

Forschungsbericht 2026-15

Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen

Maik Halbach, Robert Oertel,
Robert Markowski, Robert Alms

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Verkehrssystemtechnik
Braunschweig/Berlin



DLR

Deutsches Zentrum
für Luft- und Raumfahrt

Forschungsbericht 2026-15

Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen

Maik Halbach, Robert Oertel,
Robert Markowski, Robert Alms

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Verkehrssystemtechnik
Braunschweig/Berlin

66 Seiten
29 Bilder
1 Tabellen
11 Literaturstellen



Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2026-15
Erscheinungsjahr 2026
DOI: [10.57676/y2r8-qk14](https://doi.org/10.57676/y2r8-qk14)

Erklärung des Herausgebers

Dieses Werk – ausgenommen anderweitig gekennzeichnete Teile – ist lizenziert unter den Bedingungen der Creative Commons Lizenz vom Typ Namensnennung 4.0 International (CC BY 4.0), abrufbar über <https://creativecommons.org/licenses/by/4.0/legalcode>

Lizenz



Creative Commons Attribution 4.0 International

*Verkehrssteuerung, Lichtsignalanlage, Schaltzeitprognose, C-ITS, V2X, SPATEM
(in Deutsch veröffentlicht)*

Maik Halbach
DLR, Institut für Verkehrssystemtechnik, Braunschweig
Robert Oertel, Robert Markowski, Robert Alms
DLR, Institut für Verkehrssystemtechnik, Berlin

Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen

Der Forschungsbericht dokumentiert die Erarbeitung eines Konzepts zur Qualitätsbewertung von Schaltzeitprognosedaten an Lichtsignalanlagen.

Moderne Lichtsignalanlagen können heutzutage unter Nutzung von C-ITS-Technologien Informationen sowohl über ihren aktuellen Signalzustand, aber auch prognostizierte, zukünftige Signalzustände, mit anderen Verkehrsteilnehmern teilen. Um den Anforderungen an zukunfts- und leistungsfähige C-ITS-Services wie beispielsweise beim GLOSA-Service (Green Light Optimal Speed Advisory) gerecht zu werden, wurde ein Konzept zur Beurteilung der Prognosequalität derartiger Anlagen entwickelt. Es umfasst drei zentrale Bewertungsschritte: Ein Dynamikindex, der die Vergleichbarkeit einzelner Bewertungen über verschiedene Knotenpunkte und Steuerungen hinweg ermöglicht; eine separate Prüfung der Integrität und Plausibilität der Prognosedaten; und eine inhaltliche Beurteilung der Prognosegüte durch die gezielte Analyse von Schaltzeitdaten. Eine abschließende Gesamtwertung in Form von Qualitätsstufen kann aus diesen Einzelbewertungen abgeleitet werden.

Das Konzept liefert als Ergebnis den Projektpartnern und Stakeholdern aus Städten und Kommunen eine erste Grundlage für eine fundierte Qualitätsbewertung von Prognosedaten und -verfahren. Es soll zukünftig in praktischen C-ITS Anwendungen detaillierter erprobt und weiterentwickelt werden, um somit ein verlässliches Werkzeug zur Beurteilung von Schaltzeitprognosen bereitzustellen.

Traffic signal control, traffic light, forecasts signal timings, C-ITS, V2X, SPATEM

Maik Halbach
German Aerospace Center (DLR), Institute of Transportation Systems, Braunschweig
Robert Oertel, Robert Markowski, Robert Alms,
German Aerospace Center (DLR), Institute of Transportation Systems, Berlin

Quality of Forecasts for Signal Timings at Traffic Lights

The report documents the development of a concept for the evaluation of forecast quality for signal timings at traffic lights.

Nowadays, modern traffic light systems are capable of using C-ITS technologies to share information with other road users not only about their current signal state but also about predicted future signal states. To meet the requirements for future-oriented and high-performance C-ITS services, such as the GLOSA service (Green Light Optimal Speed Advisory), a concept for evaluating the forecast quality of such prediction data has been developed. It comprises three central evaluation steps: a signal control dynamics index that enables the comparability of individual evaluations across different junctions and signal control systems; a separate check of the integrity and plausibility of the forecast data; and a substantive evaluation of forecast quality based on the analysis of signal timing data. A final overall rating in the form of quality levels can be derived from these individual evaluations.

The proposed concept provides project partners and stakeholders from cities and local authorities with an initial basis for a systematic assessment of the quality of forecast data and prediction methods. It is intended to be tested in greater detail and further developed in practical C-ITS applications in the future, thereby establishing a reliable tool for evaluating the forecasts of traffic signal timings.



Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen

Ergebnisbericht

Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen

Ergebnisbericht

Verfasst von:

Deutsches Zentrum für Luft- und Raumfahrt e.V.
Institut für Verkehrssystemtechnik
Web: <https://www.dlr.de/ts/>
Maik Halbach, Robert Oertel, Robert Markowski & Robert Alms
Mail: maik.halbach@dlr.de

Im Auftrag und Zusammenarbeit mit:

Landesbetrieb Straßen, Brücken
und Gewässer Hamburg (LSBG)



LSBG
Landesbetrieb Straßen,
Brücken und Gewässer
Hamburg

Hamburg Verkehrsanlagen
GmbH (HHVA)



Im Rahmen des Projekts:



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Dokument-Identifikation:

Titel	Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen
Thema	Ergebnisbericht
Autoren	Maik Halbach, Robert Oertel, Robert Markowski & Robert Alms
Version 1.0	22-01-2026

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	3
1.2	Vorgehensweise	4
2	Anforderungsdefinition	6
2.1	Datengrundlage	6
2.2	Experteninterviews	11
2.3	Anwendungsbezogene Konfigurierbarkeit	15
2.4	Kernanforderungen	17
3	Bewertungskonzept	19
3.1	Steuerungsdynamik	20
3.2	Nachrichtenintegrität und -plausibilität	24
3.3	Prognosegüte	27
3.4	Gesamtwertung	31
4	Bewertungsmetrik	33
4.1	Steuerungsdynamik	33
4.2	Nachrichtenintegrität und -plausibilität	36
4.3	Prognosegüte	38
4.4	Übersicht zur Ermittlung der Bewertungsmetrik	44
5	Anwendung	46
5.1	Anwendungsbeispiel	46
5.2	Konfigurations- und Anwendungshinweise	54
6	Fazit	57
	Literatur	59
	Glossar	61
	Anhang	63

1 Einleitung

Moderne Lichtsignalanlagen (LSA) besitzen heute zunehmend technische Möglichkeiten, um Informationen an Verkehrsteilnehmende in ihrem Umfeld zu übermitteln. Unter Nutzung von C-ITS-Technologien, wie der Fahrzeug-Infrastruktur-Kommunikation (V2X), lassen sich beispielsweise Informationen zu Signalschaltzeiten übertragen. Neben dem aktuell signalisierten Zustand an einem Knotenpunkt, umfasst dies auch eine Prognoseinformation über zukünftige Signalzustände. Die Verkehrsteilnehmenden können diese Information mit entsprechender Kommunikationstechnik empfangen und das eigene Fahrverhalten gegebenenfalls adaptieren, um z. B. unnötige Verzögerungs-, Beschleunigungs- und Wartevorgänge zu vermeiden. Gerade im Bereich des automatisierten und vernetzten Fahrens stellt die LSA-Schaltzeitprognose einen wichtigen Baustein für zukünftige Weiterentwicklungen dar, da in bestimmten Situationen eine infrastrukturelle Unterstützung bei der Optimierung von Abläufen notwendig sein kann.

Für die Bereitstellung von LSA-Schaltzeitprognosedaten sind heute verschiedene Verfahren verfügbar und im Einsatz. Je nach Art sowie Eigenschaften eines Knotenpunktes und des eingesetzten LSA-Steuerungsverfahrens kann die Schaltzeitprognose dabei eine Herausforderung darstellen. Dies gilt insbesondere für die Signalzeitprognose innerhalb verkehrsabhängiger LSA-Steuerungsverfahren [1]. Diese Verfahren erfassen die aktuelle Verkehrssituation fortlaufend über lokale Detektionstechnik und können ihre Signalisierungszustände je nach verkehrlichem Erfordernis kurzfristig anpassen, was eine unmittelbare Auswirkung auf die Schaltzeitprognose haben kann. Eine qualitativ hochwertige Schaltzeitprognose sollte mit derartigen Ereignissen angemessen umgehen können, da eine sprunghafte Prognoseveränderung zu einer mangelnden Akzeptanz der ausgesendeten Informationen und darauf basierender C-ITS-Services führen kann. Allen voran ist dabei der GLOSA-Service (Green Light Optimal Speed Advisory) zu nennen, welcher Annäherungsprozesse an einem Knotenpunkt unterstützen kann, schematisch dargestellt in Abbildung 1.1, und dessen Einführung bereits in aktuellen, praxisrelevanten Hinweispapieren [2], [3] beschrieben ist. Hier könnten Prognosesprünge zu einem Vertrauensverlust und Komforteinbußen führen, speziell beim automatisiert vernetzten Fahren gar zu gefährlichen Situationen aufgrund abrupten Bremsmanöver.



Abbildung 1.1: **Schematische Darstellung der Aussendung einer Schaltzeitprognose im Rahmen eines GLOSA-Services durch eine Roadside Unit (RSU) an vernetzte Fahrzeuge.**

LSA-Schaltzeitprognosen sollten vor diesem Hintergrund qualitativ hochwertig und verlässlich sein. Dies gilt sowohl aus Sicht einer Stadt beziehungsweise eines LSA-Betreibers als verantwortliche Instanz für die Bereitstellung der Schaltzeitprognosedaten, als auch für alle anderen Stakeholder, die diese Prognosedaten in der Anwendung nutzen. Welche Anforderungen dabei zu erfüllen sind und wie eine transparente Qualitätsbewertung von LSA-Schaltzeitprognoseverfahren sowie -daten in der Praxis erfolgen kann, ist bislang nicht in Richtlinien oder Normen definiert. Eine anschauliche und vergleichende Bewertung der Prognosegüte ist für Städte, LSA-Betreiber und alle weiteren Stakeholder aufgrund einer fehlenden, einheitlichen Bewertungsmetrik daher bisher schwierig bzw. nicht möglich, weshalb hier Handlungs- und Innovationsbedarf besteht.

Die Freie und Hansestadt Hamburg verfügt mit aktuell rund 300 (Stand: 01/2026) über die bundesweit höchste Anzahl an modernen Lichtsignalanlagen mit C-ITS-Ausstattung, die die Aussendung von LSA-Schaltzeitprognosedaten unterstützen. Verschiedene neuartige C-ITS-Services befinden sich hier aktuell in der Erprobung und im Roll-out, welche auf einer hohen Qualität und Verlässlichkeit dieser Prognosedaten aufbauen. Die Anlagenanzahl mit einer standardmäßigen Unterstützung von C-ITS-Services wird sich im Stadtgebiet weiter erhöhen. Im Sinne der Innovationsförderung könnten daher zukünftig auch planungs-

und beschaffungstechnische Aspekte in den Fokus rücken. Eine Prognosequalitätsanforderung im Rahmen der LSA-Planung und des Einsatzes von LSA-Prognoseverfahren ist dabei denkbar. Die Qualitätsbewertung von LSA-Schaltzeitprognosedaten ist für die Freie und Hansestadt Hamburg somit von elementarer Bedeutung. Vor diesem Hintergrund wurde vom Landesbetrieb Straßen, Brücken und Gewässer (LSBG) zusammen mit der Hamburg Verkehrsanlagen GmbH (HHVA) und dem Deutschen Zentrum für Luft- und Raumfahrt e.V. (DLR) ein praxistaugliches Konzept für die Qualitätsbewertung von LSA-Schaltzeitprognosedaten entwickelt. Dieses Konzept soll zukünftig auf Grundlage historischer Schaltzeitdaten eine transparente und vergleichende Qualitätsbewertung von LSA-Schaltzeitprognoseverfahren und -daten an Knotenpunkten in der Praxis ermöglichen.

1.1 Motivation

Für Praxisanwender stellt die transparente und vergleichende Qualitätsbewertung von Schaltzeitprognoseverfahren und -daten bislang eine Herausforderung dar. Für sie ist nicht ohne Weiteres ersichtlich, welche Kenngrößen dafür genutzt werden sollten, ob eine Kenngrößenkombination bei der Bewertung sinnvoll ist und welche Werteausprägung für welches Qualitätsniveau steht. Zur visuellen Darstellung und manuellen Analyse von einzelnen Kenngrößen stellen Prognoseanbieter in der Regel eigene Tools zu Verfügung, die einen Zugriff auf historische Datenbestände erlauben. Für die Interpretation der so abgerufenen Einzelkenngrößen ist jedoch meist viel Expertenwissen notwendig, sodass unmittelbare Rückschlüsse auf die Prognosequalität schwierig sind.

In ersten, darauf basierenden Bewertungsansätzen werden vorhergesagte und tatsächliche Schaltzeitpunkte über einen bestimmten Zeithorizont miteinander verglichen und daraus ein prozentualer Übereinstimmungsgrad abgeleitet. Die Höhe des Übereinstimmungsgrades wird dabei mit der Verlässlichkeit der Schaltzeitprognose in Beziehung gesetzt. Die Aussagekraft ist allerdings begrenzt, da dabei in der Regel eine Vielzahl an Einzelwerten entstehen, welche mit viel Expertise gedeutet werden müssen. Des Weiteren wird oft versucht, die Prognosequalität mit Hilfe der Abweichung zwischen vorhergesagten und tatsächlichen Schaltzeitpunkten zu bewerten. Dafür werden verschiedene Fehlermaße verwendet, die allerdings zeitliche Abhängigkeiten nicht ausreichend berücksichtigen. In der Praxis hängt die qualitative Bewertung einer Schaltzeitprognose zudem stark von weiteren Größen wie z. B. Zeitpunkt und Prognosehorizont sowie räumlichen Gegebenheiten eines Knotenpunktes ab. Diese und weitere Randbedingungen werden bisher noch nicht adäquat in der Qualitätsbewertung von Prognoseverfahren adressiert.

Das Ziel des Vorhabens war es daher, ein Konzept für eine umfassende Qualitätsbewer-

tungsmetrik auf Basis realer Prognosedaten und aktueller C-ITS-Standards zu erarbeiten, welches alle praxisrelevanten Randbedingungen herstellerunabhängig berücksichtigt. Die Metrik soll es ermöglichen, Schaltzeitprognosen transparent und möglichst aufwandsarm überprüfen zu können, um diese zukünftig mit hoher Zuverlässigkeit in der Praxis einsetzen zu können. Die Qualität der Schaltzeitprognose ist eine zentrale Voraussetzung für die praktische Nutzbarkeit und letztlich die Akzeptanz neuartiger C-ITS-Services. Die einheitliche Bewertung und Vergleichbarkeit unterschiedlicher Schaltzeitprognoseverfahren könnte sowohl zukünftige C-ITS-Services selbst, als auch die Weiterentwicklung der eigentlichen Prognose- und Steuerungsverfahren wesentlich unterstützen.

1.2 Vorgehensweise

Das Vorgehen zur Erarbeitung des Konzeptes lässt sich grundsätzlich in zwei Phasen einteilen. In der ersten Phase wurde eine Analyse der Anforderungen an eine Schaltzeitprognose durchgeführt. Damit sollten alle wesentlichen Merkmale zusammengetragen werden, die für eine praxisrelevante Qualitätsbewertung von Bedeutung sind. Folgende Schritte beinhaltete die Anforderungsanalyse:

- I. Zunächst erfolgte eine Analyse der relevanten Datengrundlage, d. h. in welchem Datenformat Schaltzeitprognosen tatsächlich vorliegen bzw. verfügbar sind, welche Informationen die Datensätze enthalten und inwieweit diese den heutigen C-ITS-Standards entsprechen.
- II. Anschließend folgte eine Anforderungsanalyse in Bezug auf die Schaltzeitprognose sowohl aus Anwendersicht als auch aus Perspektive der Endnutzer. Hierfür wurden Experteninterviews mit verschiedenen Stakeholdern aller relevanten Domänen durchgeführt.

Darauf aufbauend wurde in der zweiten Phase ein Bewertungskonzept erarbeitet, welches die erhobenen Anforderungen aus der ersten Phase aufgreift und in geeigneten Kenngrößen und Metriken für die Qualitätsbewertung abbildet. Das Konzept wurde dabei in einem iterativen Prozess auf Basis von realen Schaltzeitprognosedaten entwickelt, getestet und überprüft. Folgende Schritte wurden dafür durchgeführt:

- III. Zunächst wurden alle relevanten und geeigneten Kenngrößen gemäß Anforderungsanalyse ausgewählt, konzeptionell beschrieben sowie deren Anwendbarkeit mit Hinblick auf die angestrebte Bewertung untersucht. Es wurden Wechselwirkungen und

Abhängigkeiten zwischen den einzelnen Kenngrößen sowie Gewichtungsanteile mit Hinblick auf eine kombinierte Gesamtbewertung betrachtet.

- IV. Ausgehend von dieser Kenngrößenauswahl und deren Verknüpfung wurden die entsprechenden Berechnungsgrundlagen für die Qualitätsbewertung aufgestellt.

Diese grundsätzliche Vorgehensweise vorangestellt ist der vorliegende Bericht wie folgt aufgebaut: Kapitel 2 trägt die Anforderungen an die Qualitätsbewertungsmetrik zusammen. Kapitel 3 stellt das Konzept für eine solche Metrik auf, während in Kapitel 4 konkrete Vorschläge zur Umsetzung der Bewertungsmetrik in Form mathematischer Beschreibungen und Konfigurationen diskutiert werden. Kapitel 5 illustriert anhand eines Anwendungsbeispiels die möglichen Ergebnisse einer solchen Qualitätsbewertung. Abschließend trägt Kapitel 6 die gesammelten Erkenntnisse zum Bewertungskonzept zusammen, diskutiert offene Fragen und liefert einen Ausblick auf zukünftige Forschungsaktivitäten.

2 Anforderungsdefinition

Das folgende Kapitel fasst die grundlegenden Anforderungen, welche sich sowohl aus der Datengrundlage als auch aus den Erkenntnissen der Experteninterviews ergeben, zusammen.

2.1 Datengrundlage

An einer vernetzten Lichtsignalanlage können Informationen eines Knotenpunkts mit anderen Verkehrsteilnehmern via Fahrzeug-Infrastruktur-Kommunikation (V2X) geteilt werden. Dafür sind vor allem zwei V2X-Nachrichtentypen relevant: die *Map Extended Message* (MAPEM) und die *Signal Phase And Timing Extended Message* (SPATEM). Diese beiden Typen sind eng miteinander verknüpft, um geografische Informationen in Kombination mit aktuellen Signalzuständen eines Knotenpunkts versenden zu können. In Europa werden diese V2X-Nachrichten von der ETSI-Organisation sowie der C-Roads Plattform definiert und spezifiziert, um innerhalb des europäischen Verbundes eine Harmonisierung der Entwicklung und Einführung von Nachrichtenstandards zu ermöglichen. Abbildung 2.1 gibt einen Überblick über die beteiligten Standardisierungsebenen, die verwendeten Nachrichtentypen und die in dieser Arbeit genutzte Datengrundlage, welche nachfolgend genauer beschrieben werden.

2.1.1 Standardisierung

ETSI: Das European Telecommunications Standards Institute (ETSI) ist eine führende europäische Normungsorganisation, welche Standards für Informations- und Kommunikationstechnologien entwickelt. Durch das ETSI sind die im Vorhaben relevanten V2X-Nachrichtentypen MAPEM und SPATEM im Detail spezifiziert.

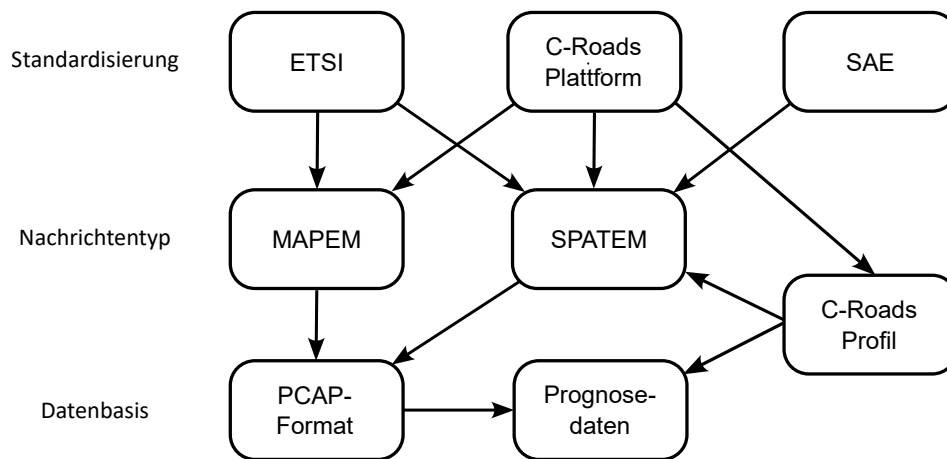


Abbildung 2.1: **Übersicht über Standardisierungsebenen, V2X-Nachrichtentypen und die verwendete Datengrundlage.**

C-Roads: Über die sogenannte C-Roads Plattform haben sich Behörden und Straßeninfrastrukturbetreiber zu einer europäischen Initiative zusammengeschlossen, um die Einführung von kooperativen intelligenten Verkehrssystemen (C-ITS) zu begleiten und diese Entwicklungen durch eine Harmonisierung der Funktionalitäten und Kommunikationssysteme voranzubringen. Hierzu gehören auch V2X-gestützte Dienste im Straßenverkehr, wie die im Vorhaben adressierte Signalschaltzeitprognose. Aufsetzend auf dem ETSI-Standard spezifiziert die C-Roads Initiative, wie die Datenfelder bestimmter V2X-Nachrichtentypen je nach Anwendungszweck befüllt werden müssen. C-Roads definiert dabei keine eigenen, neuen Nachrichtenstandards, sondern spezifiziert Nutzungskonzepte von C-ITS Diensten, sogenannte Facility Layer Services (FLS), welche auf die Standards des European Telecommunications Standards Institute (ETSI) oder der International Organization for Standardization (ISO) verweisen.

