

# **TRANSFER OF RECONSTRUCTED REAL-WORLD ACCIDENT DATA INTO SCENARIO CATALOGUES FOR THE DEVELOPMENT AND TEST OF ADAS AND ADS**

**Marcus Petzold**

**Thomas Unger**

**Henrik Liers**

Institute for Traffic Accident Research at Dresden University of Technology (VUFO)  
Germany

Paper Number 23-0320

## **ABSTRACT**

The development and test of ADAS and Automated Driving Systems (ADS) require appropriate scenario data. To ensure the correct functionality and functional safety of such systems, an incredible amount of scenarios is necessary, containing normal, critical, and accident situations. These scenarios are usually used for virtual simulations. However, selected scenarios should be also physically tested on proving grounds. We developed a method to extract and describe maneuver-based and parameterized scenario catalogues for development and test of ADAS and ADS.

We used real accident data from GIDAS (German In-Depth Accident Study). The focus was on car accidents in urban areas as the complexity in urban traffic is much higher than on highways (heterogeneous infrastructure, large variety of road users and behavior).

At first, we clustered the (weighted) GIDAS accidents into different scenario groups. Then, we identified relevant parameters that are necessary for the description of the static and dynamic content of scenarios. The static content was extracted within the “environment analysis”. With this, the scenarios can be parameterized in terms of weather and lighting conditions, road layout (e.g. number of lanes, road width etc.).

For the “dynamic analysis” we additionally used the GIDAS-PCM, containing reconstructed maneuvers, time- and location-resolved trajectories, accident sequences. Here, we generated statistical descriptions about speeds, trajectories, braking or steering maneuvers.

Finally, some concrete example scenarios have been transferred to IPG CarMaker and OpenDRIVE / OpenSCENARIO files.

With the developed method it is possible to transfer thousands of single traffic events and/or accidents with concrete characteristics into generic (test) scenarios. Within the project, scenario groups have been created using a maneuver-based approach. There are currently four main categories (following in one lane, crossing scenarios, turning scenarios, and lane change) which are further divided into sub-maneuver groups.

The created parameter sets per scenario group contain several static and dynamic parameters. These distributions can be used by system engineers for virtual simulation runs (e.g. with randomly varied scenarios) but also by test engineers to parameterize physical tests. The approach was already tested with partners with demonstrations in physical tests.

The implementation in concrete formats (IPG CarMaker, OpenX) showed that an automated transfer is not possible at the moment due to the complexity and multitude of implementation options. The developed method works for accident data out of GIDAS and was already tested in physical tests. However, the method was not yet applied to normal/critical situations but this should also work with the presented static and dynamic parameter sets. Another limitation is the lack of automatic data transfer from the PCM format into the open ASAM standards (OpenX).

As scenario catalogues are essential for virtual simulations as well as for physical tests of ADAS and AD functions the presented method helps to provide appropriate scenario data out of real-world accidents. The big advantage is that the created parameter sets and scenarios base on reconstructed accident data and can be used independently from certain software solution or format.

## PREFACE

The methodology of the scenario catalogue generation was developed as part of the BMVI-funded project „ErVast - Einsatz dynamischer Verkehrselemente für die Prüfung automatisierter Fahrfunktionen“. Project goals were:

- Development of test technologies and tools
- Manufacturer and vehicle model-independent testing of automated and connected assistance and driving functions
- Verification of correct and reliable environment detection, e.g. by means of dynamic traffic elements

## RESEARCH QUESTIONS / OBJECTIVES

For the development and testing of ADAS and automated driving systems (ADS), appropriate scenario data is required. To ensure the correct functionality and functional safety of such systems, an incredible amount of scenarios is required, including normal, critical and accident situations. These scenarios are typically used for virtual simulations. However, selected scenarios should also be physically tested on a test site. We have developed a method for extracting and describing maneuver-based and parameterized scenario catalogues based on real accident data, which can be used to answer the following research questions:

- Which speed ranges should be considered?
- Which trajectories of motion should be tested?
- Which space requirements are needed for different scenarios?
- Which dynamic requirements can be derived for the target?

## OVERVIEW OF DATA SOURCES

In order to adequately test automated and connected driving functions, a vehicle model-specific design of the test scenario is necessary. These test scenarios must be generated for the investigation of the system's or function's effectiveness. In the project, in addition to existing homologation regulations (e.g. ECE regulations) or test protocols (e.g. NCAP), real driving data and accident data were also applied and processed for scenario generation.

