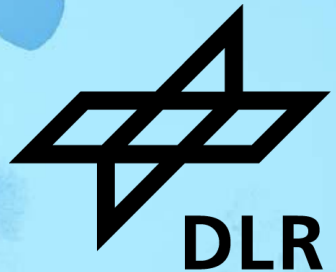


TOWARDS VIRTUAL TESTING OF PERCEPTION IN THE RAILWAY DOMAIN

Michael Wild, Jan Steffen Becker, Anda Buinoschi-Tirpescu, Eike Möhlmann



Content



- Motivation and Scope
- Background & Related work
- Challenges → Requirements
- Test creation framework
- Conceptual Architecture of Testing Framework
- ASAM OSI-Based Integration & Gaps
- Conclusions & Roadmap

Motivation & Scope



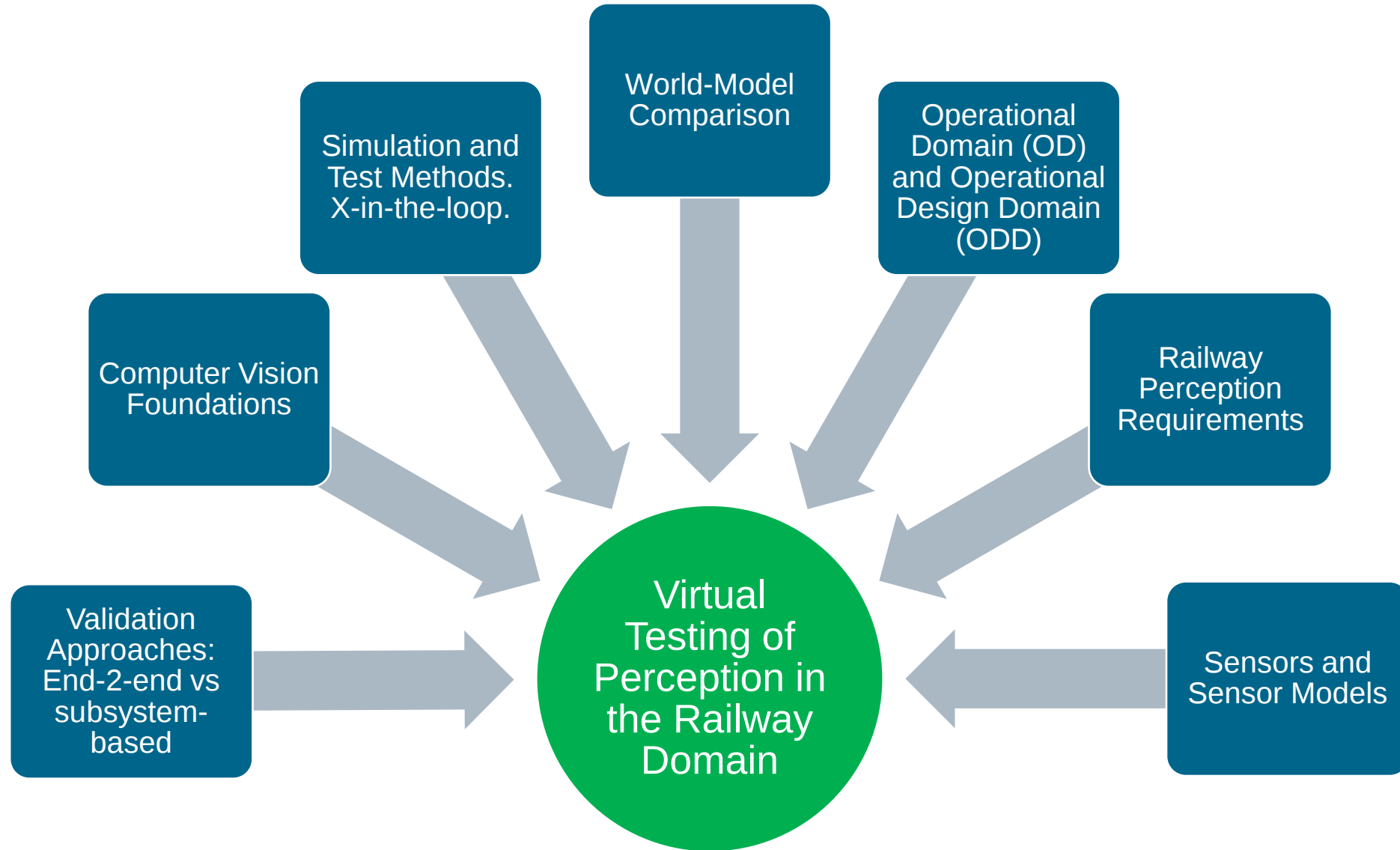
- To achieve **higher levels of automation** in the railway domain (GoA-3/4), an automated system must take over responsibilities currently allocated to human operators
- While macroscopic / technical safety exists (e.g., ATO over ETCS), operational-level **perception in open context** remains a challenge.
- Focus: Testing **perception** within the **autonomy pipeline [1]**