SAE: Der Standard J2735 [4] der Society of Automotive Engineers (SAE) ist eine weitere Norm zur V2X-Nachrichtenspezifikation, welche hauptsächlich im nordamerikanischen Raum Anwendung findet. Im Vergleich zu ETSI ist das eine über die EU hinaus agierende internationale Organisation, welche stark durch die Automobilbranche gestützt wird. In Teilen verweisen ETSI oder auch C-Roads auf die ISO-Standards der SAE. Der SAE-Standard spezifiziert, wie die Datenfelder einer SPATEM-Nachricht für Prognosen befüllt werden sollen. Da sich dieser Standard jedoch teilweise von dem in der EU etablierten ETSI-Standard unterscheidet und im europäischen Raum in der Praxis wenig eingesetzt wird, berücksichtigt dieser Bericht ausschließlich den ETSI-Standard bzw. das damit verknüpfte C-Roads Profil.

2.1.2 V2X-Nachrichtentypen

MAPEM: Die MAPEM-Nachricht (MAP (topology) Extended Message) beschreibt die Topologie eines Knotenpunktes und ist im ETSI-Standard definiert [5]. Sie erfasst insbesondere die räumliche Lage der Fahrstreifen und die logische Struktur der Fahrbeziehungen innerhalb des Knotenpunkts.

SPATEM: Die SPATEM-Nachricht (Signal Phase And Timing Extended Message) beschreibt die aktuellen Signalzustände einer LSA und kann gegebenenfalls auch eine Prognose der Signalzustände enthalten. Die SPATEM ist ebenfalls im ETSI-Standard definiert [5]. Typischerweise referenziert eine MAPEM-Nachricht die Fahrstreifen einer einzelnen Lichtsignalanlage, deren Signalzustände entsprechend in einer SPATEM enthalten sind. Das Handbuch zur MAPEM- und SPATEM-Erstellung der C-Roads Initiative [6] greift diese Verflechtung der beiden Nachrichtentypen gezielt auf und liefert eine detaillierte Beschreibung aller Nachrichtenelemente, welche für eine spätere Anwendung von Bedeutung sind (siehe C-Roads Profil). Die Prognose in einer SPATEM ist gemäß des Anwendungszwecks nicht obligatorisch, weshalb die entsprechenden Datenfelder nicht immer befüllt werden.

2.1.3 Datenbasis

C-Roads Profil: Die C-Roads Plattform spezifiziert die technischen Anforderungen für bestimmte Nachrichtentypen und deren Inhalte detailliert, strukturiert nach funktionalen Facility Layer Services (FLS) [7]. Diese FLS definieren die notwendigen Nachrichteninhalte, um z. B. Warnungen zu sicherheitsrelevanten Ereignissen, Informationen zu Straßenbauarbeiten sowie Hinweise auf Geschwindigkeits- oder Verkehrszeichenänderungen bereitstellen zu können. Prognosen zu LSA-Signalzuständen finden sich im FLS "Traffic Light Maneuver (TLM) and Road and Lane Topology (RLT) Service (TLM FLS and RLT FLS)" wieder. In diesem Bericht wird dieser FLS aus den „Message Profiles“ der C-Roads Initiative [7] (Release Version 2.2.0) folgend als „C-Roads Profil“ bezeichnet.

Die Abbildung 2.2 veranschaulicht die logischen Zusammenhänge der erforderlichen Datenfelder einer SPATEM-Nachricht für ein solches C-Roads Profil, welche für eine Prognose relevant ist. Die Kerninformationen für die Prognose eines zukünftigen Signalbildwechsels, definiert als `movementEvent`, lassen sich aus den Werten `minEndTime`, `maxEndTime`, `confidence` und `likelyTime` ableiten. Das Signalbild selbst wird in der SPATEM durch den `eventState` repräsentiert.

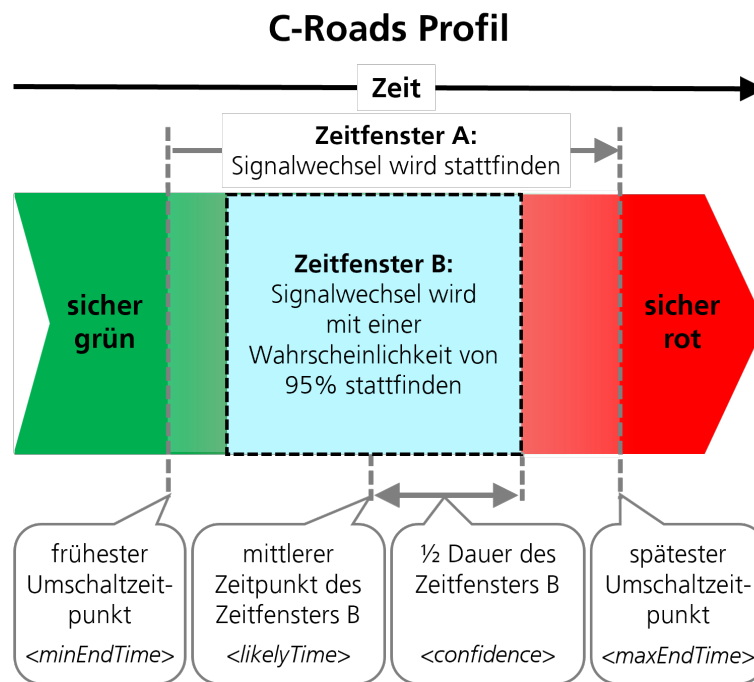


Abbildung 2.2: Illustration des C-Roads message profiles mit den vier Kerninformationen eines prognostizierten Signalbildwechsels: *minEndTime*, *likelyTime*, *confidence* und *maxEndTime*.

In der Folge werden die Begriffe, welche aus dem C-Roads Profil abgeleitet sind, näher beschrieben:

- **Zeitfenster A:** Beschreibt den Zeitbereich, in dem ein Signalwechsel sicher stattfinden wird. Dieser Bereich wird durch die Werte *minEndTime* und *maxEndTime* für Start und Ende des Zeitfensters definiert.
- **minEndTime:** Beschreibt den frühesten Umschaltzeitpunkt aus dem aktuellen *eventState*. Vor diesem Zeitpunkt ist der aktuelle *eventState* noch sicher aktiv.
- **maxEndTime:** Beschreibt den spätesten Umschaltzeitpunkt aus dem aktuellen *eventState*. Nach diesem Zeitpunkt ist der aktuelle *eventState* sicher nicht mehr aktiv.
- **Zeitfenster B:** Beschreibt den Zeitbereich, in dem der Umschaltzeitpunkt des aktuellen *eventState* mit einer Wahrscheinlichkeit von mindestens 95% liegen wird. Das Fenster B ergibt sich aus den Werten der *likelyTime* und *confidence*, d.h. es

erstreckt sich von `likelyTime - confidence` bis zur `likelyTime + confidence` und entspricht damit dem Doppelten des jeweiligen `confidence`-Werts. Hinweis: `confidence`-Werte werden dabei als Zeitwerte interpretiert (siehe Tabelle 2.1).

- **likelyTime:** Referenzzeitpunkt für das Zeitfenster B, welcher idealerweise den wahrscheinlichsten Umschaltzeitpunkt darstellt.
- **confidence:** Die `confidence` spannt um die `likelyTime` das Zeitfenster B auf. Der eigentliche Wert der `confidence` wird als eine Zeitintervallklasse zwischen 0 bis 15 angegeben (siehe Tabelle 2.1). Je kleiner der Wert der `confidence` ist, desto größer ist das aufgespannte Zeitfenster B.

Tabelle 2.1: **Konfidenz: Zuordnung zwischen Klasse und Zeitintervall**

Wert	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Zeitintervall [s]	>15.0	13.5	12.0	10.5	9.0	7.5	6.5	5.5	4.5	3.5	2.5	2.0	1.5	1.0	0.5	0

Aufzeichnungsformat PCAP: Die C-Roads Plattform hat zur Verbesserung der Interoperabilität und einer herstellerübergreifenden Harmonisierung ein Format zur Aufzeichnung der V2X-Kommunikation, welches auch die versendeten oder empfangenen V2X-Nachrichten enthält, etabliert, genannt Packet Capture (PCAP) [8]. Dieses dient generell dem Mitschnitt, speziell der Aufzeichnung der V2X-Nachrichten in einem einheitlichen Format und dem Austausch von entsprechend erhobenen Log-Dateien. Im Vorhaben wurden SPATEM-Nachrichten im PCAP-Format genutzt.

Verwendete Realdaten: Es wurde im Vorhaben mit realen Signalisierungs- bzw. Prognosedaten gearbeitet. Diese Daten wurden an fünf signalisierten Hamburger Knotenpunkten über einen Zeitraum von 24 Stunden erhoben und in Form von SPATEM-Nachrichten als PCAP-Dateien aufgezeichnet. Die Prognosebewertung fand ausschließlich auf Basis des zum aktuellen Signalzustand gehörenden `movementEvents` innerhalb einer SPATEM-Nachricht statt. Alle Knotenpunkte verwendeten das C-Roads Profil. Die Auswahl der Knotenpunkte berücksichtigte verschiedene verkehrstechnische Ausprägungen, z. B. hinsichtlich geometrischer und steuerungstechnischer Komplexität, um ein möglichst breites Spektrum an Randbedingungen in der Datenanalyse und bei der Konzeptentwicklung für eine Qualitätsbewertungsmetrik abdecken zu können. Neben den in den SPATEM-Nachrichten enthaltenen Informationen wurden keine anderweitigen Datenquellen genutzt. Dies gilt insbesondere für Mitschnitte von LSA-Steuergeräten, die nicht immer unmittelbar verfügbar

sind und im Aufzeichnungsformat keinem allgemein festgelegten Standard folgen, was eine Berücksichtigung erschweren würde.

2.2 Experteninterviews

Im Rahmen der Erarbeitung der Qualitätsbewertungsmetrik wurden leitfadengestützte Experteninterviews durchgeführt. Insgesamt wurden 16 deutsche und europäische Expertinnen und Experten aus unterschiedlichen Stakeholdergruppen befragt, darunter Vertreter von Städten und Kommunen, Infrastruktur- und ÖPNV-Betreibern, LSA-Herstellern, der Automobilindustrie, Forschungseinrichtungen sowie einzelner Prognoseanbieter.

2.2.1 Zielsetzung und Vorgehen

Ziel der Interviews war es, Anforderungen an Schaltzeitprognosen und eine mögliche Bewertung zu erfassen. Als Grundlage diente ein Fragenkatalog mit 18 Fragen (siehe Anhang Fragenkatalog). Da die Stakeholder aus unterschiedlichen Domänen stammen, wurden für die Interviews zusätzlich ein Glossar sowie eine schematische Abbildung vorbereitet, die als gemeinsame Begriffsgrundlage und Diskussionsbasis dienen.

2.2.2 Schwerpunktthemen und zentrale Ergebnisse der Expertenbefragung

Der Fragenkatalog wurde so konzipiert, dass sowohl Anforderungen an die Bereitstellung und Nutzung von Schaltzeitprognosen als auch eine darauf aufbauende Qualitätsbewertungsmetrik abgedeckt werden. Inhaltlich lassen sich die Fragen vier Schwerpunktthemen zuordnen. Die Auswertung der Interviews zeigt ein heterogenes, insgesamt aber konsistentes Bild zu Einsatz, Bewertung und Anforderungen an Schaltzeitprognosen. Die wichtigsten Erkenntnisse zu den vier Schwerpunktthemen sind im Folgenden zusammengetragen:

(1) Einsatz von Prognosen: Wie werden heute Schaltzeitprognosen eingesetzt und welche Nutzergruppen, das heißt Verkehrsteilnehmer, die als Endnutzer die Prognoseinformationen erhalten (z. B. MIV, ÖPNV, Radverkehr), sollen davon profitieren?

- Prognosen werden vor allem genutzt, um sie über V2X-Nachrichten (SPATEM) für fahrzeugeitige Dienste bereitzustellen. Im Vordergrund stehen GLOSA-ähnliche Anwendungen für den MIV; Radfahrer und ÖPNV spielen eine untergeordnete Rolle.
- Unter den Stakeholdern dominiert die Nutzung des C-Roads-Profiles; der SAE-Standard wird teilweise verwendet. In der praktischen Nutzung einer Prognose stehen insbesondere `likelyTime` und `confidence` im Fokus; `minEndTime` und `maxEndTime` werden deutlich seltener verwendet.

(2) Bewertung von Prognosen: In welcher Form werden Prognosen derzeit bewertet, welche Methoden, Werkzeuge oder Kennzahlen kommen dabei zum Einsatz und wo bestehen Defizite?

- Eine systematische, datenbasierte Bewertung der Prognosequalität erfolgt bislang nicht. Ein einheitliches Vorgehen oder etabliertes Bewertungstool existiert nicht; vereinzelt existieren herstellereigene, technische Reports oder es werden einfache Visualisierungen genutzt, um Kennwerte zu betrachten.
- Häufig basiert die Einschätzung der Prognosequalität auf subjektivem Expertenwissen, sodass eine reproduzierbare Bewertung aktuell kaum möglich ist.

(3) Erwartungen an eine Prognose: Welche funktionalen Erwartungen stellen Anwender und Endnutzer an Schaltzeitprognosen, und welche Kriterien lassen sich daraus für Auswertungen auf Basis vorhandener Realdaten ableiten?

- Die in einem Profil definierten Vorgaben an die Werte `minEndTime`, `maxEndTime`, `confidence` und `likelyTime` sollten zuverlässig eingehalten werden.
- Prognosen sollten dauerhaft verfügbar sein und mit Annäherung an den tatsächlichen Umschaltzeitpunkt genauer werden, d. h. die Abweichung sollte abnehmen und die Eintreffwahrscheinlichkeit zunehmen. Gleichzeitig ist bei wechselnden Verkehrsbedingungen und daraus resultierenden Steuerungseingriffen (durch z. B. Verkehrsabhängigkeit, ÖPNV- oder Einsatzwagenbevorrechtigungen) eine Aktualisierung erforderlich, ohne dass es außerhalb solcher Situationen zu sprunghaften Änderungen kommt.

(4) Vorgaben an eine Qualitätsbewertungsmetrik: Welche Eigenschaften muss eine Qualitätsbewertungsmetrik besitzen, um den Bedürfnissen der Anwender gerecht zu werden, und in welchem Anwendungskontext soll sie eingesetzt werden können?

- Von einer Qualitätsbewertungsmetrik wird erwartet, dass sie mehrere aussagekräftige Kennwerte bereitstellt und diese verständlich mit Schulnoten oder Stufen aufbereitet. Die Metrik sollte maschinenlesbar implementierbar sein und eine Auswertung nach Zeitbereichen oder Signalgruppen ermöglichen.
- Genannte Anwendungskontexte umfassen die Überprüfung der Prognosequalität, die Vergleichbarkeit von Prognosen und Prognoseanbietern, die Unterstützung bei Entscheidungen zum Einsatz bzw. der Standortauswahl von Prognosen, die Optimierung von LSA-Steuerungen sowie die Formulierung von Anforderungen in Ausschreibungen.

Abbildung 2.3 fasst die wichtigsten Antworten der Experteninterviews zur Übersicht in einer Mindmap zusammen:

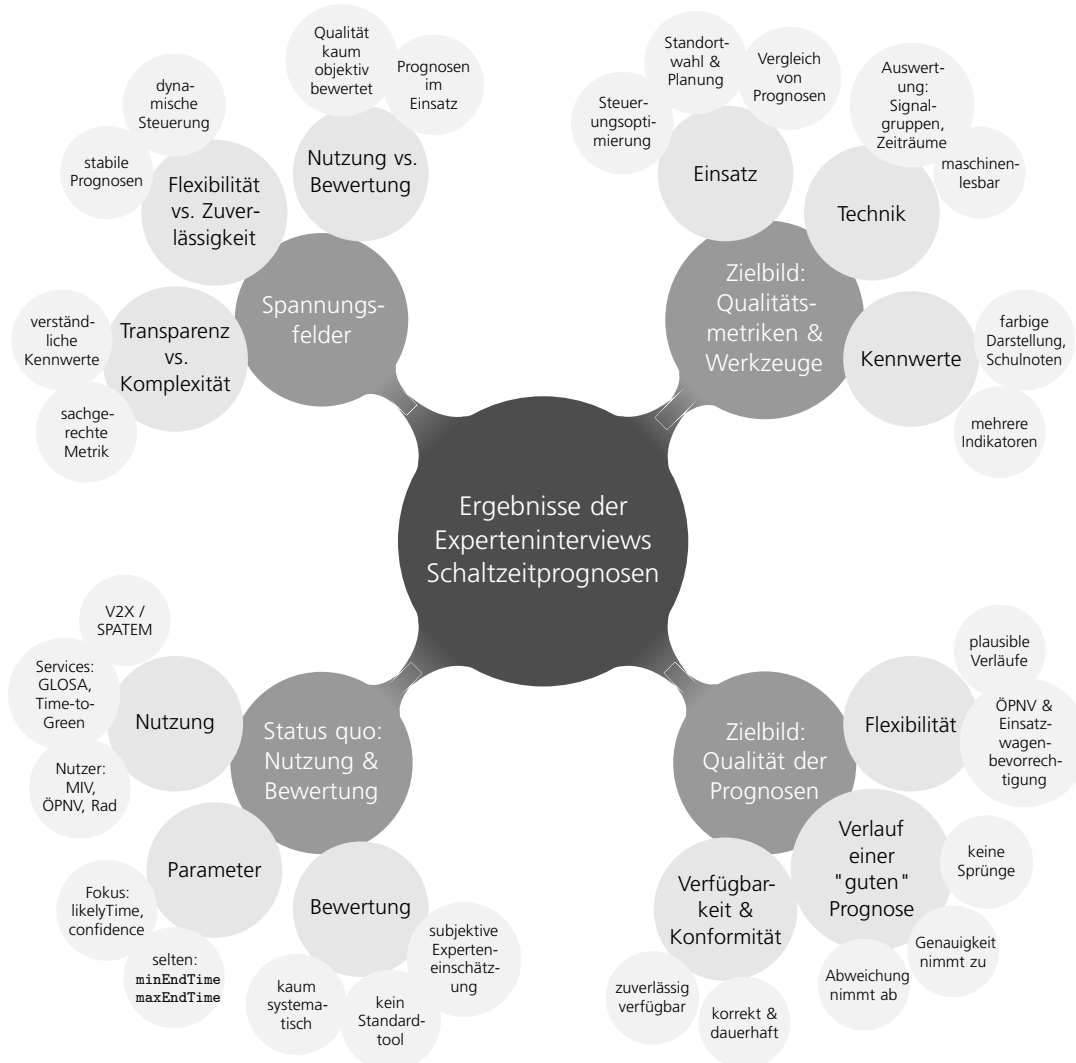


Abbildung 2.3: Schematische Zusammenfassung der Experteninterviews zu Status quo, Anforderungen und Spannungsfeldern bei Schaltzeitprognosen und deren Qualitätsbewertung.

Mit den Experteninterviews konnten wichtige Einblicke in das Themenfeld „Schaltzeitprognosen“ sowie deren Einsatz und gegenwärtige Bewertung gesammelt werden. Hierzu zählt zum Beispiel, welche Verhaltensweisen eine vermeintlich gute Prognose aufweisen sollte, oder welche Kriterien abgeleitet werden könnten, um diese für die Bewertung einer Prognose heranzuziehen. Des Weiteren konnten erste Aspekte aufgedeckt werden, wie eine Qualitätsbewertungsmetrik selbst aufgebaut sein könnte. Es wurden außerdem in-

haltliche Spannungsfelder identifiziert, welche bisher bei der Einschätzung und Bewertung von Prognosen nicht adressiert werden. Nach Auswertung der Experteninterviews können folgende Anforderungen an eine Prognose und eine Qualitätsbewertungsmetrik abgeleitet werden.

- **Profilkonformität:** Eine Prognose und deren Daten sollten die im Standard sowie in einem Profil definierten Vorgaben erfüllen. Nur standard- und profilkonforme Daten ermöglichen die Nutzung einer Prognose.
- **Verfügbarkeit:** Wenn eine Prognose zum Einsatz kommt, sollte diese stetig verfügbar sein. Nur so können Verkehrsteilnehmer bei der Annäherung und dem Passieren einer LSA dauerhaft unterstützt werden.
- **Zuverlässigkeit:** Eine Prognose sollte genauer und zuverlässiger werden, je kleiner der entsprechende Prognosehorizont ist bzw. wird.
- **Vergleichbarkeit:** Eine Qualitätsbewertungsmetrik sollte eine Bewertung und den herstellerunabhängigen Vergleich von Prognoseverfahren und Prognosen ermöglichen.
- **Konfigurierbarkeit:** Der Einsatz einer Prognose in verschiedenen Anwendungsfällen erfordert eine konfigurierbare Bewertungsmetrik. Insbesondere sollte die Metrik eine Auswahl und Gewichtung von Zeiträumen und Signalgruppen ermöglichen, welche von besonderem Interesse sein könnten.
- **Verständlichkeit und Nachvollziehbarkeit:** Die Qualitätsbewertungsmetrik sollte die Bewertung einer Prognose mit mehreren aussagekräftigen Kennwerten bereitstellen und diese verständlich in Form von Schulnoten oder Stufen darstellen.
- **Technische Umsetzung:** Eine Qualitätsbewertungsmetrik sollte maschinenlesbar sein und automatisiert umgesetzt und eingesetzt werden können.