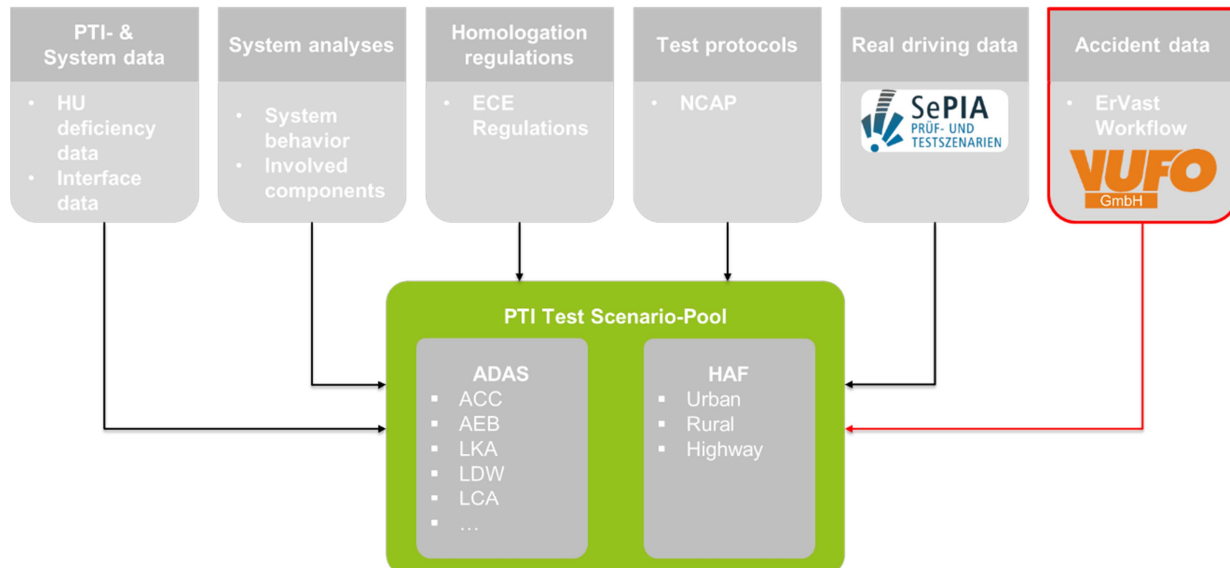


Figure 1: Overview of data sources

The goal of the scenario generation is to develop relevant and representative scenarios for the PTI test scenario pool. It is important that on the one hand a parameter set is created for a test scenario, which the testers can apply in a real test scenario. For example, do define the speed range in which the test takes place, or should a steering and/or braking maneuver must be initiated. On the other hand, it is necessary to simulate the scenarios virtually, for example to determine the space requirements for selected scenarios, to virtually secure the test scenario (feasibility), or to control the target carrier. In this paper we present some methodologies for the transfer of test scenarios from different data sources. The main focus is the scenario generation based on real accident data.

### **Real driving data**

For the development of scenarios for testing automated driving functions, the normal driving behavior of today's drivers in real road traffic is the relevant reference value. In the SePIA project, standardized processes for generating test scenarios based on real driving data were generated in this context. The results obtained in this process provided valuable input for the creation of the scenarios in the ErVast project. Examples of simple scenarios for automated driving functions are intersection scenarios without traffic signals or traffic signs, or approaching traffic signals.

### **Test protocols, homologation regulations and standards**

Some test procedures for assisting vehicle systems already exist. In both, the type approval process and in consumer protection programs (e.g. EuroNCAP), scenarios are described and test conditions are formulated in protocols. Within the scope of the project, these data were also used as a basis. An analysis of the test protocols was carried out with the aim of obtaining the relevant parameters for the test scenarios and test conditions and transferring them to the project format. One difficulty for the ErVast approach here was the large variance in functionality of systems from different vehicle manufacturers. Due to the increasing regulation of system functionalities and the resulting standardized functionality, test scenarios can also be derived across vehicle models.

### **PTI system data, System analyses**

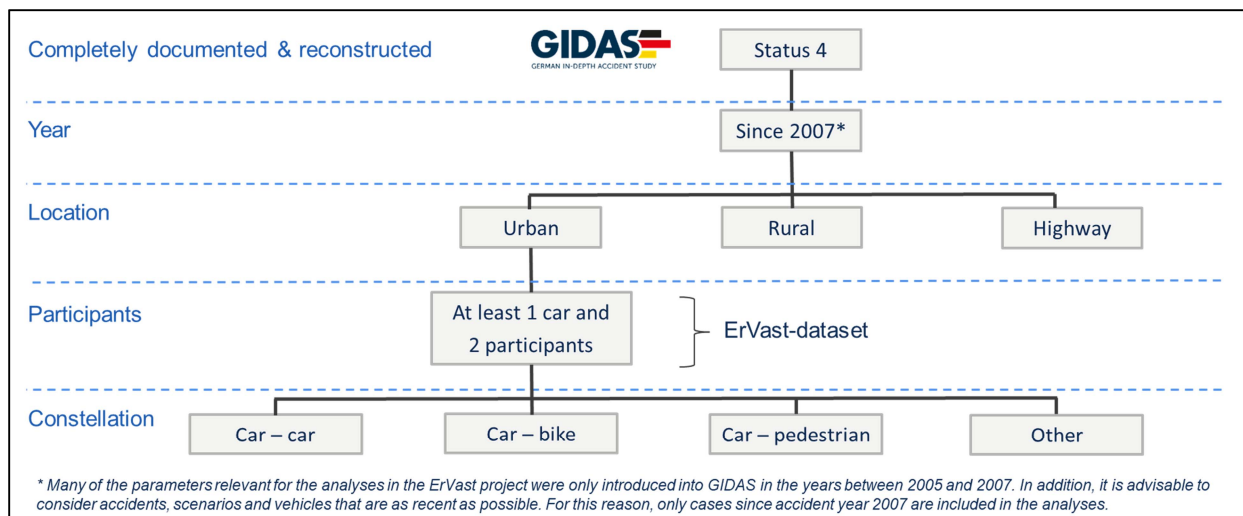
In order to be able to provide the testing organizations with specifications for the performance of the general inspection, detailed knowledge of the electronic systems installed in the vehicle, vehicle diagnostic data, and information about the installed components is essential. The provision of this data by vehicle manufacturers is ensured throughout Europe via regulation. The expertise in handling this data could be used in the project. A meta-analysis of component location plans and usable diagnostic data via the electronic vehicle interface was carried out as a basis for the development of test concepts.