- Contribution: integrate existing pieces into a holistic, simulation-based, scenario-driven **testing concept** for railway perception.

[1] J. Peleska, A. E. Haxthausen, and T. Lecomte, "Standardisation Considerations for Autonomous Train Control," in Leveraging Applications of Formal Methods, Verification and Validation. Practice, ser. LNCS. Springer, 2022.

Background & Related Work



Challenges → Requirements



Challenge	Requirement
OD/ODD spans participants, conditions, and route idiosyncrasies	R1 Cover OD/ODD
Realistic vehicle–environment dynamics	R2 Realistic Physics
Multiple sensors must align in time/space/semantics	R3 Consistent multi-sensor data
Perception must align with geo-referenced route maps and asset semantics	R4 Map data
Toolchain fragmentation across vendors	R5 (Sensor) co-simulation
Uncertainty exists throughout perception pipeline	R6 Handle uncertainty
Object-level metrics miss scene topology, occlusions, and uncertainty	R7 World model comparison
Perception benefits from active sensing	R8 Active perception
DNN overconfidence harms safety	R9 Classifier calibration

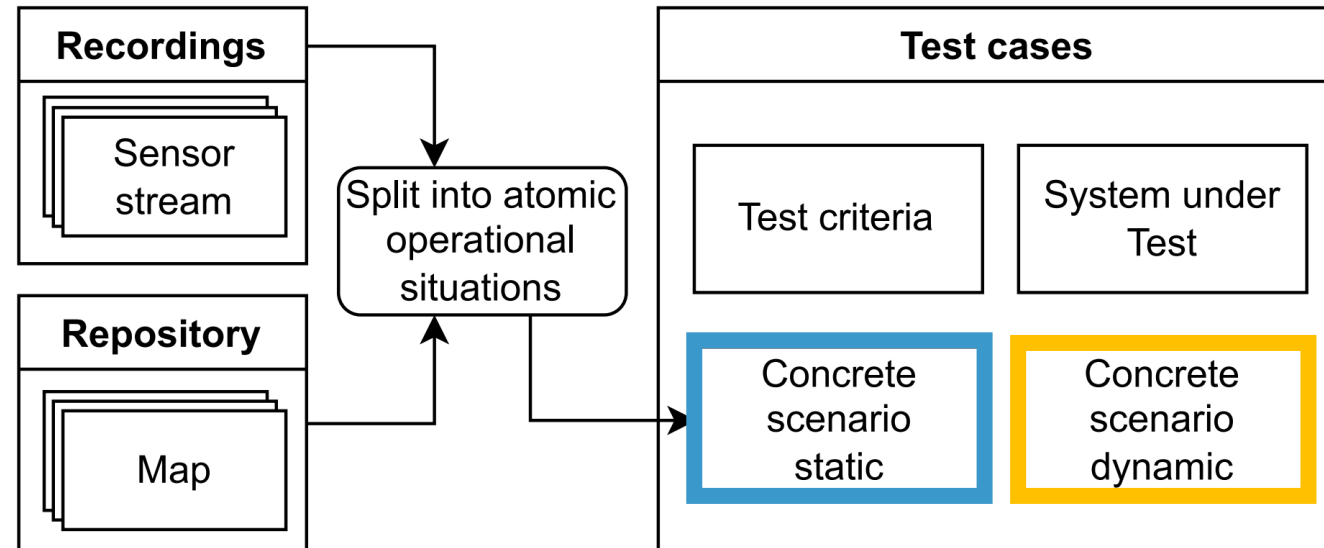
Challenges → Requirements



Challenge	Requirement
OD/ODD spans participants, conditions, and route idiosyncrasies	R1 Cover OD/ODD
Realistic vehicle–environment dynamics	R2 Realistic Physics
Multiple sensors must align in time/space/semantics	R3 Consistent multi-sensor data
Perception must align with geo-referenced route maps and asset semantics	R4 Map data
Toolchain fragmentation across vendors	R5 (Sensor) co-simulation
Uncertainty exists throughout sensing pipeline	R6 Handle uncertainty
Object-level metrics miss scene topology, occlusions, and uncertainty	R7 World model comparison
Perception benefits from active sensing	R8 Active perception
DNN overconfidence harms safety	R9 Classifier calibration

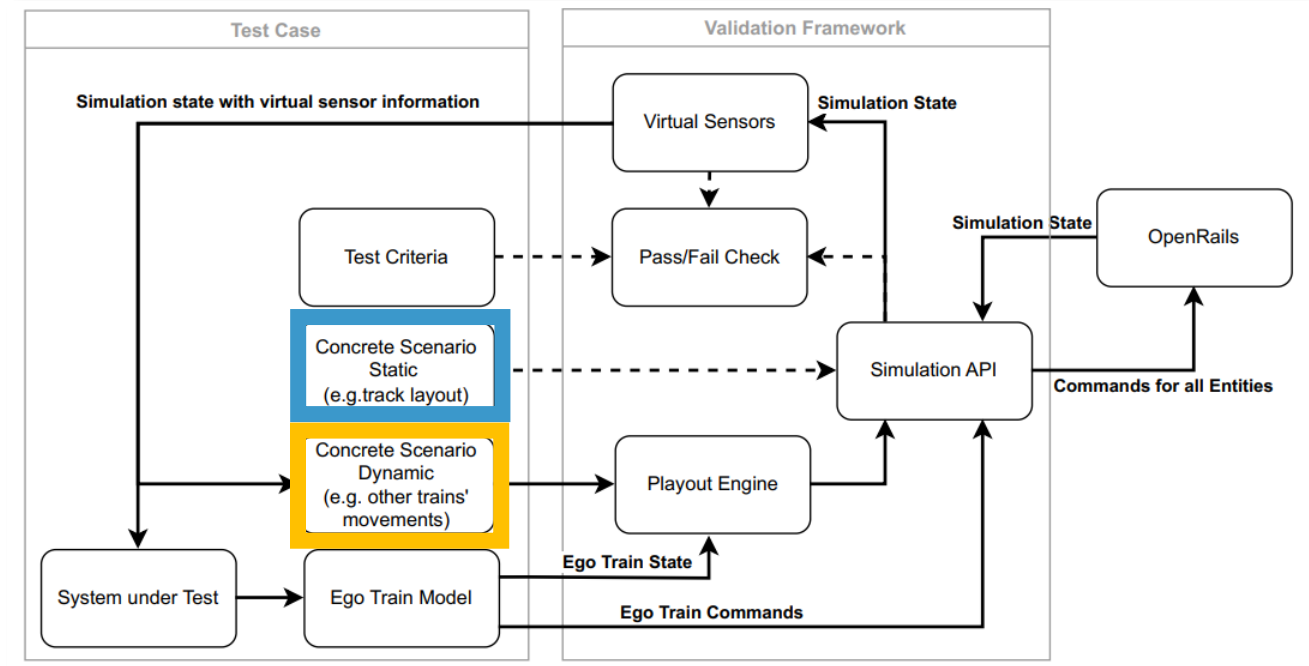
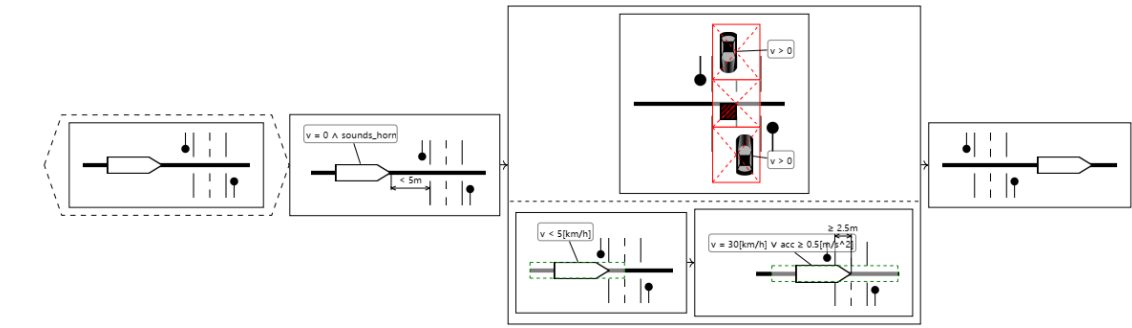
Test Creation Framework