2.3 Anwendungsbezogene Konfigurierbarkeit

Je nach Anwendungszweck können sich sehr unterschiedliche Anforderungen an eine Prognose ergeben. Eine Kategorisierung nach Endnutzern, Services und Anwendern stellt nachfolgend anhand von Beispielen dar, wie sich die Prognoseanforderungen entsprechend unterscheiden können:

- I. **Endnutzer:** Endnutzer sind Verkehrsteilnehmer bestimmter Verkehrsarten, zum Beispiel im MIV, Schwerlastverkehr, ÖPNV oder Fuß- und Radverkehr, die auf einer Prognose basierende Services während ihrer Fahrtvorgänge im Umfeld einer LSA verwenden.

Beispiel für die Relevanz des Prognosehorizonts: Die Fahrgeschwindigkeit eines Verkehrsteilnehmers hat große Auswirkungen darauf, wie viel Zeit und Fahrweg verbleiben, um auf eine Prognose noch rechtzeitig reagieren und das Fahrverhalten entsprechend anpassen zu können.

- II. **Services:** Services sind Dienste, die Endnutzer unter Verwendung einer Prognose bei der Annäherung und dem Passieren einer LSA unterstützen können, zum Beispiel GLOSA, Time-to-Green, Time-to-Red sowie Fahrerassistenzsysteme (ACC, Ampelanfahrassistent, Bremsassistent).

Beispiel für die Relevanz des Nachrichteninhalts einer Prognose: Ein Time-to-Green Service erfordert präzise Vorhersagen, wann die Freigabezeit beginnt. Der Startzeitpunkt der Sperrzeit ist hierfür irrelevant.

- III. **Anwender:** Anwender sind Akteure, die Prognosen bereitstellen oder diese für eigene Services nutzen, zum Beispiel Kommunen und Städte, Unternehmen des Automobilsektors, Verkehrsunternehmen, LSA-Betreiber, LSA- und Prognosehersteller oder App-Anbieter.

Beispiel für den Fokus auf die verkehrsartbezogene Prognose: Verkehrsunternehmen interessieren sich primär für die Prognose von ÖPNV-Signalgebern; die übrigen Signalgeber sind wenig relevant.

Um dieser Vielfalt der teils sogar widersprüchlichen Anforderungen an eine Prognose gerecht zu werden, sollte die Qualitätsbewertungsmetrik konfigurierbar sein. Das Konzept sollte vorsehen, dass der Fokus bei der Bewertung gezielt gesetzt werden kann. Dies könnten z. B. bestimmte Nachrichteninhalte, Zustände, Signalgeber oder Prognosehorizonte sein, die je nach Anwendungszweck für einen Endnutzer, Service oder Anwender von besonderem Interesse sind.

2.4 Kernanforderungen

Aus den technischen Rahmenbedingungen, welche in Abschnitt 2.1 zusammengefasst sind, und den unterschiedlichen Zielkonflikten sowie Erwartungen in Bezug auf Schaltzeitprognosen, welche sich aus den Experteninterviews in Abschnitt 2.2 identifizieren lassen, werden folgende Kernanforderungen an ein Qualitätsbewertungskonzept abgeleitet:

[A] Herstellen einer Vergleichbarkeit: Sowohl Prognoseverfahren als auch LSA-Steuerungen können sich stark voneinander unterscheiden. Es stellt sich daher die Frage, wie stark sich eigentlich Steuerungen in ihrer Dynamik unterscheiden und wie herausfordernd sich dementsprechend eine Vorhersage in Bezug auf die Steuerungen darstellt. Dynamische Steuerungen sind im Vergleich zu sehr statischen Steuerungen deutlich schwieriger zu prognostizieren.

⇒ Es bedarf daher eines Maßes zur Bewertung von Steuerungsdynamiken, um herstellerübergreifend eine Vergleichbarkeit zwischen Prognosedaten herstellen zu können.

[B] Prüfen der Nachrichten und Daten auf Verfügbarkeit und inhaltliche Konsistenz: Es wurde die klare Erwartung formuliert, dass Prognosen, wenn sie denn in der Praxis zur Anwendung kommen, stetig und dauerhaft gemäß der SPATEM-Definition und konsistent zum C-Roads Profil, verfügbar sein sollten.

⇒ Aus diesen genannten Anforderungen der Verfügbarkeit, Zuverlässigkeit und Profilkonformität von Prognosen lässt sich ableiten, dass eine systematische Prüfung der Nachrichten auf ihre Integrität und Plausibilität erforderlich ist.

[C] Bewerten der Prognosequalität: Als besondere Herausforderung und gleichzeitig Anforderung wurden die Zuverlässigkeit und Genauigkeit von Prognosen in Abhängigkeit von ihrem Vorhersagehorizont identifiziert. Im Kern wird erwartet, dass für kürzere Vorhersagen eine höhere Präzision vorliegt als für Prognosen mit sehr langen Vorausschauhorizonten. Deckungsgleich zu dieser Erwartung sollte auch die *confidence* für kleinere Horizonte höher ausfallen als für längere Vorhersagezeitfenster. Von einer zuverlässigen Prognose wird gefordert, dass diese nur möglichst kleine Abweichungen und keine sprunghaften Änderungen aufweist.

- ⇒ Um die eigentliche Qualität einer Prognose bewerten zu können, bedarf es eines Maßes zur Abbildung der Güte, welches die Abhängigkeiten zwischen Prognosehorizont, Genauigkeit und *confidence* berücksichtigt.

[D] Verständlichkeit und gleichzeitige Flexibilität einer Bewertungsmetrik: Angesichts der Komplexität von Schaltzeitprognosen und der bisher vorausgesetzten Fachexpertise zu deren Bewertung wurde die Vereinheitlichung, Vereinfachung und Verständlichkeit einer Bewertungsmetrik hervorgehoben. Zudem hat Abschnitt 2.3 die Anforderung einer anwendungsorientierten Konfigurierbarkeit dargestellt, welche sich unter Anderem aus den sehr unterschiedlichen Einsatzzwecken einer Qualitätsbewertung zwischen den sehr heterogenen Stakeholdergruppen ergeben hat.

- ⇒ Um eine herstellerunabhängige und zugleich für unterschiedliche Anwendergruppen leicht nachvollziehbare Bewertung anhand von Qualitätsstufen zu ermöglichen, sollte die Bewertungsmetrik komplexe Einzelkennwerte auf wenige, klar definierte Qualitätsklassen abbilden, die für verschiedene Anwendungsprofile konfigurierbar sind und hierdurch zugleich die notwendige Flexibilität bereitstellen.

Das folgende Bewertungskonzept in Kapitel 3 leitet sich in seinen Grundzügen aus diesen Kernanforderungen ab. Weitere Detailanforderungen aus den Experteninterviews und, in Teilen, aus den Erfahrungen im Umgang mit den Realdaten im Verlauf der Datenanalyse fließen zusätzlich in bestimmte Teilaspekte der Gesamtbewertung ein und sind Kapitel 4 eingebunden. Kapitel 5 greift insbesondere die Konfigurierbarkeit der Einzelbewertungen auf.

3 Bewertungskonzept

Das Vorhaben zur Entwicklung eines Bewertungskonzepts verfolgt das Ziel, eine Gesamtbewertung in Analogie zu einem Schulnotensystem abzuleiten, ähnlich wie es aus dem Konzept zur Einteilung der Qualitätsstufen des Verkehrsablaufs im Handbuch für die Bemessung von Straßenverkehrsanlagen (HBS) [9, Part S4.2] bekannt ist. Das Maß zur Bewertung der Prognosequalität setzt sich dabei aus mehreren Teilaspekten und Arbeitsschritten zusammen, wie in Abbildung 3.1 dargestellt.

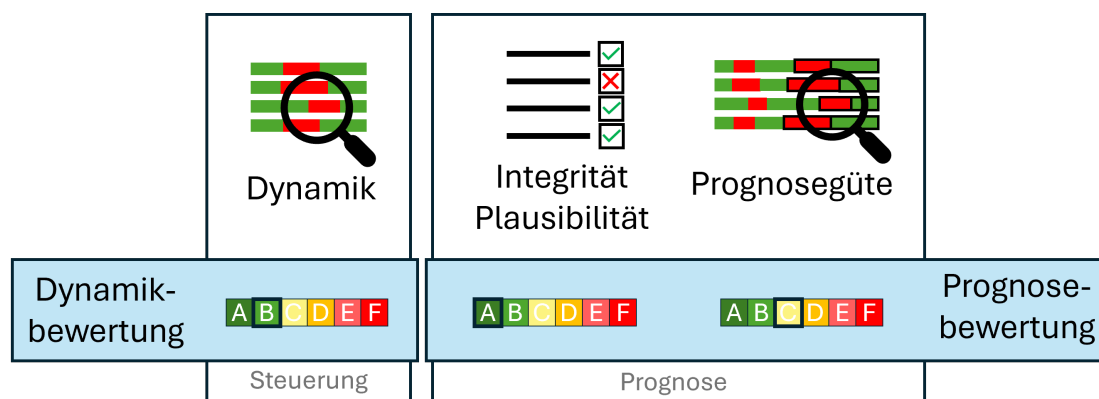


Abbildung 3.1: Illustration des Konzepts zur Bewertung von Schaltzeitprognosedaten. Eine Vergleichbarkeit zwischen Steuerungen wird durch die Bewertung der *Dynamik* sichergestellt; die Qualitätsbewertung einer Prognose erfolgt über die Teilaspekte *Integrität und Plausibilität* sowie *Prognosegüte*.

Um eine gezielte Bewertung je nach Anwendungsfall, etwa für eine Signalgruppe, durchführen zu können, muss zunächst eine geeignete Auswahl der zu betrachtenden Merkmale getroffen werden, die sich an den spezifischen Eigenschaften der jeweiligen Signalgruppe orientiert. Soll beispielsweise zunächst nur die grundsätzliche Tauglichkeit eines Datensatzes für eine spätere, detailliertere Auswertung geprüft werden, kann bereits die Betrachtung einzelner Merkmale im Rahmen einer Plausibilitätsprüfung ausreichen, um eine erste Bewertung der Datenqualität in diesem Einzelfall zu erhalten. Eine solche Abwägung einzelner

Merkmale ist im Bewertungskonzept durch eine flexible Konfiguration der Bewertungsmetrik vorgesehen. Hingegen erfordert eine deutlich umfassendere Analyse, etwa zum Vergleich zweier Knotenpunkte, eine umfangreichere, gezielte Auswahl von Signalgruppen, Signalzuständen und Prognosemerkmalen bis hin zu spezifischen Gewichtungen dieser Merkmale im Datensatz. Eine solche Konfiguration der Bewertungsmetrik setzt daher eine klare Festlegung der Analyseziele voraus, um dem jeweiligen Anwendungsfall und der Komplexität des Bewertungskontexts gerecht zu werden. In Kapitel 5 wird hierzu eine Beispielformatung für ein mögliches Anwendungsszenario diskutiert. Unabhängig von der gewählten Konfiguration sind für eine sachgerechte Bewertung grundsätzlich folgende Schritte erforderlich:

- I. In einem ersten Schritt wird die **Dynamik** einer betrachteten Steuerung ermittelt, da diese die Prognostizierbarkeit von Steuerungen maßgeblich beeinflusst.
- II. Im zweiten Schritt werden die **Integrität** und **Plausibilität** der Daten aus einer Nachricht bzw. einem Datensatz geprüft. Hierbei werden formale Mindestanforderungen an die Prognose untersucht, welche sich logisch aus den Daten ergeben müssen. Ebenso wird die Verfügbarkeit der Prognose selbst untersucht.
- III. In Schritt drei wird die **Prognosegüte** bestimmt, indem der eigentliche Inhalt der Prognose mit den tatsächlichen Umschaltzeitpunkten abgeglichen wird.
- IV. Im letzten Schritt wird die Steuerungsdynamik als eigenständiger Kennwert ausgewiesen. Davon getrennt werden die Ergebnisse der Nachrichtenintegrität und -plausibilität als auch der Prognosegüte mit definierten Gewichtungen für ihre jeweiligen Teilaspekte zu einer **Gesamtwertung** für die Prognose zusammengeführt.

In den folgenden Abschnitten 3.1–3.4 werden die entsprechenden Betrachtungsgrößen im Detail erläutert, definiert und anhand von Beispielen näher diskutiert.

3.1 Steuerungsdynamik

Das Eintreten zufälliger Ereignisse in einem betrachteten System beeinflusst maßgeblich dessen Vorhersagbarkeit. Beispielsweise sind die Signalzustände einer Festzeitsteuerung (schematisch dargestellt in Abbildung 3.2 A), geradezu trivial zu prognostizieren. Eine leicht dynamische Steuerung (siehe Abbildung 3.2 B) ist gegebenenfalls noch gut prognostizierbar, während im Vergleich hochflexible Steuerungen (siehe Abbildung 3.2 C) durch die

Anpassung an zufälligen Verkehr gegebenenfalls kein klares Muster erkennen lassen. Die aktuelle Verkehrslast hat dabei einen wichtigen Einfluss auf den Grad der beobachtbaren Dynamik. So kann sich beispielsweise eine hochflexible Steuerung im Bereich einer hohen Verkehrslast tatsächlich sehr dynamisch verhalten, während im Bereich einer sehr geringen Verkehrslast bzw. bei Sättigung ein starres Verhalten auftreten kann. In den letzteren Fällen werden minimal bzw. maximal definierte Freigabezeiten geschaltet, die mit einer geringen Dynamik gleichzusetzen sind. Verschiedene weitere Einflussfaktoren sind hier denkbar, wie beispielsweise häufige Priorisierungen des ÖPNV oder starke Fuß- und Radverkehrsströme, die durch entsprechende Anforderungen innerhalb einer dynamischen Steuerung flexibel berücksichtigt werden.

Die Prognosen zweier Steuerungen im Vergleich zu bewerten, erscheint somit nur sinnvoll, wenn beide Steuerungen in vergleichbaren Verkehrslastsituationen eine ähnliche Dynamik und damit eine ähnliche Prognostizierbarkeit aufweisen. Die Dynamik einer Steuerung beschreibt daher im Folgenden einen Index, der die Zufälligkeit des beobachteten Verhaltens einer Steuerung charakterisiert. Dieser Dynamikindex dient als Indikator der Vergleichbarkeit unterschiedlicher Steuerungen im Hinblick auf ihre Prognosequalität.

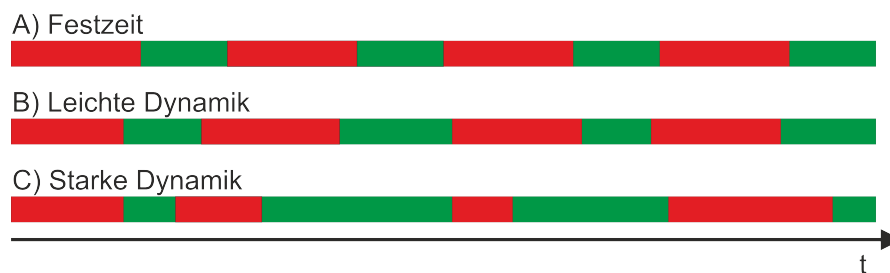


Abbildung 3.2: **Schematische Darstellung verschiedener Dynamik-Ausprägungen von Steuerungen von A) mit sich exakt wiederholenden Schaltvorgängen; über B) mit geringen Schwankungen der Signallaufzeiten; bis C) mit großen Schwankungen der Signallaufzeiten.**

Bei der Bewertung der Steuerungsdynamik liegt der Fokus auf dynamischen Signalzuständen. Als dynamische Signalzustände werden solche Signalzustände verstanden, deren Signallaufzeiten variieren können, etwa durch das Dehnen einer Phase. Typischerweise betrifft dies nur die Signalzustände *Gesperrt* (Rot) und *Freigabe* (Grün). Statische Signalzustände weisen demgegenüber feste Laufzeiten auf; hierzu zählen vor allem *Räumung* (Gelb) und *Vorbereitung* (Rot-Gelb). Da statische Signalzustände nichts zur Dynamik einer Steuerung beitragen, werden sie im Folgenden nicht weiter betrachtet.

Otto et al. [10] beschäftigten sich bereits mit der Untersuchung der Steuerungsdynamik auf Basis von Signalbildern aus realen Daten. Dabei analysieren Sie für Grünzeiten Start, Ende

und Laufzeit sowie deren Verteilungen, um Stabilitätsindikatoren bestimmen zu können. Mithilfe dieser Stabilitätsindikatoren lässt sich die Prognostizierbarkeit und Eignung einer LSA-Steuerung für den Einsatz in GLOSA-Services beurteilen. In Anlehnung daran wird hier die Steuerungsdynamik anhand der Start- und Endzeitpunkte eines `eventState` bewertet.

In den SPATEM-Nachrichten, die der folgenden Auswertung zugrunde liegen, werden die Signalzustände durch die Informationen in einem `eventState` abgebildet. Signalzustand und `eventState` können daher in der weiteren Analyse weitgehend gleichgesetzt werden. Ein Besonderheit ist die zusätzliche Unterscheidung des `eventState` für *Räumung* und *Freigabe* in *gesichert* und *ungesichert*; diese Differenzierung lässt sich am optischen Signalbild allein nicht erkennen.

Dynamische Zustände: Zur Bewertung der Steuerungsdynamik werden mehrere Aspekte betrachtet. Abbildung 3.3 veranschaulicht diese Faktoren wie folgt:

- Es werden die beobachteten Signallaufzeiten, also die Länge jeweiliger Zeitintervalle, untersucht (siehe Abbildung 3.3 A). Je mehr unterschiedliche Signallaufzeiten und je zufälliger ihre Verteilung, desto schwieriger ist die korrekte Prognose der Zeitintervalle.
- Rahmenpläne können in der Steuerung gewisse Verhalten erzwingen und so Muster erzeugen. Diese Muster werden meist durch feste Ein- oder Ausschaltzeitpunkte der Signaltbilder sichtbar. Aus diesem Grund werden auch die Abstände zwischen den Start- und Endzeitpunkten der Signallaufzeiten (siehe Abbildung 3.3 B und C) untersucht.

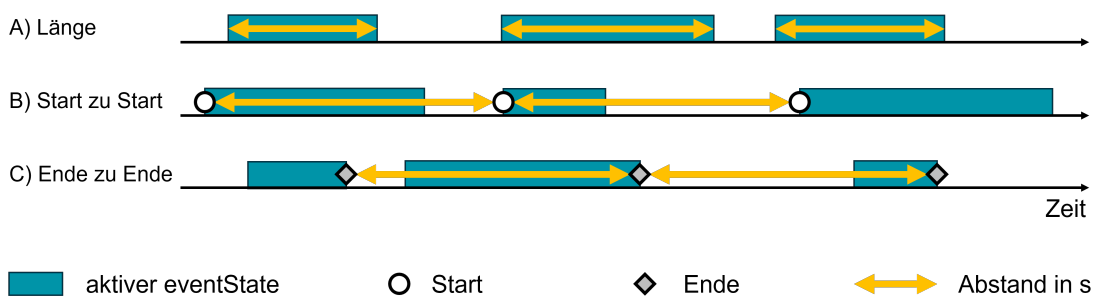


Abbildung 3.3: **Veranschaulichung der Hauptaspekte zur Analyse der Steuerungsdynamik: A) Länge der aktiven `eventStates`, B) Start-zu-Start-Beziehung der `eventStates` und C) Ende-zu-Ende-Verhältnis der `eventStates`.**

Informationsentropie: Zur Bestimmung der Zufälligkeit der beobachteten Werte wird die Informationsentropie nach Shannon [11] eingesetzt; die detaillierten Berechnungsschritte sind in Abschnitt 4.1 beschrieben. Sie ist ein Maß für die Unvorhersehbarkeit der beobachteten Zeitabstände. Die Informationsentropie steigt tendenziell mit der Anzahl der beobachteten Zeitabstände, sofern diese ausreichend häufig auftreten und damit zur Gesamtunsicherheit beitragen. Bei wenigen beobachteten Zeitabständen oder einer hohen Wahrscheinlichkeit für einige wenige Zeitabstände sinkt die Informationsentropie hingegen. Sie eignet sich damit für die Quantifizierung der Abstände bei Signallaufzeitbeginn und Signallaufzeitende (siehe Abbildung 3.3 B und C).

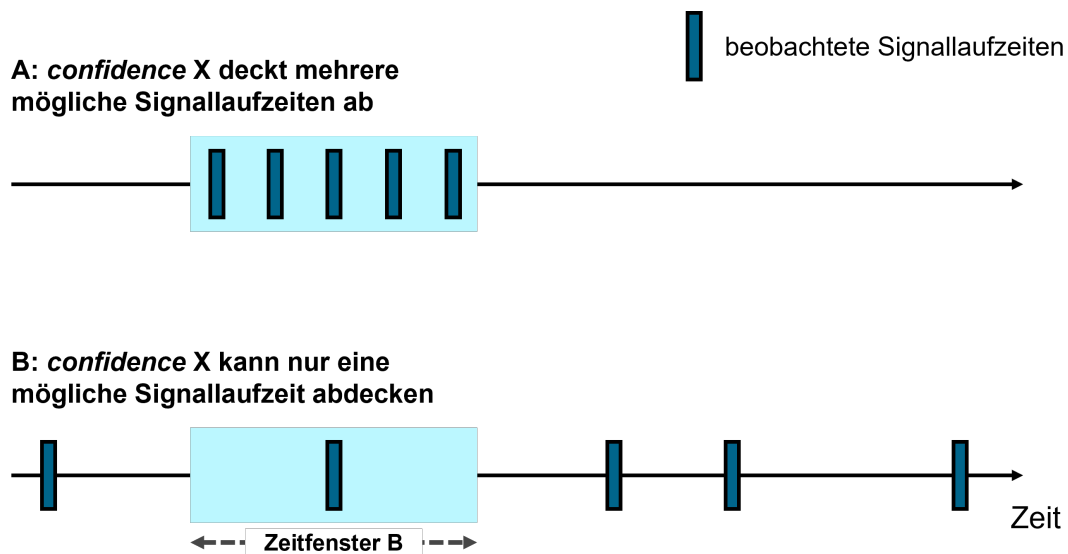


Abbildung 3.4: Die Konfidenz (*confidence*) erlaubt die Abdeckung mehrerer Signallaufzeiten, wenn die auftretenden Signallaufzeiten eng beieinander liegen.