### **Real accident data from in-depth investigations**

The main focus of this paper is the scenario generation based on real accident data. The GIDAS (German In-Depth Accident Study) database was chosen as basis, since this database has a very high level of detail with approximately 38,300 accidents (as of June 30, 2020) and about 3,500 single information per traffic accident. The database contains, for example, information on the accident environment (e.g. location, road type and markings, weather data), as well as reconstruction data for the participants (e.g. speeds, accelerations, maneuvers) which are of crucial importance for the scenario generation. Furthermore, it is possible to derive representative statements for the German accident scenario by applying a weighting procedure. The coupling to the simulation database GIDAS-PCM (Pre-Crash-Matrix) offers the possibility to analyze detailed time- and location-resolved information about the trajectories of the accident participants [2].

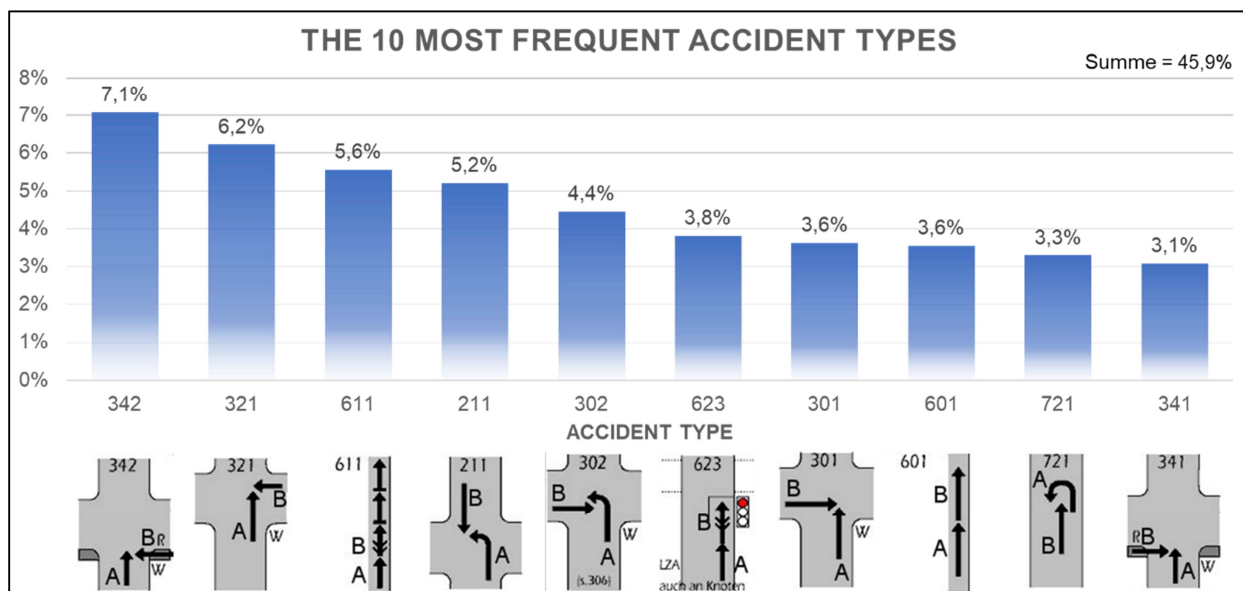
## **METHODS**

At the beginning, the accidents in GIDAS were weighted towards the German accident occurrence of 2019. Then, an analysis was performed based on the accident type, as well as a classification into scenario groups. The scenario groups were composed as follows. The project focused on urban accidents. As the scenarios should be used for PTIs of passenger cars, at least one passenger car must be involved in the accident. Furthermore, assistance systems that react to another road user should be examined. Thus, at least one other road user must be involved in the accident. These filter criteria limit the GIDAS data set within the project and lead to a first approach of an accident-based test scenario catalogue, which is illustrated in *Figure 2*.



