are required, along with a detailed analysis of the variation in τ_{RTK} throughout the entire route. These tasks remain as future work in this study.

4 Conclusion

In this study, we combined inexpensive RTK-GNSS, IMU, and 1D LiDAR sensors to achieve sufficiently reliable localization without the need for ground equipment. Our proposed RTK-GNSS delay compensation method can effectively reduce position errors, even in rainy conditions. The lack of surrounding structures may degrade the performance of the proposed method; however, the accuracy of GNSS improves in open-sky areas. We have to consider this trade-off between GNSS accuracy and the opportunity for Structure Identification. One possible solution is to use pre-determined τ_{RTK} until the first structure identification is executed. It should also be noted that the route selected for this experiment is an ideal environment for GNSS; therefore, applying it directly to other urban railways or subway lines would not be straightforward.

The subsequent objective is to enhance the localization accuracy and robustness by compensating for the communication delay of RTK-GNSS and integrating it with image information, encoders, motor current, and other parameters.

Acknowledgment

We are grateful to the Yamagata Railway Co., Ltd. for their support of the experiment. This work was supported by JSPS KAKENHI Grant Numbers JP24K07521 and JP25KJ1125.

Literature

[1] Besl, Paul. J. ; McKay, Neil D.: A method for registration of 3-D shapes. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (1992), 2, 239-256. http://dx.doi.org/10.1109/34.121791. – DOI 10.1109/34.121791. – ISSN 0162–8828

[2] Chen, Yidi ; Jiang, Wei ; Wang, Jian ; Cai, Baigen ; Liu, Dan ; Ba, Xiaohui ; Yang, Yang: An LSTM-Assisted GNSS/INS Integration System Using IMU Recomputed Error Information for Train Localization. In: *IEEE Transactions on Aerospace and Electronic Systems* 60 (2024), S. 2658–2671. http://dx.doi.org/10.1109/TAES.2023.3328318. – DOI 10.1109/TAES.2023.3328318. – ISSN 1557–9603

[3] Chiang, Kai-Wei ; Chang, Hsiu-Wen ; Li, Yu-Hua ; Tsai, Guang-Je ; Tseng, Chung-Lin ; Tien, Yu-Chi ; Hsu, Pei-Ching: Assessment for INS/GNSS/Odometer/Barometer Integration in Loosely-Coupled and Tightly-Coupled Scheme in a GNSS-Degraded Environment. In: *IEEE Sensors Journal* 20 (2020), S. 3057–3069. http://dx.doi.org/10.1109/JSEN.2019.2954532. – DOI 10.1109/JSEN.2019.2954532. – ISSN 1558–1748

[4] Daoust, Tyler ; Pomerleau, Francois ; Barfoot, Timothy D.: Light at the End of the Tunnel: High-Speed LiDAR-Based Train Localization in Challenging Underground Environments. In: *2016 13th Conference on Computer and Robot Vision (CRV)* (2016), 93-100. http://dx.doi.org/10.1109/CRV.2016.54. – DOI 10.1109/CRV.2016.54. ISBN 978–1–5090–2491–9

[5] Goor, Pieter van ; Wickramasinghe, Punjaya ; Hampsey, Matthew ; Mahony, Robert: Constructive synchronous observer design for inertial navigation with delayed GNSS measurements. In: *European Journal of Control* 80 (2024), S. 101047. http://dx.doi.org/10.1016/j.ejcon.2024.101047. – DOI 10.1016/j.ejcon.2024.101047

[6] Hansen, Jakob M. ; Fossen, Thor I. ; Johansen, Tor A.: Nonlinear observer design for GNSS-aided inertial navigation systems with time-delayed

GNSS measurements. In: *Control Engineering Practice* 60 (2017), S. 39–50. http://dx.doi.org/10.1016/j.conengprac.2016.11.016. – DOI 10.1016/j.conengprac.2016.11.016

[7] HEIRICH, Oliver: Bayesian Train Localization with Particle Filter, Loosely Coupled GNSS, IMU, and a Track Map. In: *Journal of Sensors* 2016 (2016), 2672640. http://dx.doi.org/10.1155/2016/2672640. – DOI 10.1155/2016/2672640. – ISSN 1687–725X

[8] JIANG, Wei ; YU, Yizhi ; ZONG, Kaibin ; CAI, Baigen ; RIZOS, Chris ; WANG, Jian ; LIU, Dan L. ; SHANGGUAN, Wei: A Seamless Train Positioning System Using a Lidar-Aided Hybrid Integration Methodology. In: *IEEE Transactions on Vehicular Technology* 70 (2021), S. 6371–6384. http://dx.doi.org/10.1109/TVT.2021.3080393. – DOI 10.1109/TVT.2021.3080393. – ISSN 1939–9359

[9] LIM, Cheolsoon ; PARK, Byungwoon ; YUN, Youngsun: L1 SFMC SBAS Message for Service Expansion of Multi-Constellation GNSS Support. In: *IEEE Access* 11 (2023), S. 81690–81710. http://dx.doi.org/10.1109/ACCESS.2023.3300580. – DOI 10.1109/ACCESS.2023.3300580. – ISSN 2169–3536

[10] MAYER, Philipp ; MAGNO, Michele ; BERGER, Armin ; BENINI, Luca: RTK-LoRa: High-Precision, Long-Range, and Energy-Efficient Localization for Mobile IoT Devices. In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021), S. 1–11. http://dx.doi.org/10.1109/TIM.2020.3042296. – DOI 10.1109/TIM.2020.3042296. – ISSN 1557–9662

[11] MIKHAYLOV, Denis ; AMATETTI, Carla ; POLONELLI, Tommaso ; MASINA, Enea ; CAMPANA, Riccardo ; BERSZIN, Kai ; MOATTI, Charles ; AMATO, Davide ; VANELLI-CORALLI, Alessandro ; MAGNO, Michele ; BENINI, Luca: Toward the Future Generation of Railway Localization Exploiting RTK and GNSS. In: *IEEE Transactions on Instrumentation and Measurement* 72 (2023), 1-10. http://dx.doi.org/10.1109/TIM.2023.3272048. – DOI 10.1109/TIM.2023.3272048. – ISSN 0018–9456

[12] NAGAI, Kensuke ; CHANG, Haw-Shyang ; OHNISHI, Wataru ; KOSEKI, Takafumi ; SETOGUCHI, Yusuke ; KIYOSAWA, Daichi ; MORITA, Shunji ; TANAKA, Kazuhiro: Onboard Train Localization Assisted by Surrounding Structure Identification Using One-dimensional LiDAR Sensor. In: *36th IEEE Intelligent Vehicles Symposium*, 2025, S. 502–507

[13] NAGAI, Kensuke ; OHNISHI, Wataru ; KOSEKI, Takafumi: Absolute Train Localization Based on the Identification of Surrounding Structures Using 1D LiDAR Sensor. In: *Electronics and Communications in Japan* 107 (2024), 12, e12464. http://dx.doi.org/10.1002/ecj.12464. – DOI 10.1002/ecj.12464. – ISSN 1942–9533

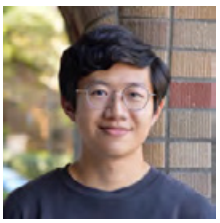
[14] OTEGUI, Jon ; BAHILLO, Alfonso ; LOPETEGI, Iban ; DIEZ, Luis E.: Evaluation of Experimental GNSS and 10-DOF MEMS IMU Measurements for Train Positioning. In: *IEEE Transactions on Instrumentation and Measurement* 68 (2019), 1, 269-279. http://dx.doi.org/10.1109/TIM.2018.2838799. – DOI 10.1109/TIM.2018.2838799. – ISSN 0018–9456

[15] WANG, Yusheng ; LOU, Yidong ; SONG, Weiwei ; TU, Zhiyong ; WANG, Yapeng ; ZHANG, Shimin: Simultaneous Localization of Rail Vehicles and Mapping of Surroundings With LiDAR-Inertial-GNSS Integration. In: *IEEE Sensors Journal* 22 (2022), S. 14501–14512. http://dx.doi.org/10.1109/JSEN.2022.3181264. – DOI 10.1109/JSEN.2022.3181264

Authors



Nagai, Kensuke
Master of Engineering
Ph.D. student at the University of Tokyo
JSPS Research Fellowship for Young Scientists (DC1)
Contact: k.nagai@ctl.t.u-tokyo.ac.jp



Chang, Haw-Shyang
Bachelor of Science
Master's student at the University of Tokyo
Contact: h.chang@ctl.t.u-tokyo.ac.jp



Ohnishi, Wataru
Ph.D. in Engineering
Associate professor at the University of Tokyo
Contact: ohnishi@ctl.t.u-tokyo.ac.jp



Koseki, Takafumi
Ph.D. in Engineering
Professor at the University of Tokyo
Contact: koseki@ctl.t.u-tokyo.ac.jp



Miyoshi, Shota
Ph.D. in Engineering
Researcher at the National Traffic Safety and Environment Laboratory
Contact: s-miyoshi@ntsel.go.jp



Hasegawa, Tomonori
Ph.D. in Engineering
Executive researcher at the National Traffic Safety and Environment Laboratory
Contact: hasegawa@ntsel.go.jp

Session 16
Jens Koenig, Thomas Mueller, Rickard Persson, Libor Lochman, Pauline Guicheney

Small, Lightweight, Driverless Rail Vehicle for Rural Areas in Europe

Rolling Stock

Koenig, Jens[1]; Mueller, Thomas[1]; Persson, Rickard[2]; Lochman, Libor[3]; Guicheney, Pauline[4]

[1]German Aerospace Centre (DLR), Institute of Vehicle Concepts, Germany
[2]KTH Royal Institute of Technology, Rail Vehicles, Sweden
[3]Wabtec, Italy
[4]SNCF, France

Summary

With the support of EU's key funding program Horizon Europe, the Europe's Rail Joint Undertaking (EU-Rail) aims to deliver faster deployment of innovations in the railway design. The EU-Rail FutuRe project aims at significantly reducing the Total Cost of Ownership (TCO) of European regional lines by means of providing new innovative solutions to make services on these lines competitive and attractive for customers. Setting the right level of requirements and specification for the regional rail rolling stock plays therefore an important role in meeting the FutuRe objectives. The ultimate objective is to define a concept of a driverless (fully automated) vehicle for up to 100 passengers or corresponding freight with significant vehicle mass savings compared to existing rolling stock. The key research areas are: the mechanical arrangement avoiding unnecessary weight and reducing maintenance costs on vehicle and track, the emission free propulsion system designed for services on non-electrified lines and the control-command system enabling the highest grade of automation – driverless operation.

Keywords: Regional railway, rolling stock, low weight, emission free, automation, cost of ownership.

1 Overall Approach

As the ecological transition becomes a global priority, regional mobility is experiencing renewed strategic interest. Too often financially struggling or even threatened of being shut down, regional railway lines have been suffering for several decades [1]. However, they could soon become an essential pillar of more sustainable, inclusive and efficient mobility. In order to ensure this transformation, trains designed specifically for these short and demanding journeys, low passenger numbers and often ageing infrastructure, are needed. With this goal in mind, a new generation of vehicles is emerging, driven by innovative projects such as Europe's Rail Joint Undertaking (EU-Rail) FutuRe, which aims to radically transform the current offering.

The industry is also facing other challenges such as little to no accessibility due to a lack of accommodations for people with reduced mobility, prams or wheelchairs, a lack of electrification while simultaneously seeking to phase out diesel trains, a steady increase in the price per seat over the last 40 years due to the combined effects of improved on-board comfort, the introduction of safety equipment and inflation, making it difficult to produce affordable trains [2], even crash safety rules for ultra-light vehicles not covered by the current Technical Specifications for Interoperability (TSIs).

This list illustrates all the main technical, regulatory and economic constraints that new regional rolling stock must meet. However, the specific needs of regional lines have already been addressed in Europe in the past. It would also be ambitious to embark on a fundamental review of regional rolling stock without considering the experience we have gained over more than half a century.

1.1 Lessons from the past: strengths and weaknesses of the regional rolling stock (1950 – 2000)

The first vehicles designed for regional lines were light, simple and inexpensive. First generation models such as the Uerdinger Schienenbus (VT95/VT98), shown in Figure 1, helped to maintain rail links in low-traffic rural areas when automobiles were not massively owned. This kind of rolling stock were affordable to produce and operate, and became common in Europe.

After several serious accidents involving railbuses and heavy trains, the development of new rolling stock with proper safety standards was initiated, which resulted in increased longitudinal strength of the vehicle bodies and, consequently, a higher overall vehicle mass.



Figure 1: Uerdinger Schienenbus (VT98) / Wikipedia Commons

What could be described as the second and third generations of regional rolling stock were developed with bogies replacing axles and continuous improvements in passenger comfort (notably with the addition of air conditioning). The X2100 (France) and X73500 (France/Germany), shown in Figure 2, are examples of these successive developments.



Figure 2: X2100 (left) and X73500 (right) / Wikipedia Commons

These vehicles, some of which are still in service, are not viable solutions for the future because, in addition to their advancing age, they have design flaws and do not longer meet current standards.

For example, those trains are commonly diesel powered and would require complete refurbishment and even engine retrofitting, which significantly increases costs that must be weighed against the purchase of new equipment at a controlled cost. They are often difficult to use for people with reduced mobility and for people with wheelchairs, prams or bicycles. Older models are also often ill-equipped to handle increasingly hot summers.

Taking these needs and shortcomings into account, along with market research conducted in Germany and Poland [2], other solutions are emerging.

1.2 Plans for the future

At the time of writing, no manufacturer in Europe offers a ready-made solution specifically designed to meet the needs of local lines. Examples of smaller available vehicles include the Alstom Coradia Lint (approximately 150 seats) and the Stadler RS Zero (around 75 seats). While these sizes fully meet the needs of long distance regional or interregional operations, 'oversupply' is not a satisfactory solution for smaller lines due to the purchase and operating costs of these trains (higher price, higher energy consumption, infrastructure wear and tear due to the weight, etc.).

But across Europe, several projects are underway to offer right-sized and affordable solutions with no compromise on safety nor comfort.

These projects share various goals, such as replacing diesel with zero-emission technologies without the need for a massive infrastructure electrification, lightweight body structures to adapt to low-traffic lines and ageing infrastructures. For this, fully accessible trains are developed to allow inclusive and active mobility. Also, different levels of automation are explored to reduce operating costs and offer a better service while reducing investment costs through standardisation and modularity.



Figure 3: DRAISY / SNCF & Lohr [3] (left), TELLi / SNCF & Yellow Window [4] (right)

For example, in France, the DRAISY [3] and TELLi [4] projects, shown in Figure 3, are two initiatives led by SNCF in the form of consortia involving several partners. DRAISY is a new-generation train capable of carrying up to 80 passengers (including 30 seats). Measuring 14 meters long, weighing 20 tons and powered by battery, it is part of a specially redesigned railway system - from vehicle to track to operating system - to address low-traffic lines and offer reduced investment and operating costs. TELLi, meanwhile, is a two-car train set concept (30 metres long, 74 seats and capacity for up to 150 passengers), battery and pantograph powered, aiming to offer economic operation

and improved performance, thanks in particular to the signalling system developed alongside the train.

Additional initiatives have been developed to revitalize regional railway services, including the NGT-Taxi by DLR [5] (Germany), the Revolution VLR [6] and Coventry VLR [7] (UK), and the Aachen Rail Shuttle [8] (Germany).

Across Europe, there is growing recognition of the specific needs of regional lines, and EU-Rail FutuRe has been launched to unify approaches and bring together countries, infrastructure managers, manufacturers and research institutes. It aims to develop a joint proposal for rolling stock that meets the needs of G1 (connected to the main network). For G2 (disconnected from the main network) lines, a reduced approach is in development.

2 Rolling Stock

The main focus of research for this new rolling stock is to combine low total cost of ownership (TCO) with sustainability, accessibility, comfort and flexibility. This requires a balance using a small, lightweight and simple design while still offering enough space and complexity to fulfil a variety of different requirements for partially diverging use cases.

As the trains are intended to run automatically without a driver on rural lines with rather low passenger counts, smaller vehicles offer low operation costs and high flexibility. Therefore, a single-unit car equipped with two single axle running gears with a 16 t maximum axle load is the simplest option. The comparably low axle weight reduces wear on the track and also the effort to revitalise decommissioned and possibly damaged tracks. Also, it allows the train to run on all line categories. As a high level of flexibility and adaptability to different use cases is mandatory, the vehicle features a modular architecture. Subsystems like the car body, energy system, the Heating, Ventilation, and Air Conditioning (HVAC) system and the interior are adaptable to different lines, user groups, climate zones, etc. enabling the vehicle to be used in a wide variety of scenarios while minimizing the number of design changes for each scenario. This reduces development and manufacturing costs per unit and increases usability and attractiveness for both users and operators.

To find solutions for both main and branch lines, two different vehicle configurations are investigated:

1. The version mainly used on mainlines (G1) is primarily tailored to longer lines in remote, rural areas with lower population density for example in Sweden. It is 16.5 m long, has a tare weight of 21.5 t and a maximum weight of 32 t using the maximum allowed axle weight. The maximum seat capacity is dependent on the chosen setup and can reach 54 seats, which is two-thirds the amount of the next bigger train currently available.
2. A shorter version exclusively used on tracks separated from the main lines (G2). Possible use cases for this vehicle are shorter lines as they are found plentiful in Germany [9]. The train is 11.3 m long, has a tare weight of 15 t and a maximum weight of 19.5 t. The maximum seat capacity can reach 24 seats.

2.1 Car body

Research on structural components focuses on enhancing both sustainability and economic viability. This is necessary to reduce the environmental impact in a way that is easily applicable in an area with high financial pressure. But to achieve this goal, a compromise has to be found between low development and manufacturing cost and low environmental impact, specifically low energy consumption in manufacturing, low weight for low energy consumption during operation and good recyclability. To find a solution which fulfils these partially contrasting requirements, two different building styles are evaluated against each other:

- A welded car body using steel pre-products like beams and plates.
- A composite sandwich car body with a high level of function integration.

Both options offer different advantages in TCO. Steel welded structures tend to be easier to manufacture, maintain and dispose of leading to lower acquisition and maintenance cost. Composite structures tend to be lighter, decreasing energy consumption and therefore operation cost. Concerning their environmental impact, both building styles also have different advantages. The lower weight of a composite car body reduces CO_2 emissions during operation, while steel is easier to recycle. [10], [11]

Initially, it is unclear, which building style offers the best suitable compromise. To inform a decision based on solid evidence, both options are developed to a state, where they can be compared in detail. For testing and demonstration purposes, the modular architecture enables a mix of building styles using steel welded structures in the front modules and composite structures in the middle sections.

One the one hand, a steel car body is developed as shown in Figure 4. It consists of laser cut and bent sheets as well as various profiles. Possible materials are different steel alloys, varying in corrosion resistance, processability and strength. To balance cost, weight and longevity, a detailed analysis is planned to be conducted during the detailed design phase.

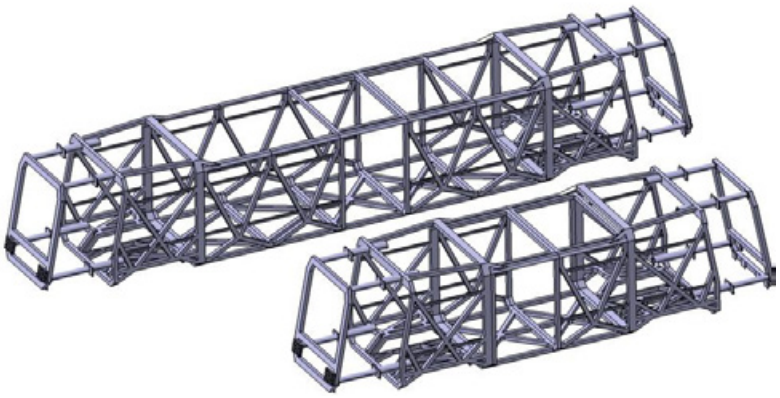


Figure 4: Long and short car body configurations in steel construction

Since the car body has to satisfy different vehicle lengths, it is also designed to be modular using segmentation into three different modules:

1. Two front modules including the running gear, vehicle front and crash systems
2. Two door modules equipped with one door on each side
3. Up to three additional modules to increase interior space

These modules can be configured differently during conceptualization. However, during manufacturing the car body is built as one inseparable unit. The modular architecture reduces the number of different parts, decreasing development, manufacturing and certification cost while still offering flexibility to adapt to specific use cases.

To utilize the material to its full potential as well as to minimize the amount of material used, a topology optimization is conducted. This optimization process uses a static linear stress analysis on a finite element model of the car body loaded with masses and forces specified in the EN 12663-1 [12] to iteratively vary the relative material density on each element with respect to the local stress in this element. As a result, in this case a car body with the ideal material distribution is created. The result is shown in Figure 5 with high material density coloured red and low material density coloured blue. The car body generated by the topology optimization cannot yet be easily manufactured as the optimization algorithm creates complex, organic looking structures difficult to realise with common materials and tools. They have to be reengineered into a simplified construction optimized for an appropriate and efficient manufacturing process.

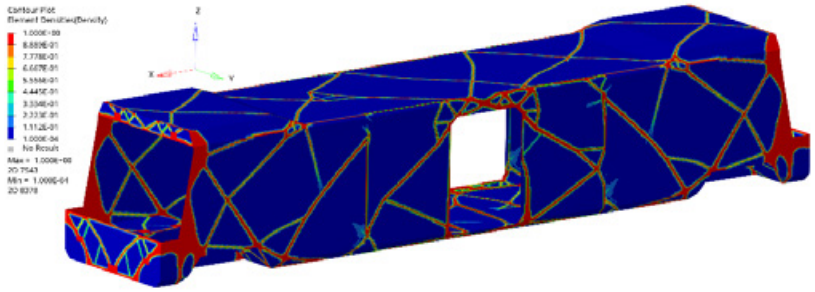


Figure 5: Topology optimization result of the car body

The sandwich car body on the other hand is built using sandwich panels. These panels consist of two thin, outer sheets out of a high-strength material like aluminium or fibre reinforced plastics and an inner low-density core material like aluminium honey comb or foam. This setup provides a high area moment of inertia which leads to high structural integrity and low weight. Additionally, sandwich panels offer the ability for function integration like sound and thermal insulation. The resulting lower weight of the car body decreases energy consumption during operation reducing operating cost and greenhouse gas emissions. As the operation phase contributes between 96.7 and 99.8% to the environmental impact [10], increased cost in material acquisition and manufacturing could be compensated by the lower operation cost. A concept for the composite sandwich car body is shown in Figure 6.