Confidence: Zusätzlich kann bei der Betrachtung der Signallaufzeiten für die Bewertung der Steuerungsdynamik die *confidence* einbezogen werden (siehe Abbildung 3.4). Liegen die beobachteten Signallaufzeiten in einem relativ engen Wertebereich, kann dieser Bereich gegebenenfalls durch eine geeignete *likelyTime-confidence*-Kombination vollständig abgedeckt sein, wie in Abbildung 3.4 A dargestellt. Treten dagegen größere Lücken im Band beobachteter Signallaufzeiten auf, kann dieselbe *likelyTime-confidence*-Kombination nicht mehr alle Signallaufzeiten abdecken (Abbildung 3.4 B). Aus diesem Grund wird die Shannon'sche Informationsentropie dahingehend angepasst, dass die Intervallabstände in den beobachteten Signallaufzeiten zu einer weiteren Erhöhung der Entropie beitragen und damit die Unvorhersehbarkeit der Steuerung in solchen Fällen besser abgebildet wird.

3.2 Nachrichtenintegrität und -plausibilität

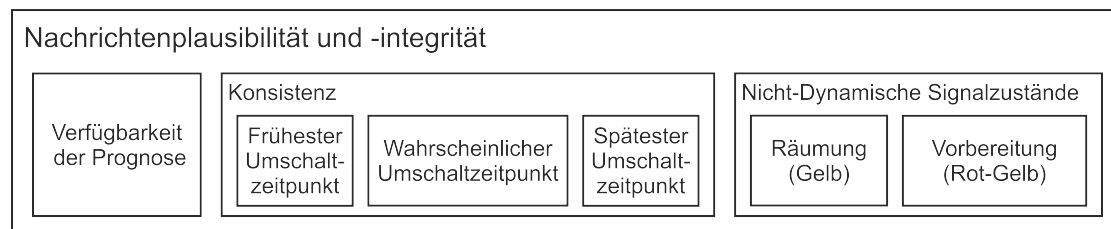


Abbildung 3.5: Überblick über die Zusammensetzung der Plausibilitäts- und Integritätsbewertung.

Zur Bewertung der Nachrichtenintegrität und -plausibilität werden die Nachrichten ohne Bezug auf die tatsächlichen Umschaltzeitpunkte geprüft. Einen Überblick dazu bietet Abbildung 3.5. Die Grundlage bildet die Verfügbarkeit der Prognose, da eine hohe Verfügbarkeit die Voraussetzung für die Nutzung der Prognose in unterschiedlichen Anwendungen (z. B. GLOSA oder Time-to-Green) darstellt. Das C-Roads Profil und Rahmenbedingungen aus der Verkehrstechnik verlangen zudem gewisse Abhängigkeiten der Nachrichteninhalte untereinander, insbesondere Konsistenzbeziehungen der Umschaltzeitpunkte und eine feste Dauer nicht-dynamischer Signalzustände. Die einzelnen Aspekte werden im Folgenden näher erläutert.

Verfügbarkeit: Die Verfügbarkeit der Prognose ergibt sich aus dem Anteil der SPATEM-Nachrichten, die eine profilkonforme Prognose enthalten. Zusätzlich wird mindestens eine SPATEM-Nachricht pro Sekunde des Beobachtungszeitraums erwartet, da LSA-Steuerungen typischerweise in diesem festen Zeitintervall arbeiten und Zustandsänderungen vornehmen können.

Verschiebung frühester und spätester Umschaltzeitpunkte: Die Werte `minEndTime` und `maxEndTime` aus einem `eventState` beschreiben jeweils den frühesten beziehungsweise spätesten Umschaltzeitpunkt. Ein frühester Zeitpunkt `minEndTime` darf in aufeinanderfolgenden Nachrichten unverändert bleiben bzw. sich nur nach hinten verschieben. Ein spätester Zeitpunkt `maxEndTime` darf ebenfalls unverändert bleiben oder sich nur nach vorne verschieben. Abbildung 3.6 illustriert diese Plausibilitätsprüfung schematisch für diese beiden Fälle. Andere Verschiebungen als die beiden genannten werden als nicht plausibel betrachtet.

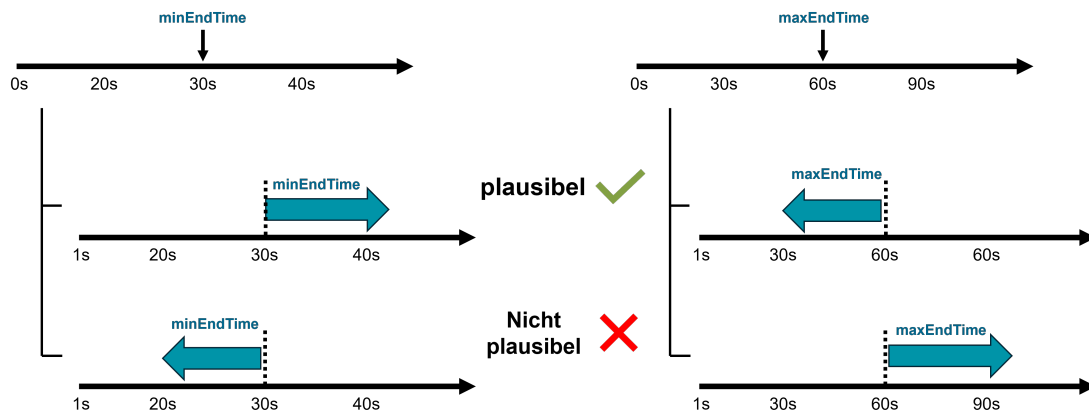
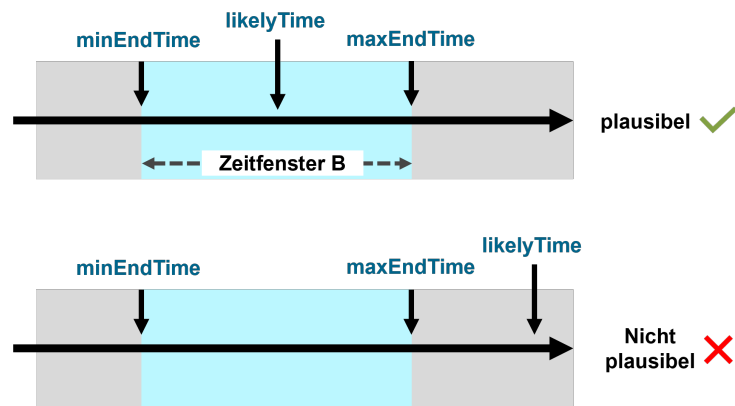


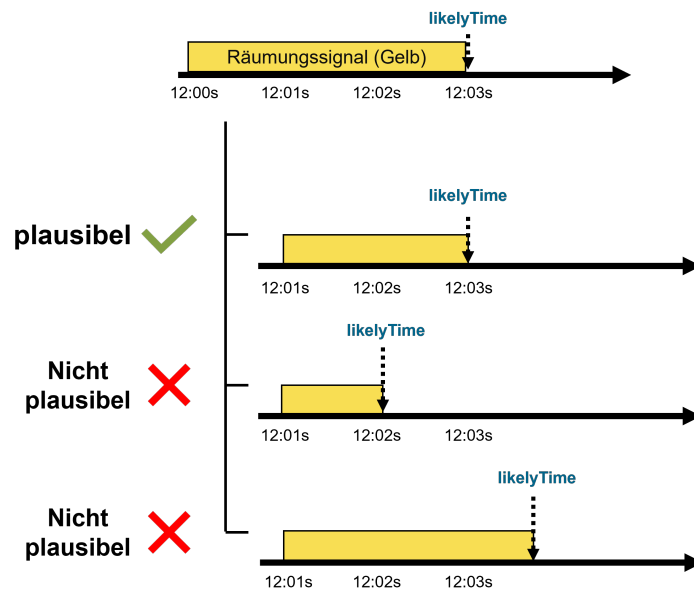
Abbildung 3.6: Beispielhafte Darstellung der Plausibilitätsprüfung für die Werte $minEndTime$ und $maxEndTime$.

Logische Konsistenzbedingungen: Zusätzlich wird geprüft, ob die Werte $minEndTime$, $maxEndTime$ und $likelyTime$ in einer logisch konsistenten zeitlichen Relation zueinander stehen. Abbildung 3.7a verdeutlicht diese Plausibilitätsprüfung für die $likelyTime$, welche zwischen der $minEndTime$ und $maxEndTime$ liegen muss. Außerdem muss geprüft werden, ob die $minEndTime$ kleiner als die $maxEndTime$ ist.

Signallaufzeit: Nicht-dynamische Signalzustände, also insbesondere Räumung (Gelb) und Vorbereitung (Rot-Gelb) haben eine feste Signallaufzeit und somit einen wohldefinierten und unveränderlichen Umschaltzeitpunkt. Verschiebt sich der prognostizierte Umschaltzeitpunkt nicht-dynamischer Signalzustände während der Prognose, das heißt während man sich im nicht-dynamischen Signalzustand selbst befindet, ist dies unplausibles Verhalten. Abbildung 3.7b veranschaulicht eine solche Plausibilitätsprüfung beispielhaft für eine Räumung (Gelb).



(a) Plausibilitätsprüfung der *likelyTime*-Prognose in Relation zu *minEndTime* und *maxEndTime*.



(b) Prüfung der Prognose eines Räumungssignals auf Plausibilität anhand unterschiedlicher *likelyTime*-Verschiebungen.

Abbildung 3.7: Beispielhafte Illustrationen zur Plausibilitätsprüfung der *likelyTime*.

3.3 Prognosegüte

Da bisher auf kein etabliertes Maß für die Bewertung einer Prognose zurückgegriffen werden kann, wird in der Folge ein Konzept für eine datengetriebene Beurteilung beschrieben. Insbesondere wird das Zusammenspiel aus Prognosehorizont, Prognosegenauigkeit und der zugehörigen *confidence* untersucht, da diese drei Aspekte sich untereinander stark beeinflussen, wie Abbildung 3.8 verdeutlicht. Zunächst wird dabei der maximale, C-Roads-konforme Prognosehorizont bestimmt. In dessen Rahmen werden die durchschnittlichen *confidence*-Werte beurteilt und einer Gewichtung nach ihrem zugehörigen Prognosehorizont unterzogen. Die einzelnen Schritte werden nachfolgend erläutert.

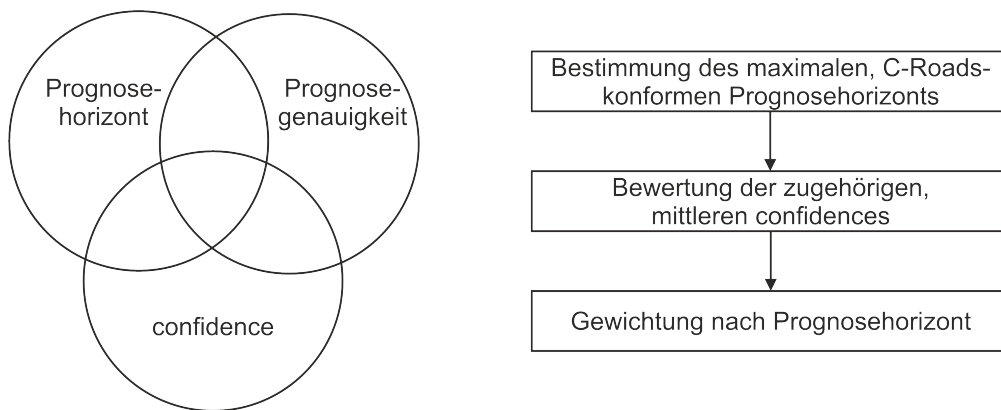


Abbildung 3.8: **Übersicht über die Zusammenhänge zwischen Prognosehorizont, Prognosegenauigkeit und *confidence* sowie die Ablaufschritte zur Bestimmung der Prognosegüte.**

Abweichung der Prognosegenauigkeit: In einem ersten Schritt wird auf Basis der vorliegenden Daten ermittelt, bis zu welchen Prognosehorizonten die Prognosen die im C-Roads Profil geforderte Mindestgenauigkeit von 95% erreichen. Der Prognosehorizont wird dabei durch die zugehörigen *likelyTimes* der Prognosen repräsentiert. Abbildung 3.9 zeigt zwei Beispiele für eine solche Datensatzanalyse im Hinblick auf die Prognosegenauigkeit. Für jede der Nachrichten wird bestimmt, ob die tatsächlichen Umschaltzeitpunkte innerhalb des akzeptablen Zeitfensters B liegen, das durch *likelyTime* und *confidence* aufgespannt wird. Liegen mehr als 95% der tatsächlichen Umschaltzeitpunkte im aufgespannten Zeitfenster, werden für diese *likelyTime* die Anforderungen des C-Roads Profils erfüllt (siehe Abbildung 3.9 Oben). Umgekehrt erfüllen *likelyTimes* außerhalb dieses Akzeptanzbereichs nicht die Anforderungen des C-Roads Profils (siehe Abbildung 3.9 Unten).

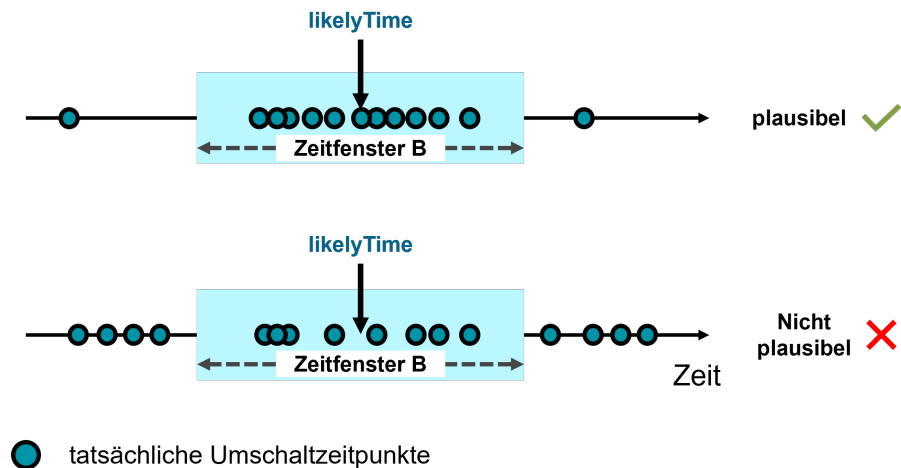


Abbildung 3.9: Erlaubte und unerlaubte Abweichung in der Prognosegenauigkeit.
Oben: 95% der tatsächlichen Umschaltzeitpunkte liegen innerhalb ihrer Prognosezeitfenster.
Unten: Mehr als 5% der tatsächlichen Umschaltzeitpunkte liegen außerhalb ihrer Prognosezeitfenster.

Prognosegenauigkeit versus Aussagekraft: Die Beurteilung einer Prognose allein anhand der Genauigkeit ist nicht besonders aussagekräftig. Dies soll mit Hilfe von Abbildung 3.10 verdeutlicht und diskutiert werden: Teilbild A1 veranschaulicht eine gute Prognose, da hier die `likelyTime` nahe am tatsächlichen Umschaltzeitpunkt bei gleichzeitig hoher `confidence` liegt. Eine solche Prognose wäre für einen Anwender sehr gut nutzbar. Teilbild A2 zeigt einen Fall, der im Vergleich zum später dargestellten Teilbild B2 eine deutlich bessere Prognose illustriert, da die `likelyTime` näher am tatsächlichen Umschaltzeitpunkt liegt. Allerdings kann ein Endnutzer den tatsächlichen Umschaltzeitpunkt in einer Live-Situation nicht kennen; aus dem sichtbaren Datenmuster allein ist daher nicht erkennbar, ob die entsprechende `likelyTime` tatsächlich genau oder ungenau ist. Die Interpretation von A2 ist damit ähnlich schwierig wie bei Teilbild B2. Teilbild B1 zeigt analog eine falsche Prognose: Die `likelyTime` weicht deutlich vom tatsächlichen Umschaltzeitpunkt ab, obwohl die `confidence` hoch ist. Die Prognose erscheint zunächst vertrauenerweckend, ist aber inhaltlich unzutreffend und damit für den Anwender problematisch. Teilbild B2 stellt schließlich eine Prognose dar, bei der die `likelyTime` sehr weit vom tatsächlichen Umschaltzeitpunkt entfernt ist. Bei gleichzeitig niedriger `confidence` wird ein so breites Zeitfenster für mögliche Umschaltzeitpunkte aufgespannt, dass der tatsächliche Umschaltzeitpunkt noch in dieses potenzielle Zeitfenster fällt. Formal wäre eine derartige Prognose damit noch korrekt. Ein solcher Fall kann bei Prognosen mit großem Vorhersagehorizont durchaus akzeptabel sein, ist für Prognosen in näherer Zukunft jedoch wenig hilfreich.

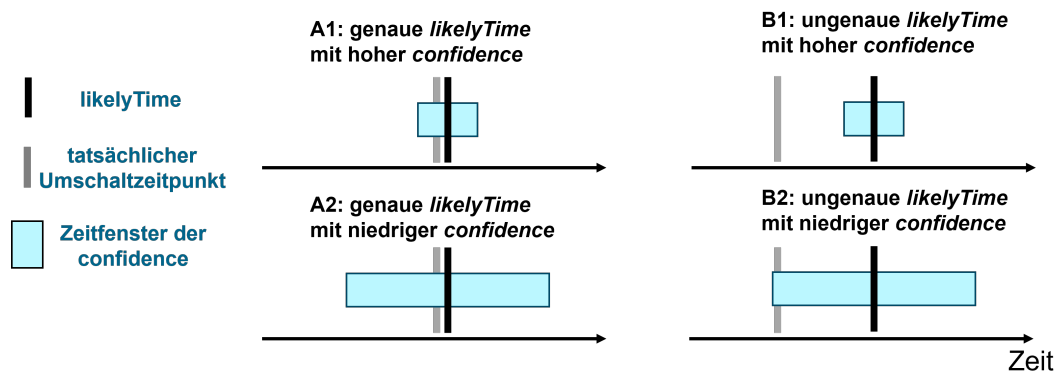


Abbildung 3.10: Zusammenspiel von *likelyTime* und *confidence*:

- A1) der tatsächliche Umschaltzeitpunkt wird mit hoher *confidence* getroffen,
- A2) der tatsächliche Umschaltzeitpunkt wird mit niedriger *confidence* getroffen und hat deswegen wenig Aussagekraft,
- B1) eine falsche Prognose und
- B2) eine Prognose mit ungenauer *likelyTime* wird durch eine niedrige *confidence* trotzdem richtig.

Diese Erwägungen zu Abbildung 3.10 gehen auf ein bei der Analyse der bereitgestellten Realdaten beobachtetes Phänomen zurück, das sich wie folgt zusammenfassen lässt:

Häufig wurde beobachtet, dass die *confidence* mit sinkender *likelyTime*, das heißt bei Prognosen in näherer Zukunft, deutlich abnimmt. Entgegen der Erwartung der Endnutzer werden die Prognosen damit immer unsicherer, je näher der prognostizierte Umschaltzeitpunkt rückt. In diesen Fällen steigt zugleich scheinbar die Genauigkeit der jeweiligen Prognosen, weil durch das breitere Prognosezeitfenster automatisch mehr Umschaltzeitpunkte in dieses Zeitfenster fallen.

Da unter anderem die Nutzerbefragung gezeigt hat, dass von einer Prognose erwartet wird, dass diese umso sicherer wird, je kürzer ihr Vorhersagehorizont ist, sollten Genauigkeit, *likelyTime* und *confidence* nicht isoliert, sondern stets gemeinsam betrachtet werden.

Prognosegüte als gewichtetes Maß:

Aus den zuvor dargelegten Überlegungen sowie den Anforderungen der Expertenbefragungen wird für die Bewertung der Prognosegüte ein gewichtetes Maß eingeführt:

Die Prognosegüte wird als gewichtete Summe über eine aggregierte *confidence*-Bewertung je *likelyTime* beschrieben.

Die Bewertung der *confidence* basiert auf der Breite des aufgespannten Prognosezeitfensters: Je breiter das Zeitfenster, desto niedriger fällt die Wertung aus. Zu diesem Zweck wird eine Bewertungsfunktion verwendet, die flach ausläuft, da die Nutzbarkeit ab einer gewissen Fensterbreite nicht mehr signifikant abnimmt. Eine solche Funktion fällt jedoch anfangs relativ stark ab. Dies ist ein Umstand, der zukünftig gegebenenfalls weiterer Optimierungen in der Bewertungsmethodik bedarf.