**Figure 2: Filter criteria and categorization of accidents**

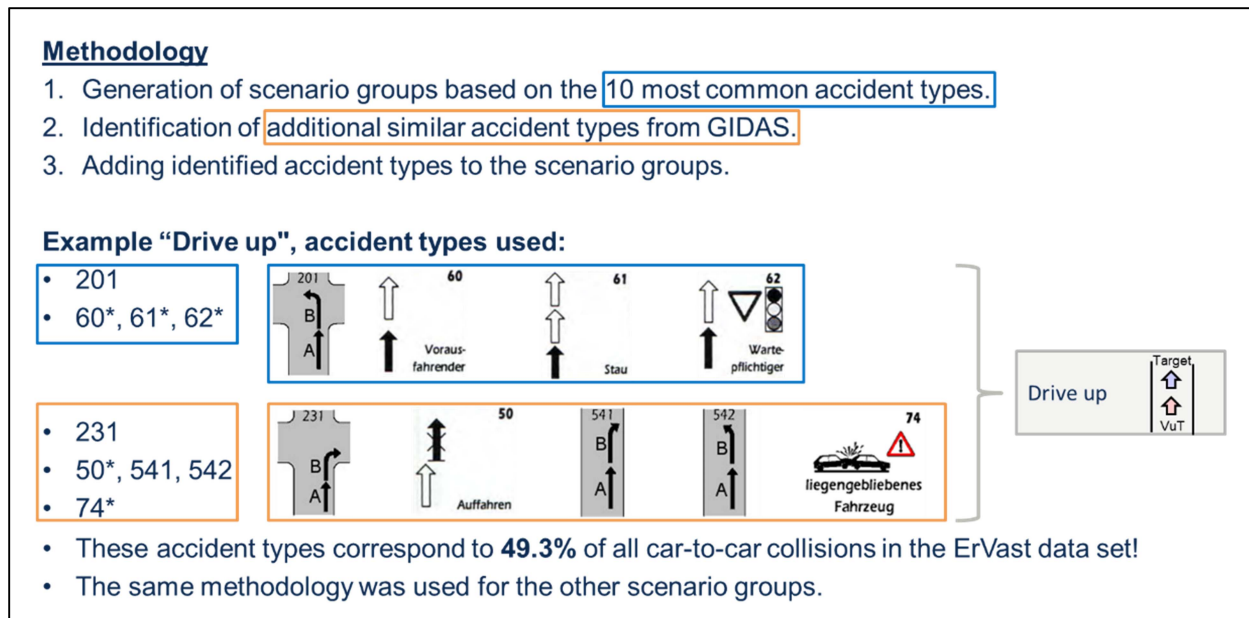
In the next step, the accident types of the ErVast data set were analyzed and evaluated. The following figure shows an overview of the 10 most frequent accident types for passenger cars in urban accidents.



**Figure 3: 10 most frequent accident types in the considered dataset of urban car accidents**

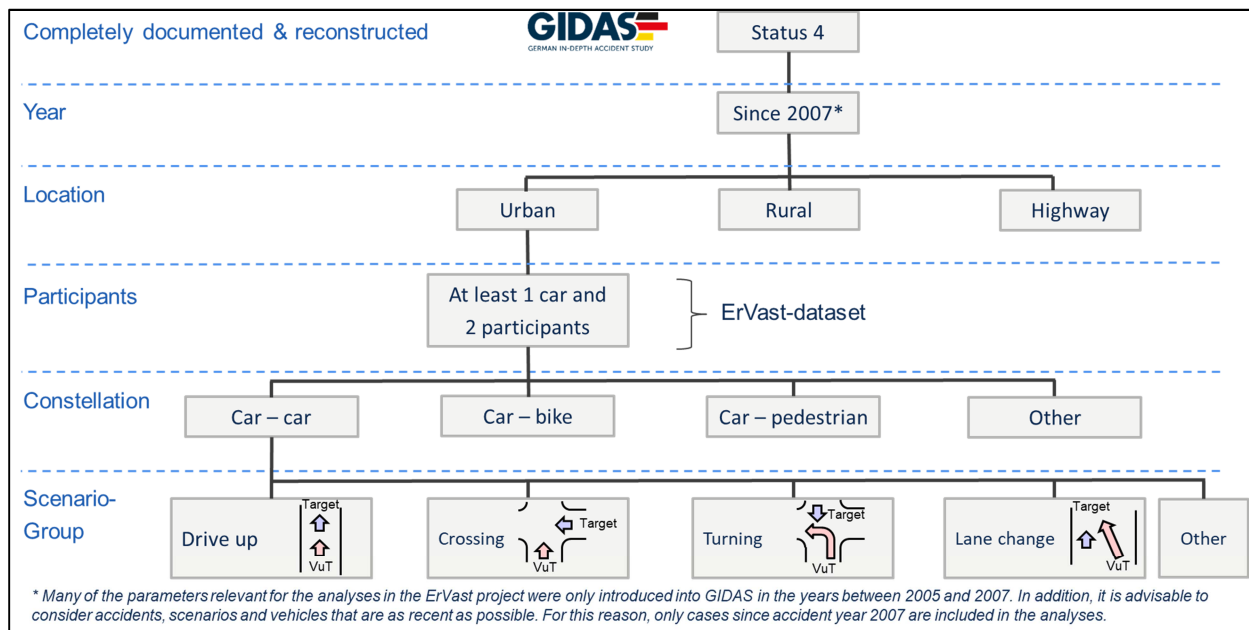
It can be seen that the analysis of the top 10 accident types already addresses almost half of all accidents (45.9%) in the ErVast dataset. In addition, the figure shows a great variety in the definitions of the accident types. For example, accident type 342 represents a passenger car and crossing bicyclist, accident type 321 includes a crossing of two participants regardless of the type of involved road users, and accident type 611 describes a longitudinal traffic accident, which is also independent of the road user types. For further analyses, it is therefore necessary to divide the data set according to the type of road user. The focus was on test scenarios for passenger cars (cars), so the data set was limited to the constellations "Car - car", "Car - bike", "Car - pedestrian" and "Other".