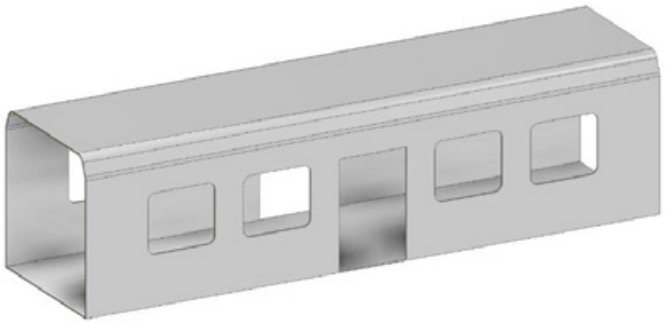


Figure 6: Conceptual design of an all-sandwich composite car body structure

2.2 Running Gear

As the vehicle is expected to operate on rural tracks that may be in degraded condition, the running gear must provide good ride comfort and the ability to negotiate tight curves. In addition, it should feature low acquisition and maintenance costs, as well as low weight to reduce traction energy consumption and track wear. To identify the optimal solution, two different single-axle running gear concepts are being researched and will be compared once their development is sufficiently advanced for detailed evaluation.

The first proposed running gear concept is a derivation of the Next Generation Train (NGT) running gear developed by DLR. This running gear is intended to reduce total mass and wear while accommodating the constraints of a compact design, employing independently rotating wheels powered by an active torque vectoring system [13], [14]. The running gear concept for the FutuRe vehicle follows the same general design, but is tailored for lower speeds and reduced cost. A state-of-the-art low floor portal axle reduces certification and acquisition effort as they are widely used e.g. in tram trains. Additionally, the absence of an axle allows a low-level floor throughout the whole vehicle without steps or ramps for maximum accessibility. Every wheel is driven independently by a permanent magnet synchronous motor with an external runner providing high torque density to obviate the need for a gear box. Using a feedback control system with sensors on the axle and individual controllers on each wheel motor, the running gear can be precisely steered which reduces wear and noise. To reduce the steering effort for the motors, the axle is suspended using a Watt's linkage. This is a lever system, that is designed to move the middle point, in this case the axle on a nearly straight, horizontal path, which results in low resistance against the steering motion. The motors and friction brakes are suspended in the running gear frame for a lower unsprung mass and less stress on the motors which leads to higher longevity and better driving dynamics. Torque transmission from the motors to the wheels is realized using a cardan shaft. An additional magnetic track brake increases braking performance for emergencies. For service braking, the motors are used in energy recuperation mode. In conclusion, this running gear concept offers a highly compact and lightweight design with few parts promising low wear, low noise, and low certification effort.

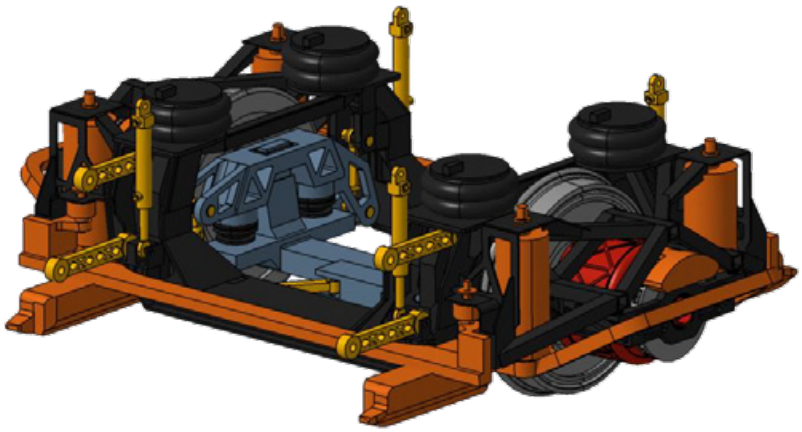


Figure 7: Single wheel running gear concept

The second proposed running gear concept is a single-axle conventional wheelset running gear. The starting point for the design is a wheelset paired with a coil spring suspension

with high spring rates to limit displacements due to load and vehicle dynamics. This setup is similar to the design used for freight waggons, Figure 8 (left). Besides its simplicity, it also supports designing a vehicle without a compressed air system. The suspension will be soft enough to ensure safe running but with deficiencies when it comes to ride comfort and curving performance.

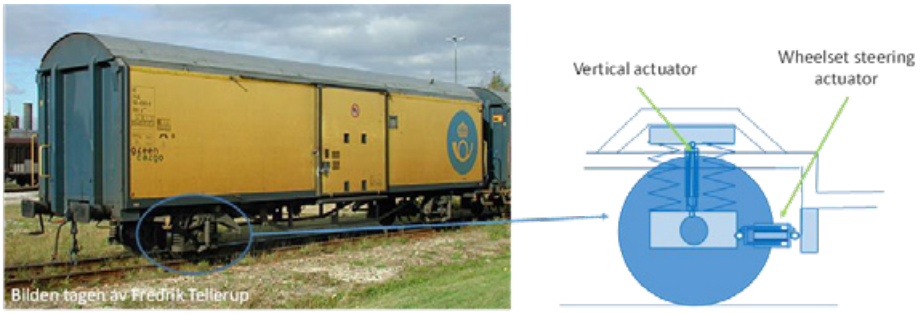


Figure 8: Freight waggon with single axle running gear (left), active suspension elements (right)

To address the identified deficiencies, two types of active suspension elements are proposed. A low-force, active vertical suspension based on electro-dynamic principles will be installed in parallel with two coil springs for each axle box. In terms of the freight wagon design, this means that an actuator will replace the existing hydraulic vertical damper, as shown in Figure 8 (right). The four actuators on each vehicle will be capable of dynamically controlling the car body's bounce, pitch, and roll motions. To limit the force requirement for the actuator, control of quasi-static motions is excluded. The other foreseen active suspension is an active wheelset steering, where a longitudinal actuator between axle box and body will control the wheelset to a radial position in curves. With reference to the freight waggon design, this means that an actuator replaces the longitudinal link between axle box and body. As the actuator replaces the longitudinal link, it can steer the wheelset to the radial position with low force and that only one actuator per wheelset is needed. The setup for the vehicle is shown in Figure 9, similar setups as have been proofed in on-track tests by [15] and [16], but with hydraulic actuators in a bogie running gear.

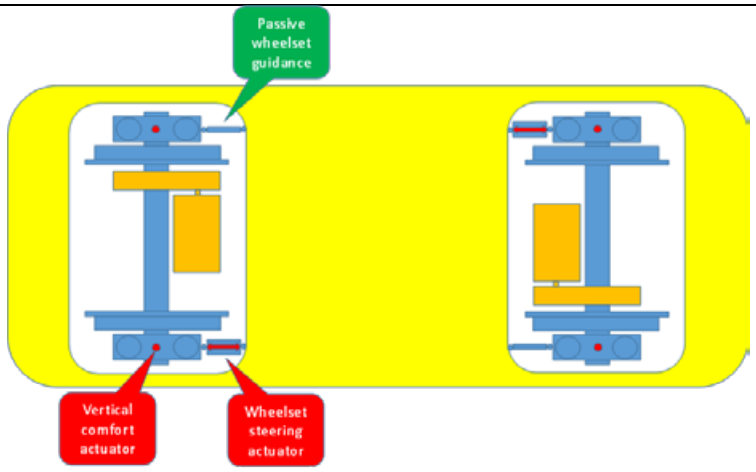


Figure 9: Running gear setup (top view)

A multibody simulation (MBS) model of the vehicle, with its two different running gears, has been developed to optimize performance and validate the feasibility of the proposed approach. These simulations are also used to define the requirements for the actuators, particularly in terms of force capacity and bandwidth. The effectiveness of active wheelset steering, measured as wear number of the leading wheelset, is shown in Figure 10.

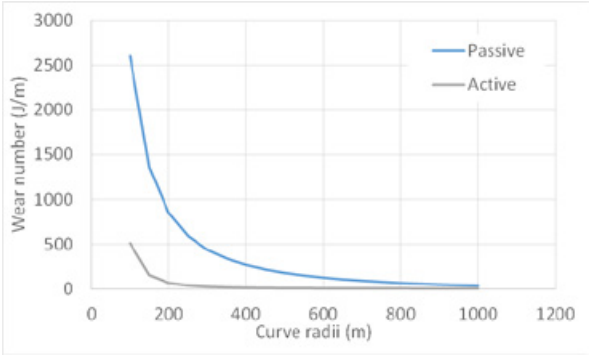


Figure 10: Wear number comparison for the leading wheelset with passive and active wheelset steering

Wear numbers higher than about 200 J/m are known to cause significant wear on rails and wheels, whereas very low wear numbers increase the risk of rolling contact fatigue (RCF). For longest wheel life, there should be some wear to limit RCF crack initiation. Hence, optimal wheelset steering depends on the several factors, where the curve radii distribution is the most important one. Regional lines have often curves with smaller curve radii than mainlines and the vehicle must be able to negotiate these without too much wear. For the active wheelset steering, Figure 10 therefore shows the minimum

achievable wear number. The system is capable of reducing wear across all curve radii, but it can also be tuned to allow a controlled amount of wear to mitigate the risk of RCF.

2.3 Traction and Energy System

Depending on the line length and gradient, different energy storage solutions are possible utilizing a modular equipment concept on the roof of the vehicle. The main power storage system is a battery. Depending on the use case, different cell chemistries are considered to find a balance between range, longevity, sustainability, cost and weight. For longer lines, a hydrogen fuel cell can be added to act as a range extender which is especially feasible for the bigger vehicle versions. In this case, the battery handles the higher power fluctuations occurring during acceleration or regenerative braking while the fuel cell is used to recharge the battery. This allows the fuel cell to run with constant output power at its optimum efficiency. Different combinations of batteries and fuel cells can be optimised for the best compromise in range, weight, cost and sustainability. Different lines and their respective battery and fuel cell capacity distribution can be seen in Figure 11.

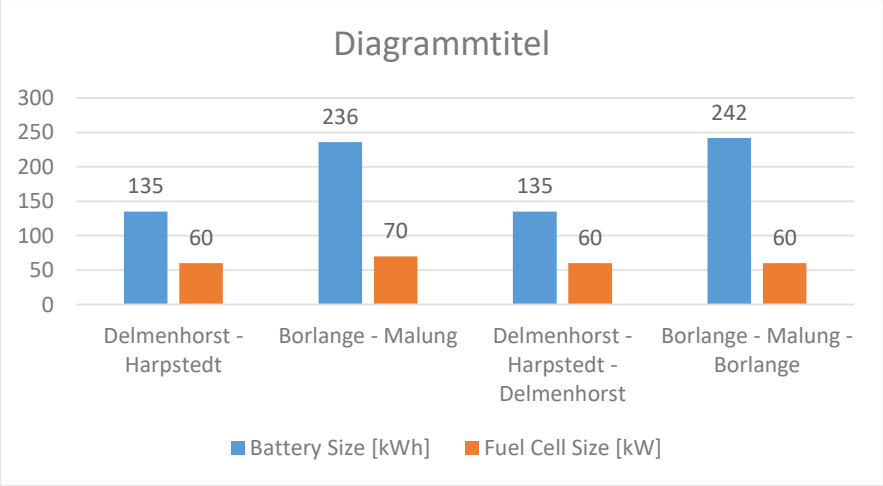


Figure 11: Fuel Cell and battery power distribution for different lines [2]

2.4 Interior

The interior is also tailored to maximize flexibility and adaptability. Therefore, components like seats, desks, hand rails, luggage and bike racks are mounted on C-rails traversing the vehicle. As the components are swappable and removable, the operator can easily adapt the interior to different user groups such as commuters or tourists, to different operation areas including bigger cities or more isolated regions and to different operation

times during working days or weekends or even during day and night. Among others, the following three main setups are considered:

- Commuter setup, offering office spaces with desks, privacy areas and also open spaces for standing people, wheelchairs, prams and bicycles.
- Tourism setup, accommodating a higher amount of comfortable seating, additional luggage racks, space for bicycles, vending machines and possibly a toilet. A version of this setup is shown in Figure 12 as an example.
- Cargo setup, adapted for transporting different goods, freight, mail, etc. The interior space of this setup is completely emptied out making room for up to 23 euro-pallets being loaded using a pellet truck from the station platform or even a tail lift from a truck. The C-rails act as cargo mounting points. Whenever the user demand is low and vehicles are underutilized, this setup can offer additional usage, increasing utilization and consequently revenue.

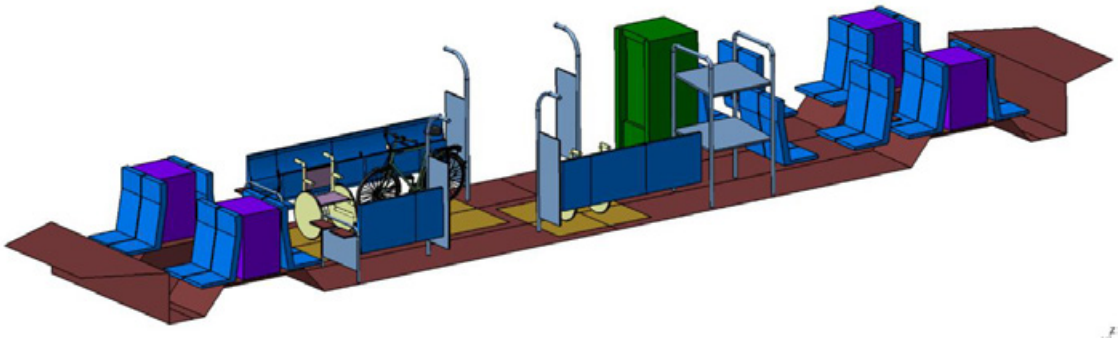


Figure 12: Vehicle interior with a tourism setup as example

As the vehicle can be configured to be completely empty, it offers space for additional use-cases to bring various services to remote villages in rural areas. A mobile kiosk setup with vending machines and mail stations could replace stationary kiosks or post offices that cannot be run competitively. A rolling doctor's office could counteract the scarcity of health care facilities in rural areas.

To maximize accessibility especially for people with reduced mobility, the vehicle has a low floor design tailored a platform height of 550 mm. The doors are offset to each other, so that the accommodation of prams, wheelchairs and bicycles are located directly opposite of the doors, obviating the need to navigate tight corners. Depending on the setup, the vehicle can be equipped with a large and flexible multi-use area surrounded by folding seats with integrated bike racks and fastening belts to secure luggage, prams and wheelchairs.

3 Control-Command and Signalling

Both vehicles, the small version for G2 lines and the big version for G1 lines are to be operated autonomously using advanced Automatic Train Operation (ATO) technology. As no driver is needed, this reduces operation cost and enables more flexible use cases like on demand service. While G1 vehicle is fully compliant with Control-Command and Signalling (CCS) TSI to ensure compatibility with mainline services, the G2 vehicle uses a different approach to further reduce TCO.

The CCS low density lines architecture being developed within the FutuRe project is based on four main inter-related building blocks which can be decomposed from an Operational, Functional and System point of view. The building blocks are Communication data network, On-board equipment, Wayside equipment, Traffic Management and Control Centre – see Figure 13.

An initial implementation is intended on G2 lines, where the compliance with TSIs is not obligatory and thus approvals and authorisations are fully in hands of respective National Safety Authority

That will allow demonstrating significant cost reduction and system efficiency gains achieved thanks to two principal factors: drastic reduction of trackside equipment and utilisation of industrial, COTS components instead of tailor-made rail specific ones.

The individual elements, building blocks or subsystems may then find its way into the G1 CCS solutions, resulting in progressive improvement of the regional rail competitiveness.

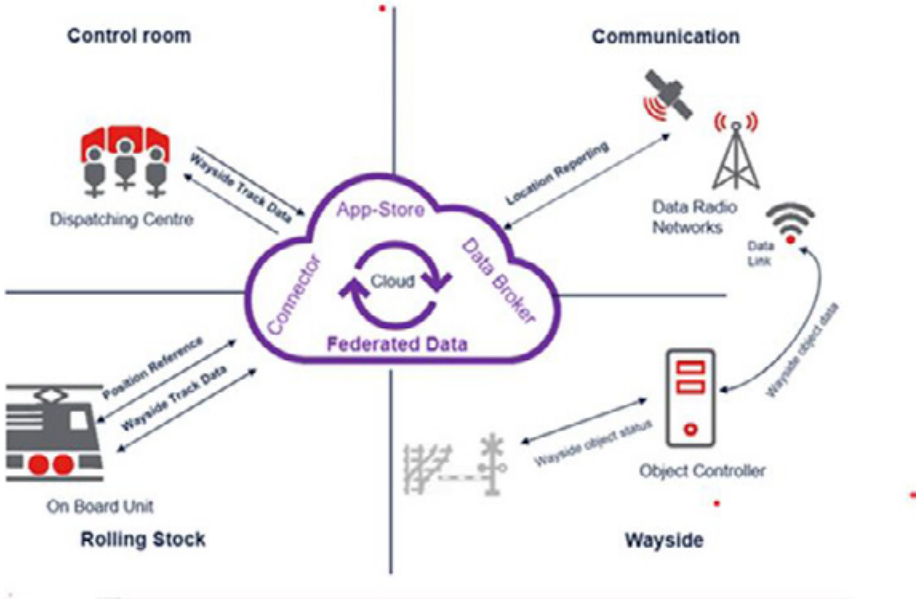


Figure 13: Low density lines high level architecture

3.1 Onboard assembly

On-board segment of the CCS low density lines architecture uses the well-known concept of Communication Based Train Control (CBTC), with which positioning based on radio communication technologies can now be integrated. Modern train control solutions are generally radio based (such as various CBTC systems for automatic metros, railways using ERTMS L2/L3 and Positive Train Control). Information among traffic control rooms, trains and wayside elements is exchanged wirelessly.

In the proposed concept, radio and satellites provide (augmented for high precision, where required) train positioning data to the on-board unit, which can use this to index an onboard digital track map (resulting from a standardised validation process), thus removing the need for wayside balises/transponders along the track. For safety reasons, to eliminate the need for high-cost track circuits and axle counters, the system requires the on-board unit to determine train length and train integrity.

Positioning technology performance is mature, currently to a precision of a few centimetres using satellites in conjunction with publicly available correction services. Soon, 6G & 7G will make localization cheaper and easier to implement (integrating cellular with satellite technology).

Such technologies will evolve more quickly than railway technology, and the CCS architecture must allow for this evolution, separately from the railway applications.

The CCS low density lines approach will enable the implementation of cost-effective, interoperable, fail-safe, highly accurate train positioning (including dead reckoning where required) based on data fusion of information from hybrid, multi sensor technologies, digital maps, onboard database.

This will not only increase operational efficiency and decrease TCO in the context of regional lines. In addition, it will allow to adopt smoothly higher grades of Automatic Train Operation, enabling driverless services on rural lines.

4 Conclusion

The EU-Rail's FutuRe project develops key technologies to increase attractiveness, accessibility and utilization of rolling stock on rural lines while simultaneously decreasing TCO and environmental impact. Its focus of research is a small, lightweight, locally emission-free, driverless vehicle concept tailored to enhance economic and ecological viability on main and branch lines in areas with lower passenger volume. Due to its modular design, a wide variety of different use cases can be covered. Its focus on a locally emission free propulsion system, low wear running gear, lightweight design and a high level of automatization provides a safe and comfortable mode of transport with a minimum of cost for the acquisition and operation of the vehicle and the maintenance of the track.

One advantage of the FutuRe project is its diverse member group. The different project partners all contribute their unique perspective as manufactures, operators, suppliers, researchers and traffic associations into the development. Also, the setup of the project allows for multiple approaches to be developed and compared in detail. In summary, the project provides the ideal basis to research and develop different approaches and inform decisions for comprehensive beneficial solutions.

Acknowledgements

This project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation program under grant agreement (No. 101101962). The content of this paper reflects only the author's view and the JU is not responsible for any use that may be made of the information it contains.

Sources

[1] F. Philizot, „Petites lignes ferroviaires, Des plans d'actions régionaux," 2020. [Online]. Available: https://www.ecologie.gouv.fr/sites/default/files/documents/20200220_JBD_DP_Petites_lignes_vf.pdf. [Accessed on 30.06.2025].

[2] EU Rail FP6 FutuRe, „Deliverable D5.1: State-of-the-Art Report for Regional Lines Rolling Stock," 18 June 2025. [Online]. Available: https://rail-research.europa.eu/wp-content/uploads/2025/10/FP6_D5.1_FINAL-1.pdf. [Accessed on 28.10.2025].

[3] „DRAISY: le train sur-mesure pour les plus petites lignes," SNCF, 30.05.2025 [Online]. Available: https://www.groupe-sncf.com/fr/innovation/mobilite-territoires/draisy. [Accessed on 10.11.2025].

[4] „Revitalizing rural rail lines with TELLi," SNCF, 20.01.2025 [Online]. Available: https://www.groupe-sncf.com/en/innovation/local-mobility/telli. [Accessed on 10.11.2025].

[5] DLR, „NGT TAXI," DLR, [Online]. Available: https://www.dlr.de/de/fk/forschung-transfer/projekte/innovative-fahrzeugkonzepte/ngt-taxi. [Accessed on 11.11.2025].

[6] „Revolution VLR," [Online]. Available: https://revolutionvlr.com/. [Accessed on 11.11.2025].