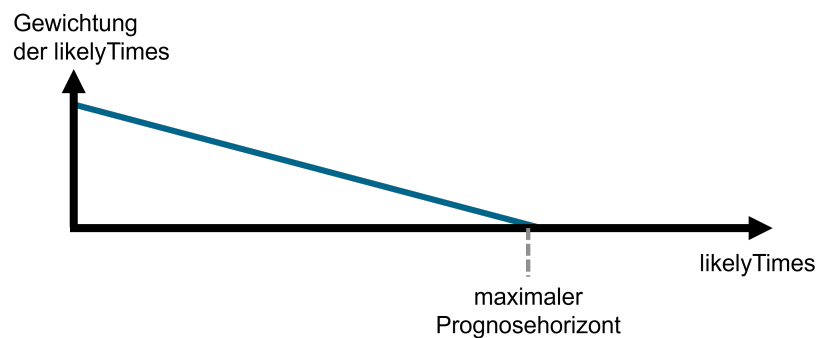


Abbildung 3.11: **Die Gewichtung der *likelyTime* nimmt bis zum maximalen Prognosehorizont kontinuierlich ab, um Prognosen der nahen Zukunft stärker zu gewichten.**

Abbildung 3.11 veranschaulicht, wie eine solche Gewichtung je *likelyTime* für eine Vielzahl von Datenpunkten zu interpretieren ist. Je höher dabei die relative *likelyTime* ist, desto weiter in der Zukunft liegt der zugehörige prognostizierte Umschaltzeitpunkt. Die Gewichtung erfolgt, indem die Bewertungen der *confidence* je zugehöriger relativer *likelyTime* aggregiert und mit der zugrunde gelegten Funktion (in diesem Fall einer linearen Funktion) gewichtet werden. Zusätzlich wird nach der Schwelle im Prognosehorizont gesucht, ab der die Prognosegenauigkeit nicht mehr die 95%-Genauigkeit nach C-Roads erfüllt. Alle Bewertungen der *confidence* oberhalb dieser Schwelle werden auf den Wert Null gesetzt, um die C-Roads-Konformität zu gewährleisten. Damit gilt: Je näher der prognostizierte Umschaltzeitpunkt rückt, desto stärker wird die zugehörige aggregierte *confidence* gewichtet. Eine niedrige *confidence* in der nahen Zukunft wird somit stärker

bestraft als eine niedrige *confidence* in ferner Zukunft. Außerdem führt eine Verletzung der geforderten Genauigkeit aus dem C-Roads Profil zu signifikanten Einbußen in der Qualitätsbewertung.

3.4 Gesamtwertung

Das Bewertungskonzept folgt in Anlehnung an die Qualitätsstufen des Verkehrsablaufs aus dem Handbuch für die Bemessung von Straßenverkehrsanlagen (HBS) dem Prinzip von Bewertungsstufen nach einem Qualitätsindex. Ein solcher Qualitätsindex ergibt sich aus der Zuordnung von Einzelbewertungen in Abhängigkeit vom jeweilig betrachteten Kriterium. Abbildung 3.12 veranschaulicht beispielhaft das Prinzip einer solchen Qualitätsbewertung in sechs Stufen. Um eine bestimmte Qualitätsstufe zu erreichen, muss je nach betrachtetem Kriterium ein Grenzwert über- bzw. unterschritten werden. In Abhängigkeit des zu bewertenden Kriteriums müssen sowohl die Bedeutung als auch die Grenzwertskalierung der Qualitätsstufen spezifisch und passend festgelegt werden. Eine Beurteilung nach diesem einheitlichen Stufenprinzip ermöglicht somit eine Vergleichbarkeit zwischen verschiedenen Steuerungen.

Die Gesamtbewertung einer Prognose kann sich dabei aus Teilaspekten zusammensetzen, indem jeder Teilindex zur Gesamtwertung beiträgt.

Qualitätsstufe	Bewertung	Wert
A	ausgezeichnet	...
B	sehr gut	...
C	gut	...
D	mittel	...
E	schlecht	...
F	sehr schlecht	...

Abbildung 3.12: Der Qualitätsindex umfasst 6 Stufen von A - *ausgezeichnet* bis F - *sehr schlecht*. Die Grenzwerte für die Erreichung einer Stufe hängen vom jeweiligen Kriterium ab und sind anwendungsspezifisch festzulegen.

Die Qualitätsstufe einer Prognose ergibt sich in diesem Fall aus den drei Teilaspekten, welche zuvor in den Abschnitten 3.1–3.3 beschrieben wurden:

1. Index der Steuerungsdynamik
2. Index der Nachrichtenintegrität und -plausibilität
3. Index der Prognosegüte

Der Index der Steuerungsdynamik wird dabei als eigenständiger Kennwert ausgewiesen und dient vor allem dazu, Prognosen nur zwischen Signalgruppen oder Knotenpunkten mit vergleichbarer Dynamik sinnvoll gegenüberzustellen. Die Teilindizes der Nachrichtenintegrität und -plausibilität sowie der Prognosegüte können dagegen mit definierten Gewichtungen zu einem Qualitätsindex zusammengeführt werden, der die eigentliche Prognosequalität beschreibt. Um verschiedene Anwendungsfälle abbilden zu können, besteht zusätzlich die Möglichkeit, bestimmte Teilaspekte stärker zu gewichten oder die Bewertung auf ausgewählte Teilaspekte zu begrenzen. Beispielsweise ist für eine Restgrünanzeige vielleicht nur ein Qualitätsindex für Freigaben interessant, sodass Sperrzustände nicht berücksichtigt werden müssen.

Das HBS nutzt für die Formalisierung der Berechnung der Qualitätsstufen definierte Formblätter [9, Part S4.7]. Diese tabellarischen Berechnungsvorlagen werden verwendet, um Kennzahlen und Qualitätsstufen einzelner Verkehrsströme zu berechnen und um diese Ergebnisse für jede Zufahrt zusammengefasst darstellen zu können. In Anlehnung an die Formblätter des HBS wird für die Gesamtbewertung der Prognosequalität eine ähnliche Zusammenführung für einen Knotenpunkt und entsprechender Prognose vorgesehen. Eine derartige tabellarische Formalisierung ist für alle drei Teilaspekte des Bewertungskonzepts umsetzbar. Kapitel 5 führt ein Anwendungsbeispiel für die Qualitätsbewertung der drei Teilaspekte in Anlehnung an die Formblätter des HBS auf.

4 Bewertungsmetrik

In diesem Kapitel werden konkrete Vorschläge zur Umsetzung des Bewertungskonzepts in einer Bewertungsmetrik präsentiert. Folgend werden dazu Berechnungswege aufgeführt und die Anwendung konkreter Formeln auf Datensätze diskutiert, welche bei der Entwicklung anhand von Beispieldaten getestet wurden. Das Kapitel orientiert sich dabei an den im Kapitel 3 eingeführten Konzepten und Beschreibungen.

4.1 Steuerungsdynamik

Der Dynamikindex einer Steuerung setzt sich aus den Dynamikindizes der Start- und Endzeitpunkte der Signallaufzeiten sowie dem Teilindex der Zeitintervalle der Signallaufzeiten zusammen. Um diesen zu bestimmen, werden zunächst die Abstände der Start- und Endzeitpunkte der Signallaufzeiten für jeden `eventState` untersucht. Dies erfolgt für jede Signalgruppe einzeln, wobei der Index einer Signalgruppe die gewichtete Summe aller bewerteten `eventStates` anzeigt. Gewichtet wird hierbei nach dem Anteil des `eventStates` am Beobachtungszeitraum. Zur Bewertung eines gesamten Knotenpunktes werden am Ende die Dynamikindizes der einzelnen Signalgruppen als Mittelwert zu einem Gesamtindex zusammengeführt. Folgend werden die Teilindizes der Gesamtdynamik einzeln beschrieben.

Teilindex der Startzeitpunkte: Der Teilindex der Startzeitpunkte beschreibt die Zufälligkeit der Abstände der Startzeitpunkte von Signallaufzeiten. Eindeutige Muster können hier auf das Vorhandensein von Rahmenplänen hinweisen, welche die Prognostizierbarkeit der Steuerung erhöhen können.

Für den Teilindex der Startzeitpunkte der Signallaufzeiten werden alle Abstände zwischen aufeinanderfolgenden Signallaufzeiten innerhalb des Beobachtungszeitraums bestimmt.

Damit lässt sich für jeden dieser Abstände eine Wahrscheinlichkeit bestimmen, welche zur Entropie H_e zusammengeführt werden:

$$H_e(X) = - \sum_i p_{i,e} \cdot \log_2 p_{i,e} \quad (4.1)$$

Hier beschreibt H_e die Shannon'sche Informationsentropie eines `eventStates` e , X die Gesamtheit aller Abstände der Startzeitpunkte $t_{i,e}$ und $p_{i,e}$ deklariert die Wahrscheinlichkeit für einen Abstand $\Delta t_{i,e}$ zwischen zwei Startzeitpunkten. Für den Teilindex D_s einer Signalgruppe s werden anschließend die Entropien der einzelnen `eventStates` gewichtet nach ihrer Gesamtlaufzeit an der Signalgruppe aufaddiert:

$$D_s(X) = \sum_e a_{e,s} \cdot H_{e,s}(X) \quad (4.2)$$

Hier beschreibt $a_{e,s}$ den Anteil des `eventStates` e an der Gesamtlaufzeit aller betrachteten `eventStates` aus Signalgruppe s .

Ein wichtiger Hinweis: Da der Indexwert D_s nach oben offen ist, wird hier als Maximum für die Normierung der Wert $\beta = 4$ angenommen. Die Annahme besteht hier darin, dass der Maximalwert so hoch gewählt wird, dass sich die Prognostizierbarkeit der Steuerung bei noch höheren Werten nicht mehr signifikant ändert und diese Steuerungen damit vergleichbar bleiben. Somit ergibt sich der normierte Index $D_{s,n}$ folgendermaßen:

$$D_{s,n} = \frac{1}{\beta} \cdot \min(D_s, \beta) \quad (4.3)$$

Teilindex der Endzeitpunkte: Dieser Teilindex beschreibt die Zufälligkeit der Abstände der Endzeitpunkte von Signallaufzeiten. Ähnlich wie beim Teilindex der Startzeitpunkte können auch hier eindeutige Muster auf das Vorhandensein von Rahmenplänen hinweisen, welche die Prognostizierbarkeit der Steuerung erhöhen können.

Der Teilindex für die Abstände der Signallaufzeitenden der verschiedenen Signalgruppen und `eventStates` wird analog zum Teilindex der Startzeitpunkte (siehe Gleichung 4.3) berechnet, mit dem Unterschied, dass $\Delta t_{i,e}$ in diesem Fall den Abstand zwischen den Endzeitpunkten beschreibt.

Teilindex der Zeitintervalle: Der Teilindex der Zeitintervalle beschreibt die Zufälligkeit der beobachteten Signallaufzeiten. Viele verschiedene beobachtete Signallaufzeiten weisen auf eine erschwerte Prognostizierbarkeit hin, während eine Beschränkung auf nur einige wenige Signallaufzeiten die Prognose erleichtern sollte.

Für den Teilindex der Zeitintervalle der Signallaufzeiten werden die Intervallabstände der beobachteten Signallaufzeiten der `eventStates` betrachtet. Dazu wird die ursprüngliche Entropieformel um einen zusätzlichen Gewichtungsfaktor w_e erweitert:

$$H_e(X) = - \sum_i w_e(i) \cdot p_{i,e} \cdot \log_2 p_{i,e} \quad (4.4)$$

Der Gewichtungsfaktor $w_e(i)$ ist dabei abhängig von der Größe des Intervallabstands zur nächstkleineren und nächstgrößeren beobachteten Signallaufzeit. Auf Basis der ersten Datenanalyse hat sich folgender Gewichtungsfaktor in der Anwendung bewährt:

$$w_e(i) = \frac{1}{2} \cdot \log_{10}(\Delta^- x_{i,e} + \alpha) + \frac{1}{2} \cdot \log_{10}(\Delta^+ x_{i,e} + \alpha) \quad (4.5)$$

wobei $\Delta^- x_{i,e} = x_{i,e} - x_{i-1,e}$ und $\Delta^+ x_{i,e} = x_{i+1,e} - x_{i,e}$ die Rückwärts- bzw. Vorwärtsdifferenz zur nächstkleineren oder -größeren beobachteten Signallaufzeit des `eventStates` e darstellen. Die Addition des Werts $\alpha = 9$ innerhalb der Logarithmusfunktion bewirkt, dass bei einem minimalen Zeitintervall von 1s ein Gewichtungsfaktor von $w_e = 1$ zustande kommt. Bei größeren Zeitlücken steigt w_e entsprechend an.

Die restliche Akkumulation der Zeitintervalle der Signallaufzeiten erfolgt analog zu den vorher beschriebenen Teilindizes.

Dynamikindex der Steuerung: Zunächst ergibt sich der Index der Dynamik $I_{D,s}$ der Signalgruppe aus der Summe der normierten Teilindizes für Startzeitpunkte $D_{\text{Start},s,n}$, Endzeitpunkte $D_{\text{Ende},s,n}$ und Zeitintervalle $D_{\text{Intervall},s,n}$ mit den zugehörigen Gewichten w :

$$I_{D,s} = \frac{w_{\text{Start},s} \cdot I_{\text{Start},s,n} + w_{\text{Ende},s} \cdot I_{\text{Ende},s,n} + w_{\text{Intervall},s} \cdot I_{\text{Intervall},s,n}}{w_{\text{Start},s} + w_{\text{Ende},s} + w_{\text{Intervall},s}} \quad (4.6)$$

Zusammenfassend ergibt sich der Dynamikindex einer Steuerung I_D für einen Knotenpunkt dann als Summe der Teilindizes für die Signalgruppen normiert auf die Anzahl der

Signalgruppen N_s . Er dient als Indikator der Vergleichbarkeit der Prognose verschiedener Steuerungen:

$$I_D = \frac{1}{N_s} \cdot \sum_s I_{D,s} \quad (4.7)$$

4.2 Nachrichtenintegrität und -plausibilität

Zur Prüfung der Nachrichten hinsichtlich ihrer Integrität und Plausibilität werden diese zunächst unabhängig von den tatsächlichen Umschaltzeitpunkten ausgewertet. Es wird untersucht, ob die Nachrichten in sich konsistent sind und die Vorgaben des C-Roads Profils sowie grundlegenden Rahmenbedingungen der Verkehrstechnik erfüllen.

Verfügbarkeit: Unter der Verfügbarkeit der Prognose wird der Anteil der SPATEM-Nachrichten verstanden, die überhaupt eine Prognose bereit stellen. Eine hohe Verfügbarkeit ist dabei die Grundvoraussetzung dafür, dass Prognosen in Anwendungen zuverlässig genutzt werden können.

Formal ergibt sich die Verfügbarkeit aus dem Prozentsatz der SPATEM-Nachrichten, in denen die notwendigen Datenfelder `minEndTime`, `maxEndTime`, `likelyTime` und `confidence` gemäß der Profilvergaben befüllt sind und in denen das Feld `confidence` einer Zeitintervallklasse einen Wert > 0 aufweist. Zeitintervallklassen mit einem Wert von Null werden nach dem C-Roads Profil als Nichtverfügbarkeit der Prognose gewertet.

Plausibilität: Für die Größen `minEndTime`, `likelyTime` und `maxEndTime` gelten bestimmte Rahmenbedingungen. Die `minEndTime` darf sich in aufeinanderfolgenden Nachrichten nur weiter in die Zukunft verschieben, die `maxEndTime` darf sich nicht weiter in die Zukunft verschieben, und die `likelyTime` muss stets zwischen `minEndTime` und `maxEndTime` liegen. Entsprechend werden folgende Kriterien getrennt ausgewertet:

- Die `minEndTime` der aktuellen Nachricht ist kleiner oder gleich der `minEndTime` der Folgenachricht desselben `eventStates` e .

$$\text{minEndTime}_{t,e} \leq \text{minEndTime}_{t+1,e} \quad (4.8)$$

- Die `maxEndTime` der aktuellen Nachricht ist größer oder gleich der `maxEndTime` der Folgenachricht desselben `eventStates` e .

$$\text{maxEndTime}_{t,e} \geq \text{maxEndTime}_{t+1,e} \quad (4.9)$$

- Die `likelyTime` einer Prognose liegt zwischen `minEndTime` und `maxEndTime` dieser Prognose.

$$\text{minEndTime}_t \leq \text{likelyTime}_t \leq \text{maxEndTime}_t \quad (4.10)$$

Auf dieser Grundlage lässt sich für jedes Kriterium der Prozentsatz der SPATEM-Nachrichten mit verfügbarer Prognose bestimmen, die das jeweilige Kriterium erfüllen.

Prüfung auf statische `eventStates`: Nicht-dynamische `eventStates` haben eine feste Signallaufzeit und damit einen fest definierten, unveränderlichen Umschaltzeitpunkt. Die `likelyTime` gibt diesen Umschaltzeitpunkt an. Die Unveränderlichkeit des Umschaltzeitpunkts wird hier dadurch geprüft, dass paarweise aufeinanderfolgende Nachrichten nach dem folgenden Kriterium auf Übereinstimmung untersucht werden:

$$\text{likelyTime}_{t,e} = \text{likelyTime}_{t+1,e} \quad (4.11)$$

Ist dieses Kriterium erfüllt, gelten die entsprechenden Nachrichten als profilkonform. Aus dem Anteil der profilkonformen Nachrichten ergibt sich ein Prozentsatz, der in den Teilindex eingeht. Konkret werden die `eventStates` `gesichertes Räumen (Gelb)`, `bedingt verträgliches Räumen (Gelb)` und `Vorbereitung (Rot-Gelb)` betrachtet.

Ermittlung der Teilindizes: Der Teilindex für Nachrichtenintegrität und -plausibilität F_s einer Signalgruppe s setzt sich entsprechend aus der gewichteten Summe der Teilindizes zusammen:

$$F_s = \frac{1}{\sum_x w_x} \cdot \sum_x w_x \cdot F_{x,s} \quad (4.12)$$

Die verschiedenen Teilindizes $F_{x,s}$ einer Signalgruppe s mit dem Laufindex $x = 1, \dots, 7$ lauten wie folgt:

- „1“ Prozentsatz der Verfügbarkeit,
- „2“ Prozentsatz profilkonform aufeinanderfolgender `minEndTimes`,
- „3“ Prozentsatz profilkonform aufeinanderfolgender `maxEndTimes`,
- „4“ Prozentsatz profilkonformer `minEndTime-likelyTime-maxEndTime`-Beziehungen,
- „5“ Prozentsatz profilkonformer Vorhersagen für `gesichertes Räumen` (Gelb),
- „6“ Prozentsatz profilkonformer Vorhersagen für `bedingt verträgliches Räumen` (Gelb),
- „7“ Prozentsatz profilkonformer Vorhersagen für `Vorbereitung` (Rot-Gelb).

Gesamtindex: Der Gesamtindex eines Knotenpunktes für die Nachrichtenintegrität und -plausibilität F wird abschließend aus der Summe über die Teilindizes F_s aller Signalgruppen gebildet und auf die Anzahl der Signalgruppen N_s normiert:

$$F = \frac{1}{N_s} \cdot \sum_s F_s \quad (4.13)$$

4.3 Prognosegüte

Um die Güte einer Prognose bewerten zu können, wird zunächst das Zusammenspiel aus Prognosehorizont, Prognosegenauigkeit und der `confidence` für den Teilindex der Prognosegüte aller Signalgruppen untersucht. Dazu wird die maximale, C-Roads-konforme `likelyTime` bestimmt, die zugehörige `confidence` bewertet und diese Bewertung in Abhängigkeit vom Prognosehorizont gewichtet. Abschließend werden die Teilindizes für `minEndTime` und `maxEndTime` bestimmt und dann für einen Knotenpunkt ein Gesamtindex aus der normierten Summe aller Signalgruppen ermittelt.

Bestimmung der maximalen, profilkonformen `likelyTime`: Für die Korrektheit einer einzelnen Prognose in einer SPATEM-Nachricht muss folgende Bedingung erfüllt sein:

$$L_{p,T} - c_{p,T,C-Roads} \leq T \leq L_{p,T} + c_{p,T,C-Roads} \quad (4.14)$$

Dabei bezeichnet T den konkreten Umschaltzeitpunkt einer Signalgruppe, $L_{p,T}$ die von einer konkreten Prognose p ausgegebene `likelyTime` für diesen Zeitpunkt und $c_{p,T,C-Roads}$ die zugehörige `confidence` gemäß dem C-Roads Profil.

Diese Bedingung wird für jede Prognose im Betrachtungszeitraum geprüft, sodass für jeden `eventState` e und für jede zugehörige relative `likelyTime` L die durchschnittliche Genauigkeit $G_{L,e}$ der entsprechenden Prognosen bestimmt werden kann. Da das C-Roads Profil vorschreibt, dass insgesamt eine Prognosegenauigkeit von 95% erreicht werden soll, wird diese Bedingung für jede Genauigkeit $G_{L,e}$ separat geprüft. Infolgedessen kann für jeden `eventState` e die maximale relative `likelyTime` $L_{\max,e}$ bestimmt werden, für die folgende Bedingung gilt:

$$\forall i \in \{1, \dots, L_{\max,e}\} : G_{i,e} \geq 0.95 \quad (4.15)$$

Diese Bedingung besagt, dass für alle relativen `likelyTimes` mit Index i im Bereich von 1 bis $L_{\max,e}$ die durchschnittliche Genauigkeit $G_{L,e}$ mindestens 95% beträgt. $L_{\max,e}$ beschreibt dementsprechend den maximalen Prognosehorizont, bis zu dem die Prognosen eines `eventStates` e insgesamt C-Roads-konform sind.