In a further step, additional corresponding accident types were identified in GIDAS for the identified scenario groups and added to the groups. The method is summarized in the following figure.



**Figure 4: Methodology for the generation of the accident-based test scenario catalogue**

The analyses within the scope of the project lead to the following test scenario catalogue:



**Figure 5: Accident-based test scenario catalogue**

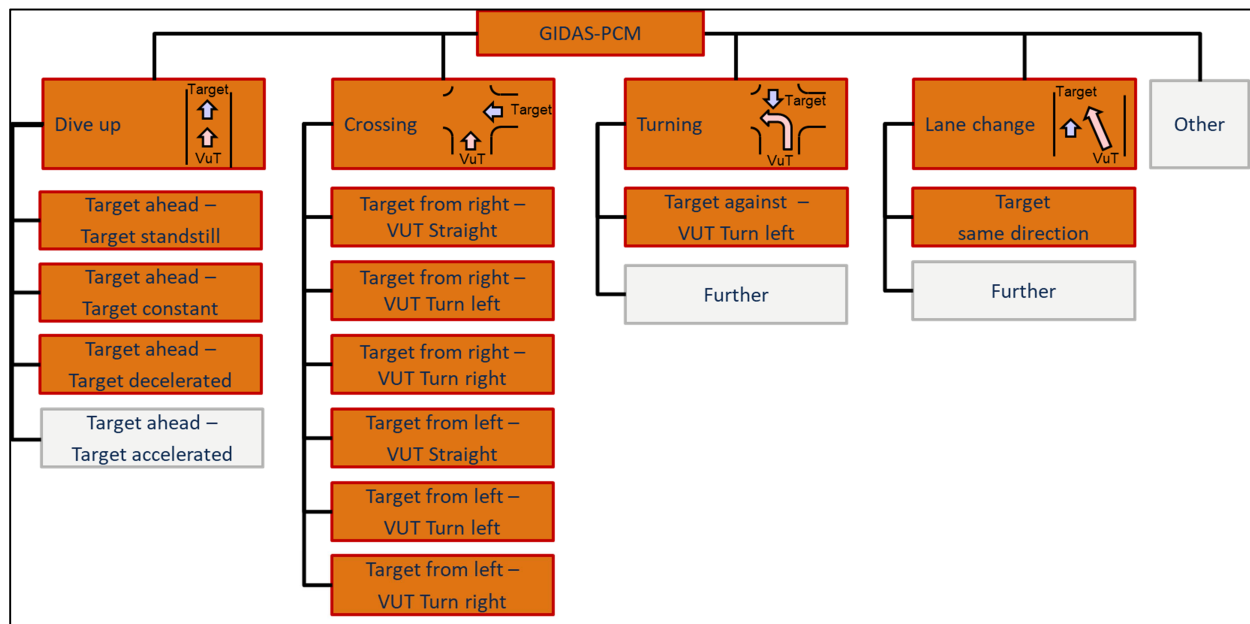
For in-depth analyses and to be able to derive test scenarios (especially maneuvers and associated dynamics data) based on real accident data, it is necessary to use another database, the GIDAS-PCM. This database contains reconstructed, time-dependent motion trajectories of the accident participants based on the accidents of the GIDAS database. The accident trajectories of the involved participants in the initial collision are considered. Based on the time- and location-resolved trajectories and the corresponding stored data of the accident participants, this database is suitable for a detailed dynamics analysis of the scenario groups. At the time of the project, the GIDAS-PCM contained approximately 10,300 accident scenarios. For car-vehicle collisions, a total of 1,951 cases could be identified for the analysis.

For the respective scenario groups, it was necessary to create further clusters for a more detailed description of the accident scenarios to analyze them with regard to characteristic features. For this step it was necessary to develop a maneuver catalogue with defined limits. In *Figure 6* the maneuver catalogue is shown.