[7] „Coventry Very Light Rail," Coventry City Council, [Online]. Available: https://www.coventry.gov.uk/coventry-light-rail. [Accessed on 11.11.2025].

[8] „Der Aachner Rail Shuttle," Institut für Schienenfahrzeuge und Transportsysteme der RWTH Aachen, [Online]. Available: https://aachener-rail-shuttle.de/. [Accessed on 11.11.2025].

[9] VDV, Auf der Agenda, Reaktivierung von Eisenbahnstrecken, Bd. 4, Verband Deutscher Verkehrsunternehmen e.V. (VDV) (Hrsg.), 2024.

[10] A. Özdemir und A. Önder, „An environmental life cycle comparison of various sandwich composite panels for railway passenger vehicle applications," *Environmental Science and Pollution Research,* Nr. 27, p. 45076–45094, 10 August 2020.

[11] R. A. Witik, J. Payet, V. Michaud, C. Ludwig und J.-A. E. Månson, „Assessing the life cycle costs and environmental performance of lightweight materials in automobile applications," *Composites Part A: Applied Science and Manufacturing,* Nr. 42, p. 1694–1709, 2011.

[12] EN 12663-1:2010+A2:2023, „Railway applications - Structural requirements of railway vehicle bodies Locomotives and passenger rolling stock (and alternative method for freight wagons)".

[13] A. Heckmann, „A high-tech running gear for the train of the future," DLR e.V., [Online]. Available: https://www.dlr.de/en/latest/news/2022/03/a-high-tech-running-gear-for-the-train-of-the-future. [Accessed on 16.06.2025].

[14] D. Lüdicke, D. Krüger, C. Weber, B. Goetjes und A. Heckmann, „DLR Forschungsinfrastruktur," *ETR – Eisenbahntechnische Rundschau,* Nr. 3, pp. 62-68, 2021.

[15] A. Qazizadeh, R. Persson und S. Stichel, „On-track tests of active vertical suspension on a passenger train," *Veh Syst Dyn,* Nr. 53(6), pp. 798-811, 2015.

[16] R. Schneider, „Wankkompension und aktive Radsatzsteurerung Flexx Tronic von Bombardier," *Eisenbahn-Revue,* Nr. 4, 2010.

Authors:



Jens König

Jens König studied Mechanical Engineering (Dipl.-Ing.) at the Universität Stuttgart from 2001 till 2008. He made his Doctorate in Engineering (Dr.-Ing.) till 2016 at the Universität Stuttgart at the German Aerospace Center (DLR) in the field of lightweight design for car bodies of highspeed trains. He was leader of the team "Lightweight Design Concepts and Methods Rail Vehicles" at the Institute of Vehicle Concepts, DLR, from 2014-2020. He was acting Professor, head of the Chair of Rail Vehicle Technology, Deputy Head of the Institute, Institute of Machine Elements, University of Stuttgart from 2020 till 2022. Since 2022 Dr. König is Senior Researcher and Manager Product and Research Portfolio Major Rail Vehicle Projects at the Institute of Vehicle Concepts,

DLR. He is Lecturer at the University of Stuttgart in the field of Rolling Stock and rail vehicle dynamics.



Thomas Müller

Thomas Müller studied aerospace engineering (M. Sc.) at the university of Stuttgart from 2011 till 2017. From 2018 till 2023 he worked on mechanical ground support equipment at the IABG Space Center in Ottobrunn. Since 2023 he is working at the German Aerospace Center (DLR) in the fields of structural design, running gear development and packaging of rolling stock.



Rickard Persson

Rickard Persson studied Mechanical Engineering at KTH, The Royal Institute of Technology Stockholm, from 1977 till 1982. He joined ASEA, which after several steps become Alstom, where he worked with rail vehicle dynamics for more than 40 years. The focus area has been active mechanical systems, like carbody tilting and active suspension. He has also led several vehicle development projects. In parallel to industrial employment, Rickard have worked closely to KTH, where he made his Doctorate in Engineering 2011 in the field motion sickness in tilting trains. He is Lecturer at KTH in the area of Rolling Stock



Libor Lochman

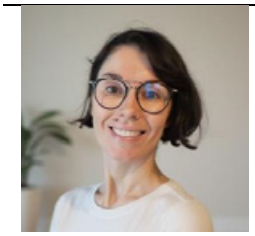
Libor Lochman has long-term experience in control-command and signaling domain. He was a member of the ERRI A200 project team, developing the very first set of the ETCS specification in the early nineties. Later he covered wider scope of the railway technology when serving as the Director of the largest European rail Test Centre – VUZ in Velim (Czechia).

Between 2011 and 2020, he has been the Executive Director of CER (Community of European Railway and Infrastructure Companies) – the principal rail lobby in the EU. In 2021, he was the project manager of the Connecting Europe Express – the flagship project of the European Year of Rail.

Today he works with clients on various rail related topics, mainly related to EU regulation, rail research and innovation. Thus, amongst other activities, Libor is acting as Ambassador for the digital automatic coupler in Central, Eastern and South European regions or working on the solutions aiming to increase low density lines competitiveness through Europe's Rail Joint Undertaking projects.

Session 16

Jens Koenig, Thomas Mueller, Rickard Persson, Libor Lochman, Pauline Guicheney



Pauline Guicheney

Coming from an operational background in train control and traffic management, Pauline Guicheney progressively specialised in railway safety, interoperability and EU regulatory frameworks. At SNCF, she leads Safety & TSI topics for the TELLi project and represents the company within the Europe's Rail Joint Undertaking FutuRe project. Her role is to bridge operational experience and system-level regulatory design, with a focus on pragmatic innovation for regional rail systems.

IRSA 2025

SESSION 17

BAHNBETRIEB

RAILWAY OPERATIONS

Session 17
Paula von der Heide, Lars Schnieder

A SUMO-Based Study of Urban Rail Operations on Frankfurt‘s Corridor A

Paula von der Heide, M.Sc.[1] und Prof. Dr.-Ing. Lars Schnieder[2]

[1]Institut für Eisenbahnwesen und Verkehrssicherung, TU Braunschweig
[2]Cassis Railway Experts GmbH

November 13, 2025

Abstract

Simulation is a valuable method for understanding and improving rail operations, as it allows the testing of operational or technological innovation in a virtual environment. For easy adaptation of new concepts, we explore the open-source simulator SUMO (Simulator of Urban MObility). In a case study, we build a SUMO scenario to simulate rail operation on Frankfurt's light rail corridor A and examine delay propagation as well as the stabilizing effects of Communications Based Train Control (CBTC) in the tunnel segments. The findings highlight SUMO's potential for modeling urban rail systems while identifying areas for further improvement.

Keywords: Simulation, SUMO, urban rail operations, virtual innovation, delay

1 Introduction

Urban rail systems play an important role in modern mobility with rising demand as they provide an efficient and sustainable transit in densely populated areas. Thus, public transport operators must employ short headways between vehicles to provide high capacity while maintaining steady and punctual operations. As a result, operators must regularly review operations, implement improvements to the operational program and evaluate new assisted or automated driving strategies and advanced technologies.

A key innovation is the introduction of Communications Based Train Control (CBTC). This employs the migration from conventional fixed block to moving block train separation. When not stuck to block boundaries, line throughput (expressed in number of trains per hour and direction) could be increased or accumulated delays could be better recovered. Another notable enhancement is the deployment of an automated, driverless reversal system, which can significantly shorten turnaround times. [13]

Simulations can be used to understand and evaluate the impact of any operational or technological modification in a virtual environment before bringing actual changes into operation [9]. Established commercial railway simulation tools such as OpenTrack, LUKS and RailSys provide highly specialized railway modeling capabilities and are used for planning railway lines, creating or verifying timetables and analyzing the capacity [14]. As they are proprietary tools, they come with licensing costs and are not easily adaptable to evaluate novel signaling or automation concepts. This limits their suitability for rapid prototyping. A flexible and open simulation framework capable of implementing and testing new concepts, as well as quantifying their potential benefits, is therefore indispensable.

This study employs the open-source simulator SUMO (Simulation of Urban MObility) [8] as such a playground. To verify the feasibility of this, we explore the SUMO simulator and run a case study on the tunnel section of Frankfurt´s light rail corridor A. The baseline scenario reproduces present-day fixed block signaling with non-automated reversal operations. We then introduce moving block and a driverless reversal, measuring key performance indicators such as average arrival delay and the 3-min-punctuality rate. Comparative analysis across scenarios reveals the mechanisms that influence network performance and confirms that SUMO can serve as a versatile, low-cost testbed for digital transformation in urban railways.

2 State of the Art

We give a short introduction of currently established operational modes for urban railways and the improvements that could be made with CBTC. Additionally, we introduce to the simulator SUMO and compare it to commonly used simulators in the railway domain.

2.1 Urban Railway Systems

The widely established approach in urban railway systems is to drive on-sight on surface tracks whenever interaction with other types of traffic is to be expected and then transitioning to fixed block signaling and train control within tunnel sections.

Lately, the introduction of CBTC technology, as outlined by IEEE standard 1474.1 [7], enables a shift toward moving block signaling in the tunnel sections. This allows for significantly reduced headway between rail vehicles as they may follow each other in braking distance plus a safety margin. This leads to an increased system efficiency which can be used as additional buffer time or to accommodate more frequent traffic [2].

Urban rail systems in mixed-traffic environments often experience delays from interaction with road traffic, traffic lights, and other operational disruptions. While delays accumulate on outer branches of the network, moving block train separation in the tunnels may help stabilize operations [2]. Moreover, CBTC systems facilitate progressive automation, like Automatic Train Operation (ATO). A key initial step involves implementing a driverless train reversal, which effectively reduces turnaround times [13]. This is a significant enhancement, reducing a currently existing bottleneck in the system.

2.2 Modeling Delay in Urban Rail Systems

Ostermann et al. [9] highlight the potential of simulation tools to evaluate signaling and control technologies without affecting the real system. They present a case study on the Korean National Railroad, demonstrating how the OpenTrack simulator compares conventional multi-block signaling with ETCS Level 1.

Becker et al. [1] use cumulative distribution functions to introduce random primary delay in several simulation runs. They use average lateness per train and punctuality as key parameters for comparison between diverse operational and infrastructure scenarios.

Büker et al. [2] contributed specifically to this field through a railway operations study that demonstrates how implementing moving block signaling with a CBTC system significantly enhances buffer times. These increased buffer times, if not allocated to additional traffic, can substantially mitigate delays by reducing their propagation across the network, thus contributing positively to operational stability and service reliability.

2.3 SUMO Simulator

SUMO is an open-source, synchronous, microscopic, and multi-modal traffic simulation tool developed by the German Aerospace Center (DLR) [8]. It simulates individual cars, pedestrians, bicycles and public transport across urban and or regional networks offering precise control over each element's behavior and paths. [4]

SUMO started in 2001 as a freely available tool for large-scale traffic modeling. Since 2018, its governance moved under the Eclipse Foundation. It is widely used in transport research for traffic forecasting, exploration of automated driving strategies, vehicle communication systems (V2V) and traffic light optimization. [4] Historically, it has been designed for road traffic modeling, but its capabilities for railway simulation and public transport modeling have been expanding recently (since SUMO version 1.22) [5].

Like other simulation frameworks, it implements a nodes-edges-model. Tracks are represented as edges and are assigned with behavioral properties and restrictions. Junctions are implemented as nodes at which traffic may branch or merge. For this reason, they may have a traffic light, priority or rail signal attached to it. Public transport stops are added to their respective edge and indicate the stop position.

2.4 Comparison of SUMO to Established Railway Simulation Tools

Classical railway simulation tools such as *OpenTrack* (ETH Zurich), *LUKS* (Quattron) and *RailSys* (RMCon) provide comprehensive environments dedicated for simulating railway operations, as compared in [14].
These commercial platforms typically cover four main aspects:

1. **Train movement simulation**, incorporating detailed physical modeling of traction, resistance, and braking curves.
2. **Train dispatching simulation**, using rule-based or algorithmic approaches to simulate dispatcher decisions.
3. **Timetabeling**, facilitating the creation, modification or verification of timetables.
4. **Capacity analysis**, assessing the maximum number of train movements that can be accommodated by the infrastructure both along tracks and at network nodes.

Conventional railway simulators incorporate anti-deadlock strategies. However, deadlock avoidance in synchronous simulation remains challenging, as no positive strategy has yet been established [10]. Like other simulators, SUMO implements mechanisms to prevent deadlocks; the relative effectiveness of these remains to be evaluated.

The following Table 1 compares SUMO and conventional railway simulation tools:

Feature	**Conventional Tools**	**SUMO**
Driving dynamics	detailed	detailed
Layout types	abstract	geometric or abstract
Signal modeling	full (main, distant)	main signals only
Anti-deadlock strategies	yes	yes
Route locking	certified interlocking	basic
Dispatching algorithms	rule-based or algorithmic	rule-based only
Timetable integration	yes	via GTFS or manual creation
Timetable construction	yes	no
Rolling stock database	extensive	minimal
Capacity assessment	yes	no
Delay simulation / evaluation	built-in	python scripts
Moving block	with small blocks	yes

Table 1: Comparison between SUMO and other railway simulators based on [6, 11, 14]

There are also some features of SUMO that go beyond the classical railway simulation tools which motivate us to work with it:

- Multi-modality: SUMO supports road, rail, bicycle and pedestrian traffic in one environment. This allows for interaction between trains and other traffic, especially in a light rail context. Also there may be passengers riding the trains.
- Moving block operation: SUMO implements a realistic following behavior using the absolute braking distance and a minimum gap instead of virtual or small blocks.
- TraCI (Traffic Control Interface) is a powerful API that enables external programs to interact with SUMO and control the running simulation or integrate it into a co-simulation framework.
- Network creation from OpenStreetMap (OSM): With given data scarcity, generating networks with OSM data may serve as a valuable basis for scenario generation.

3 Methodical Approach for Simulating Urban Networks

Our goal is to explore how urban rail systems can be modeled and analyzed in SUMO. To achieve this, we create a SUMO scenario of Frankfurt's light rail corridor A and add traffic with defined headway and delay. We examine variations in traffic flow and systematically analyze the impact of adding delay and how it is recovered.

3.1 Creation of SUMO scenario

The subway lines U1, U2, U3, and U8 have a common section between Südbahnhof and Heddernheim, referred to as corridor A. Then the lines diverge into branches serving different destinations.

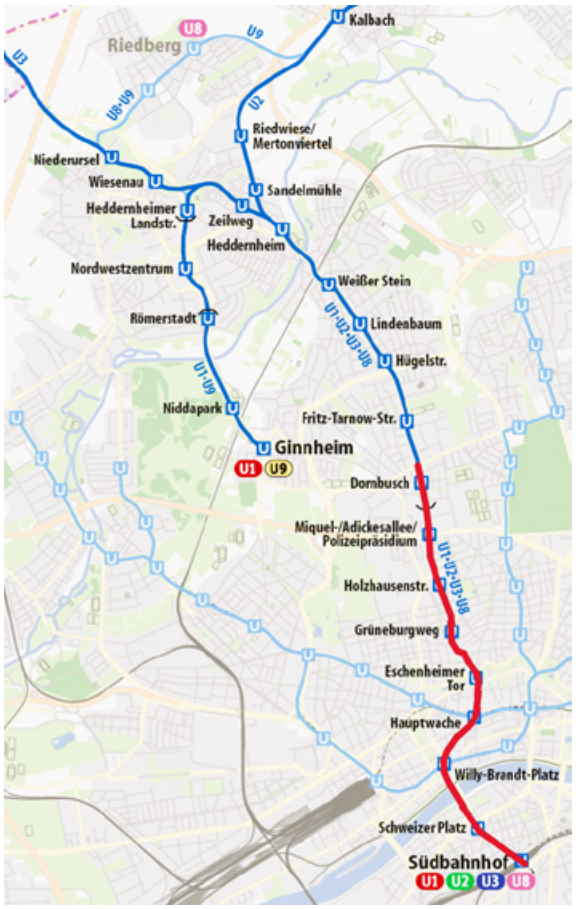


Figure 1: Light rail corridor A in Frankfurt [19]

In the overground sections, operation is on-sight with the driver responsible for safe train movement and maintaining a safe distance to the preceding train. At intersections it is protected by F-signals (tram signals). Delays may occur due to interactions with surface traffic and traffic signal phases. Trains transition from overground to underground between Dornbusch and Miquel-/Adickesallee and continue through the tunnel to Südbahnhof, where they reverse direction and return via the opposite track. Tunnel operations are governed by Automatic Train Protection (ATP) systems employing either fixed block or in the future, moving block train separation [16].

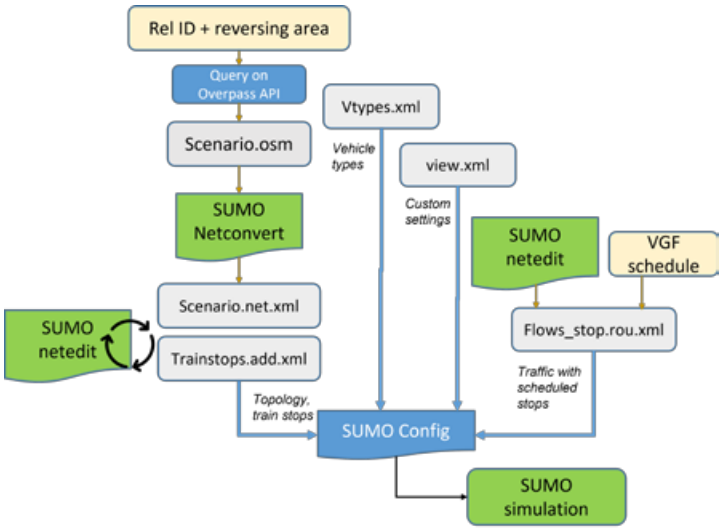
As the distinction between fixed block and moving block operation primarily affects the tunnel section, this study focuses on that part of the corridor. To simulate a potential delay upon entering the tunnel, we include the station of Dornbusch where we add delay to a

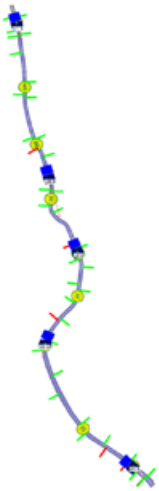
vehicle. Accordingly, the SUMO scenario includes the station of Dornbusch, the tunnel to Südbahnhof, the reversing yard and the return path to Dornbusch (as highlighted in red in Figure 1). The one-way distance is approximately 5 km. To obtain the SUMO scenario, we use the steps as depicted in Figure 2a. The scenario has been created with publicly available data. All created scripts are provided on Github [17].

3.1.1 Track Topology

The topology is queried from OpenStreetMap (OSM) via the Overpass API. We use the relation ID of all four subway lines and additionally include all tracks in its surroundings which are tagged with operator VGF. The OSM data contains railway switches, signals and speed limits. Due to the nature of OSM data, it is not always complete and correct. Nevertheless, it serves very well as a starting point to create a SUMO scenario.

This data is converted to a SUMO network via *SUMO netconvert* and is subsequently edited in *SUMO netedit* to create a representative simulation. The obtained scenario is shown in Figure 2b. Stations are marked with an circled "H" (German "Haltestelle"). Main signals appear as bars, either green ("proceed") or red ("stop"). Switches are marked in yellow. Vehicles may be enlarged for visualization and are colored by their current delay; blue indicates no delay and it changes to green, yellow and red according to the amount of delay.





(a) scenario generation *(b)* snapshot of SUMO GUI

Figure 2: SUMO scenario generation process and resulting scenario

We verify and extend the placement of signals with the help of cab videos [3], this does not allow for exact positioning of signals, but gives some indication. We ensure that at minimum signals are placed at the entry and exit of each station. For the reversing area, we add signals so that all realistic driveways are possible in SUMO.

We set the permissible speed along the main track in both directions to 60 km/h (approx. 16.67 m/s), whereas in the reversing area, we restrict the speed to 20 km/h (approx. 5.56 m/s). These values are assigned as *"maxSpeed"* to the edges in SUMO.

Elevations are not used in this simulation, because for this scenario no reliable elevation data was available and except for the tunnel ramps we expect no significant changes in altitude. SUMO provides this option, and when given appropriate topological input, elevation and gradient effects could be included into the simulation.

3.1.2 Vehicle data and driving dynamics

SUMO has a set of predefined rail vehicles to use. Alternatively, a custom behavior can be described via parametrized curves. This allows implementing driving dynamics of a specific rail vehicle as theoretically described in [11]. For simplification, we only use one vehicle type. We chose the U5 vehicle of VGF with two coaches (U5-50) [15] as this is the vehicle to be equipped with CBTC by VGF [16]. We create a custom vehicle named "VGF_U5" that matches with vehicle length, mass, tractive effort, and maximum speed:

```
<vType id="VGF_U5" vClass="subway" trainType="custom" carFollowModel=
    "Rail" maxPower="1040" maxTraction="211" decel="1.7" length="49"
    mass="86000" maxSpeed="19.44" minGap="25" resCoef_quadratic=
    "0.03312" resCoef_linear="0.01512" resCoef_constant="0.0099"/>
```

The tractive effort is calculated with Equation 1. For low speeds, the traction is limited by the friction ($F_R = \mu \cdot m \cdot g$) between steel wheels and the tracks. For higher speeds, the maximum traction is speed-dependent and limited by the power of the vehicle:

$$\text{traction [kN]} = \min\left(\text{maxTraction[kN]}, \frac{\text{maxPower [kN]}}{\text{speed [m/s]}}\right) \tag{1}$$

The vehicle's resistance forces are described using the commonly used Davis formula. The parameters $\text{resCoef}_{\text{quadratic}}$ (α), $\text{resCoef}_{\text{linear}}$ (β), and $\text{resCoef}_{\text{constant}}$($\gamma$) are obtained empirically (we use values of Wende [18] for subways).

