DESIGN HOURLY VOLUME ESTIMATION AT FREEWAY NODES FROM SHORT-TERM TRAFFIC COUNTS

Marvin V. Baumann, Corresponding Author
Karlsruhe Institute of Technology, Institute for Transport Studies
Kaiserstrasse 12, D-76131 Karlsruhe, Germany
Email: marvin.baumann@kit.edu
ORCiD: 0000-0001-5952-7112

Magdalena Schilling
University of Stuttgart, Institute for Road and Transport Science
Pfaffenwaldring 7, D-70569 Stuttgart, Germany
Email: magdalena.schilling@isv.uni-stuttgart.de
ORCiD: 0000-0001-6515-1908

Markus Friedrich
University of Stuttgart, Institute for Road and Transport Science
Pfaffenwaldring 7, D-70569 Stuttgart, Germany
Email: markus.friedrich@isv.uni-stuttgart.de
ORCiD: 0000-0003-0225-025X

Sebastian Reichert
PTV Transport Consult
Harffstrasse 43, D-40591 Düsseldorf, Germany
Email: sebastian.reichert@ptvgroup.com
ORCiD: 0000-0001-7081-4171

Peter Vortisch
Karlsruhe Institute of Technology, Institute for Transport Studies
Kaiserstrasse 12, D-76131 Karlsruhe, Germany
Email: peter.vortisch@kit.edu
ORCiD: 0000-0003-1647-2435

Volker Wassmuth
PTV Transport Consult
Stumpfstrasse 1, D-76131 Karlsruhe, Germany
Email: volker.wassmuth@ptvgroup.com

Word count: 6,370 words text + 2 table(s) × 250 words (each) = 6,870 words

Submission Date: March 20, 2023

ABSTRACT
This paper extends the concept of a design hourly volume (DHV) which is derived from the 'n^{th} hour' to a concept based on the n^{th} highest saturated hour. To calculate this n^{th} highest saturated hour at each ramp junction of a node, permanent traffic counts (PTC) are necessary on all ramps and the main lanes. In practice, such counts are often not available. For such cases, the German HCM proposes a method that enables the estimation of the design hourly volume through short-term traffic counts (STC) and the extrapolation of the results using available PTC in the vicinity. Within the scope of this study, it is examined how accurate the required n^{th} highest saturated hour can be estimated with this method and similar concepts. Furthermore, it is investigated to what extent the number and the location of the available PTC affect the accuracy of the estimation. Scenarios without PTC are also considered. The evaluation is based on a database with a total of 72 freeway nodes for which PTC data from three years (2017-2019) are processed.

The results show that the estimation of the n^{th} highest saturated hour with the method of the German HCM works accurately, even if only one PTC is available on each inflowing approach. The results further indicate that STC are crucial to achieve accurate results when only few PTC are available. Acceptable results are also obtained by STC of one week, even without a projection at a PTC.

Keywords: Highway Capacity Manual, Design Hourly Volume Estimation, Short-term Traffic Counts, Permanent Traffic Counts

INTRODUCTION

Estimating the design hourly volume (DHV) is an essential step when it comes to estimating the Level-of-Service (LOS) of traffic facilities. In the Highway Capacity Manual (HCM, (*1*)) and the German HCM ("Handbuch für die Bemessung von Strassenverkehrsanlagen", (*2*)), the DHV is determined based on a concept for DHV estimation known as the n^{th} hour or respectively the hour of the year with the n^{th} highest traffic volume.. This implies designing the traffic facility in such a way that it is only oversaturated in n-1 hours per year, with typical values for n ranging between 30 and 200 hours per year. To calculate this n^{th} hour precisely, a permanent traffic count (PTC) is necessary at the corresponding traffic facility, since the traffic volume for all 8,760 hours of the year must be known. If no PTC is available, the DHV can be alternatively estimated using short-term traffic counts (STC). Appropriate methods are provided by the Federal Highway Administration's (FHWA) Traffic Monitoring Guide (TMG, (*3*)).