Bewertung der `confidence`: Zu jeder Genauigkeit $G_{L,e}$ kann zusätzlich eine entsprechende, durchschnittliche `confidence`-Zeitintervallklasse $c_{L,e}$ bestimmt werden. Das zugehörige Zeitintervall $c_{L,e,C-Roads}$ beschreibt die Breite des durchschnittlichen Prognosezeitfensters für die jeweilige `likelyTime`. Im Idealfall hat $c_{L,e}$ den Wert 15, was gemäß Tabelle 2.1 im C-Roads Profil einem Prognosezeitfenster der Breite Null entspricht. Daraus wird die Bewertung $C_{L,e}$ definiert als:

$$C_{L,e} = \frac{1}{\log_{\gamma}(c_{L,e,C-Roads} + \gamma)} \quad (4.16)$$

Der Parameter γ wurde auf Basis von Erfahrungen aus den Beispieldatensätzen auf $\gamma = 4$ festgelegt. Der Nenner $\log_{\gamma}(c_{L,e,C-Roads} + \gamma)$ steigt logarithmisch mit der Breite des Prognosezeitfensters, das heißt $C_{L,e}$ wird dementsprechend kleiner. Bei maximaler `confidence` ist $C_{L,e} = 1$, im schlechtesten Fall beträgt $C_{L,e} \approx 0,48$ (näherungsweise $1/2$).

Einen weiteren Einfluss auf die Bewertung hat die maximale, profilkonforme *likelyTime* $L_{\max,e}$. Abbildung 4.1 veranschaulicht diesen Einfluss auf die *confidence*-Bewertung $C_{L,e}$ anhand zweier Beispiele unter Berücksichtigung von $L_{\max,e}$. In Abbildung 4.1 A gehen alle Wertungen der durchschnittlichen *confidence*-Zeitintervallklasse $c_{L,e}$ ein. In Abbildung 4.1 B wird demgegenüber der Fall gezeigt, dass die Bewertungen $C_{L,e}$ für alle *likelyTimes* annulliert werden, deren zugehörige *likelyTime* größer als die maximale, profilkonforme *likelyTime* ist. „Annulliert“ bedeutet in diesem Zusammenhang, dass für diese Werte die minimale Wertung Null angenommen wird. Damit gehen nur Bewertungen der *confidence* in die Prognosegüte ein, deren relative *likelyTimes* auch den Anforderungen des C-Roads Profils genügen.

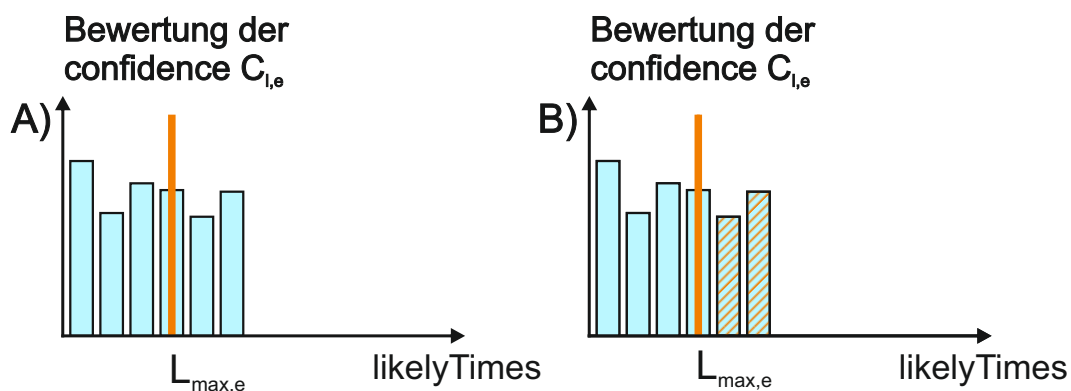


Abbildung 4.1: Einfluss der maximalen, profilkonformen *likelyTime* $L_{\max,e}$ auf die Bewertung der *confidence*; A: Zeigt alle Wertungen der *confidence*. B: Es werden alle Wertungen, die zu einer größeren *likelyTime* als $L_{\max,e}$ gehören, annulliert.

Gewichtung der *confidence*-Wertung: Es wird erwartet, dass Prognosen für die nahe Zukunft eine höhere *confidence* aufweisen als Prognosen für weiter in der Zukunft liegende Zeitpunkte. Dazu wird eine Gewichtungsfunktion eingeführt, die ihr Maximum bei einer relativen *likelyTime* von 1s annimmt und linear bis zum maximalen Prognosehorizont $T_{P,\max}$ auf den Wert Null abfällt. Abbildung 4.2 illustriert die Umsetzung einer solchen Gewichtung unter Berücksichtigung vom Verhältnis $T_{P,\max}$ zu $T_{S,\max,e}$. Der maximale Prognosehorizont $T_{P,\max}$ beschreibt die Zeitgrenze, bis zu der Prognosen bewertet werden; $T_{S,\max,e}$ ist die maximal beobachtete Signallaufzeit. Zur Normierung wird die Gewichtungsfunktion so skaliert, dass die Summe aller Teilgewichte gleich 1 ergibt; entsprechend besitzt die in Abbildung 4.2 A dargestellte Normierungsfläche einen Flächeninhalt von 1. Sollte jedoch $T_{P,\max}$ größer sein als die maximale beobachtete Signallaufzeit $T_{S,\max,e}$, wird nur auf die Summe der Gewichte bis zu dieser maximal beobachteten Signallaufzeit

$T_{S,max,e}$ normiert (siehe Abbildung 4.2 B). In diesem Fall gehen die Gewichte für höhere `likelyTimes` nicht in die Normierung ein, sodass niedrigere `likelyTimes` entsprechend stärker zur Gesamtgewichtung beitragen.

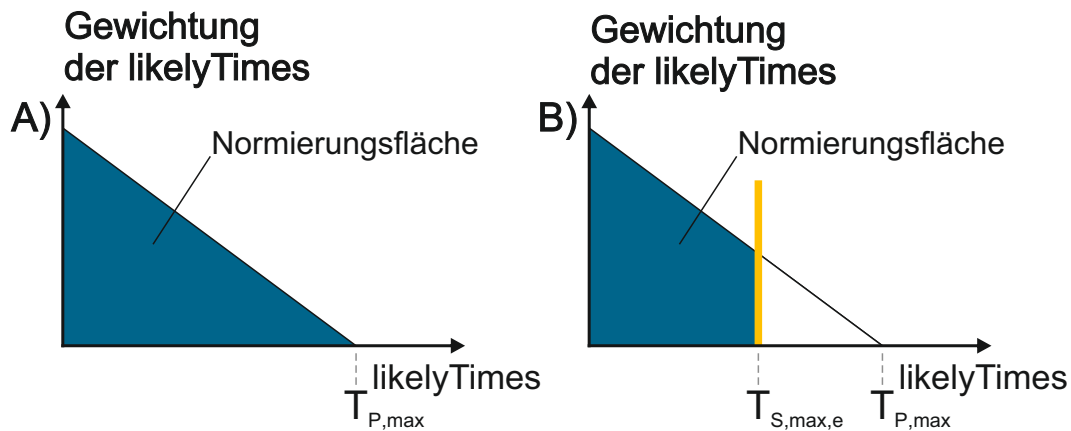


Abbildung 4.2: Gewichtung der *likelyTimes*, die zur *confidence*-Wertung beitragen. **A:** Alle Gewichte bis zum festgelegten maximalen Prognosehorizont $T_{P,max}$ werden gewertet. **B:** Es fallen alle Gewichte, die über der maximal beobachteten Signallaufzeit $T_{S,max,e}$ liegen, aus der Normierung heraus.

Teilindex der Prognose: Der Teilindex der Prognose beschreibt die Bewertung für das Zusammenspiel aus Prognosehorizont, Prognosegenauigkeit und *confidence* ab. Die Bewertung der *confidence* dient dabei als Grundlage des Index. Die Prognosegenauigkeit fließt in die maximale, C-Roads-konforme `likelyTime` $L_{max,e}$ ein. Der Prognosehorizont wird ebenfalls durch $L_{max,e}$ sowie die Gewichtung nach Prognosehorizont abgebildet, welche im Folgenden erläutert wird.

Für den Teilindex der Prognose $I_{P,e,s}$ eines `eventStates` e und einer Signalgruppe s wird die Bewertung der *confidence* mit der Gewichtung zusammengeführt. Abbildung 4.3 veranschaulicht diese Vorgehensweise. Die Bewertungen der *confidence* $C_{L,e,s}$ werden mit den zugehörigen Gewichtungen $w_{L,e,s}$ der relativen `likelyTimes` L multipliziert und aufsummiert:

$$I_{P,e,s} = \sum_L w_{L,e,s} \cdot C_{L,e,s} \quad (4.17)$$

Dabei ist jedoch Folgendes zu beachten: Alle Bewertungen $C_{L,e}$ mit `likelyTimes` größer der maximalen, profilkonformen `likelyTime` $L_{\max,e}$ werden annulliert und mit Null angenommen (siehe Schraffur in Abbildung 4.3). Die Normierung erfolgt in dem Fall bis zum kleineren Wert aus der maximal beobachteten Signallaufzeit $T_{S,\max,e}$ und dem maximalen Prognosehorizont $T_{P,\max}$, unabhängig davon, wie groß $L_{\max,e}$ ist. Diese Art der Bewertung sieht vor, dass der Fall bestraft wird, dass die Genauigkeit der Prognose nicht ausreichend im vom C-Roads Profil geforderten Wertebereich bis 95% liegt.

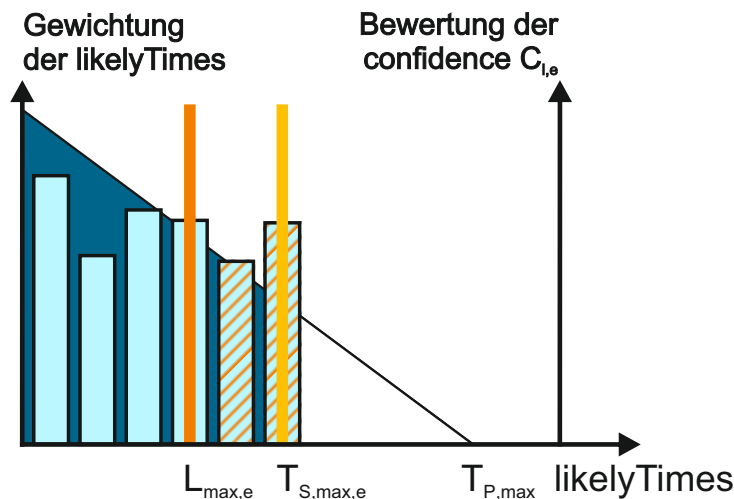


Abbildung 4.3: Veranschaulichung des Zusammenspiels von Normierung, *confidence*-Bewertung und der maximalen, profilkonformen *likelyTime* $L_{\max,e}$; der Normierungsbereich ist reduziert, da $T_{P,\max} < T_{S,\max,e}$, zusätzlich werden aber zwei *confidence*-Wertungen annulliert, weil die maximale, profilkonforme *likelyTime* $L_{\max,e}$ zu niedrig ist.

Die Zusammenführung zum Teilindex der Prognose $I_{P,s}$ einer Signalgruppe s über alle `eventStates` e erfolgt analog zur Teilindexbildung der Steuerungsdynamik mit einer Gewichtung $a_{e,s}$, die den Anteil des `eventStates` e an der Gesamtlaufzeit der Signalgruppe widerspiegelt:

$$I_{P,s} = \sum_e a_{e,s} \cdot I_{P,e,s} \quad (4.18)$$

Die Maximalwertung für den Teilindex der Prognose für einen `eventState` e wird erreicht, wenn die Prognose bis zum maximalen Prognosehorizont beziehungsweise der maximal beobachteten Signallaufzeit des `eventStates` C-Roads-konform ist und dabei stets die maximale *confidence* erreicht wird. Eine niedrigere *confidence* hat insbesondere in der

nahen Zukunft starke negative Auswirkungen auf die Prognosegüte, ebenso wie die Nicht-Konformität zum C-Roads Profil ab einer niedrigen relativen `likelyTime`. Ungenauigkeiten in der fernerer Zukunft haben im Umkehrschluss deutlich geringere negative Auswirkungen auf die Prognosegüte.

Teilindizes für `minEndTime` und `maxEndTime`: Die Teilindizes $I_{\min E}$ und $I_{\max E}$ erfassen jeweils den Prozentsatz der Prognosen, bei denen der tatsächliche Umschaltzeitpunkt T konsistent zu den prognostizierten `minEndTime` und `maxEndTime` liegt. Für die prognostizierte `minEndTime` $t_{\min,p,T}$ einer Prognose p für einen konkreten Umschaltzeitpunkt T muss demnach gelten:

$$t_{\min,p,T} \leq T \quad (4.19)$$

Entsprechend muss für die prognostizierte `maxEndTime` $t_{\max,p,T}$ einer Prognose p für einen konkreten Umschaltzeitpunkt T gelten:

$$t_{\max,p,T} \geq T \quad (4.20)$$

Die Teilindizes $I_{\min E}$ und $I_{\max E}$ entsprechen somit dem Prozentsatz der Prognosen, für den die jeweilige Bedingung erfüllt ist.

Gesamtindex der Prognosegüte: Zunächst ergibt sich die Prognosegüte Q_s einer Signalgruppe s aus dem Teilindex der Prognose $I_{P,s}$ und der Teilindizes $I_{\min E,s}$ und $I_{\max E,s}$ für `minEndTime` und `maxEndTime` mit den zugehörigen Gewichten w zu:

$$Q_s = \frac{w_{P,s} \cdot I_{P,s} + w_{\min E,s} \cdot I_{\min E,s} + w_{\max E,s} \cdot I_{\max E,s}}{w_{P,s} + w_{\min E,s} + w_{\max E,s}} \quad (4.21)$$

Die Gesamtprognosegüte I_P für einen Knotenpunkt ergibt sich abschließend als auf die Anzahl der Signalgruppen N_s normierte Summe der Prognosegüten Q_s der Signalgruppen s :

$$I_P = \frac{1}{N_s} \sum_s Q_s \quad (4.22)$$

Die in den Abschnitten 4.1–4.3 errechneten Werte der Teilindizes können abschließend tabellarisch gemäß des Bewertungskonzepts zusammengeführt und als Qualitätsstufen dargestellt werden.

4.4 Übersicht zur Ermittlung der Bewertungsmetrik

Abschließend soll noch einmal eine Gesamtübersicht über das in Kapitel 3 beschriebene Bewertungskonzept und die in Kapitel 4 ausgeführte Bewertungsmetrik gegeben werden.

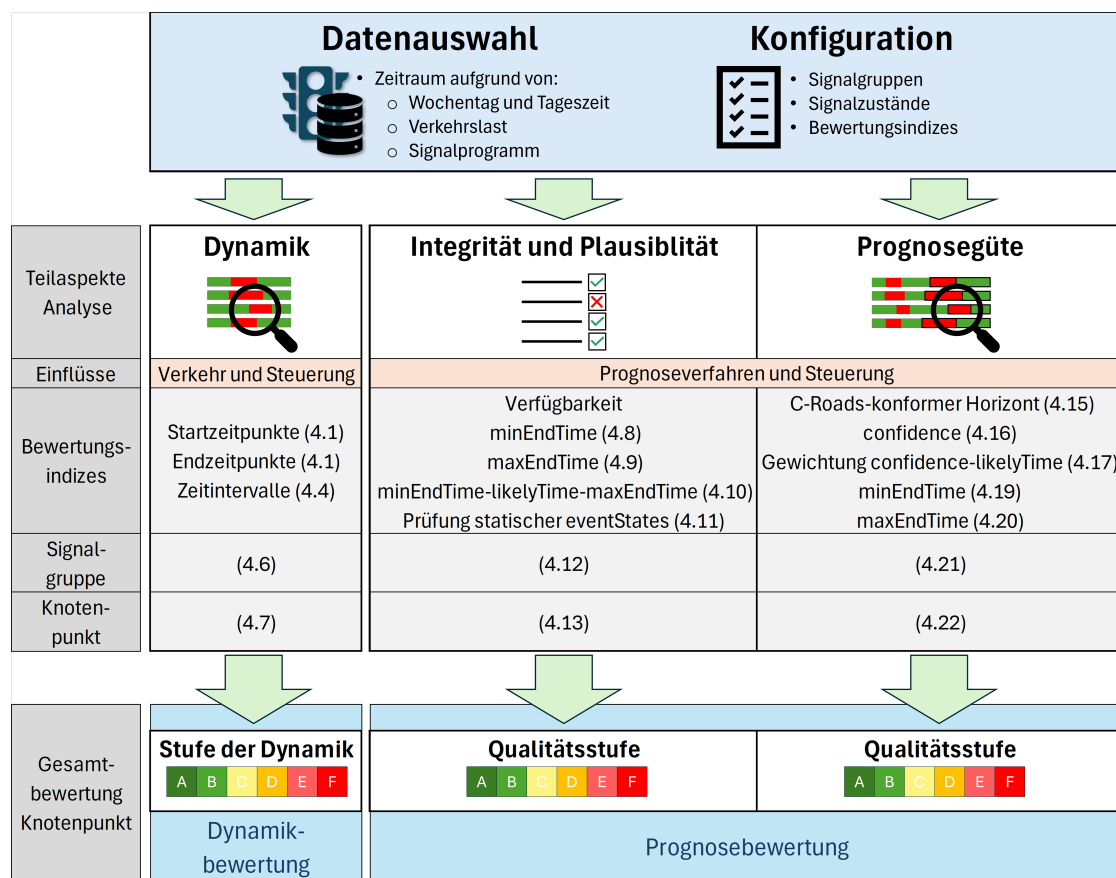


Abbildung 4.4: Übersicht des Bewertungskonzepts zur Ermittlung einer Gesamtbewertung für einen Knotenpunkt. Die Klammerangaben (4.1 – 4.22) referenzieren auf die entsprechenden Gleichungen, welche in Kapitel 4 aufgeführt sind.

Abbildung 4.4 fasst dazu die grundsätzlichen Voraussetzungen, Teilaspekte und Arbeitsschritte grafisch zusammen und ordnet detailliert die einzelnen Bewertungsindizes den jeweiligen Bewertungskriterien zu, welche schließlich zu einer abschließenden Gesamtbewertung führen. Ausgehend von der dargestellten Struktur beginnt die Ermittlung der Bewertungsmetrik mit der Auswahl einer geeigneten **Datengrundlage** und **Konfiguration**. Zunächst ist festzulegen, welche Zeiträume, Verkehrsbelastungen und Signalprogramme in die Analyse einfließen und welche Signalgruppen, Signalzustände und Bewertungsindizes für den jeweiligen Anwendungsfall zu berücksichtigen sind. Auf dieser Grundlage wird die Bewertungsmetrik so konfiguriert, dass sie sowohl den Analysezielen als auch der gewünschten Vergleichbarkeit mit anderen Knotenpunkten gerecht wird.

Im Anschluss daran wird die **Dynamik** der betrachteten Steuerung ermittelt. Hierzu werden die Start- und Endzeitpunkte sowie die Zeitintervalle der Signallaufzeiten analysiert und zu einem Index der Steuerungsdynamik zusammengeführt, aus dem sich eine Dynamikstufe für den Knotenpunkt ableiten lässt. Dieser Kennwert beschreibt, wie stark sich die Steuerung im Zeitverlauf verändert und wie herausfordernd eine verlässliche Prognose grundsätzlich ist.

Unabhängig davon werden die **Integrität und Plausibilität** der vorliegenden Prognosedaten geprüft. Im Fokus stehen dabei insbesondere die Verfügbarkeit und Profilkonformität der Prognosen sowie die zeitliche Konsistenz von `minEndTime`, `likelyTime` und `maxEndTime`. Die Ergebnisse der Prüfungen werden in einem Teilindex zusammengefasst. Zusätzlich wird die **Prognosegüte** bewertet, indem die profilkonformen `likelyTimes` und die zugehörige `confidence` mit den tatsächlich beobachteten Umschaltzeitpunkten abgeglichen und über den Prognosehorizont gewichtet werden. Aus dieser Bewertung entsteht ein Teilindex der Prognosegüte, der die inhaltliche Qualität der Prognose beschreibt.

Die Teilindizes werden abschließend zu einer Dynamikbewertung und einer Prognosebewertung zusammengeführt, aus denen sich in Form von **Qualitätsstufen** eine kompakte Gesamtbewertung ableiten lässt.

5 Anwendung

Die Qualitätsbewertungsmetrik ermöglicht Schaltzeitprognoseverfahren und -daten aus verschiedenen Blickwinkeln und für unterschiedliche Anwendungsfälle zu analysieren. So ist beispielsweise eine vergleichende Bewertung zwischen verfügbaren Prognoseverfahren für den potentiellen Einsatz an einem Knotenpunkt möglich. Ebenso kann die Prognosequalität für ein und dasselbe Verfahren über mehrere Knotenpunkte hinweg analysiert werden. Auch eine Eignungsprüfung von ausgesendeten Prognosedaten für die geplante Umsetzung von C-ITS-Services an einem Knotenpunkt ist denkbar. Die Anwendungsmöglichkeiten sind vielfältig, doch für eine verlässliche Prognosebewertung ist dabei stets eine geeignete Datengrundlage von entscheidender Bedeutung. Hinsichtlich des zu wählenden Datensatzes sollte dieser eine typische Verkehrslastsituation abbilden und einen ausreichend zeitlichen Umfang haben. Des Weiteren muss die Bewertungsmetrik entsprechend des Analyseziels konfiguriert werden. Nur so kann eine passgenaue und aussagekräftige Qualitätsbewertung erfolgen.

Nachfolgend soll anhand eines fiktiven Anwendungsbeispiels die Nutzung der erarbeiteten Qualitätsbewertungsmetrik demonstriert werden. Ergänzend werden allgemeine Hinweise gegeben, die für die Anwendung der Bewertungsmetrik hilfreich sein können. Dazu zählen z. B. Hinweise für die Wahl des Analysezeitraums, für die Konfiguration der Metrik sowie für die Festlegung einer Gewichtung und von Grenzwerten für die Qualitätsstufen.