Braking curves are not explicitly modeled, and no distinction is made between operational modes. Instead, a simplified approach is used, assuming a constant deceleration of 1.7 m/s². This does not account for jerk softening as performed by an experienced driver or automated braking. Consequently, the actual braking time and distance are expected to be longer. This increases driving time, slightly reducing scheduled buffer times.

3.1.3 Schedule

In SUMO, vehicles can be introduced into the simulation either as individual *trips* or as part of vehicle *flows*. For the present work, we employ vehicle flows, as they allow the specification of a recurring sequence of vehicles following a predefined *route* and *period* making them particularly suitable for modeling public transportation systems.

The route is defined for the reversing operation (see 3.1.4), as the network contains no additional branches or turning alternatives. We run a service interval of 600 s for lines U1 and U2, and 900 s for lines U3 and U8, aligning with the scheduled intervals of current operation by VGF [12]. The following specification creates a series of vehicles (e.g. U1.1, U1.2, U1.3 ..) with a headway of 600 s that perform a reversal in the upper track.

```
<flow id="U1" type="VGF_U5" begin= "0" route="Langkehre_oberesGleis"
    end="3600" period="600"/>
```

Furthermore, it is necessary to specify the scheduled stops along the route. For each stop, we define both a *duration* (s), indicating the dwell time at the stop, and an *until* value (s), representing the scheduled departure time from that stop.

```
<stop busStop="Dornbusch_0" duration="20" until="56"/>
```

Departure intervals in the simulation are established based on the 2025 timetable data from VGF [12]. The schedule indicates an 11 min travel duration from Dornbusch to Südbahnhof and a shorter duration of 9 min for the return trip. In the simulation model, a total driving time of 5.7 min and cumulative stopping time of 2.7 min is computed, as we are assuming a passenger exchange duration of 20 s per stop. To reconcile with the real-world schedule durations, buffer times of 2 min (equivalent to 15 s per stop) are incorporated into the inbound route, whereas a buffer of 40 s (equivalent to 5 s per stop) is applied to the outbound route.
Table 2 summarizes the scheduled departure times (s) for the first vehicle, which enters

the simulation at 0 s and departs Dornbusch after 56 s. Subsequent vehicles are inserted after the defined service period and have identical travel times between the stations. If the insertion is obstructed by a delayed vehicle, the following vehicle is inserted once the corresponding position in the network becomes available, while ensuring at least the required braking distance or block separation from its predecessor.

Table 2: Train stops and scheduled departure (dwell time at each stop: 20 s, reversing time: 300 s)

Train stop	inbound until (s)	return until (s)
Dornbusch	56	1561
MiquelAdickesallee	137	1489
Holzhausenstraße	214	1421
Grüneburgweg	292	1355
EschenheimerTor	371	1283
Hauptwache	441	1224
Willy-Brandt-Platz	519	1156
SchweizerPlatz	617	1068
Südbahnhof	698	998

We assume an unlimited vehicle pool, ignoring possible circulation constraints due to the inserted delay; upstream operations are truncated.

3.1.4 Reversing Area

Two types of reversing maneuvers are modeled in the simulation, utilizing either the upper or lower track in the reversing area after leaving the platform (highlighted in red and blue in Figure 3). Each vehicle is assigned a dedicated route through the reversing area to ensure simulation accuracy. The third track is not used during scheduled operations; instead, it is reserved for vehicle pausing or for managing operational disturbances. The minimum duration of a reversing maneuver observed in the simulation is:

- 82 s for the upper track,
- 91 s for the lower track.

In scenarios with non-automated reversals, an additional 120 s is modeled to account for the time required for the driver to close the driver‘s desk, walk to the other side of the train and open the other driver‘s desk. This is implemented as a mandatory stop at the reversing track before the vehicle may resume its route. If CBTC-controlled driverless operation is

assumed, the additional stopping time can be omitted, and the reversing process is limited to its minimum physical duration. In this case, the driver leaves and accesses the vehicle before and after the reversal at the station platform. [13]

To accommodate possible variations and ensure operational robustness, a total scheduled reversing time of 300 s is allocated in the timetable.

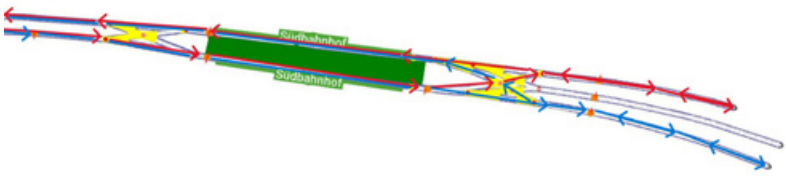


Figure 3: Reversing area at Südbahnhof as simulated in SUMO

3.1.5 Train Sequencing

On-sight driving is represented as a moving block operation in our SUMO simulation, as it corresponds to manually keeping a braking distance. This applies to the section at Dornbusch, which is still overground.

For the conventional scenario, we then transition to fixed block train control upon entering the tunnel. The main signals divide the track in blocks, and only one train may enter into one block. As SUMO does not know distant signals, trains behave as if they always had perfect sight upon the signals or as if they had in-cab signaling.

In the CBTC scenario, we continue moving block operation also within the tunnel. Trains may proceed with a headway corresponding to their absolute braking distance plus a safety margin. SUMO implements a following model which spaces trains accordingly and uses a *minGap* parameter to specify the safety margin. We set this to 25 m, comprising of a minimum inter-train spacing of 5 m and an additional positional uncertainty of the CBTC system of 10 m per train, as specified in [7].

3.1.6 SUMO Configuration

A configuration file (*.sumocfg*) is used for general settings of the simulation and to bundle the needed xml files which specify the topology, train stops and vehicles.

We set the duration of simulation for 1 hour (3600 s) plus simulation steps needed for all vehicles to leave the simulation after that. So the duration of the simulation may vary depending on the introduced delay.

Heavily delayed vehicles are by default removed or teleported after a certain time. Since we especially want to examine the delay, it is important to deactivate the teleports `(-1)`.

SUMO generates various types of simulation output. In this work, we enable output for:

- FCD output (floating car data): position, speed, acceleration of every vehicle in every time step
- Stop output: actual arrival and departure times for every vehicle and station
- Tripinfos: Information on every journey of every vehicle in the simulation (start and end time and place, duration, waiting times and total stopping duration)

3.2 Adding Delay

We look at the microscopic implications when adding delay to one specific vehicle after it was inserted to the simulation, but before entering the tunnel. We apply delay only to one vehicle so that we know exactly which vehicle is causing the delay and which are only affected by it. Importantly, the first vehicles after simulation startup may experience artificially favorable conditions due to the absence of preceding traffic (e.g. in the reversing area). Thus, we deliberately apply the delay to the sixth vehicle in the flow (*U2.1*) with both preceding and succeeding vehicles present.

We examine delay durations of 120 s, 180 s, 300 s, 600 s and 900 s. These represent medium severe operational disturbances caused by e.g. a slow speed zone due to construction work, a jammed intersection or a medical emergency response. We observe how this delay can be recovered and how the subsequent vehicles are affected.

To introduce delay in SUMO, we use the SUMO interface TRACI which allows for run time changes in the simulation. Via a Python script, we connect to the SUMO environment via TRACI, start the simulation, and then add delay to a selected vehicle via the *sumolib* function *traci.vehicle.slowDown()* setting the current speed to 0 m/s for a certain duration. It will presume its journey automatically once the delay duration has past.

3.3 Scenario Verification

To ensure the viability of simulation outcomes, it is essential to verify the simulation model against real-world conditions as far as possible.

3.3.1 Verification: Driving Dynamics

We examine the stop to stop driving dynamics of a vehicle in the simulation to ensure realistic behavior. Using the vehicle parametrization described in section 3.1.2, the driving dynamics are visualized in an speed over time diagram, showing the motion profile between two consecutive stops (see Figure 4). The results are verified against the U5 data sheet [15]. A constant braking deceleration of 1.7 m/s² is observed, corresponding to the specified service braking rate. The maximum acceleration is 2.34 m/s², with an average acceleration of about 1.27 m/s² (specified with 1.3 m/s² in the data sheet). The vehicle requires roughly 15 s to accelerate from rest to its maximum speed of 60 km/h.

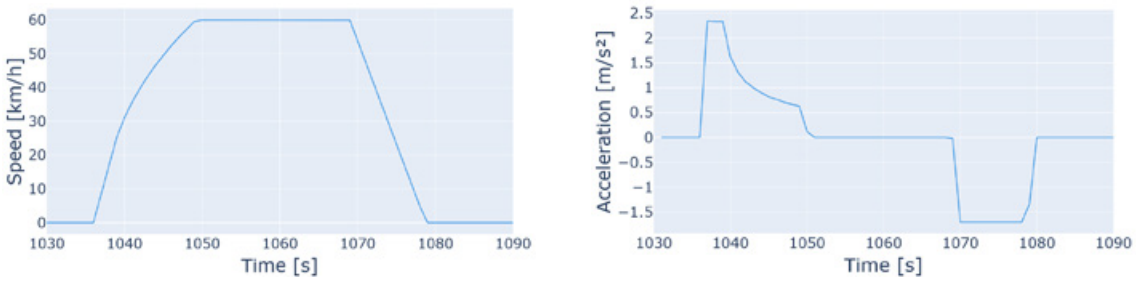


Figure 4: Speed and acceleration of vehicle *U2.1* over time

3.3.2 Verification: Undelayed Fixed Block Scenario

We conduct a baseline simulation using fixed block conventional signaling logic, the indicated schedule, and an idealized situation free of delays or disruptions. This scenario serves to check the correctness of the network topology and ensure that the infrastructure, as modeled, can accommodate the present-day VGF schedule.

In this "perfect world" scenario, vehicles operate exactly on schedule, with an average delay of 0.0 s and 100 % of stops served punctually. The results confirm that the infrastructure as modeled introduces no artificial delays and that traffic flow remains unobstructed.

4 Results

In our case study, we examine two different operational modes in SUMO: conventional fixed block operation with non-automated train reversals, and a CBTC operation implementing moving block train operation with automated, driverless reversals. For each operational mode, we introduce a fixed delay to a single vehicle examining the resulting delay behavior.

4.1 Evaluation Approach and Key Performance Indicators

We convey a detailed, microscopic examination of each vehicle's journey and stops along the way. SUMO collects extensive data on vehicle accelerations, speeds, positions, stopping places and times as well as waiting times. This extensive dataset is processed and analyzed using custom Python scripts to visualize and evaluate the observed delays. All scripts developed for this analysis are publicly available on GitHub [17].

To systematically evaluate the observed delay behavior, we define the following metrics:

- **Average Arrival Delay (s)**: We look at the average arrival delay of a vehicle across all stations during one simulation run.
- **Arrival Delay** (s): We examine the delay of each vehicle at each station and analyze how well the delayed vehicle and its successors recover from their delay.
- **3-minute Punctuality(%)**: We consider a vehicle's arrival punctual if it reaches a scheduled stop within a 3-minute threshold beyond the planned arrival time.

4.2 Evaluation of Adding a Fixed Delay

Delay values of 120 s, 180 s, 300 s, 600 s, and 900 s were considered in the analysis. In Figure 5, we see the movement over time of vehicle *U2.1* with different primary delays. The scheduled trip is shown as a black line. Notably, until the reversing area we observe identical delay recovery behavior for both operational modes, only utilizing the buffer times allocated in the schedule. A significant difference emerges only in the reversal section: driverless automated reversal under CBTC allows a faster turnaround compared to the non-automated reversal requiring a driver's cab change (see solid and dashed lines).

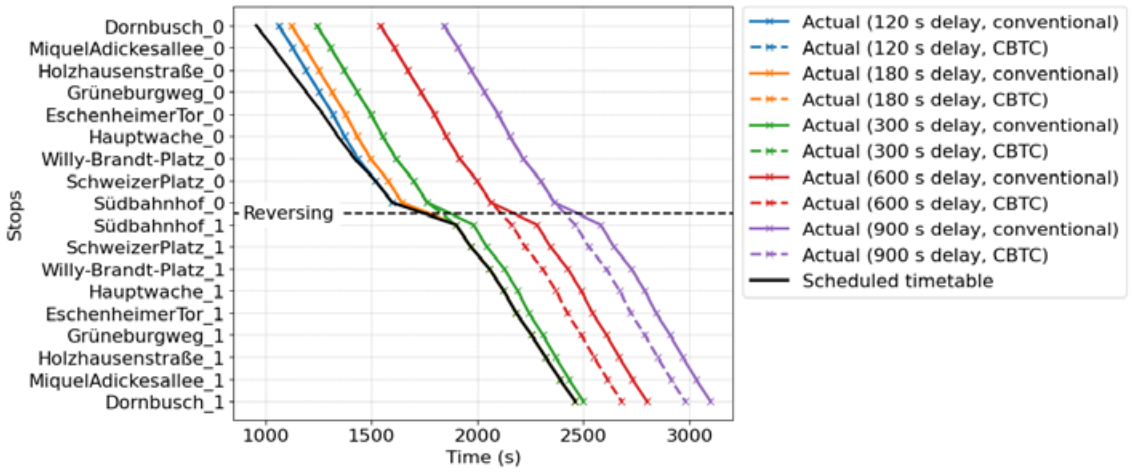


Figure 5: Space Time diagram for the delayed vehicle *U2.1*

When looking at the subsequent vehicles, results indicate that delays of 120 s and 180 s have minimal impact, as successor vehicles remain almost unaffected. Higher delays propagate to subsequent vehicles, causing more significant disruptions. Figure 6 illustrates the average delay of subsequent vehicles across all stations when a 600 s delay is introduced to vehicle *U2.1*. The blue bars illustrate the average delays measured during conventional operation, in contrast to the orange bars, which reflect the delays recorded with CBTC operation. The results indicate that delays are recovered more quickly with CBTC, especially for the subsequent vehicles.

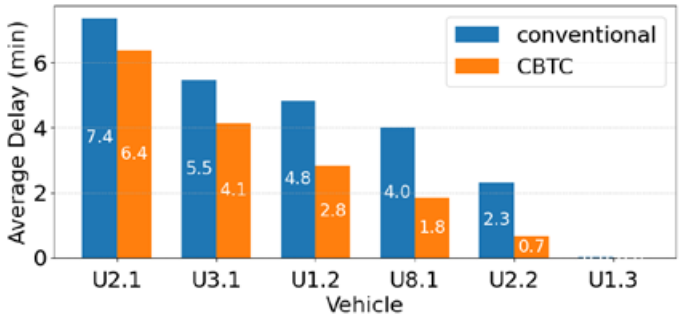


Figure 6: Average delay of each vehicle across all stations when introducing a 600 s delay to *U2.1*

To investigate this further, we look at the delay across stations for subsequent vehicles. Figure 7 shows the resulting delay when introducing a 300 s primary delay to vehicle *U2.1*. Like in Figure 5, for vehicle *U2.1* the difference emerges only from the reversal. The advantages of moving block operation become apparent when analyzing the performance of succeeding vehicles. With moving block, vehicles follow each other more

closely. As a result, CBTC operated trains can recover delay faster. We also see that vehicle *U1.2* can recover more delay than vehicle *U3.1*.

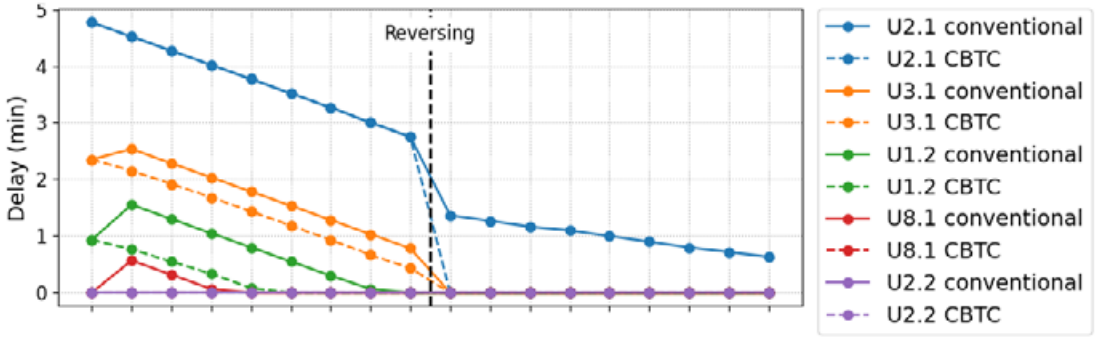


Figure 7: Delay by station for the delayed vehicle and its successors with 300 s delay introduced

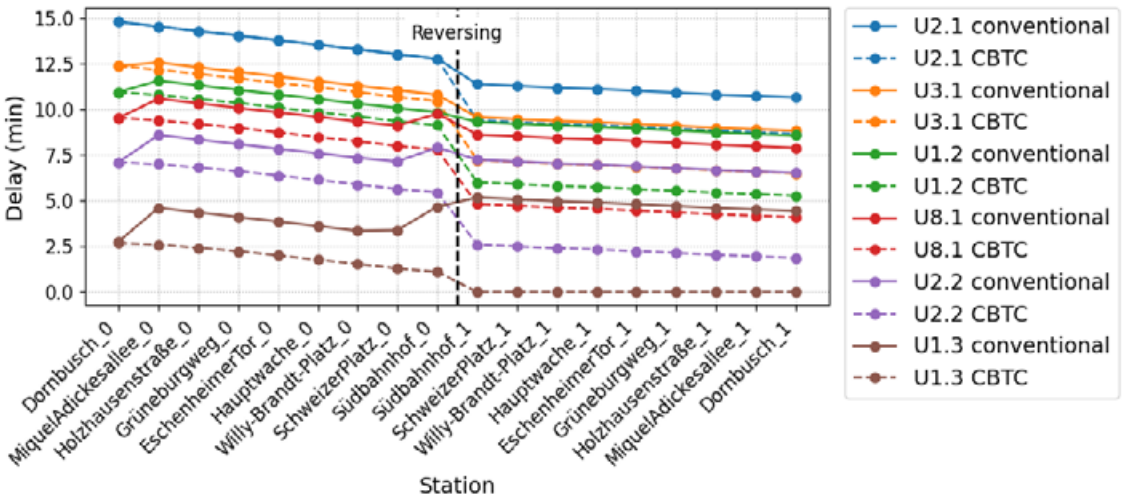


Figure 8: Delay by station for the delayed vehicle and its successors with 900 s delay introduced

Upon transitioning to fixed block operation when entering the tunnel, trains must wait for a clear block ahead, resulting in a slight increase in delay. The added delay strongly depends on the signal placement. Positioning the station entry signal closer to the platform reduces the restrictive effect of the block separation, allowing a following train to approach while the preceding one is still at the station.

When a larger delay is introduced, as shown in Figure 8 for 900 s, we find that the non-automated reversals create a system bottleneck. Vehicles bunch up, and once the first two occupy the reversing yard (blue and orange lines), all following vehicles incur additional delay (green, red, lila and brown lines). In a real-world operation, trains would likely re-route using the crossovers at Eschenheimer Tor or Willy-Brandt-Platz (which we have

not modeled in SUMO). Using the third track in the reversing yard would also reduce this bottleneck. Importantly, this jamming does not occur with automated reversals, so CBTC delay values downstream of the reversing area remain much lower. Our analysis shows that under large delay conditions, the benefit of CBTC becomes even more pronounced.

Table 3 shows the observed values for the delayed vehicle *U2.1* and its two next successors *U3.1* and *U1.2* for 180 s, 300 s, 600 s and 900 s.

	180 s		300 s		600 s		900 s	
	conv	CBTC	conv	CBTC	conv	CBTC	conv	CBTC
3 Min- Punctuality (%)								
U2.1	100.00	100.00	55.56	55.56	0.00	0.00	0.00	0.00
U3.1	100.00	100.00	100.00	100.00	0.00	50.00	0.00	0.00
U1.2	100.00	100.00	100.00	100.00	0.00	50.00	0.00	0.00
All	100.00	100.00	97.39	97.39	76.14	85.95	61.76	73.53
Average delay (s)								
U2.1	52.89	52.89	142.72	112.89	442.72	383.22	742.72	683.22
U3.1	4.00	1.67	52.06	42.50	327.39	247.83	627.39	547.83
U1.2	0.00	0.00	21.83	8.89	289.06	170.61	589.06	470.61

Table 3: Comparison of punctuality and average delay

Overall, the results demonstrate that CBTC operations provide higher resilience to delays, particularly under severe disturbances. While minor delays can be recovered in both systems, CBTC more effectively limits the propagation of major delays.

5 Discussion

It was possible to create a scenario for Frankfurt light rail corridor A from OSM relying only on public data and operate trains with fixed block as well as moving block train spacing. Additionally, we modeled the reversing area at Südbahnhof. We could observe the behavior of the vehicles when adding delay, via the SUMO GUI as well as via the simulation output xml files.

5.1 Shortcomings on the implemented scenario

We defined fixed routes in the reversing area. Studying the results showed that trains sometimes wait for a specific track to clear, even when another was already free. To address this, dynamic re-routing is needed, which SUMO supports. Additionally, the use of the third track could be enabled.

Also, the simulation would become more realistic, if we added the possibility to reverse at the cross overs at Eschenheimer Tor or Willy-Brandt-Platz and skip some stations. In heavily delayed situations, this would stabilize operations much quicker.

After the delayed vehicle, subsequent vehicles are inserted directly after each other within braking distance. Including the preceding station would yield a more realistic behavior, since the dwell time of trains there will create a more realistic headway at Dornbusch.

We have not explicitly modeled variations in braking profiles or coordinated braking and acceleration behaviors under CBTC conditions. To improve the accuracy of our analysis, especially the impact of conservative braking curves could be evaluated.

5.2 Operational findings

The results reflect the widely recognized benefits of CBTC systems, confirming the expected advantages of automation and moving block operation such as improved delay recovery and enhanced operational resilience. We especially point out the improvement that is achieved through the automated reversal. These findings are consistent with those of Büker et al. [2], who likewise emphasize the substantial positive impact of automated reversals on delay recovery compared to non-operated reversals.