- Record the target **line** with a sensor suite; derive static environment (geometry, semantics, 3D meshes). Also use Maps, if available.
- Split route into **atomic operational situations** (open track, level crossings, stations, shunting...).
- Model **abstract scenarios** (Traffic Sequence Charts) → sample **concrete** scenarios via constraint solving.
- Simulate concrete scenario on reconstructed static environment
- Aggregate test outcomes to line-level verdicts and weak-spot identification.



Simulation Framework

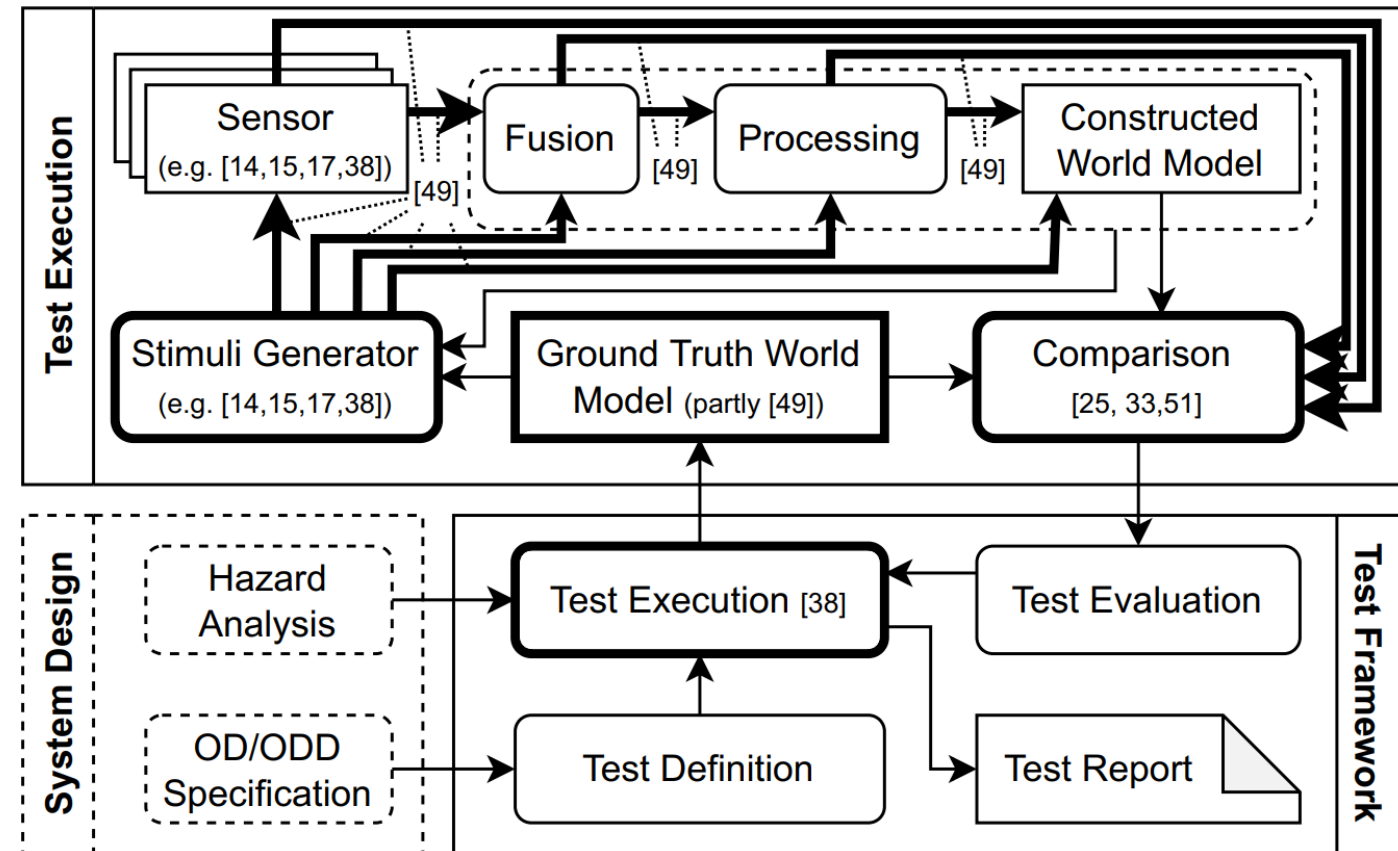
- We extended OpenRails:
 - an API for simulation control
 - enabling the generation of maps
 - the control of trains and their environment
 - as well as the reception of data such as positions, velocities, camera images, or point clouds from the simulation.
- The goal of these extensions is to make OpenRails usable as a flexible simulator in distributed simulations for research.
- <https://github.com/DLR-SE/openrails-for-simulation>