5.1 Anwendungsbeispiel

Das Anwendungsbeispiel zeigt und beschreibt beispielhaft, wie die Konfiguration und die daraus resultierende Qualitätsbewertung der Prognose vorgenommen wird. Ziel im Anwendungsbeispiel ist es, mit der Qualitätsbewertungsmetrik für einen fiktiven, verkehrsabhängig gesteuerten Knotenpunkt gemäß des Lageplans in Abbildung 5.1 herauszufinden, ob für den motorisierten Verkehr in Ost-West-Richtung der Einsatz einer GLOSA-Anwendung verlässlich umsetzbar ist. Dabei wird angenommen, dass eine geeignete Datengrundlage verfügbar ist, d. h. SPATEM-Nachrichten für einen ausreichend langen und bezüglich der

Verkehrslast charakteristischen Zeitraum vorliegen. Es stehen dabei weniger die konkret ermittelten Qualitätsstufen im Fokus, sondern viel mehr der Ablauf und die Handhabung der Qualitätsbewertungsmetrik. Der fiktive Knotenpunkt wurde vor diesem Hintergrund bewusst minimalisiert ausgeführt und nicht alle Aspekte umfassend bewertet.

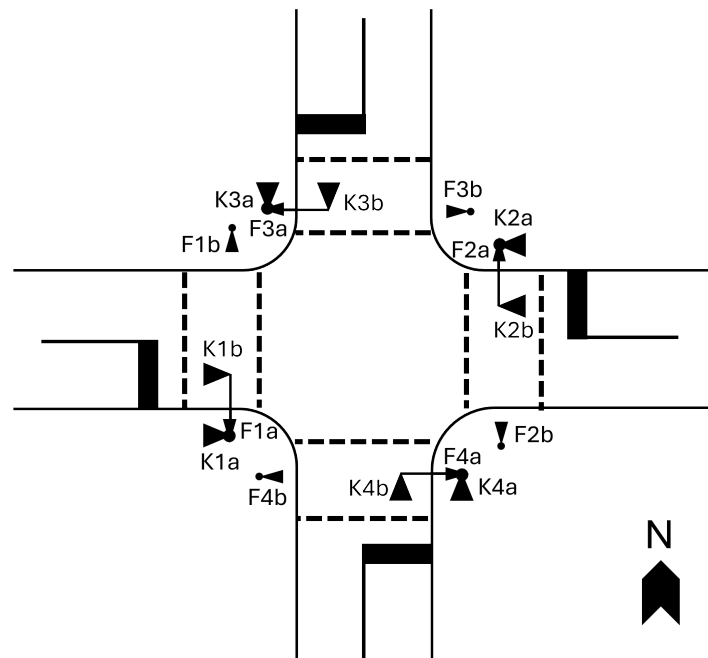


Abbildung 5.1: Lageplan für das Anwendungsbeispiel: Prüfung des möglichen Einsatzes einer GLOSA-Anwendung an einem symmetrischen Knotenpunkt.

5.1.1 Beispielkonfiguration

Bevor eine Bewertung durchgeführt werden kann, muss die Metrik zunächst konfiguriert werden. Das heißt, es muss festgehalten werden, welche Merkmale und Eigenschaften der Knotenpunkt aufweist und welche Kenngrößen innerhalb der einzelnen Teilindizes bei der Bewertung zur Anwendung kommen sollen. Die Abbildung 5.2 fasst die entsprechende Konfiguration für das Beispiel tabellarisch zusammen. Der Wert "1" symbolisiert dabei die Auswahl eines Merkmals oder einer Eigenschaft, der Wert "0" entsprechend die Abwahl. Die wesentlichen Konfigurationsmerkmale sind dabei folgende:

- Es sollen die Kfz-Signalgruppen K1 und K2 für den potenziellen Einsatz eines GLOSA-Services analysiert werden. Diese werden entsprechend auf den Wert "1" gesetzt; die restlichen Signalgruppen werden nicht betrachtet und entsprechend nicht ausgewählt, da für sie keine GLOSA-Anwendung implementiert werden soll. Dies ist in Abbildung 5.2 im Abschnitt *Signalgruppe* vollständig ausgeführt.
- Für die GLOSA-Anwendung im Anwendungsbeispiel ist der Zustand *Bedingtverträgliche Freigabe (Grün)* nicht relevant, da die GLOSA-Anwendung nur für die Unterstützung der West-Ost-Überfahrt eingesetzt werden soll. Bei Freigabe sind hier die geradeaus fahrenden GLOSA-Nutzer vorfahrtsberechtigt. Daher wird dieser Zustand in der Konfiguration im Abschnitt *Zustand* in Abbildung 5.2 abgewählt.
- Für die Bewertung der drei Teilindizes Dynamik, Nachrichtenplausibilität und -integrität sowie Prognosegüte wird eine hypothetische Standardkonfiguration verwendet, d. h. es werden alle Merkmale ausgewählt, und der maximal zu analysierende Prognosehorizont auf 15 s festgelegt. Dies ist in Abbildung 5.2 in den entsprechenden Abschnitten spezifiziert.

Signalgruppe	Auswahl	Teilindex Dynamik	Auswahl
K1 (Kfz)	1	Länge	1
K2 (Kfz)	1	Start zu Start	1
K3 (Kfz)	0	Ende zu Ende	1
K4 (Kfz)	0		
F1 (Fußgänger)	0	Teilindex Plausibilität/integrität	Auswahl
F2 (Fußgänger)	0	Verfügbarkeit	1
F3 (Fußgänger)	0	minEndTimes	1
F4 (Fußgänger)	0	maxEndTimes	1
		minEndTime-likelyTime-maxEndTime	1
Zustand	Auswahl	gesichertes Räumen (Gelb)	1
gesicherte Freigabe (Grün)	1	bedingt verträgliches Räumen (Gelb)	1
bedingtverträgliche Freigabe (Grün)	0	Vorbereitung (Rot-Gelb)	1
Gesperrt (Rot)	1		
		Teilindex Prognosegüte	Auswahl
		confidence likelyTime	1
		minEndTimes	1
		maxEndTimes	1
		maximaler Prognosehorizont [s]	15

Abbildung 5.2: **Beispielkonfiguration für den fiktiven Knotenpunkt des Anwendungsbeispiels.**

5.1.2 Beispielbewertung der Dynamik

Abbildung 5.3 zeigt eine Beispielbewertung der Steuerungsdynamik in Anlehnung an das Form- und Berechnungsblatt der Qualitätsstufen des Verkehrsablaufs des HBS. Die einzelnen Spalten der tabellarischen Darstellung sind wie folgt zu verstehen.

Signalgruppe	Auswahl	Zustand [event state]	Auswahl	Anteil Beobachtungszeitraum	Teilindex	Auswahl	Teilindex Wert	Zustand Wert	Zustand Stufe	Signalgruppe Wert [$I_{D,s}$]	Signalgruppe Stufe	Gesamt Wert [I_D]	Gesamt Stufe					
0	1	2	3	4	5	6	7	8	9	10	11	12	13					
K1	1	Gesicherte Freigabe (Grün)	1	40%	Länge	1	0,8	0,9	B	0,7	C	0,7	C					
					Start zu Start	1	0,9											
					Ende zu Ende	1	1,0											
		Gesperrt (Rot)	1	55%	Länge	1	0,5	0,4	E									
					Start zu Start	1	0,3											
					Ende zu Ende	1	0,4											
	Bedingtverträgliche Freigabe (Grün)	0	-	-	-	-	-											
	K2	1	Gesicherte Freigabe (Grün)	1	60%	Länge	1	1,0	1,0					A	0,8	B	0,7	C
						Start zu Start	1	0,9										
Ende zu Ende						1	1,0											
Gesperrt (Rot)			1	35%	Länge	1	0,6	0,7	C									
					Start zu Start	1	0,7											
					Ende zu Ende	1	0,7											
Bedingtverträgliche Freigabe (Grün)		0	-	-	-	-	-											
K3		0	-	-	-	-	-	-	-	-	-	-	-					
K4		0	-	-	-	-	-	-	-	-	-	-	-					
F1	0	-	-	-	-	-	-	-	-	-	-	-						
F2	0	-	-	-	-	-	-	-	-	-	-	-						
F3	0	-	-	-	-	-	-	-	-	-	-	-						
F4	0	-	-	-	-	-	-	-	-	-	-	-						

Abbildung 5.3: Tabellarische Darstellung einer Gesamtbewertung aus Teilindizes der Steuerungsdynamik.

In der Spalte 0 sind die Signalgruppen abgebildet; Spalte 1 zeigt, dass die Signalgruppen K3, K4 und alle Fußgängersignalgruppen nicht ausgewählt sind und somit nicht in die Bewertung einfließen. Bei den Zuständen in Spalte 2 gehen nur die **gesicherte Freigabe** (Grün) und **Gesperrt** (Rot) in die Bewertung ein; der Zustand **Bedingtverträgliche Freigabe** (Grün) ist für den Anwendungsfall nicht relevant und wurde entsprechend in Spalte 3 abgewählt. Für die Werte der Dynamik der Teilindizes wurden in Spalte 7 beispielhaft fiktive Werte im Intervall $[0; 1]$ angenommen. Grundsätzlich werden diese Teilindizes über die Gleichungen 4.1 und 4.4 bestimmt. Dieser Wert beschreibt die Entropie und Vorhersagbarkeit des jeweiligen analysierten Teilindexes. Die Werte der einzelnen Teilindizes (Spalte 7) werden zunächst innerhalb jedes Zustands in Spalte 8 als normierte Summe zusammengeführt. Somit beschreibt Spalte 8 die Dynamik jedes einzelnen Zustands. Anschließend werden die Zustandswerte mit Gleichung 4.6 nach dem Anteil des Zustands an der Gesamtlaufzeit einer Signalgruppe gewichtet und zu einem Wert für jede Signalgruppe in Spalte 10 zusammengefasst. Somit liegt in Spalte 10 für jede Signalgruppe ein Wert der Dynamik vor. Die Werte der einzelnen Signalgruppen aus Spalte 10 werden schlussendlich nach Gleichung 4.7 zu einem Gesamtwert der Steuerungsdynamik in Spalte 12 zusammengefasst. Die Dynamikstufen in den Spalten 9, 11 und 13 ergeben sich dann anhand der fiktiven Festlegung von Grenzwerten.

Abbildung 5.4 veranschaulicht eine solche Festlegung. Die Steuerungsdynamik des Beispielknotens erreicht in der Gesamtbewertung die Stufe **C**.

Stufe der Dynamik	Bewertung	Wert
A	vollkommen flexibel	$\leq 1,0$
B	sehr flexibel	$\leq 0,9$
C	flexibel	$\leq 0,7$
D	teilweise statisch	$\leq 0,5$
E	statisch	$\leq 0,3$
F	sehr statisch	$> 0,1$

Abbildung 5.4: **Beispielhafte Grenzwertzuordnung zu den Stufen der Dynamik.**

5.1.3 Beispielbewertung der Integrität und Plausibilität

Abbildung 5.5 zeigt eine Bewertung der Nachrichtenintegrität und -plausibilität. Auch in diesem Beispiel ist die Berechnung an das Form- und Berechnungsblatt des HBS angelehnt. Entsprechend der Konfiguration (siehe Abschnitt 5.1.1) findet in den Spalten 1 und 3 die Auswahl der zu bewertenden Signalgruppen und Teilindizes statt.

Signalgruppe	Auswahl	Teilindex	Auswahl	Teilindex Wert	Signalgruppe Wert $[F_s]$	Signalgruppe Stufe	Gesamt Wert $[F]$	Gesamt Stufe
0	1	2	3	4	5	6	7	8
K1	1	Verfügbarkeit	1	0,9	0,9	B	0,9	B
		minEndTimes	1	0,8				
		maxEndTimes	1	0,7				
		minEndTime-likelyTime-maxEndTime	1	1,0				
		gesichertes Räumen (Gelb)	1	1,0				
		bedingt verträgliches Räumen (Gelb)	1	1,0				
		Vorbereitung (Rot-Gelb)	1	1,0				
K2	1	Verfügbarkeit	1	0,8	0,9	B	0,9	B
		minEndTimes	1	0,9				
		maxEndTimes	1	0,8				
		minEndTime-likelyTime-maxEndTime	1	1,0				
		gesichertes Räumen (Gelb)	1	1,0				
		bedingt verträgliches Räumen (Gelb)	1	1,0				
		Vorbereitung (Rot-Gelb)	1	1,0				
K3	0	-	-	-	-	-	-	-
K4	0	-	-	-	-	-	-	-
F1	0	-	-	-	-	-	-	-
F2	0	-	-	-	-	-	-	-
F3	0	-	-	-	-	-	-	-
F4	0	-	-	-	-	-	-	-

Abbildung 5.5: Tabellarische Darstellung einer Gesamtbewertung aus den Teilindizes der Integrität und Plausibilität.

Für die Teilindizes der Nachrichtenintegrität und -plausibilität (Spalte 2) wurden beispielhaft fiktive Werte im Intervall $[0; 1]$ angenommen (Spalte 4). Die Werte der Teilindizes werden zunächst innerhalb einer Signalgruppe mit der Gleichung 4.12 in Spalte 5 zusammengeführt. Somit liegt in Spalte 5 ein Wert der Nachrichtenintegrität und -plausibilität für jede Signalgruppe vor. Anschließend wird mit der Gleichung 4.13 die Summe über die Werte aller Signalgruppen gebildet und auf die Anzahl der Signalgruppen normiert, sodass man in Spalte 7 einen Gesamtwert für den Knotenpunkt erhält. Die Qualitätsstufen in den Spalten 6 und 8 ergeben sich dann aus der Zuordnung zu fiktiven Grenzwerten, wie sie zur Illustration in Abbildung 5.6 aufgeführt sind. Die Nachrichtenintegrität und -plausibilität des Beispielknotens erreicht in der Gesamtbewertung die Qualitätsstufe **B**.

Qualitätsstufe	Bewertung	Wert
A	ausgezeichnet	$\leq 1,0$
B	sehr gut	$\leq 0,9$
C	gut	$\leq 0,7$
D	mittel	$\leq 0,5$
E	schlecht	$\leq 0,3$
F	sehr schlecht	$> 0,1$

Abbildung 5.6: **Beispielhafte Grenzwertzuordnung zu den Qualitätsstufen.**

5.1.4 Beispielbewertung der Prognosegüte

Abbildung 5.7 zeigt beispielhaft die Bewertung der Prognosegüte in Anlehnung an das Form- und Berechnungsblatt des HBS. In den Spalten 1, 3 und 6 findet die Auswahl der zu bewertenden Signalgruppen, Zustände und Teilindizes gemäß der Konfiguration aus Abschnitt 5.1.1 statt.

Für die Teilindizes `minEndTime` und `maxEndTime` werden in Spalte 8 beispielhaft fiktive Werte im Intervall $[0; 1]$ angenommen, welche grundsätzlich mit den Gleichungen 4.19 und 4.20 bestimmt werden. Der Wert des Teilindex `confidence-likelyTime` in Spalte 8 wird grundsätzlich mit der Gleichung 4.18 und unter Einbezug des maximalen Prognosehorizonts aus Spalte 7 bestimmt. Auch für diese Teilindizes wurden fiktive Werte im Intervall $[0; 1]$ festgelegt. Die Werte der einzelnen Teilindizes (Spalte 8) werden zunächst innerhalb jedes Zustands in Spalte 9 als normierten Summe zusammengeführt. Die Werte der Zustände (Spalte 9) werden anschließend innerhalb einer Signalgruppe mit der Gleichung 4.21 in Spalte 11 zusammengeführt. Somit liegt in Spalte 11 für jede Signalgruppe ein Wert für die Prognosegüte vor. Abschließend wird mit Gleichung 4.22 eine Summe über die Werte aller Signalgruppen aus Spalte 11 gebildet und auf die Anzahl der Signalgruppen normiert, um schlussendlich in Spalte 13 einen Gesamtwert der Prognosegüte für den Knotenpunkt zu erhalten.

Die Qualitätsstufen in den Spalten 10, 12 und 14 ergeben sich aus den fiktiven Grenzwerten gemäß Abbildung 5.6. Der Beispielknoten erreicht hinsichtlich der Prognosegüte eine Gesamtbewertung der Qualitätsstufe **C**.

Signalgruppe	Auswahl	Zustand [event state]	Auswahl	Anteil Beobachtungszeitraum	Teilindex			Zustand Wert	Zustand Stufe	Signalgruppe Wert	Signalgruppe Stufe	Gesamt Wert	Gesamt Stufe	
					Auswahl	maximaler Prognosehorizont [s]	Teilindex Wert							
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
K1	1	Gesicherte Freigabe (Grün)	1	40%	confidence	1	15	0,4	0,7	C	0,7	C	0,8	C
					likelyTime									
					minEndTimes									
		Gesperrt (Rot)	1	55%	confidence	1	15	0,5	0,7	C				
					likelyTime									
					minEndTimes									
		Bedingtverträgliche Freigabe (Grün)	0	-	-	-	-	-	-	-				
					maxEndTimes									
					maxEndTimes									
K2	1	Gesicherte Freigabe (Grün)	1	60%	confidence	1	15	0,6	0,7	C	0,8	C	0,8	C
					likelyTime									
					minEndTimes									
		Gesperrt (Rot)	1	35%	confidence	1	15	0,8	0,9	C				
					likelyTime									
					minEndTimes									
		Bedingtverträgliche Freigabe (Grün)	0	-	-	-	-	-	-	-				
					maxEndTimes									
					maxEndTimes									
K3	0	-	-	-	-	-	-	-	-	-	-	-	-	
K4	0	-	-	-	-	-	-	-	-	-	-	-	-	
F1	0	-	-	-	-	-	-	-	-	-	-	-	-	
F2	0	-	-	-	-	-	-	-	-	-	-	-	-	
F3	0	-	-	-	-	-	-	-	-	-	-	-	-	
F4	0	-	-	-	-	-	-	-	-	-	-	-	-	

Abbildung 5.7: Tabellarische Darstellung einer Gesamtbewertung aus den Teilindizes der Prognosegüte.

5.1.5 Beispiel der Gesamtbewertung

Die drei analysierten Teilaspekte für die mögliche Einführung einer GLOSA-Anwendung werden gemäß Abbildung 5.8 abschließend zu einer Gesamtbewertung zusammengefasst. Innerhalb dieser Gesamtbewertung werden die Dynamik der Steuerung mit der Stufe **C** als "flexibel", die Nachrichtenintegrität und -plausibilität mit der Qualitätsstufe **B** als "sehr gut" und die Prognosegüte mit der Qualitätsstufe **C** als "gut" eingestuft. Damit ergibt sich ein verständliches und vergleichbares Bewertungsergebnis, auf Grundlage dessen entschieden werden kann, ob eine GLOSA-Anwendung an diesem Knotenpunkt unter den bestehenden Randbedingungen eingesetzt werden könnte.

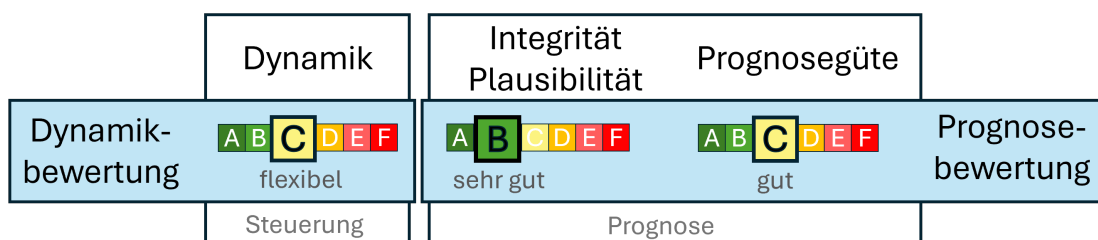


Abbildung 5.8: Darstellung der Gesamtbewertung des Anwendungsbeispiels.

5.2 Konfigurations- und Anwendungshinweise

Bevor eine Qualitätsbewertungsmetrik zur Beurteilung herangezogen werden kann, muss eine geeignete Datengrundlage zur Verfügung stehen, die eine charakteristische Verkehrslastsituation abbildet und einen ausreichend langen Analysezeitraum umfasst. Es sind eine bedarfsorientierte Konfiguration für alle Teilindizes der Metrik vorzunehmen und sinnvolle Grenzwerte für die Qualitätsstufen zu verwenden. Ziel dieser Festlegungen ist es, eine adäquate Vergleichbarkeit bei der Analyse unterschiedlicher Prognosen herzustellen. Es werden daher noch einige Hinweise gegeben, basierend auf den Erfahrungen im Umgang mit den genutzten Realdaten, welche bei der Vorbereitung, Konfiguration und Anwendung der Qualitätsbewertungsmetrik zu beachten sind.

5.2.1 Konfiguration

Auswahl und Gewichtung: Der Zweck der Konfigurierbarkeit der Bewertungsmetrik liegt primär in einer gezielten Abbildung bestimmter Anforderungen an eine Prognose, bei gleichzeitigem Erhalt der Vergleichbarkeit zu anderen Steuerungen. Wie im vorangegangenen Anwendungsbeispiel in Kapitel 5.1 beschrieben, könnte ein möglicher Anwendungsfall sein, die Prognose nur für den motorisierten Verkehr bewerten und vergleichen zu wollen. In diesem Fall können alle Fuß- und Radverkehrssignalgruppen aus der Bewertung gestrichen werden. Ein solches Vorgehen lässt sich auf andere Anwendungsszenarien und Knotenpunkte übertragen. Unterschiedliche Gewichtungen, wie beispielsweise von Haupt- und Nebenrichtung, lassen sich jedoch bedeutend schwieriger unter Erhalt der Vergleichbarkeit übertragen und sollten, wenn möglich, vermieden werden. Grundsätzlich besteht auch die Möglichkeit, nur einzelne Verkehrsströme oder Signalgruppen und ihre Prognosen zu analysieren.