5.3 Findings on SUMO

Regarding our choice of simulation tool, SUMO proved capable of effectively capturing and evaluating operational behavior. It features a detailed vehicle movement model and provides extensive output from the simulation. Moreover, it allows for configuration of scheduled train operations with predefined stopping points and dwell times, enabling a thorough assessment of the operational impacts of timetable deviations.

Some railway-specific details are simplified in SUMO. We believe that these railway specifics do not affect the operational outcomes of this study in a significant manner.

Upon configuration of the moving block operation, we observed some unrealistic behavior and even collisions when trains followed each other in the reversing area and switched tracks. After reporting this issue to the SUMO team, this issue was solved in SUMO version 1.24. We identify this as another benefit of open-source software. Any changes needed for a better simulation accuracy can be submitted or requested.

6 Conclusion

The presented case study demonstrates the feasibility of employing SUMO for realistic urban rail simulations and provides insights into operational behavior when implementing CBTC systems with moving block train separation and automated reversals.

Automation remains a dynamic and evolving topic in railway operations, raising important research questions related to system behavior, process reliability, and operational resilience. Tools like SUMO offer a valuable foundation for addressing these questions, due to its accessibility, ease of setup, and open-source nature.

In future work, we plan to investigate failure modes and their impact on system performance and resilience. This becomes particularly important when considering the implementation of fully driverless operations (GoA 3–4) and the elimination of secondary train detection systems. Under such conditions, failure scenarios such as the loss of sensor data, degraded radio communications, and passenger emergencies can significantly compromise operational stability, given the absence of an on-board driver. To effectively manage these emerging challenges, a comprehensive evaluation of fallback strategies is essential to ensure resilience and maintain service quality under degraded conditions.

Conventional railway simulation tools are limited to railway systems, whereas SUMO is also capable of simulating other traffic modes. This makes it well suited for studying light rail systems, where parts of the network interact with road based transportation at intersections and along shared roadway sections. In this context, Cooperative Intelligent Transport Systems (C-ITS) facilitate coordination between vehicles and infrastructure. Extending the present work to incorporate C-ITS in the surface sections of the corridor is a next step towards enhancing the analysis and optimization of urban rail traffic flows.

Acknowledgement

We give credit to the developers of SUMO who have provided the opportunity to work with an open-source simulator with open data.

This text was created with the assistance of ChatGPT to rephrase sentences and enhance its readability. Additionally, it was used to support with scripting Python.

Literature

[1] BECKER, Matthias ; BÜKER, Thorsten ; HENNIG, Eike ; KOGEL, Felix: Sound evaluation of simulation results. In: *8th International Conference on Railway Operations Modelling and Analysis*. Norrköping, 2019

[2] BÜKER, Thorsten ; SCHNIEDER, Lars ; HÖVELL, Madeleine van ; MEURER, Daniel: Eisenbahnbetriebswissenschaftliche Untersuchung von CBTC-Systemen. In: *Signal + Draht* (2021)

[3] DDFUBUKI: *[Metro Cab Ride] U-Bahn Frankfurt Linie 1 / Ginnheim to Südbahnhof*. `https://www.youtube.com/watch?v=ZOG8HeETC9I`. Version: April 2020

[4] DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT (DLR): *About Eclipse SUMO*. `https://eclipse.dev/sumo/about/`. Version: Juli 2025

[5] DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT (DLR): *ChangeLog - SUMO Documentation*. `https://sumo.dlr.de/docs/ChangeLog.html`. Version: Juli 2025

[6] DEUTSCHES ZENTRUM FÜR LUFT- UND RAUMFAHRT (DLR): *Railways - SUMO Documentation*. `https://sumo.dlr.de/docs/Simulation/Railways.html`. Version: Juli 2025

[7] IEEE: *IEEE 1474.1 Standard for Communications-Based Train Control (CBTC) Performance and Functional Requirements*. `https://doi.org10.1109/IEEESTD.2004.95746`. Version: 2004

[8] LOPEZ, Pablo A. ; BEHRISCH, Michael ; BIEKER-WALZ, Laura ; ERDMANN, Jakob ; FLOTTEROD, Yun-Pang ; HILBRICH, Robert ; LUCKEN, Leonhard ; RUMMEL, Johannes ; WAGNER, Peter ; WIESSNER, Evamarie: Microscopic Traffic Simulation using SUMO. In: *21st IEEE International Conference on Intelligent Transportation Systems (ITSC)*. Maui, HI : IEEE, 2018

[9] OSTERMANN, N. ; SCHÖBEL, A. ; OSTER, M. ; MESSAUER, C.: Anwendungen

der Betriebssimulation im Eisenbahnwesen. In: *e & i Elektrotechnik und Informationstechnik* 122 (2005), April, Nr. 4, 124–130. https://doi.org/10.1007/BF03054061

[10] PACHL, Jörn: Avoiding Deadlocks in Synchronous Railway Simulations. In: *2nd International Seminar on Railway Operations Modelling and Analysis*. Hannover, Germany, 2007

[11] PACHL, Jörn: *Systemtechnik des Schienenverkehrs: Bahnbetrieb planen, steuern und sichern*. Wiesbaden : Springer Fachmedien Wiesbaden, 2025 https://doi.org/10.1007/978-3-658-45732-7

[12] RMV / VGF: *RMV.DE - Fahrplantabellen*. https://www.rmv.de/c/de/fahrplan/fahrplaene/linienfahrplaene/fahrplantabellen

[13] SCHNIEDER, LARS: *Communications-Based Train Control (CBTC): Komponenten, Funktionen und Betrieb*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2024 https://doi.org/10.1007/978-3-662-68952-3

[14] TSHABALALA, Zandile: *A SUMO analysis of the railway traffic flow on the SOWETO corridor*, KTH Royal Institut of Technology, Diplomarbeit, 2020

[15] VGF: *VGF Fahrzeugtypen Flyer U5-Wagen*. https://www.vgf-ffm.de/fileadmin/VGF/Die_VGF/Fuhrpark/Documents/VGF_FB_20_06_Fahrzeugtypen_Flyer_RZ3-U5-Wagen_Web_U5_V2.pdf. Version: Juli 2021

[16] VGF ; SIEMENS MOBILITY: *Presse-Information: Vernetzter, schneller, zuverlässiger: VGF und Siemens Mobility präsentieren Meilenstein zur Streckenausstattung mit CBTC*. https://assets.new.siemens.com/siemens/assets/api/uuid:672bdccd-058a-4172-a296-324055c79000/20250924-Pressemitteilung-VGF-DE.pdf. Version: September 2025

[17] VON DER HEIDE, Paula: *Simulate Urban Rail in SUMO*. https://github.com/pvonderheide/simulate_urban_rail_in_sumo. Version: Juni 2025

[18] WENDE, Dietrich: *Fahrdynamik des Schienenverkehrs*. Wiesbaden : Vieweg + Teubner Verlag, 2003 https://doi.org/10.1007/978-3-322-82961-0

[19] WIKIPEDIA: *U-Bahn Strecke A (Frankfurt am Main)*. https://upload.wikimedia.org/wikipedia/commons/f/f7/Karte_U-Bahn_Frankfurt_A-Strecke.png. Version: November 2025

Session 17
Paula von der Heide, Lars Schnieder

Authors



von der Heide, Paula
Since 2023, Paula von der Heide is a research associate at the Institute of Railway Systems Engineering and Traffic Safety at TU Braunschweig. After studying electrical and industrial engineering, she worked for a manufacturer of critical radio systems. She was then employed as a project and team leader in software development for railway components and participated in consulting projects related to future train control systems.



Schnieder, Lars
Lars Schnieder studied civil engineering with a specialization in railroad engineering at the TU Braunschweig. Parallel to his work in industry, he completed his doctorate at the TU Braunschweig and finished his postdoctoral qualification at the TU Dresden. He teaches as a private lecturer at RWTH Aachen and is an honorary professor at TU Braunschweig. He is Chief Executive Officer (CEO) of Cassis Railway Experts GmbH.

Session 17
Nina D. Versluis, Mahnam Saeednia

Modelling Rail Carrier Assignment and Relocation in Multimodal Pod System

Versluis, Nina D.[1] and Saeednia, Mahnam[1]

[1]Department of Transport and Planning, Delft University of Technology, The Netherlands

Abstract

A multimodal pod system features modular autonomous vehicles consisting of detachable transport units mounted on mode-specific carriers. Within this system, transport units arriving at the railway network need to be assigned to available rail carriers to continue their journey. For better alignment between carrier availability and transport demand, rail carriers are allowed to relocate empty within the network. This work presents a mathematical optimisation model for the associated rail carrier assignment and relocation problem. The model's performance is demonstrated through a small case study featuring five transport units, which highlights the impact of the number of rail carriers and their initial distribution in the railway network on the number of handled transport units and total relocation time. Specifically, increasing the number of rail carriers from two to three resulted in an average of 1.25 additional transport units to be assigned, with an average 55% increase in relocation time. Furthermore, varying the initial distribution of rail carriers results in a difference of up to two transport units handled, while relocation times varied by as much as 120% of the average relocation time. To enable a more general assessment, future work will apply the model to a larger and more realistic case study.

Keywords: modular vehicles, pods, matching, fleet circulation, optimisation

Session 17
Nina D. Versluis, Mahnam Saeednia

1. Introduction

In the multimodal pod system considered in the Pods4Rail[1] project, freight and/or passengers are transported by modular autonomous vehicles, i.e., pods. A pod consists of a transport unit mounted on a carrier, as shown in Figure 1. The transport unit is detachable from the carrier, with the carrier being specific to a transport mode, e.g., road or rail.

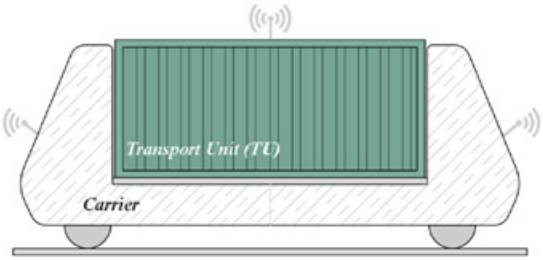


Figure 1: Illustration of a pod (3).

Transport units enter and exit the railway network at stations. To move the transport units from their pickup station to their delivery station, they must be assigned a rail carrier. The assignment of rail carriers to transport units is a challenge due to system constraints such as the limited number of rail carriers available, the pickup and delivery time windows of the transport units, and the railway network capacity. To accommodate the handling of transport units, rail carriers can be relocated, i.e., travel empty, between stations.

So far, the rail carrier assignment and relocation problem as such has been hardly considered in the railway-related literature. To the best of our knowledge, only one study directly addresses this problem. Liao et al. (3) consider a greedy heuristic approach in which assignments are made based on the earliest due date of the transport units and the earliest arrival time of the rail carriers. This approach, however, provides no guarantee in terms of the number of transport unit handled or the rail carrier relocation effort required.

In contrast, the related one-way car-sharing with relocation problem has been more extensively studied. In that context, vehicles are assigned to one-way trip requests while minimising relocation costs. For the modelling of this problem, research has explored (rolling horizon) optimisation, e.g. (4), (discrete-event) simulation, e.g. (1), or a combination of the two, e.g. (2). This car-sharing problem variant, however, differs significantly from the rail carrier problem, particularly due to the operational constraints inherent to railway systems, such as routing and safe separation on the network.

[1] https://pods4rail.eu/

In this work, we propose a mathematical optimisation model based on a multi-objective mixed integer linear programming (MILP) formulation to address the rail carrier assignment and relocation problem at the network level. Maximising the number of handled transport units with minimum relocation, the model assigns rail carriers to transport units, while accounting for constraints of the railway network. The model is demonstrated using a small case study, in which we briefly analyse the impact of the number of available rail carriers and their distribution in the network.

The paper is organised as follows. In Section 2.1, the rail carrier assignment and relocation problem is further specified, conceptually and mathematically. In Section 3, the mathematical model is demonstrated on a small case study. Section 4 concludes the paper.

2. Problem formulation

In this section, we further specify the rail carrier assignment and relocation problem. First, Section 2.1 provides a conceptual description of the problem and corresponding assumptions. Second, Section 2.2 presents a mathematical model formulation to formally describe the problem.

2.1. Problem description and assumptions

In the rail carrier assignment and relocation problem, transport units are assigned rail carriers for transportation within the railway network. To support effective and efficient service, rail carriers may be relocated within the network to better align carrier availability with the transport demand.

We consider the railway network as a set of stations connected by railway lines with limited capacity, modelled as the maximum number of rail carriers that may enter or exit a line per time unit. Accordingly, time is modelled as discrete, with time units corresponding to the minimum separation between rail carriers. We assume fixed travel times between stations, and a fixed maximum capacity in terms of rail carriers for each station.

We consider a fixed number of homogeneous rail carriers available in the network, which are initially located at a station. Each rail carrier can handle one transport unit at a time

and may travel without transport unit, i.e., be relocated, when enabling the handling of a transport unit.

For the formulation of the rail carrier assignment and relocation problem, we define transport requests to represent transport units and their characteristics, i.e., pickup and delivery stations and associated time windows. To support operational flexibility, we allow transport requests to be rejected if they cannot be handled. If handled, a transport request is assigned a single rail carrier that transports the corresponding unit from its pickup to its delivery station. The loading and unloading of a transport unit takes a non-negligible amount of time.

In addition to transport requests, we include relocation requests to allow empty rail carriers to move between stations. Relocations helps to balance carrier availability with the demand and may occur either before the first assigned transport request or between consecutive transport requests. Relocation requests are derived from the combination of transport requests and the initial locations of the rail carriers.

2.2. Mathematical model

In this section, we propose a multi-objective MILP-based optimisation model for the rail carrier assignment and relocation problem. The three main components of the model are the decision variables, the objective function and the constraints.

2.2.1. Sets, parameters and variables

The sets, parameters and variables of the model are listed in Table 1. The sets represent collections of elements used in the model notation, the parameters define fixed input values, and the variables correspond to decisions made within the optimisation model.

The main sets are the set of requests R, the set of rail carriers C, the set of stations S, and the set of time units T. The set of requests is partitioned into two subsets: the set of transport requests $R^t \subset R$ and the set of relocation requests $R^l \subset R$.

Most parameters relate to the requests. $s_r^p \in S_r$ and $s_r^d \in S_r$ indicate the pickup and delivery station of request $r \in R$. $t_r^{p,\text{early}} \in T$ and $t_r^{p,\text{late}} \in T$ indicate the earliest and latest pickup time of request $r \in R$, and $t_r^{d,\text{early}} \in T$ and $t_r^{d,\text{late}} \in T$ indicate the earliest and latest delivery

Table 1: Sets, parameters and variables for the MILP model formulation.

Sets	
R	set of requests
$R^t \subset R$	set of transport requests
$R^l \subset R$	set of relocation requests
C	set of rail carriers
S	set of stations
T	set of time units
Parameters	
$s_r^p \in S$	pickup station of request $r \in R$
$s_r^d \in S$	delivery station of request $r \in R$
$t_r^{p,\text{early}} \in T$	earliest pickup time of request $r \in R$
$t_r^{p,\text{late}} \in T$	latest pickup time of transport request $r \in R^t$
$t_r^{d,\text{early}} \in T$	earliest delivery time of transport request $r \in R^t$
$t_r^{d,\text{late}} \in T$	latest delivery time of request $r \in R$
$\tau_r^{\text{load}} \in T$	loading time of request $r \in R$ (= 0 if $r \in R^l$)
$\tau_r^{\text{unload}} \in T$	unloading time of request $r \in R$ (= 0 if $r \in R^l$)
$\tau_r \in T$	minimum travel time of request $r \in R$ from pickup to delivery station
$\tau_{s,s'} \in T$	travel time from station $s \in S$ to adjacent station $s' \in S$
$l_{c,s}^{\text{init}} \in \{0,1\}$	$= 1$ if carrier $c \in C$ is initially located at station $s \in S$
$n_s \in \mathbb{N}$	carrier capacity at station $s \in S$
$M \in \mathbb{R}_+$	a large constant
$w_t, w_l \in \mathbb{R}_+$	weight for transport/relocation requests handling in objective function
$w_r \in \mathbb{R}_+$	handling priority weight of request $r \in R$
Variables	
$t_r^p \in T$	pickup time of request $r \in R$
$t_r^d \in T$	delivery time of request $r \in R$
$x_{r,c} \in \{0,1\}$	$= 1$ if request $r \in R$ is assigned to carrier $c \in C$
$y_r \in \{0,1\}$	$= 1$ if request $r \in R$ is handled
$z_{c,s,s',t}^c \in \{0,1\}$	$= 1$ if carrier $c \in C$ is departing from station $s \in S$ to station $s' \in S$ at time $t \in T$
$z_{r,s,s',t}^r \in \{0,1\}$	$= 1$ if request $r \in R$ is departing from station $s \in S$ to station $s' \in S$ at time $t \in T$
$o_{r,r',c} \in \{0,1\}$	$= 1$ if request $r \in R$ is handled before $r' \in R$ by carrier $c \in C$
$o_{r,r',c}^d \in \{0,1\}$	$= 1$ if request $r \in R$ is handled directly before $r' \in R$ by carrier $c \in C$
$b_{r,r',r'',c} \in \{0,1\}$	$= 1$ if request $r'' \in R$ is handled between requests $r, r' \in R$ by carrier $c \in C$
$l_{c,s} \in \{0,1\}$	$= 1$ if carrier $c \in C$ is located at station $s \in S$ after handling last request

time of request $r \in R$. Note that $t_r^{p,\text{late}}$ and $t_r^{d,\text{early}}$ are only defined for transport requests $r \in R^t$. $\tau_r^{\text{load}} \in T$ and $\tau_r^{\text{unload}} \in T$ represent the loading and unloading time for transport request $r \in R^t$. $\tau_r \in T$ indicates the minimum time to travel between the request's pickup and delivery station.

Other parameters relate to various things. $\tau_{s,s'} \in T$ indicates the travel time from station $s \in S$ to adjacent station $s' \in S$. $l_{c,s}^{\text{init}} \in \{0,1\}$ indicates whether or not station $s \in S$ is the initial location of rail carrier $c \in C$. $n_s \in \mathbb{N}$ indicates the capacity of stations $s \in S$. Finally, we define $M \in \mathbb{R}_+$ to be a large constant, e.g., $|T|$, and weights w_t, w_l and w_r for the objective function. w_t and w_l are the weights for the two different objectives in the objective function, related to the handling of transport and relocation requests, respectively. w_r is the priority weight of transport request $r \in R^t$.

The model variables are mostly binary variables. The exception is formed by the timing variables $t_r^p \in T$ and $t_r^d \in T$ which determine the pickup and delivery time of request $r \in R$. The main binary variables are the assignment decision variables $x_{c,r},\ r \in R$ indicating whether request $r \in R$ is assigned to rail carrier $c \in C$, and the auxiliary handling variables $y_r,\ r \in R$ indicating whether request $r \in R$ is handled. Flow variables $z_{c,s,s',t}^c \in \{0,1\}$ and $z_{r,s,s',t}^r \in \{0,1\}$ indicate whether rail carrier $c \in C$ or request $r \in R$ starts travelling from station $s \in S$ to stations $s' \in S$ at time $t \in T$. Ordering variables $o_{r,r',c} \in \{0,1\}$ and $o_{r,r',c}^d$ are defined to indicate whether request $r \in R$ is handled (directly) before request $r' \in R$ by rail carrier $c \in C$. Auxiliary variables $b_{r,r',r'',c} \in \{0,1\}$ indicate whether request $r'' \in R$ is handled in between requests $r, r' \in R$ by rail carrier $c \in C$. Finally, $l_{c,s} \in \{0,1\}$ indicates whether rail carrier $c \in C$ is located at station $s \in S$ after handling its last request.

2.2.2. Objective function

The primary objective of the rail carrier assignment and relocation problem is to handle as many transport requests as possible. As a secondary objective, the empty travelling of rail carriers should be limited to reduce unnecessary use of network capacity. The two objectives are captured in the objective function (Equation (1)) as the maximisation of the number of handled transport requests and the minimisation of the travel time of relocation requests. The former is included as the weighted sum of non-handled transport requests $(1 - y_r)$, in which the penalty weights w_r can, for example, be interpreted as priority factors for passenger over freight requests. The latter is included as a sum of handled relocation requests $(r \in R_l)$, weighted by their minimum travel time (τ_r). Weights w_t and w_l are included to represent the importance of the two respective objectives. To allow for relocation whenever it leads to the handling of at least one additional transport request, $w_t >> w_l$.

$$\text{minimise } w_t \sum_{r \in R^t} w_r (1 - y_r) + w_l \sum_{r \in R^l} \tau_r y_r \tag{1}$$

2.2.3. Constraints

The constraints of the rail carrier assignment and relocation model are given by Equations (2) to (32). Constraints 2 ensures that each request is assigned to exactly one rail carrier, if the request is handled.

$$\sum_{c \in C} x_{r,c} = y_r \qquad \forall r \in R \tag{2}$$

The remaining constraints are categorised into the following groups: time window constraints, handling order constraints, traffic flow constraints, and network capacity constraints.