MID	lateral	longitudinal	Motion	MID	lateral	longitudinal	Motion	MID	lateral	longitudinal	Motion
1	Straight	Forward	Constant	8	Left	Forward	Constant	14	Right	Forward	Constant
2	Straight	Forward	Accelerated	9	Left	Forward	Accelerated	15	Right	Forward	Accelerated
3	Straight	Forward	Decelerated	10	Left	Forward	Decelerated	16	Right	Forward	Decelerated
4	Straight	Backward	Constant	11	Left	Backward	Constant	17	Right	Backward	Constant
5	Straight	Backward	Accelerated	12	Left	Backward	Accelerated	18	Right	Backward	Accelerated
6	Straight	Backward	Decelerated	13	Left	Backward	Decelerated	19	Right	Backward	Decelerated
7	-	-	Standstill								

**Figure 6: Maneuver catalogue**

The maneuver catalogue describes the movement and the direction (longitudinal/lateral) of a participant and assigns it to a maneuver identifier (MID). Thus, it is possible to assign a MID to each participant, at each time step, in the GIDAS-PCM. The goal is to subdivide a scenario group into maneuver variants to analyze and describe the maneuvers separately for the VUT and target. By analyzing the single maneuvers and also their combination these results could be done. Figure 7 shows an overview of the additional maneuver variants.



**Figure 7: Maneuver variants in the scenario groups**

In the figure above, the analyzed scenarios are shown in orange. Scenarios with a gray background were not considered. For all clusters and their corresponding cases a clear assignment of the VUT and the target is included. This allows for the separate dynamics and environment analyses.

In the next step, the GIDAS-PCM was analyzed with respect to the following parameters:

- Trajectories
- Speeds
  - Start speed: Start of simulation
  - Initial speed: speed before critical scenario
  - Collision speed
- Vehicle accelerations
- Vehicle alignment
- Center of gravity positions of the vehicles
- Distances covered in the GIDAS-PCM

The analysis of dynamic parameters does not provide all necessary information and general conditions for the tests. Thus, an additional analysis regarding environmental aspects was done. Therefore, the same basic (GIDAS) master-dataset was used.

With the environment analysis, general information and conditions for the entire scenario can be extracted, e.g. lighting conditions, surrounding infrastructure, road layout etc. In addition, statements can be made which only apply to the VUT or only to the target, e.g. number of lanes or road width. The following variables were investigated.

- Lighting conditions
- Infrastructure
- Number of lanes
- Lane layout
- Road and lane width
- Marking lane
- Road condition and surface
- Type of traffic regulation
- Maximum permitted speed / type of speed limit
- Visual obstruction / type of visual obstruction

## RESULTS

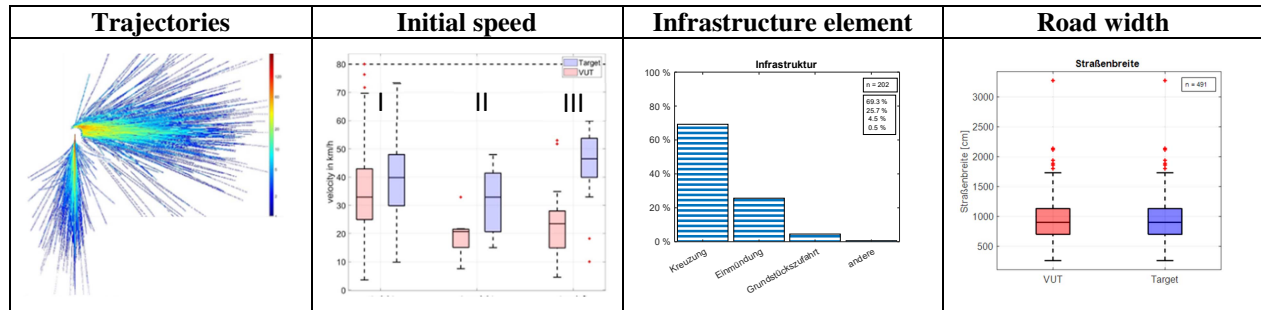
All analyzed parameters were stored in the form of diagrams and data tables and serve as parameter sets for the creation of real and virtual test scenarios. The following figure shows some evaluation examples for a crossing scenario.

### The left diagram in

*Figure 8* shows the trajectories of the two relevant road users. The colors represent the frequencies of different trajectories (red: very frequent; blue: rare). The second diagram (mid, left) shows boxplots with the relevant speed distributions of the vehicles (blue: Target vehicle, red: VUT) for three different sub-scenarios (I: VUT is driving with constant speed, II: VUT is starting and accelerating, III: VUT is decelerating prior to the collision). The third diagram (mid, right) shows a distribution of the different road elements where the considered crossing accidents happened (e.g. X-Crossing, T-junction, property exit). The right diagram displays another infrastructure parameter which is very relevant for the definition of actual test scenarios. It shows the distribution of the road width in boxplots.

These four diagrams show only a small part of the variety of analyzed parameters. Besides the dynamic parameters and infrastructure variables mentioned above, there are many more parameters characterizing available.