Ähnliche Abwägungen gelten für die Inklusion und Exklusion von `eventStates`. Für eine Time-To-Green-Anwendung kann es bei der Bewertung zum Beispiel sinnvoll sein, sich auf die Prognose der Rotzeitenden zu fokussieren und die anderen `eventStates` nicht weiter zu betrachten. Zudem sollte berücksichtigt werden, dass möglicherweise eine feinere Gewichtung, welche andere `eventStates` nicht vollumfänglich ausschließt, beim Vergleich unterschiedlicher Steuerungen zu unerwarteten Effekten führen und damit die Vergleichbarkeit beeinträchtigen kann.

Vergleichbarkeit: Für einen geeigneten Vergleich zweier Prognosen, müssen sich die zugehörigen Steuerungsdynamiken unabhängig vom Analysezeitraum grundsätzlich ähneln. Der Analysezeitraum selbst kann starke Auswirkungen im Hinblick auf die Steuerungsdynamik und dessen Prognostizierbarkeit haben. Daher sollten die ausgewählte Analysedaten unter anderem vergleichbare Erhebungslängen aufweisen, Konfigurationen möglichst gleich gesetzt und ähnliche Grenzwerte bei der Skalierung der Dynamik- und Qualitätsstufen definiert werden.

5.2.2 Analysezeitraum aus Sicht von Anwendern und Endnutzern

Auswahl des Analysezeitraums: Der Analysezeitraum beschreibt den Zeitraum oder eine Zusammenstellung ausgewählter Zeiträume, innerhalb derer die Schaltzeitprognosedaten mit der aufgestellten Metrik analysiert und bewertet werden sollen. Dessen geeignete Auswahl ist maßgeblich für die Belastbarkeit der ermittelten Qualitätsaussagen.

Grundsätzlich sind längere Analysezeiträume zu bevorzugen, um ein breites Spektrum an verkehrlichen Situationen erfassen und statistische Besonderheiten ausgleichen zu können. Dadurch soll verhindert werden, dass kurzzeitige statistische Ausreißer das Ergebnis der Analyse verfälschen. Gleichzeitig ist der Analysezeitraum so zu wählen, dass sich darin die für einen Knotenpunkt typischen bzw. gewollt zu untersuchenden Verkehrslastszenarien widerspiegeln. Die Wahl eines bestimmten Wochentages oder einer spezifischen Tageszeit können dabei von Bedeutung sein, da diese in der Regel in Relation zu bestimmten Verkehrslastszenarien stehen.

Anwender versus Endnutzer: Aus Anwendersicht kann es sinnvoll sein, unterschiedliche LSA-Signalprogramme getrennt voneinander zu analysieren, da in der Regel für jedes LSA-Signalprogramm eine andere Prognose zur Anwendung kommt. Insbesondere im Forschungskontext kann es sein, dass eine Prognose nur unter bestimmten Rahmenbedingungen für eine Forschungsanwendung, beispielsweise bei niedrigem Verkehrsaufkommen in einem bestimmten LSA-Signalprogramm, gute Ergebnisse liefern muss. In solchen Fällen kann es sinnvoll sein, entsprechende Zeiträume zu extrahieren und einzeln oder als Gruppe zu analysieren. Der SPATEM-Nachrichtentyp als Datengrundlage bietet außerdem keine Möglichkeit zur Analyse der Verkehrsbelastung. Soll die Schaltzeitprognose unter Beachtung einer bestimmten Verkehrsbelastung analysiert werden, muss der Anwender entsprechende Analysezeiträume selbst wählen.

Aus Sicht der Endnutzer hingegen sind Merkmale wie zum Beispiel der Grad der Verkehrsabhängigkeit einer LSA-Steuerung, das gewählte LSA-Signalzeitenprogramm oder die aktuelle Verkehrslage für die subjektiv wahrgenommen Qualität einer Schaltzeitprognose nicht von Relevanz. Die Endnutzer erwarten eine verlässliche Prognose, wenn sie einen signalisierten, prognosefähigen Knotenpunkt passieren.

6 Fazit

Im Rahmen des Vorhabens „Qualität von Schaltzeitprognosedaten an Lichtsignalanlagen“ wurde ein Konzept zur Qualitätsbewertung von Schaltzeitprognosedaten erarbeitet. Dabei wurden mehrere Arbeitsschritte durchlaufen, die in den einzelnen Kapiteln dieses Berichtes beschrieben sind. Kapitel 1 führt zunächst in die Thematik ein, beschreibt die Herausforderungen sowie die Zielstellung des Vorhabens. Kapitel 2 fasst die technischen Grundlagen, Ergebnisse von Experteninterviews sowie Anforderungsanalysen zusammen und leitet daraus Kernanforderungen an ein Qualitätsbewertungskonzept ab. Das Kapitel 3 beschreibt das entwickelte Bewertungskonzept mit seinen einzelnen Teilindizes, in Kapitel 4 wird das Bewertungskonzept anhand konkreter Berechnungsschritte in eine Bewertungsmetrik überführt. Kapitel 5 ergänzt diese Ausführungen um die Anwendung des Konzeptes und der Metrik auf einen fiktiven Knotenpunkt, sowie um einige Konfigurations- und Anwendungshinweise, die sich aus der Arbeit mit realen Prognosedaten ableiten lassen.

Das Konzept umfasst drei zentrale Bewertungsschritte: (i) ein Dynamikindex zur Prognostizierbarkeit von Steuerungen, der die Vergleichbarkeit der Bewertungen über verschiedene Knotenpunkte und Steuerungen hinweg ermöglicht; (ii) eine separate Prüfung der Integrität und Plausibilität der vorliegenden Prognosedaten; und (iii) eine inhaltliche Beurteilung der Prognosen, bei der die Prognosegüte durch die Analyse der Schaltzeitdaten bewertet wird. Eine abschließende Gesamtwertung der Prognosequalität lässt sich aus diesen Einzelbewertungen ableiten.

Ein wesentliches Ergebnis der Konzepterarbeitung ist, dass auf Basis der entwickelten Metrik prinzipiell eine maschinelle, also automatisiert ablaufende Bewertung der Prognosequalität erfolgen könnte. Das erarbeitete Bewertungskonzept bietet dafür die Grundlage. Gleichzeitig hat sich jedoch gezeigt, dass für die Konfiguration der Metrik weiterhin ein gewisses Maß an Fachkenntnis und Expertenwissen erforderlich ist, etwa um zu beurteilen, ob Prognosen bei unterschiedlich dynamischen Steuerungen überhaupt sinnvoll miteinander verglichen werden können. Bei der Ausgestaltung des Konzepts sind zahlreiche, teils unvorteilhafte Ausprägungen aus realen Prognosedaten berücksichtigt worden, wodurch sich die Komplexität der Bewertungsmetrik erhöht hat. Einige Umsetzungen innerhalb der Bewertungsmetrik beruhen auf begründeten Annahmen und können bei Bedarf zukünftig auch komplexer ausgestaltet werden. Die Berücksichtigung von spezifischen LSA-Steuerungseigenschaften

und Korrelationen zwischen Signalgruppen bei der Dynamikanalyse sind hier beispielsweise zu nennen. Dies müsste dann allerdings auf einer Datengrundlage basieren, die über die hier genutzten SPATEM-Nachrichten hinausgeht. Eine Vielzahl von Erweiterungen ist hierbei denkbar, die modular in die erarbeitete Bewertungsmetrik integriert werden könnten. Insgesamt liefern die bislang gewonnenen Anwendungserfahrungen einen ersten Einblick in die Praxistauglichkeit der Metrik, machen aber zugleich das weitere Entwicklungspotenzial deutlich.

Offen ist, wie sich der derzeit noch notwendige Aufwand zur manuellen Konfiguration der Metrik künftig reduzieren oder durch geeignete Standardkonfigurationen und Unterstützungswerkzeuge verbessern lässt. Dies gilt speziell für die geeignete Auswahl und Festlegung der Analysemerkmale. Hier besteht Forschungsbedarf, die Metrik anhand repräsentativer Realdatensätze in unterschiedlichen Szenarien zu erproben, um robuste Standardparametersätze abzuleiten und bessere, systematische Vergleiche über verschiedene Knotenpunkte hinweg zu ermöglichen. Idealerweise könnten so typische und in der Praxis wiederkehrende Use Cases definiert werden, wie z. B. die Prüfung eines verkehrsabhängig gesteuerten Knotenpunktes auf die Eignung für eine GLOSA-Anwendung. Die Bandbreite möglicher Use Cases ist hierbei groß, jedoch würde die Verfügbarkeit zugehöriger Standardparametersätze die Anwendung der Metrik für diese Use Cases bedeutend vereinfachen.

In diesem Zusammenhang ist die Festlegung von geeigneten Grenzwerten für die stufenhafte Dynamik- und Prognosebewertung von weiterem Interesse. Das Konzept sieht die Verknüpfung von Qualitätsstufen mit konkreten Grenzwerten vor. Die qualitative Entsprechung einer Stufe muss hierbei auf einen nominalen Wert heruntergebrochen werden. Je nach Anwendungsfall und Use Case können sich die Kombinationen aus Stufenentsprechung und zugehörigem Grenzwert unterscheiden. Für die Herstellung sinnhafter Verknüpfungen zwischen Stufe und Grenzwert ist daher künftig die Berücksichtigung von Praxiserfahrungen der Endnutzer und Anwender von großer Bedeutung.

Vor dem Hintergrund dieser beschriebenen Aspekte sollte die Bewertungsmetrik in Zukunft verstärkt in praktischen Anwendungen erprobt und gemeinsam mit unterschiedlichen Stakeholdern weiterentwickelt werden. Auf diese Weise könnte sich die Metrik als verlässliches Werkzeug zur Beurteilung von Schaltzeitprognosen etablieren und als gemeinsamer Qualitätsstandard in der Praxis angewendet werden.

Literatur

- [1] R. Hoyer und L. Schneegans, „Prognose von Schaltzeiten verkehrsabhängiger Lichtsignalanlagen in Deutschland – Herausforderungen und Ansätze,“ *at - Automatisierungstechnik*, Jg. 69, S. 931–939, Nov. 2021. DOI: [10.1515/auto-2021-0097](https://doi.org/10.1515/auto-2021-0097)
- [2] M. Gay u. a., *Nutzung der C2X-basierten ÖV-Priorisierung an signalisierten Knotenpunkten* (Berichte der Bundesanstalt für Straßen- und Verkehrswesen, Reihe V: Verkehrstechnik 353). Bremen: Fachverlag NW in der Carl Ed. Schünemann KG, 2022.
- [3] *Hinweise für den Technologiewechsel C-ITS an Lichtsignalanlagen* (FGSV), 2024. Aufl. Köln, Germany: FGSV, 2024, Bd. 321,4.
- [4] SAE International, „V2X Communications Message Set Dictionary,“ SAE International, SAE Standard J2735_202211, Nov. 2022, Revised November 2022, Issued December 2006. DOI: [10.4271/J2735_202211](https://doi.org/10.4271/J2735_202211)
- [5] ETSI, „Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Facilities layer protocols and communication requirements for infrastructure services,“ European Telecommunications Standards Institute, Technical Specification TS 103 301 V1.3.1, 2020, Version 1.3.1 (2020-02).
- [6] C-Roads Platform, Working Group 2 Technical Aspects, „C-ITS European Handbook for MAPEM and SPATEM Creation,“ C-Roads Platform, Technical Report Release 2.0.9, 2023, Working Group 2 – Technical Aspects.
- [7] C-ROADS Platform, Working Group 2 Technical Aspects, Taskforce 3 Infrastructure Communication, „C-ITS Message Profiles,“ C-ROADS Platform, Technical Report Release 2.2.0, Juli 2024.
- [8] C-ROADS Platform, Working Group 2 Technical Aspects, Taskforce 5 Cross-Testin and Validation, „C-ITS Cross-Border Testing: PCAP Exchange Specification,“ C-ROADS Platform, Technical Report Release 2.0.3, Apr. 2022.
- [9] FGSV, Hrsg., *Handbuch für die Bemessung von Straßenverkehrsanlagen: HBS 2015* (FGSV W1 - Wissensdokumente FGSV 299 B), 2015. Aufl. Köln, Germany: FGSV-Verl., 2015.

-
- [10] T. Otto, I. Partzsch und M. Klöppel-Gersdorf, „Intersection with Highly Adaptive Traffic Lights Can Still Be Suitable for C-ITS Service GLOSA,“ in *Proceedings of the 11th International Conference on Vehicle Technology and Intelligent Transport Systems - Volume 1: VEHITS*, INSTICC, SciTePress, 2025, S. 148–156. DOI: [10.5220/0013354500003941](https://doi.org/10.5220/0013354500003941)
- [11] C. E. Shannon, „A mathematical theory of communication,“ *The Bell System Technical Journal*, Jg. 27, Nr. 3, S. 379–423, 1948. DOI: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x)

Glossar

Begriff/Kurzform	Beschreibung
C-ITS	Cooperative Intelligent Transport System, kooperatives intelligentes Verkehrssystem
confidence	Zeitfenster, in dem mit einer Wahrscheinlichkeit von 95% eine Zustandsänderung einer Signalgruppe stattfindet
dynamischer Signalzustand	Signalzustände mit unterschiedlichen Signallaufzeiten
ETSI	Europäisches Institut für Telekommunikationsnormen
eventState	Zustand einer Signalgruppe in einer SPATEM-Nachricht
GLOSA	Green Light Optimal Speed Advisory
likelyTime	wahrscheinlichster Zeitpunkt der Zustandsänderung einer signalGroup; Referenzzeitpunkt für die confidence
LSA	Lichtsignalanlage
MAPEM	MAP (topology) Extended Message
maxEndTime	spätestmöglicher Zeitpunkt für eine Zustandsänderung einer signalGroup
maximaler Prognosehorizont	Zeitbereich (in Sekunden), bis zu dem eine Prognose ausgewertet wird
minEndTime	frühestmöglicher Zeitpunkt für eine Zustandsänderung einer signalGroup
nicht-dynamischer Signalzustand	Signalzustand mit einer festen Signallaufzeit
Prognose	beschreibt die prognostizierte Zustandsänderung einer Signalgruppe
Prognosehorizont	Zeitbereich (in Sekunden), über den eine Prognose abgegeben wird

Begriff/Kurzform	Beschreibung
signalGroup	Signalgruppe mit einer Signalgruppen-ID in einer SPATEM-Nachricht
Signalgruppe	Zusammenfassung gleichgeschalteter Lichtsignale an einer Knotenpunktzufahrt; ein oder mehrere Signalgeber, die bestimmte Verkehrsströme gemeinsam steuern und zu jeder Zeit das gleiche Signalbild zeigen
SPATEM	Signal Phase And Timing Extended Message
tatsächlicher Umschaltzeitpunkt	Zeitpunkt der tatsächlich stattfindenden Zustandsänderung einer Signalgruppe
V2X	Vehicle-to-everything; bezeichnet u. a. Fahrzeug-Infrastruktur-Kommunikation

Anhang: Fragenkatalog

Im Rahmen des Vorhabens wurden Experteninterviews mit Vertretern verschiedener Stakeholdergruppen durchgeführt, um zentrale Anforderungen an die Bewertung von Schaltzeitprognosen zu ermitteln. Da Stakeholder aus unterschiedlichen Domänen beteiligt waren, wurden zusätzlich hierfür eigenes ein Glossar und eine Abbildung vorbereitet, die als gemeinsame Begriffsgrundlage und Diskussionsbasis dienen. Das verwendete Glossar und die 18 Fragen sind nachfolgend aufgeführt.

Fragenkatalog

1. Für welche Anwendung(en) nutzen Sie die Prognose?
2. Auf welchem Wert in der Prognose liegt bei Ihrer Anwendung der Fokus?
3. Ist die Prognose nur zu einem bestimmten Zeitpunkt oder in einem bestimmten Zeitbereich vor dem Signalwechsel relevant?
4. Ist der Zustandswechsel von Sperrung zu Freigabe/Freigabe zu Sperrung für Ihren Anwendungsfall relevant?
5. Welche Definition (Definition nach C-Roads „C-ITS Message Profiles“ / Definition nach SAE J2735) halten Sie für sinnvoller? (siehe nachfolgende Abbildungen)
6. Welche Herausforderungen sehen Sie bei der Nutzung der Prognose?
7. Wer nutzt Ihre Anwendung?
8. Woran machen Sie die Qualität der Prognose fest?

9. Was sind die Mindestanforderungen, die eine Prognose erfüllen muss, damit sie nutzbar für Sie ist? Welche Ausprägung hat diese?
10. Darf sich eine nutzbare Prognose über die Zeit verändern? Und wenn ja, in welcher Form?
11. Welche Eigenschaft hat eine mögliche Veränderung der Prognose, damit diese nicht mehr brauchbar ist?
12. Was benötigen Sie für einen Prognosehorizont (in Sekunden) für Ihre Anwendung?
13. Welche Kriterien muss eine Prognose abhängig vom Prognosehorizont erfüllen? (bspw. max. Abweichung, Häufigkeit der Abweichung, Genauigkeit)
14. Was unterscheidet aus Ihrer Sicht eine „gute“ von einer „schlechten“ Prognose?
15. Bewerten Sie aktuell die Qualität einer Prognose? Wenn ja, mit welchem Tool?
16. Welche zusätzlichen übertragbaren Werte oder Parameter (zusätzlich zur SPaT) wären wünschenswert, um Ihre Anwendung zu verbessern?
17. Wofür soll eine Qualitätsbewertungsmetrik zukünftig angewendet werden?
18. Wie soll die Darstellung der Bewertung aussehen? Bspw. als visuelle Darstellung oder als Ausgabe von einzelnen oder mehreren Kennwerten?

Glossar für Experteninterviews

Prognose: beschreibt den zukünftigen/ prognostizierten Signalbildzustand sowie Signalbildwechsel einer Signalgruppe/ eines Signals

Prognosehorizont: beschreibt den Zeitbereich (in Sekunden), über den eine Prognose abgegeben wird

Prognosequalität: beschreibt inwiefern eine Prognose zuvor definierte Anforderungen hinsichtlich verschiedener Merkmale und Ausprägungen erfüllt)

Signalbildwechsel: die tatsächliche Zustandsänderung einer Signalgruppe/ eines Signals (z. B. gesperrt nach freigegeben; Rot → Rot/Gelb → Grün)

Umlauf: beschreibt die einmalige Abwicklung aller Phasen in einer Folge, so dass alle Verkehrsteilnehmenden einmal Grün hatten; kann fest oder dynamisch u. a. mit der Auslassung einiger Phasen sein

Umlaufsekunde: beschreibt einen Zeitpunkt innerhalb eines Umlaufs, gerechnet ab Umlaufbeginn

Zeitpunkt: definiert eine spezifische Umlaufsekunde

Zeitbereich: beschreibt einen Zeitraum (von Umlaufsekunde X bis Umlaufsekunde Y)

Ausprägung: beschreibt die Veränderung einer Prognose und Eigenschaften dieser Veränderung

Veränderung: beschreibt den Vergleich der Prognose zu zwei Zeitpunkten miteinander und dabei auftretende Veränderungen; beschreibt die Veränderung der Prognose über die Zeit bzw. über einen definierten Zeitbereich

Freigabebeginn: Signalbildwechsel von Rot auf Grün (Rot-Gelb der Sperrzeit zugerechnet, vereinfachte Darstellung)

Freigabeende: Signalbildwechsel von Grün auf Rot (Gelb wird der Sperrzeit zugerechnet, vereinfachte Darstellung)

Beispielgrafik: Abbildung 6.1 illustriert folgende Aspekte einer Prognose: Am aktuellen Zeitpunkt gibt es eine Prognose für den Zustandswechsel für das Freigabeende von Grün zu Rot, bezogen auf einen bestimmten Prognosehorizont. Prognose A prognostiziert den Signalwechsel zu früh. Prognose B prognostiziert den Signalwechsel zu spät. Ausprägung beschreibt die Veränderung einer Prognose über die verschiedenen Zeitpunkte/Umlaufsekunden.

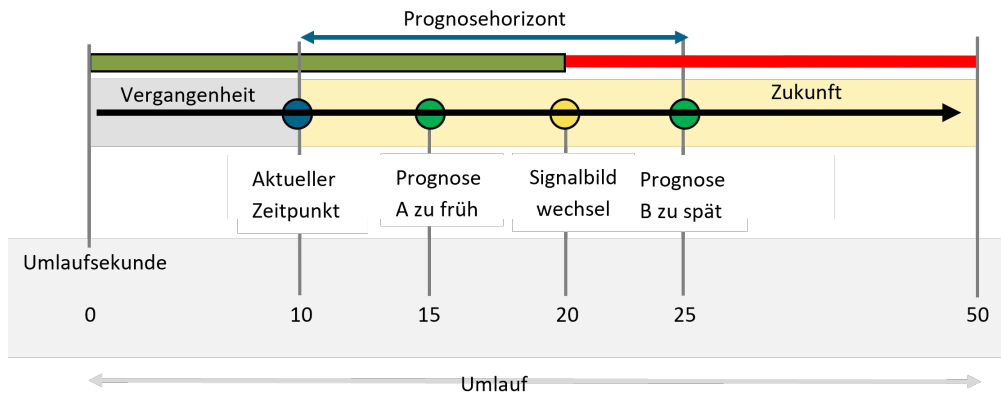


Abbildung 6.1: Diskutierte Prognosegrafik aus Experteninterviews