Time window constraints. Constraints (3), (4), (5) and (6) ensure that the pickup and delivery times of handled requests are in line with their time windows. For transport requests (Constraints (3) and (4)), this means between earliest and latest pickup/delivery times. For relocation requests (5) and (6), this means after earliest pickup and before latest delivery time.

$$t_r^{p,\text{early}} y_r \leq t_r^p \leq t_r^{p,\text{late}} y_r \qquad \forall r \in R^t \tag{3}$$

$$t_r^{d,\text{early}} y_r \leq t_r^d \leq t_r^{d,\text{late}} y_r \qquad \forall r \in R^t \tag{4}$$

$$t_r^{p,\text{early}} y_r \leq t_r^p \qquad \forall r \in R^l \tag{5}$$

$$t_r^d \leq t_r^{d,\text{late}} y_r \qquad \forall r \in R^l \tag{6}$$

Handling order constraints. Through Constraints (7), (8) and (9), it is ensured that requests that are handled by the same rail carrier are ordered. That is, the pickup time of the request handled later is after the delivery time of the request handled earlier.

$$t_r^d < t_{r'}^p + M\,(1 - o_{r,r',c}) \qquad \forall r, r' \in R,\ c \in C \tag{7}$$

$$t_{r'}^d < t_r^p + M\,\left(3 - (1 - o_{r,r',c}) - x_{r,c} - x_{r',c}\right) \qquad \forall r, r' \in R,\ c \in C \tag{8}$$

$$o_{r,r',c} \leq 0.5 x_{r,c} + 0.5 x_{r',c} \qquad \forall r, r' \in R,\ c \in C \tag{9}$$

With the general order per pair of requests handled by the same rail carrier established, the direct order of requests per rail carrier can be derived. As Constraints (10) to (16) describe, requests are directly ordered if and only if they are ordered without any requests in between.

$$t^p_{r''} > t^d_r - M\,(1 - b_{r,r',r'',c}) \qquad \forall r, r', r'' \in R,\ c \in C \qquad (10)$$

$$t^d_{r''} < t^p_{r'} + M\,(1 - b_{r,r',r'',c}) \qquad \forall r, r', r'' \in R,\ c \in C \qquad (11)$$

$$o^d_{r,r',c} + b_{r,r',r'',c} \leq 1 \qquad \forall r, r', r'' \in R,\ c \in C \qquad (12)$$

$$o^d_{r,r',c} \leq o_{r,r',c} \qquad \forall r, r' \in R,\ c \in C \qquad (13)$$

$$o^d_{r,r',c} \geq o_{r,r',c} - \sum_{r'' \in R} b_{r,r',r'',c} \qquad \forall r, r' \in R,\ c \in C \qquad (14)$$

$$\sum_{r' \in R} o^d_{r,r',c} \leq 1 \qquad \forall r \in R,\ c \in C \qquad (15)$$

$$\sum_{r' \in R} o^d_{r',r,c} \leq 1 \qquad \forall r \in R,\ c \in C \qquad (16)$$

Traffic flow constraints. Constraints (17) to (22) ensure the alignment of pickup and delivery stations along a rail carrier route. For requests that are directly ordered on a rail carrier, Constraints (17) and (18) ensure that the delivery station of the first request is the pickup station of the second one. Constraints (19) and (20) ensure that the pickup station of the first request handled by a rail carrier is its initial location. Similarly, Constraints (21) and (22) ensure that the delivery station of a rail carrier's last request is its final location.

$$s^p_{r'} \geq s^d_r - M\,(1 - o^d_{r,r',c}) \qquad \forall r, r' \in R,\ c \in C \qquad (17)$$

$$s^p_{r'} \leq s^d_r + M\,(1 - o^d_{r,r',c}) \qquad \forall r, r' \in R,\ c \in C \qquad (18)$$

$$s^p_r \geq \sum_{s \in S} l^{\text{init}}_{c,s} - M\,(1 - x_{r,c} + \sum_{r' \in R} o^d_{r',r,c}) \qquad \forall r \in R,\ c \in C \qquad (19)$$

$$s^p_r \leq \sum_{s \in S} l^{\text{init}}_{c,s} + M\,(1 - x_{r,c} + \sum_{r' \in R} o^d_{r',r,c}) \qquad \forall r \in R,\ c \in C \qquad (20)$$

$$s^d_r \geq \sum_{s \in S} l_{c,s} - M\,(1 - x_{r,c} + \sum_{r' \in R} o^d_{r,r',c}) \qquad \forall r \in R,\ c \in C \qquad (21)$$

$$s^d_r \leq \sum_{s \in S} l_{c,s} - M\,(1 - x_{r,c} + \sum_{r' \in R} o^d_{r,r',c}) \qquad \forall r \in R,\ c \in C \qquad (22)$$

Complementary to the carrier flow in terms of stations, Constraints (23) to (26) address the flow of requests in terms of times. Constraints (23) ensure that the departure times at

subsequent stations along the route of a request are at least the travelling time between the stations apart. Constraints (25) and (26) ensure the alignment of the pickup time with the departure time of the pickup station and the alignment of the delivery time with the departure towards the delivery station.

$$\sum_{t\in T}\sum_{s'\in S}(t+\tau_{s',s})z^r_{r,s',s,t} \le \sum_{t\in T}(t\ z^r_{r,s',s,t}) \qquad \forall r\in R,\ s\in S\backslash\{s^p_r,s^d_r\} \qquad (23)$$

$$\sum_{t\in T}\sum_{s'\in S} z^r_{r,s,s',t} = \sum_{t\in T}\sum_{s'\in S} z^r_{r,s,s',t} \qquad \forall r\in R,\ s\in S\backslash\{s^p_r,s^d_r\} \qquad (24)$$

$$t^p_r + \tau^{\text{load}}_r y_r \le \sum_{t\in T}\sum_{s\in S}(t\ z_{r,s^p_r,s,t}) \qquad \forall r\in R \qquad (25)$$

$$t^d_r - \tau^{\text{unload}}_r y_r \ge \sum_{t\in T}\sum_{s\in S}(t\ z_{r,s,s^d_r,t}) \qquad \forall r\in R \qquad (26)$$

As the models features timing variables for both rail carriers and requests, these are aligned by Constraints (27) and (28).

$$z^c_{c,s,s',t} \ge z^r_{r,s,s',t} - (1-x_{r,c}) \qquad \forall r\in R,\ c\in C,\ s,s'\in S,\ t\in T \qquad (27)$$

$$z^c_{c,s,s',t} \le z^r_{r,s,s',t} + (1-x_{r,c}) \qquad \forall r\in R,\ c\in C,\ s,s'\in S,\ t\in T \qquad (28)$$

Constraints (29) ensure that the number of rail carriers starting from or travelling to a certain station equals the number of rail carriers travelling from or ending at that station.

$$\sum_{c\in C} l^{\text{init}}_{c,s} + \sum_{c\in C}\sum_{s',s\in S}\sum_{t\in T} z^c_{c,s',s,t} = \sum_{c\in C}\sum_{s,s'\in S}\sum_{t\in T} z^c_{c,s,s't} + \sum_{c\in C} l_{c,s} \qquad \forall s\in S \qquad (29)$$

Network capacity constraints. Line capacity is addressed by Constraints (30) and (31): at most one rail carrier can leave or enter the same station within a time unit. Station capacity is addressed by Constraints (32): the number of rail carriers that started at or have entered but not left a station does not exceed that station's capacity at any time unit.

$$\sum_{c\in C}\sum_{s'\in S} z^c_{c,s,s',t} \le 1 \qquad \forall s\in S,\ t\in T \qquad (30)$$

$$\sum_{c\in C}\sum_{s'\in S} z^c_{c,s',s,t} \le 1 \qquad \forall s\in S,\ t\in T \qquad (31)$$

$$\sum_{c\in C} l^{\text{init}}_{c,s} + \sum_{r\in R}\sum_{s'\in S}\sum_{t'\le t} z^c_{c,s',s,t} - \sum_{r\in R}\sum_{s''\in S}\sum_{t'\le t} z^c_{c,s,s'',t} \le n_s \qquad \forall s\in S,\ t\in T \qquad (32)$$

3. Application to a small-scale case study

To illustrate the proposed model for the rail carrier assignment and relocation problem, we analyse its performance using a small-scale case study. We consider a railway network consisting of four stations which are interconnected with five double track lines with fixed travel times as shown in Figure 2. We assume a station capacity of three rail carriers for each station. In our analysis later on, we vary the number of rail carriers available in the network, along with their initial locations.

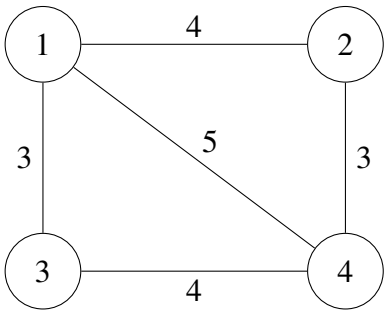


Figure 2: Schematic representation of small case study network.

A period of 30 time units is considered, during which five transport requests of equal priority are scheduled. The transport requests and their pickup and delivery characteristics are listed in Table 2. Given the transport requests, a first set of relocation requests is generated, i.e. relocation requests connecting transport requests that could be handled subsequently. Based on the time windows and stations, relocation requests are generated to connect transport requests 4 and 5 with transport requests 2 and 3. These four 'intermediate' relocation requests and their characteristics are added to Table 2 as fixed requests. Additional 'initial' relocation requests are defined per analysis scenario as they depend on the varying initial carrier locations.

As said, we define a set of scenarios by varying the number of rail carriers (two or three), and their initial locations at a station. With two rail carriers, there are ten possibilities to distribute them over the four station, including the four options to have the same initial location. In the same way, there are twenty possibilities to distribute three rail carriers over the four stations. An overview of the possible scenarios is given in Table A1, together with the results in terms of handled transport requests and total empty travel times.

Table 2: Fixed requests for test case study.

Request	Pickup station	Delivery station	Pickup time window	Delivery time window
Transport requests				
1	4	2	[11, 18]	[14, 21]
2	1	3	[12, 18]	[15, 21]
3	3	2	[14, 19]	[21, 26]
4	1	4	[1, 5]	[8, 13]
5	3	4	[1, 7]	[5, 11]
Intermediate relocation requests				
6	4	1	[8, $\rightarrow$)	($\leftarrow$, 18]
7	4	3	[8, $\rightarrow$)	($\leftarrow$, 19]
8	4	1	[5, $\rightarrow$)	($\leftarrow$, 18]
9	4	3	[5, $\rightarrow$)	($\leftarrow$, 19]

With two rail carriers, two to four transport requests (3.4 on average) can be handled with three to 11 relocation time units (6.5 on average). The exact number as well as which of the requests that are handled depends on the initial carrier locations, though request 1 is always handled. In five scenarios, four requests (1, 3, 4 and 5) are handled, with three to 11 time units of relocation. Requests 4 and 5 are handled in parallel, after which one rail carrier can directly pick up request 1. For the second rail carrier, request 3 is preferred over request 2 due to its relative proximity to station 4 and its later time windows. In four scenarios, three requests are handled, always including requests 1 and 2, with four or eight relocation time units. Request 4 or 5 is handled before request 1 depending on the initial carrier location. Only if there is a rail carrier at station 1, request 4 is handled. In the last scenario, in which both rail carriers are initially located at station 2, only two requests are handled. This is because there is no transport request with pickup station 2, in combination with the limited time. We note that for all scenarios with two rail carriers, a computation time of around two seconds can be reported.

With three rail carriers, three to five transport requests (4.65 on average) are handled with three to 15 relocation time units (10.05 on average). Requests 1, 2, and 3 are handled in all scenarios. In 14 out of 20 scenarios, all transport requests can be handled - with relocation time units varying between three and 15. Of the remaining six scenarios, five involve handling four requests and one involves handling three requests. In the scenarios where four requests are handled, either request 4 or 5 is rejected, depending on whether

a rail carrier is initially located at station 1. The scenario in which only three requests are handled is the one with all rail carrier initially located at station 2, where no transport requests are to be picked up. On that note, when all rail carriers are initially located at either station 1 or station 3, where two transport requests are to be picked up at different times, all requests are handled. With all rail carriers initial located at station 4 with one scheduled pickup at a later time, four request are handled. It is worth noting that for some scenarios the computation time is significant. While typical computation times lie around five seconds, for several scenarios require between 15 and 20 seconds, and one scenario even takes around 40 seconds.

Based on the results of these scenarios, we conclude that both the number of rail carriers and the initial location of the rail carriers are key factors for the model objectives of, first, maximising the number of handled transport requests and, second, minimising the total relocation time. As shown in Figure 3, there is no clear relation between the two objectives. For the two-carrier scenarios, the relation is slightly positive, while for the three-carrier scenarios, the trend is slightly negative.

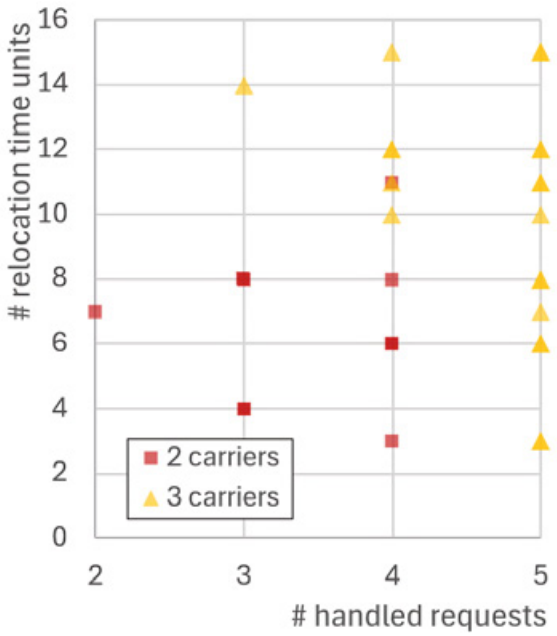


Figure 3: Total relocation time versus number of handled transport requests.

To conclude, we acknowledge the potentially strong effects of the considered case study, e.g., by the small network and short time period. To obtain more general results, application of the model to a variety of realistic case studies should be considered.

4. Conclusions

In this paper, we proposed a mathematical optimisation model for the rail carrier assignment and relocation problem in the context of a multimodal pod system. The model assigns rail carriers to transport units with specified pickup and delivery stations and time windows, while accounting for constraints of the railway network. The objective is to maximise the number of handled transport units while minimising the total relocation time of rail carriers.

The model is demonstrated using a small case study involving a four-station railway network and five transport units. A preliminary analysis showed the significant impact of the number of available rail carriers and their initial distribution within the network on both the (number of) handled transport units and the total relocation time. While these findings illustrate the model's capabilities and provide initial insights into its behaviour, they are limited by the simplified nature of the consider case study. To asses the model's practical applicability, it is essential to evaluate it on more realistic scenarios.

In future steps, we will explore the use of soft time windows, in combination with the possibility to form platoons of rail carriers. Furthermore, the model will be applied to a more practical case study in terms of both network size and time horizon. Given the already observed increase in computational effort, this will require the adoption of a more advanced solution method.

Acknowledgment

This research has received funding from the Europe's Rail Joint Undertaking under grant agreement No. 101121853 (Pods4Rail). The project is supported by the Europe's Rail Joint Undertaking and its members, funded by the European Union. Views and opinions expressed are those of authors only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Literature

[1] Di Febbraro, A. ; Sacco, N. ; Saeednia, M.: One-way carsharing: solving the relocation problem. In: *Transportation Research Record*, 2319(1): 113-120, 2012.

[2] Kek, A. G. ; Cheu, R. L. ; Meng, Q. ; Fung, C. H.: A decision support system for vehicle relocation operations in carsharing systems,. In: *Transportation Research Part E: Logistics and Transportation Review*, 45(1): 149-158, 2009.

[3] Liao, X. ; Han, J. ; Saeednia, M. ; Paz Martinez, A.: Unlocking the Potentials of Modularity in Railways, a Heuristic Framework for Pods Scheduling. In: *IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)*, 2024.

[4] Weikl, S. ; Bogenberger, K.: Relocation strategies and algorithms for free-floating car sharing systems. In: *IEEE Intelligent Transportation Systems Magazine*, 5(4): 100-111, 2013.

A. Detailed case study results

Table A1: Handled transport requests and total empty travel time per scenario.

Initial carrier locations (stations)	**Handled transport requests** (number: list)	**Total empty travel time** (time units)
2 carriers	*3.4*	*6.5*
1, 1	4: 1, 3, 4, 5	6
1, 2	3: 1, 2, 4	4
1, 3	4: 1, 3, 4, 5	3
1, 4	4: 1, 3, 4, 5	8
2, 2	2: 1, 2	7
2, 3	3: 1, 2, 5	4
2, 4	3: 1, 2, 5	8
3, 3	4: 1, 3, 4, 5	6
3, 4	4: 1, 3, 4, 5	11
4, 4	3: 1, 2, 5	8
3 carriers	*4.65*	*10.05*
1, 1, 1	5: 1, 2, 3, 4, 5	6
1, 1, 2	5: 1, 2, 3, 4, 5	11
1, 1, 3	5: 1, 2, 3, 4, 5	3
1, 1, 4	5: 1, 2, 3, 4, 5	8
1, 2, 2	5: 1, 2, 3, 4	11
1, 2, 3	5: 1, 2, 3, 4, 5	12
1, 2, 4	5: 1, 2, 3, 4, 5	3
1, 3, 3	5: 1, 2, 3, 4, 5	8
1, 3, 4	5: 1, 2, 3, 4, 5	12
1, 4, 4	5: 1, 2, 3, 4, 5	14
2, 2, 2	3: 1, 2, 3	10
2, 2, 3	4: 1, 2, 3, 5	15
2, 2, 4	4: 1, 2, 3, 5	10
2, 3, 4	5: 1, 2, 3, 4, 5	15
2, 4, 4	4: 1, 2, 3, 5	12
3, 3, 3	5: 1, 2, 3, 4, 5	6
3, 3, 4	5: 1, 2, 3, 4, 5	11
3, 4, 4	5: 1, 2, 3, 4, 5	15
4, 4, 4	4: 1, 2, 3, 5	12

Session 17
Wolf-Dietrich Geitz, Upendra Atre

Impact of incorrect and incomplete transport statistical data: A Case Study of the Rhineland – ZARA Ports Corridor

Geitz, Wolf-Dietrich[1], Atre, Upendra[1]

[1]Railistics GmbH

Session 17
Wolf-Dietrich Geitz, Upendra Atre

Abstract

This paper is based on a study commissioned by IHK-Initiative Rheinland GbR (IIR) and conducted by Railistics GmbH. It highlights how incorrect, and incomplete are the transport statistics, used in the Federal Transport Infrastructure Plan (BVWP), which may in turn lead to a significant underestimation of future rail freight demand between the Rhineland and the ZARA ports (Zeebrugge, Antwerp, Rotterdam, Amsterdam). The major Rhineland corridors that we are taking into consideration are Aachen-Düsseldorf, Aachen-Bonn-Rhine-Sieg, Venlo-Köln, Venlo-Duisburg and Emmerich-Duisburg. Capacity bottlenecks will persist despite current infrastructure investment plans because of the miscalculated mode transitions and underestimated train numbers. These errors in the federal statistics (Destatis) were previously uncovered by the Bundesnetzagentur (Federal Network Agency) and Railistics GmbH which is also further supported by the press release by the Verkehrsforum. The Rhineland corridors are a vital part for German and European freight transport. Without a strategic expansion of rail infrastructure, the region would have risks in losing competitiveness, missing climate targets, and suffering environmental setbacks. A strong modal shift from road to rail, especially considering rising intermodal traffic and port investments, is crucial for future-proofing the corridor. The federal government and the state of North Rhine-Westphalia must strengthen the Rhineland region in such a way that it can secure its strategic position in European freight transport by expanding the rail infrastructure. This is the only way to continue to benefit from the positive effects of international trade.

This paper summarizes the fact that even with the planned infrastructure measures, there would still be bottlenecks in the network due to the increasing volumes from ZARA.

Keywords: ***Incorrect statistics; Rhineland – ZARA corridor; consequences of inaccurate and incomplete statistics; forecast***

Session 17
Wolf-Dietrich Geitz, Upendra Atre

1 Introduction & Background

Transportation is the backbone of economic growth, sustainability, and public interaction, for which accurate data forecasting is crucial. On the contrary, inaccurate forecasts can have detrimental effects on certain areas. This is the case for the Federal Transport Infrastructure Plan (BVWP 2030), where there is an underestimation of rail freight demand in western NRW due to outdated and incomplete statistics. Studies by IVV and others confirm that rail corridors between the Rhineland and the ZARA ports (Zeebrugge, Antwerp, Rotterdam, Amsterdam) are almost up to their capacity, with forecasts for 2030 having already been reached by 2019. The study was ordered by IHK IIR and conducted by Railistics GmbH in 2024 in corporation with SSP Consult.

According to Dr. Jörg Mosolf, Chairman of the Presidium of the German Transport Forum and Chairman of the Board of Directors, "The explosive thing about the false statistics is that the volumes in rail freight transport are needed for the benefit-cost assessments of infrastructure projects as well as for the examination of funding applications, for example for rail sidings and combined transport projects." [1]

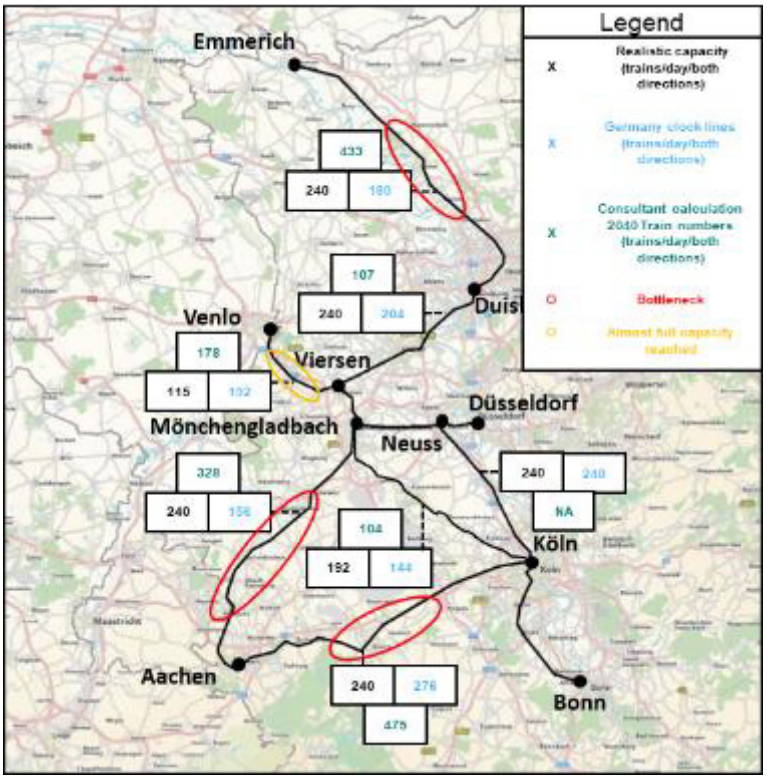


Figure 1: Demand-Capacity forecast comparison 2040 (Source: Railistics GmbH)

In the corridors between Rhineland – ZARA Ports, the Position's paper[2] published by IIR assumes that the traffic demand will rise and even after implementing all planned infrastructure measures in BVWP 2030, there will still not be sufficient capacity by 2040. On the other hand, in the fall 2024, the federal government forecasted

[1] Whistleblower DVF
[2] Bedeutung der ZARA-Häfen für die Infrastruktur im Rheinland Schieneninfrastruktur ertüchtigen, IHK IIR, Schieneninfrastruktur ertüchtigen IIR Positionspapier - IHK Düsseldorf, 24th March, 2025

lower volumes from ports of Rotterdam and Antwerp for 2040.[3] Both these ports have actually set ambitious targets for increasing their rail modal share. Moreover, the number of freight trains running in 2019 along the ZARA corridors was already higher than expected projection in the forecast data for 2030.[4] Additionally, rail companies relevant to this area also stated that transit from Benelux through Germany to other countries has developed even more strongly than the source and destination market to and from Germany. All the analysis and calculations are carried out based on the data available till September 2024.