M. Wild, J. Becker, C. Schneiders, and E. Möhlmann, "A Scenario-Based Simulation Framework for Testing of Highly Automated Railway Systems," presented at the 11th International Conference on Vehicle Technology and Intelligent Transport Systems, Jun. 2025, pp. 88–99. Accessed: Jun. 18, 2025. [Online]. Available: <https://www.scitepress.org/Link.aspx?doi=10.5220/0013286600003941>

Conceptual Architecture of Testing Framework

- Inputs: system design: hazard analysis + OD/ODD Specification
- Test definition and Test execution
- Simulation maintains **ground-truth world model**; drives **stimuli generator** (physical/virtual sensors, fused).
- SUT's perception constructs **perceived world model**; runtime feedback allows **active sensing**.
- **Evaluation**: compare perceived vs. ground truth; generate reports.



Bold elements have existing solutions given as literature references.

- The ASAM Open Simulation Interface (OSI) standard enables **modular co-simulation** (environment, agents, sensors) and standardizes FMU packaging; **Ground truth and perceived models share a schema**; For perceived world models OSI adds **uncertainty**.
- **Gaps for rail**: missing rail infrastructure classes (rails, switches, balises, railway signals), mission data (stops, schedules), and some sensor types (odometry, GNSS, sound).
- High fidelity sensor models need **detailed environment models** not provided out-of-the-box.

- We described key challenges for simulative testing of perception in the railway domain and compiled requirements (R1–R9) which were mapped to sources and partial solutions.
- We provide an approach for creating realistic test cases as well as a conceptual architecture of a testing framework
- **Next steps:**
 - detailed OSI gap analysis + rail extensions,
 - modularize existing sensor simulations for OSI co-simulation
 - integrate extended OSI specification into simulators/test execution frameworks
 - advance test generation/aggregation methods
 - specify SUT & record a test track for evaluation.
- Goal: credible, scenario-driven, virtual testing path for perception toward GoA-3/4 deployment.