With such results it is now possible to derive a variety of relevant test scenarios. On the one hand side, legislation and homologation authorities or test organizations can choose the “median scenario”, using the median values from the boxplots and the “average conditions”. On the other hand side, corner cases could be defined by using the outermost values of each distribution, e.g. by consumer protection agencies that want to promote the most effective systems and functions in the market.



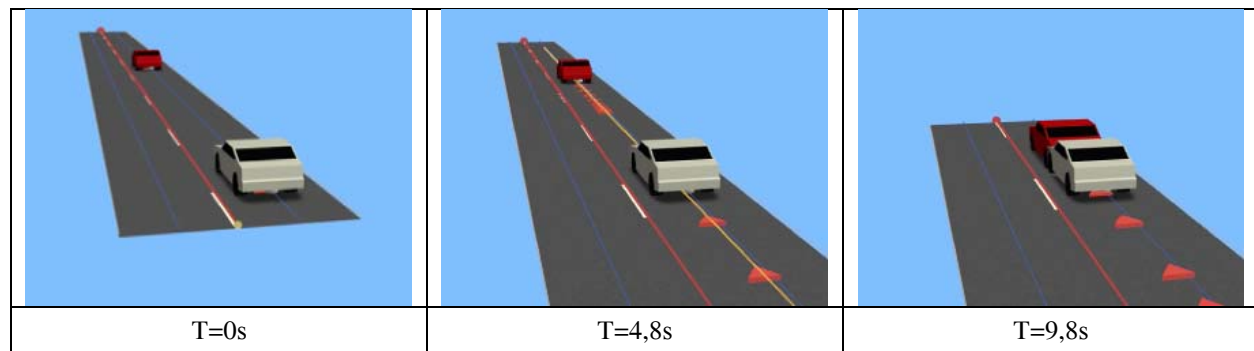
**Figure 8: Results of parameter analyses for a crossing scenario**

Based on the created parameter sets, the transfer from concrete (real) accidents to generic test scenarios was done. Additionally, some selected scenarios have been further processed to demonstrate the complete tool chain from scenario extraction to the transfer in usable format for test beds.

Therefore, three different scenarios were selected for a transfer to CarMaker and OpenDRIVE / OpenSCENARIO:

- Rear-end collision - Target ahead, target decelerating
- Lane change - Target in same direction
- Left turn across path (LTAP) - Target in oncoming lane, VUT turs left

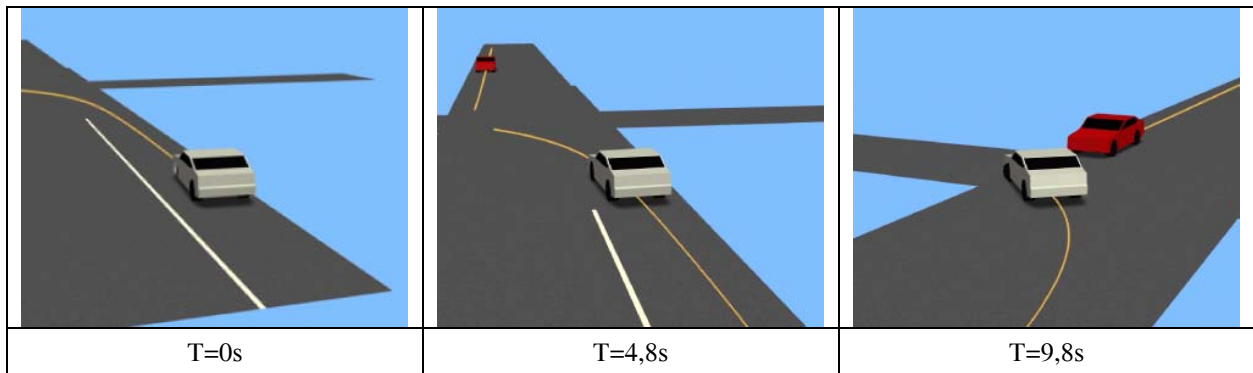
The following figure shows the rear-end collision scenario as OpenDrive / OpenSCENARIO in ESMINI.



**Figure 9: “Rear-end collision”-Scenario w/o ADAS / ADS as OpenDRIVE / OpenSCENARIO in ESMINI**



The following figure shows the LTAP scenario as OpenDRIVE / OpenSCENARIO in ESMINI.



**Figure 10: “LTAP”-Scenario w/o ADAS / ADS as OpenDRIVE / OpenSCENARIO in ESMINI**

Based on the median values of the data analyses, CarMaker, OpenDRIVE and OpenSCENARIO files could be created and successfully embedded in the further tool chains for all three scenarios. Furthermore, a variation of the dynamic values for the scenario "Rear-end collision" was carried out in CarMaker. Three variations were made to the scenario, with one parameter being changed for each variation:

- Deceleration of the target vehicle
- Initial speed of the VUT
- Variation of the trajectories

The decisive factor after the variation is the testing of the newly generated scenario. The goal of this test is collision avoidance through a specific system intervention. The simulation could be carried out successfully for the three variations in CarMaker.