Acknowledging the fact that the current congestion on these rail routes is already threatening to slow down the economic development in the region as freight transport connections are a key factor for companies in the Rhineland. When any kind of major infrastructure works occur in the Aachen – Köln route, both the freight and consequently, insufficient capacity will lead to increasing competition between passenger and freight transport by rail. This will also particularly have an impact on the international connections to the ZARA ports, which play the most important role for the export industry in North Rhine-Westphalia. Additionally, traffic on the roads will also gradually increase, which not only pollutes the environment but also impairs the quality of life of people in the Rhineland. The stakeholder consultation, which included the port of Rotterdam, Port of Antwerp, go.Rheinland, ProRail, Bundesnetzagentur, etc., showed that the biggest bottleneck is infrastructure and capacity and there were huge discrepancies between the forecasts from different sources. Demand assessment also showed a higher number of trains in 2019 than those forecasted in BVWP for 2030. The current forecasts by the federal government have systematically underestimated the intermodal trains and the trains of foreign railroad companies on the routes in the Rhineland in the past. Moreover, the truck-trailer business on the railways and the associated investments in recent years are also underestimated. Apart from that, the federal government underestimates the impact of past, present and future investments in the ZARA ports. Sections such as Venlo–Viersen are already overloaded, and freight forwarders are rerouting traffic via France and Belgium due to German corridor constraints, posing risks to the Rhineland's competitiveness.

In Figure 1, different approaches for capacity increase with planned investments are compared with the realistic demand increase. It shows that the current planned investments still produce bottlenecks. Projects like the Euregio Railport in Stolberg[5],

[3] Port of Rotterdam freight volumes rise 15% as economy recovers | Reuters
[4] Federal Transport Infrastructure Plan 2030 | BMV
[5] From road to rail: Freight transshipment point in Stolberg | We. Life. Eifel.

being a structurally strengthening project for the Rhineland and expansion of the Betuwe route further highlights the need for enhanced capacity.

The development forecasts should be corrected and further developed over the period under consideration up to 2040 using the current data. The estimates for the development of train numbers to and from the ZARA ports are also decisive for this. So, this paper helps to assess the current and forecasted traffic volumes in key corridors, validates official projections, and stresses the urgent need for updated planning to avoid long-term disadvantages and ensure alignment with climate goals.

2 Objectives and Scope of study

The primary objective of this study is the assessment of the capacity of the cross-border freight infrastructure between Germany, especially the region of North-Rhein-Westphalia (NRW) region and the ZARA ports. Ensuring adequate rail connection to these ports will be vital for maintaining the long-term growth and economic competitiveness of freight transportation as these ports have a major influence for managing import and export of commodities.

The study aims to address the following queries;

- What is the current status of the infrastructure and the existing number of trains operating on the main cross-border routes that link the ports of ZARA and NRW?
- What are the main infrastructure constraints that limit the growth of rail-freight transport and also the constraints for capacity?
- To what extent does the predictions and forecasts made by the Federal Transport Infrastructure Plan (Bundesverkehrswegeplan, BVWP) on freight volumes vary in the above-mentioned corridors?

The study makes it clear that the growth in rail freight traffic flows from the ports of Zeebrugge, Antwerp, Rotterdam and Amsterdam (ZARA ports) will be significantly higher than assumed in the federal government's current transport forecast for 2040. So, forecast evaluation, demand assessment and plausibility check are made for the corridors which are the potential connections to the ZARA ports. For Antwerp and Rotterdam alone, the authors forecasts at least two thirds more freight volumes relevant to Germany than the federal government for all foreign seaports. The expected increases particularly observed to be affected in the following central corridors:

- Emmerich - Duisburg
- Venlo – Viersen – Mönchengladbach - Neuss - Köln

- Venlo – Viersen – Duisburg
- Aachen – Mönchengladbach – Neuss - Düsseldorf
- Aachen – Köln – Bonn

3 Methodology

3.1 Data collection and Analysis

As a part of the data collection process, stakeholders were consulted in two different ways. Initially an online survey was handed out to the stakeholders which included a large group of railway companies, port authorities, intermodal operators, logistics companies, freight forwarders, etc., and responses were noted according to the nations in which they operate. Subsequently, the ports of Emmerich, Rotterdam, go.Rheinland, ProRail, The Federal Network Agency and the Dutch province of Limburg were specifically interviewed after they completed the online survey. Each of these companies were able to submit the survey reactions according to their industry, which spans over nine European and Asian countries.



Figure 2: List of Stakeholders

In total 16 different stakeholders were interview/surveyed for a time period that started from July up until September 2024. The Survey was designed in a way to focus on the major concerns of the stakeholders regarding the study corridors with an emphasis on key bottlenecks identified, areas of improvement, ineffective share of rail, traffic flow in last 5 years, their willingness to support and expectations about the corridors mentioned. The results of the survey reflected that, for most of the stakeholders involved, infrastructure

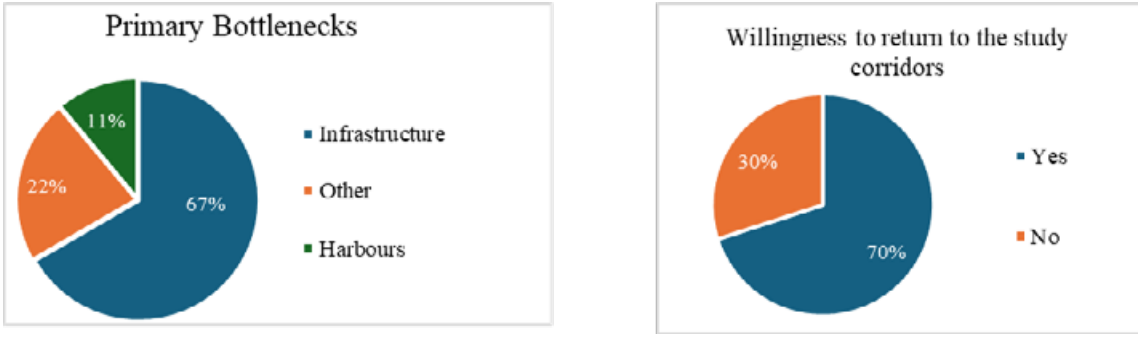


Figure 3: Stakeholder concerns and willingness (Source: Railistics GmbH)

was identified as the biggest bottleneck and the area where the greatest improvement is needed.

The survey also showed that, for the majority of participants, traffic volumes (all modes of transport) had remained at least constant or even increased over the past five years (only very few reported a negative trend) and that there was a willingness to use the German rail corridors once the bottlenecks had been eliminated.

In the light of the findings from the stakeholder interviews, following conclusions were determined;

- In 2023, the momentarily decline in the freight volumes due to temporary bottlenecks is slightly getting recovered by now as the numbers are already indicating a rebound in 2024.
- Long distance traffic on the relevant corridors has increased notably.
- Infrastructure and Capacity are considered as the biggest bottlenecks.
- Stakeholders have made it apparent that they have the willingness to use the corridors again, once the prevailing issues are resolved.
- A revision has been made by the DB Cargo in the calculation method for freight handling from minimum distance to the actual distance, which leads to a rise of 10-15% in the tonne-km.
- Statistical number of freight trains are underestimated.

The selective surveys have shown that the decline in 2023 is attributable to temporary shortcomings. (e.g. numerous construction sites on the rail infrastructure in the Rhineland, strikes in France and Germany, disrupted the rail freight transport operations, the consequences of accidents in the Gotthard Tunnel, system-related congestion and capacity bottlenecks, the allocation of diversion capacity during the closure at Rastatt, the effects of the energy crisis, etc. are all included as the contributing factors to the issue).

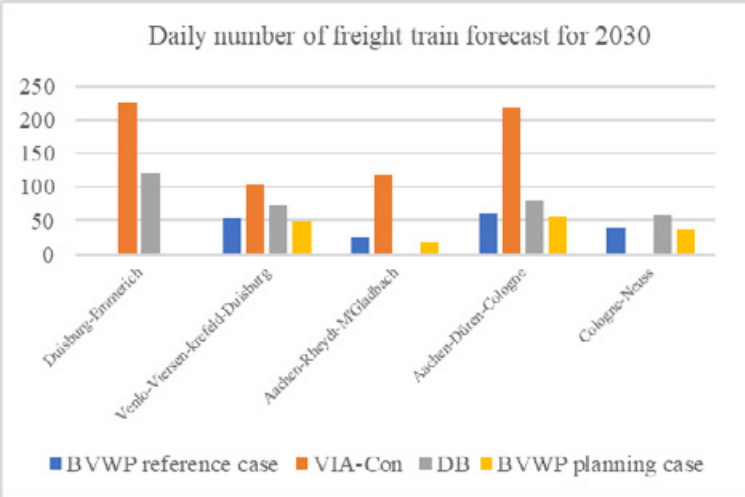


Figure 4: Comparison of the forecasts (Source: Railistics GmbH)

Although the market decline in the years of 2022/23 is attributable to these temporary disruptions, long-distance traffic in the Benelux region and east-west traffic are picking up again, which indicates that performance will improve again. Intermodal NL-PL and NL-IT traffic is increasing despite capacity and infrastructure problems. The railway companies relevant to the area under review have increased their traffic more than the market as a whole. Transit from the Benelux countries through Germany to other countries has developed even better than the source and destination market to or from Germany. This also calls into attention whether long-term capacity planning is enough.

However, recent estimates show notable differences in the anticipated number of freight trains in 2030 along the important western German corridors. The following comparison illustrates the 2030 forecast numbers from three main sources: DB, ViaCon and BVWP. Although the BVWP forecast show the low numbers, ViaCon and DB estimations show significantly higher demand. For the Projection analysis of BVWP, a reference case as well as a planning case is considered. The discrepancies among the three underline the danger of underestimating future capacity needs and the necessity of bringing the federal projections into line with current and projected increase in the number of freight trains.

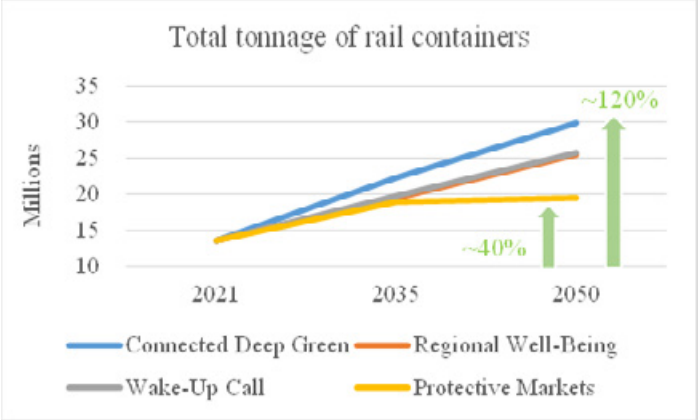


Figure 5: Forecast for the port of Rotterdam

The forecasts from various sources for the number of freight trains in 2030 vary widely. All four forecast scenarios developed by the Port of Rotterdam[6] anticipate an increase from 11.5% in 2021 to a peak of 15.2% in 2050. Even with this modest growth in the share of rail freight transport, the number of container trains in 2050 will increase by at least

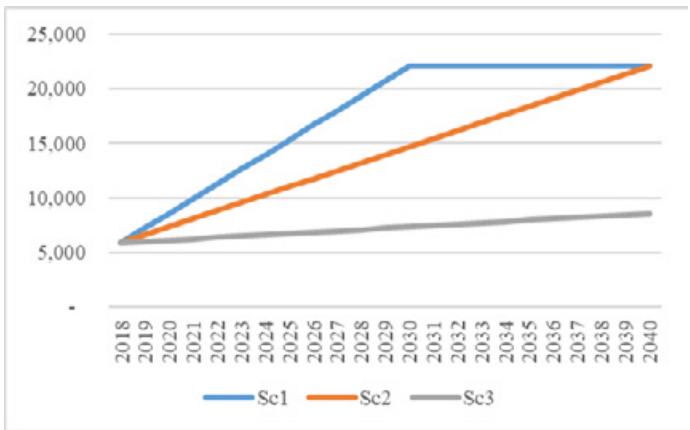


Figure 6: Forecast for the port of Antwerp

[6] Scenarios for 2050 | Port of Rotterdam Authority

40% and to a maximum of 120% compared to the 2021 scenario. Making an assumption that the share of trailer trains in total container trains is 30%, these figures could be even higher. The reason for this is that almost all trailers are intended for international transport. An even higher growth in the number of trains is possible if the political target scenario with a 25% of modal share for rail is achieved in 2040.

On the other hand, for the port of Antwerp, three scenarios were considered in which scenario 1 assumes that the port will achieve its target of 15% rail freight transport share by 2030, whereas for scenario 2 (a more realistic one), it states that the target will be achieved by 2040. And scenario 3 (the most pessimistic) assumes that the target will not be achieved by 2040. Even if the maximum growth scenario seems to be a little too optimistic at the moment, it is not regarded as purely theoretical. The authors believes that the reality lies somewhere between scenario 1 and scenario 2.

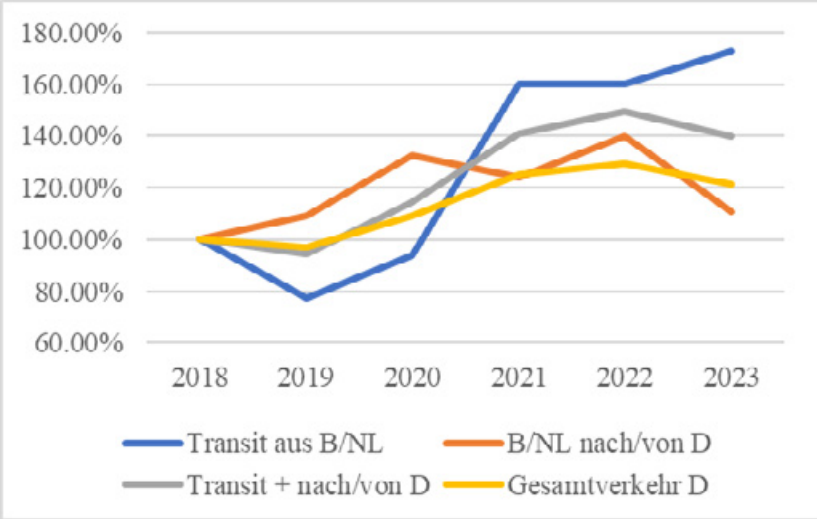


Figure 7 : BNetzA figures for transit traffic

Based on the UIRR (International union for road-rail combined transport) Report of 2023-24[7], it is also visible that the intermodal freight in Europe has faced a number of operational difficulties in 2023. Due to the numerous construction sites on the railway infrastructure, there were capacity bottlenecks and had a major effect on the freight train punctuality. Along with that, scheduled flows across the key corridors were disrupted by the strikes in France and Germany. The network was further strained by crisis-related problems like accident investigations, industrial site congestion, capacity bottlenecks and also by the effects on energy crisis. The statistics from the Bundesnetzagentur shows a similar trend on the performance of relevant energy supply companies in the corridors that has been analysed, as that of the UIRR report. While the overall rail transit across Germany declined in 2023, the transit from Belgium and Netherlands to Germany performed noticeably better than the overall market which is depicted in the Figure 7.

[7] UIRR Report 2023-24

Although the data related to the foreign railway undertakings are not recorded by Destatis, transit through Germany even better than in/out, similar to the UIRR.

All the data gathered from the forecasts shown in Figure 4, Figure 5, Figure 6 and Figure 7 were then used for the calculation of train numbers in each of the scenarios mentioned below.

3.2 Assessment of Demand

The freight train loads on the key German corridors connecting to the ZARA ports are thoroughly examined in order to gain a better understanding of the capacity issues and future demands of international rail freight transport. Three stages are included in this comparison, the actual traffic situation in 2019, BVWP forecasted values for 2030, and a trend-based development scenario with modified framework conditions by using a traffic model and planning assumptions into account in cooperation with SSP Consult GmbH. To identify bottlenecks, capacity gaps and potential reasons for forecast errors, each of the five corridors is analysed separately. For the Rhineland's economy, the ZARA ports are extremely significant. They serve as North Rhine-Westphalia's entry point to international markets and account for around 82% of seaport hinterland transportation out of all forms of transportation in NRW. As an industrial and logistical hub, the Rhineland region is an important source and destination area for freight transport. The Rhineland is also one of the most significant centres for the movement of products within Europe and beyond the borders. Due to its advantageous geographic location and current logistical capabilities, it is a vital hub for both domestic and international trade. The region's economic strength is greatly enhanced by its proximity to the most prominent European waterways, road and rail connections, and the trans-European networks (TEN-T).

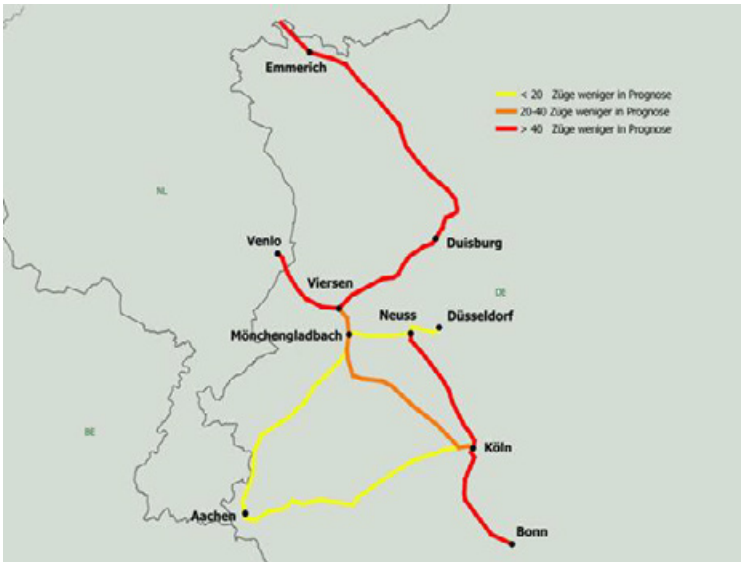


Figure 8: Comparison of BVWP values with the actual values along the corridors

Back in 2017, the Rhineland chambers of Industry and commerce and go.Rheinland criticized the fact that the figures used as a basis for the BVWP 2030 maritime traffic forecast did not match the expected growth rates of the ZARA ports on the basis of the authors' report on the Aachen Rail Hub. There was a systematic underestimation of the traffic volumes to and from the ZARA ports. The current assessment of demand confirms the fear that this is also the case in the 2040 traffic forecast, which was published in Autumn 2024. This analysis concludes that the train numbers in the Rhineland corridors were already higher in 2019 than those predicted for 2030 in the BVWP analysis. This indicates that, using the Aachen-Köln route as an example, the BVWP 2030 assumed ten fewer freight trains per day in each direction between Aachen and Köln than actual number in 2019. This clearly outlines the fact that more goods and trains are currently using the Rhineland's corridors than were anticipated in the 2030 prediction, which serves as the foundation for the BVWP. Below are the corridors that we have analysed and taken into consideration for the comparison with the BVWP forecast.

Demands on politicians

The Chambers of Industry and Commerce in Rhineland are calling on the federal government to correct the traffic forecasts and to realign the federal transport infrastructure planning. The capacities on the affected rail corridors urgently need to be expanded in order to meet the actual demands. In concrete terms, this means:[8]

- Adjustment of the traffic forecast for 2040 to the real growth figures of the ZARA ports
- Prioritization of the expansion of central rail corridors in federal transport infrastructure planning
- Close cooperation with Belgium and the Netherlands to strengthen cross-border infrastructures
- Accelerated implementation of already planned projects, in particular under the Coal Regions Investment Act

The analysis provides significant insights into the possible underestimation in policy planning and the immediate need for capacity enhancements on these strategically critical corridors.

3.3 Infrastructure Capacity

[8] Federal Transport Infrastructure Plan | Chamber of Industry and Commerce Düsseldorf

The total capacity of the network with and without the planned investments was calculated using three different scenarios. The first method used DB Netz standards to determine capacity based on line speeds; the second method (which the authors considered realistic) assumed that a double-track main line has a total capacity of 240 trains[9] per day and that capacity doubles when converted from a single-track to a double-track line, but only increases by 20%[10] when a third track is introduced; and in the last method, the train paths provided by Deutschlandtakt were used, that is the capacity based on the Germany's synchronized paths are used. All figures are general values; the actual achievable capacity and operating quality depends on the specific timetable and must be specified with the help of Tool based scientific simulation of timetables (**Eisenbahnbetriebswissenschaftliche Untersuchungen**). The total capacity of the network for all three methods is shown in the Figure 9 below. When assessing the planned infrastructure investments, the capacity based on DB standards appears too high, while the train paths provided by Deutschlandtakt appear too low. So, the reality (if all infrastructure measures are implemented) lies somewhere in between. The three scenarios used for the 2040 capacity forecasts, using data collected from various stakeholder groups and some assumptions, are explained in a bit more detail below.