Finally, it was checked whether the values of the dynamics analysis could be realized with the target carrier. For this purpose, the technical data of the target carrier were compared with the boxplot values of the dynamics analysis of the entire ErVast data set and then with the entire GIDAS-PCM 2020-1 data set. For this purpose, the maximum of the following parameters was considered:

- Speed
- Deceleration
- Longitudinal acceleration
- Lateral acceleration

The figure below shows the results of the analysis of the ErVast dataset. In the "Feasibility" column, the variable "0" stands for not feasible and "1" for feasible in terms of the technical abilities of the target.

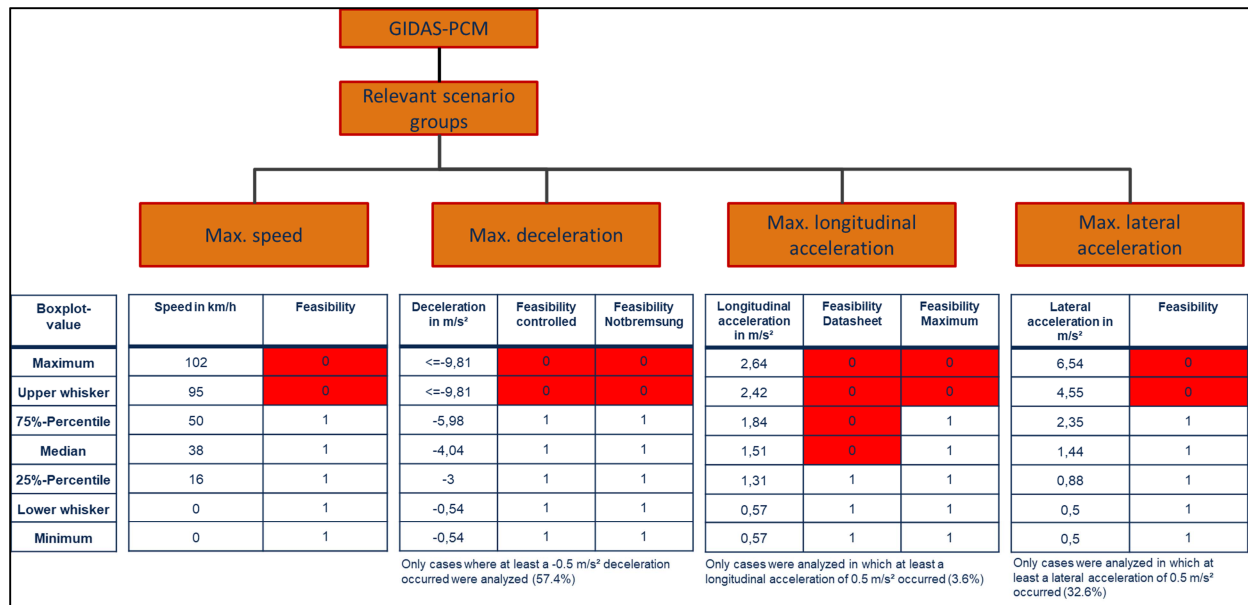


Figure 11: Overview about the dynamic feasibility of the target

## DISCUSSION

The results show that not all possible dynamics quantities can be mapped with the target carrier, but a large portion of the data can.

It was shown that a broad data basis could be used to derive test scenarios for several applications, e.g. test and validation of ADAS and AD functions but also for scenario testing within future PTI. The various data sources partly address overlapping focal points of the test, so that it is not possible to derive the required scenarios unambiguously.

In the field of accident analysis, it was shown that the methodology used is suitable for developing scenarios for testing critical driving conditions. These can be used, for example, for testing ADAS systems such as Autonomous Emergency Braking (AEB) systems.

## LIMITATIONS

It became apparent that a fully automated transfer of logic or concrete scenarios is not possible yet. This is due to the large complexity and the large number of implementation options. This means that some manual effort in combination with expertise is necessary for a meaningful application and further processing of the parameter sets with regard to the transfer into simulation formats (CarMaker, OpenDRIVE, OpenSCENARIO).

Currently, VUFO is further working on a method to transfer accident scenarios automatically from GIDAS-PCM into OpenDRIVE and OpenSCENARIO. Some hundred (less complex) cases (accidents from GIDAS, based on in-depth investigation and accident reconstruction) could be already transferred.

## REFERENCES

- [1] **The Research Association of Automotive Technology (FAT), Federal Highway Research Institute (BAST).** German In-Depth Accident Study (GIDAS). Retrieved from [www.gidas.org](http://www.gidas.org). 2022
- [2] **Schubert, A., Liers, H., Petzold, M.** The GIDAS Pre-Crash-Matrix 2016 - Innovations for standardized pre-crash-scenarios on the basis of the VUFO simulation model VAST. 7<sup>th</sup> International Conference on ESAR "Expert Symposium on Accident Research", Hannover, June 09/10, 2016.