Scenario 1: Maximum Capacity

The DB standards were designed in an effort to direct traffic flows and thereby harmonise the speed and flow of traffic as much as possible. These capacity standards are absolute **'Maximum values'**. They can differ according to the train types, degree of delay, real timetables, route layouts, etc. The purpose of these is to design the routes according to their affiliation to one of the three network types, Passenger Transport (P), Freight

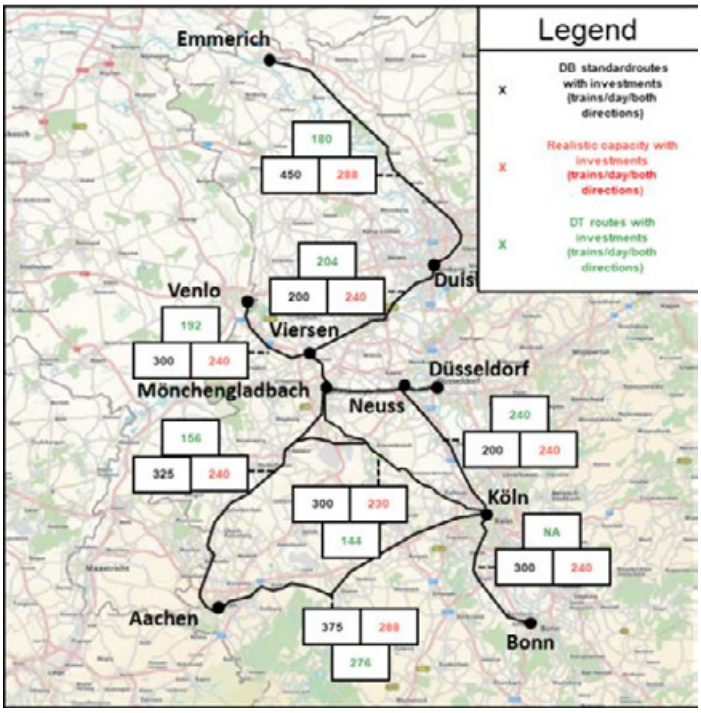


Figure 9: Various capacity scenarios (Source: DB standard routes, Railistics GmbH, DT Trassen 2030)

[9] Episode 1: The capacity of the existing line, Brenner North Approach
[10] 20 Percent | Die Güterbahnen

Transport (G) and Mixed Transport (M) in the case of expansion and new construction measures.

There are 2 cases considered for the scenario 1;

- Maximum capacity with no investments in the infrastructure. In this instance, DB standards are used to determine the maximum train movement on the current infrastructure. It is predicted on the absence of significant improvements or capacity-boosting initiatives. The outcome takes into account the highest possible theoretical use of the current infrastructure under ideal operating circumstances.
- While the former case is without any investments, a situation is taken into account in the other way around with strategic infrastructure investments. Here the values are calculated based on the planned infrastructure investments and expansion projects such as new routes, double tracking and signal system modifications.

Scenario 2: Realistic Capacity

In the realistic scenario, the authors assumed a smaller increase in the share of rail transport in the port of Rotterdam, extended the target year of the port forecast for Antwerp from 2030 to 2040 and assumed 65 TEU per train. This was done for reasons of realism and caution. There are 2 cases considered for this scenario;

- **Realistic capacity based on the number of tracks - Without investments:** Capacity was calculated based on current track infrastructure with the existing number of tracks. The authors' calculation is considered more realistic as it takes into account the quality of operation and the stability of the timetable as well as the network effect. In addition, the authors have made some changes to the calculation if the track section is not double or triple track throughout.

These capacity standards are based on a reliable, realistic and economical timetable. These are the international standard values.

- **Realistic Capacity based on number of tracks with investments:** Here a realistic capacity is calculated after incorporating the future and ongoing infrastructure modernization projects.

Certain assumptions were also considered while calculating the capacity;

Assumption 1: Capacity increases by 100% when switching from a single-track to a double-track line.

Assumption 2: Capacity increases by 20% if a double-track line is extended by one additional track.

List of infrastructure projects considered for Infrastructure capacity;

1. Continuation of the Betuwe line on the German side to Oberhausen
2. Expansion of the Köln - Mönchengladbach - Venlo line to two tracks throughout; construction of the third track Mönchengladbach - Rheydt
3. Aachen - Düren (Eilendorf section) third track extension and thus extension of the Burtscheid viaduct.
   a) Track plan adjustment at Aachen railway station
   b) Construction of an intersection in Langerwehe
   c) Additional block signal in each direction between Aachen Hbf-Aachen West
   d) Construction of an intersection in Stolberg
   e) Fast lanes in the Aachen- Mönchengladbach area
   f) Passing track between Aachen-Köln
   g) Track layout adjustment and 2nd platform Aachen West
4. Double track S6 Köln - Grevenbroich - Mönchengladbach
5. Rheydt curve

Scenario 3: Capacity based on the Deutschlandtakt (Germany's synchronized paths)

This scenario outlines the capacity derived from the Deutschlandtakt planning model, which provides a clockface timetable for coordinated, fast and reliable connections in local, long-distance and freight transport.

There is also a list of infrastructure projects considered in this scenario;

1. 2-track new construction Köln - Hansaring - Köln South - Hürth - Kalscheuren (Westspange Köln)
2. Creation of parallel boarding and alighting facilities at the western end of Köln Central Station to and from Köln West.
3. Aachen Central Station: Construction of a double switch connection to track 28 in the west and connection of track 26 from the direction of Aachen West.
4. Construction of two central overtaking tracks in Aachen main station as a waiting position for Goods trains with a useful length of 740 m
5. Oberhausen junction (expansion of the at-grade connecting curves)
6. Capacity expansion Oberhausen-Emmerich (ESTW Emmerich, block consolidation) 7. Oberhausen - Emmerich - border (3-track extension, removal of 55 level crossings)
7. Conversion of passing tracks for 740 m long goods trains in the stations Dinslaken, Wesel, Mehrhoog, Empel Rees and Emmerich
8. Extension with ETCS

9. Aachen-Rothe Erde railway station is the extension of the passing tracks for freight traffic to improve the quality of the service.

The realisation of the Rheydt curve can alleviate the bottlenecks on the Aachen-Köln route, but the "Rheydt curve" was not included in the DT measures. This will however create overloading on the Aachen-Rheydt section.

4 Plausibility Check

To ensure the reliability and validity of the demand scenarios and capacity projections, a thorough plausibility check was carried out. This involved a careful analysis of the recent key assumptions:

- DB standard route capacity: Based on Handbook – the Railway System (Handbuch das System Bahn)
- Maximum capacity on double tracks: 240 trains/day and Coal train gross weight: 2,800 tonnes (Industry knowledge)
- Iron ore train gross weight: 5,400 tonnes (die.Zugfolgestelle.de)
- 1 container train = 75 TEU: (Port of Antwerp/Appraisal)
- 1 TEU = 11 tonne: (Industry knowledge)
- Trailer traffic share in container traffic: 30% (Railway Gazette international – 09/24)
- Traffic distribution from Antwerp: 30% via Aachen–Mönchengladbach and 70% via Aachen–Köln (Port of Antwerp)

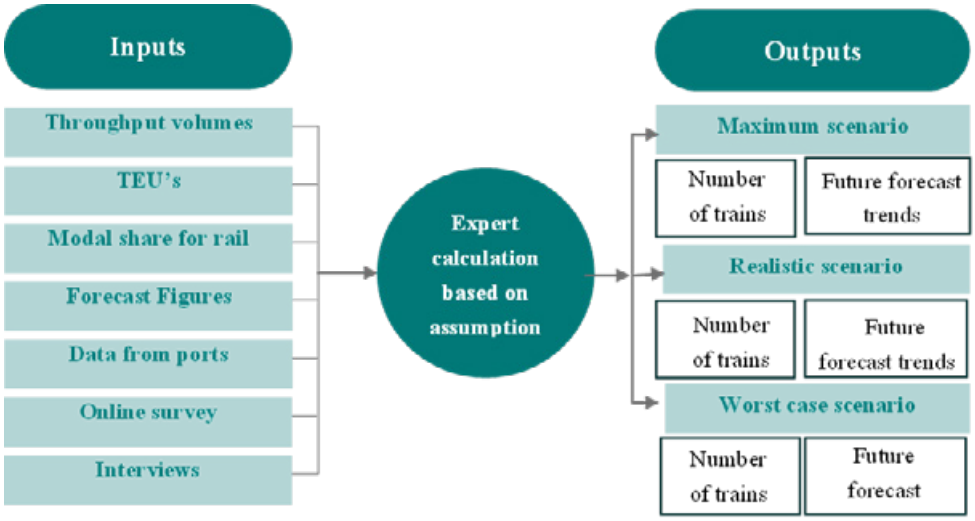


Figure 10: Methodology for calculating the number of trains

- Traffic distribution from Rotterdam: 36% via Venlo 50% via Emmerich (ProRail) advancements in the European rail freight with sector specific assumptions. Using the methodology outlined, the analysis carried out can be summarized below, authors created three scenarios for the 2040 demand forecast using data collected from various stakeholders and some assumptions (also listed above).

The anticipated number of trains on the major rail corridors for 2030 by ViaCon and the three authors'-calculated scenarios for 2040 are given in the table below. The reality of the demand scenarios lies somewhere between the realistic and the maximum demand scenario.

Table 1: Forecasted number of trains as per ViaCon and authors Calculation

Corridor	ViaCon figures 2030	2040 authors calculation		
		Maximum	Realistic	Minimum
Aachen–Düren–Köln	394	529	475	401
Duisburg–Emmerich	345	455	433	345
Venlo–Viersen–Krefeld	142	195	178	153
Aachen–Rheydt– Mönchengladbach	287	360	328	287
Köln–Mönchengladbach	90	111	104	94
Viersen–Duisburg	85	117	107	92

Source for the ViaCon figures: Upgrading of the Aachen railway junction, 2017

Key highlights:

- The number of trains on the corridors would increase in all scenarios by 2040.
- The realistic scenario does not include any increase in trailer traffic.
- The authors are also of the opinion that the 7% additional trains used in the current long-term forecast to balance out the peaks are too low and that this percentage will also increase.
- The reality lies somewhere between the realistic and the maximum scenario.

4.1 Capacity utilization without investments

With the current "**realistic**" capacity conditions, there will be numerous **bottlenecks.**

The intra- and intermodal price war is so fierce, particularly on the Rhine corridor, that shippers switch providers even at price differences of 1-2 cents per TEU-km. There are

bottlenecks identified on Aachen-Neuss, Aachen-Köln and Emmerich-Duisburg corridors. Routing the trains between Aachen West and Köln via the Rheydt curve increases the costs per container by €10 to €20 per container and will therefore be avoided by the railway companies at all costs.

4.2 Capacity utilization with investments

Even if the investments are made, major bottlenecks will remain unsolved. Based on the comparison conducted between the realistic capacity, the authors calculation of capacity forecasts and the capacity computed from the Deutschland takt, following notes can be inferred;

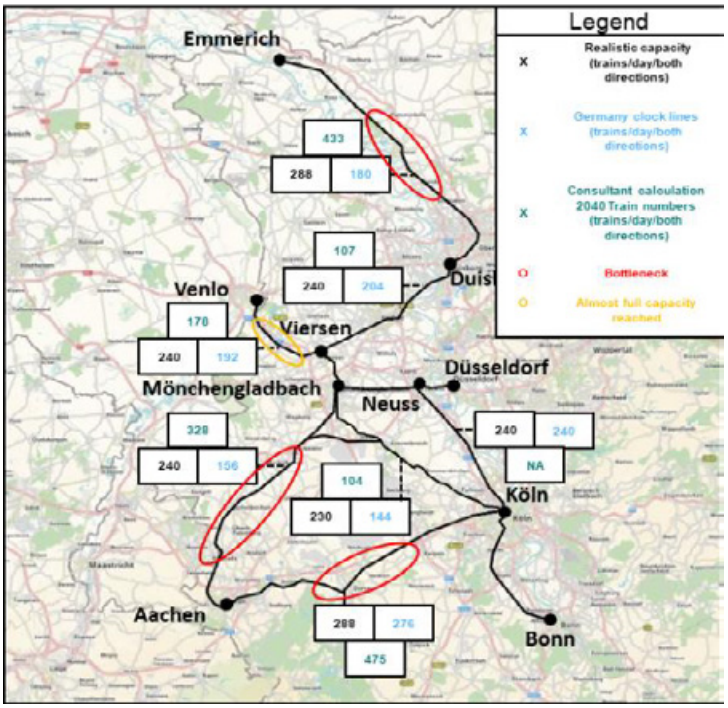


Figure 11: Capacity utilization on the corridors with investments under the three methods

- In the realistic scenario, the authors assumed a lower increase in the share of rail transport in the port of Rotterdam, extended the target year of the port forecast for Antwerp from 2030 to 2040 and assumed 65 TEU per train. This was done for the reasons of realism and caution.
- However, the maximum scenarios, that is the 25% rail share of the port of Rotterdam, or the port of Antwerp reaching the forecast earlier than 2040 can also materialise.
- The maximum scenario is not only theoretical but may also materialise in part. In other words, our horizon of expectations lies between the realistic and the maximum case rather than between the realistic and the minimum case.

5 Conclusion

Session 17
Wolf-Dietrich Geitz, Upendra Atre

The development of the ZARA port's hinterland rail traffic is severely underestimated in the official forecasts, which also ignores the figures from foreign Railway Undertakings. The main risk associated with this is a conflict between freight and passenger transport. Either the number of passenger trains would fall or there would be more lorries on the road, which would have a negative impact on the economy of the region. A lot of reasons were cited for the underestimation of the rail freight volume. Until recently, Destatis had consistently understated the number of intermodal trains as well as the trains operated by the non-German Railway Undertakings which influenced the forecasts which are valid today. Trail business on rail and its associated investments has been also underestimated in the previous years.

Reasons for the underestimations in the volumes;

- Until recently, methodological underestimation of **intermodal trains** and trains operated by **non-German Railway Undertakings** by Destatis which influenced the forecasts which are valid today. Both these segments are over proportionally running in this area
- Underestimation of **trailer business** on rail and related investments in all past years
- Undervaluation of the impact of significant rail investments made at the ports of Antwerp and Rotterdam in the **past, present, and future.**
- Insufficient assessment of the development in volumes at **hinterland hubs**.
- Underestimation of development of rail volumes in **east west axis**

Even with planned steps to handle the new traffic, there will still exist the bottlenecks because of the growth of volume in the port of Antwerp for example. Forecasts for rail developments at the port of Rotterdam and the port of Antwerp appear to be underestimated. While the port of Antwerp hopes to double the rail volumes, DB predicts that the number of freight trains in the Belgium-Viersen-Ruhr section will decline. It is because they ignored the volumes from the outside of Germany, so the official German forecasts were always too low. The primary danger posed by these bottlenecks is the potential conflict between the passenger and the freight operations. It would have a detrimental effect on the local economy because either there would be fewer passenger trains or there would be a greater number of trucks on the road. The situation in 2040 will be worse than that in 2030, and there would be even more bottlenecks in the most extreme scenarios, such as the port of Rotterdam having a 25% rail modal share and the port of Antwerp having a 15% share. The predictions undervalue the importance of the corridors and fail to sufficiently account for the demand for rail freight transportation.

Session 17
Wolf-Dietrich Geitz, Upendra Atre

Authors



Geitz, Wolf-Dietrich

Mr. Geitz studied at RWTH Aachen university, Mechanical engineering (focus on rolling stock) to a Dipl-Ing. degree. He has been working in Railway Consulting since 1993, and since 1995 he managed several national and international railway projects. As a railway operations expert he concentrated on optimizing railway infrastructure operations as well as installations and the respective signalling and train protection equipment. He was the leader of a perennial European project on the improvement of international rail freight transport. The optimization of processes and the efficiency increase of railway operations and infrastructure constitute one of his main focuses. With his international experience he works on business plans for railway operators, evaluation of investment in infrastructure, market research for the railway industry, analysis of railway operations and the development of IT systems. He published various articles in international journals and lectures in the fields of rail operations planning and capacity issues at THM university of Giessen, Germany. He is a member of the steering committee for rail transport of Deutsches Verkehrsforum and member of the International Association of Railway Operations Research IAROR.



Atre, Upendra

Mr. Atre completed his study in Railway systems engineering at RWTH Aachen university in 2021, where he has obtained various critical thinking and research related skills. All rail related topics including principles of rail technology, rail vibration dynamics, track guiding technology, railway systems and control systems, railway capacity management and operations were part of his study. He has since been working with Railistics GmbH and has gained experience in rail projects from around the world. He has gained vast expertise of the rail industry and transportation policy as a result of these undertakings. He also has a wealth of technical knowledge and expertise with specialist software.

IRSA 2025

AUTORENVERZEICHNIS
AUTHOR INDEX

A

Atre, Upendra 740

B

Barteldes, Sören 288
Beck, Michael J. 11
Beckmann, Leon 429
Benz, Alexandra 11
Bernd, Schmidt 166
Bertl, Melanie 133
Bittner, Andreas 288
Büker, Thorsten 11
Buschbacher, Harald 84

C

Caimi, Gabrio 626
Cao, Yan 256
Chang, Haw-Shyang 666
Chen, Gang 361
Cichon, Martin 256, 459

D

Dickenbrok, Björn 537
Ding, Xin 361
Doose, Felix 537

E

Eisold, Jan 65
Emunds, Tamme 44
Enning, Manfred 166

F

Feng, Joseph 596
Fernández-Bobadilla, Héctor Alberto 580
Fischer, Marcus 237
Fuchs, Kurt 181

G

Geitz, Wolf-Dietrich 740
Gerhardus, Andreas 568
Gerstl, Erik 302
Gil, Pascal 192
Gläser, Ute 596
Goersch, Jannik 221
Götz, Dennis 609
Graffagnino, Thomas 313
Guicheney, Pauline 682
Güthenke, Heiko 638

H

Hack, Linus 11
Hagemann, Oliver 429
Hänni, Vanessa 626
Hasegawa, Tomonori 496, 666
Heimann, Kai 395
Heizler, Felix 149
Hellings, Gustav 412
Hempel, Thomas 221
Herwartz-Polster, Sebastian 149
Hiura, Toshihiro 328
Hofmeier, Tobias 256
Horn, Florian 133

J

Jacke, Tobias 553
Jakubek, Stefan 205
Janecek, David 345
Janßen, Mirco 192
Jelinski, Matthias 476
Jendritzki, Felix 133
Jenne, Sven 192
Jing, Guozhen 361
Joch, Martin 205
Jordanski, Christina 133
Jovanovic, Mirjana 596

K

Kaneko, Takuya 328
Kekes, Cédric 100
Kliem, Maximilian 133
Knabl, Dieter 650
Koch, Susanne 133
König, Jens 682
Körner, Tobias 638
Koseki, Takafumi 666
Kraft, Dietmar 395
Kramer, Urs 11
Kreisel, Norman 237
Kröger, Felix 237
Kuckelberg, Alexander 31
Kuhlins, Martin 429

L

Lahoud, Charbel 272
Lampe, Felix 537
Lang, Carolin 459
Leibner, Philipp 221
Leicht, Lukas 459
Lenze, Wiebke 133
Lipinski, Stefan 412
Lochmann, Libor 682
Lux, Beate 65

M

Marschnig, Stefan 650
Martin, Ullrich 580
Miyoshi, Shota 496, 666
Moder, Jakob 302
Mohn, Tobias 638
Möhring, Martin 192
Mößner, Klaus 272
Müller, Thomas 682
Münchow, Sam 272
Murawa, Franz 288

N

Nachtigall, Karl 100
Nagai, Kensuke 666
Narita, Kiyoshi 328
Neubauer, Benedikt 221
Neumann, Thorsten 568
Nießen, Nils 11, 44, 133
Ning, Zheng 116
Niu, Yan 395

O

Ohnishi, Wataru 666
Olbrich, Margit 444
Otto, Kilian 256

P

Pagenkopf, Johannes 149
Pauwels, Katrien 345
Paz-Martinez, Aaron 149
Pelz, Johannes 510
Persson, Rickard 682
Pfaff, Raphael 221
Pfeifer, Andreas 11, 133
Pickl, Stefan 476
Plotho, Alexander 650
Popescu, Oana-Iuliana 568

Q

Quaglietta, Egidio 116

R

Rabel, Martin 568
Rasel, Thomas 237
Reiterer, Alexander 510
Rhein, Alexander 553
Rodríguez-Bayona, Laura Tatiana 580
Rohrberg, Danny 638

S

Saeednia, Mahnam 116, 725
Satitrattanashewin, Sirikorn 65
Schaper, Christian 537
Scherer, Philipp 11
Schindler, Christian 192, 221, 412
Schirrer, Alexander 205
Schmidt, Bernd 166
Schmidt, Hauke 361
Schmitz, Walter 476
Schnell, Toni 444
Schnieder, Lars 703
Seeger, Henning 192
Sparsbrod, Marc 149
Stumm, Sebastian 237
Sugahara, Yoshiki 378

T

Theurich, Franziska 100
Tomii, Norio 328
Trappein, Mario 609

U

Ude, Sebastian 553

V

Van de Mergel, Anneleen 345
Van Leeckwyck, Quentin 345
Versluis, Nina D. 725
von der Heide, Paula 703
von Heel, Emanuel 11

W

Watanabe, Shuho 328
Weymann, Frédéric 345
Willgardt-Levin, Sascha 553

X

Xue, Chutan 361

Y

Yabuki, Hideyuki 328
Yagi, Makoto 496
Yamaguchi, Daisuke 496
Yonemoto, Kazushige 328

Z

Zdarsky, Aaron 459
Zelinka, Matthias 205
Ziegler, Niklas 11
Ziesel, Patrick 459
Zollinger, Ernst 626
Zschiesche, Kira 510