However, these methods refer exclusively to freeway segments with only one traffic stream. A method for the estimation of the DHV at nodes (freeway exits and interchanges on freeways) or ramp junctions (merging, diverging, and weaving segments) using STC (and its extrapolation at nearby PTC) is proposed by the German HCM.

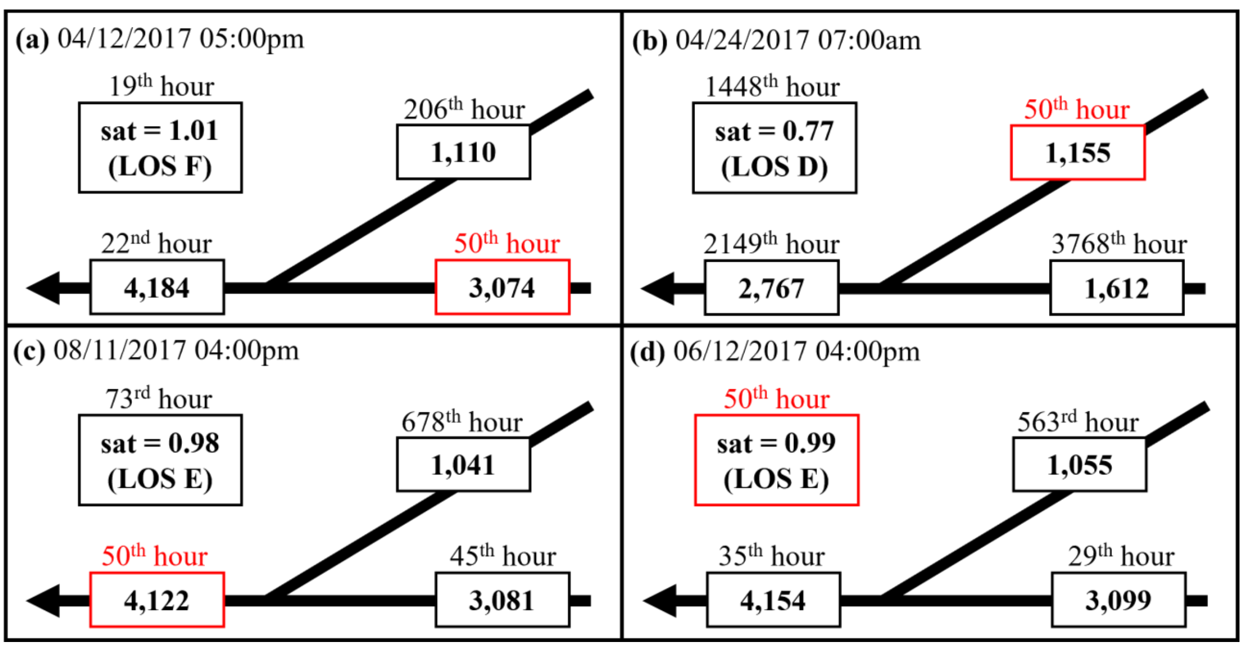


Figure 1: **Saturation rate (sat), LOS, and traffic volume (veh/h) of all streams in the 50^{th} hour of the main lane upstream (a), the on-ramp (b), the main lane downstream (c), and the saturation rate (d).**

At ramp junctions, determining the n^{th} hour for each of the streams separately would lead to an inconsistent demand-situation since these volumes would most likely not occur at the same hour. Thus, an artificial demand-situation would be created, which did not occur in reality. Figure 1 illustrates this as an example for the 50^{th} hour of a merging segment: the 50^{th} hour of the traffic stream on the main lane upstream is measured at 05:00pm in the afternoon peak-hour (a), whereas the 50^{th} highest traffic volume on the on-ramp occurs at 07:00am in the morning peak-hour (b). The 50^{th} hour of the main lane upstream corresponds to the 206^{th} hour of the on-ramp (a), whereas the 50^{th} hour of the on-ramp appears at the time of the 3768^{th} hour of the main lane upstream (b).

Calculating the LOS for each of these different hours (using the methods of the German HCM), leads to considerably different results: in the 50th hour on the main lane upstream, the result is LOS F, whereas in the 50th hour of the on-ramp, it results in LOS D. Considering the traffic volume downstream of the merging segment (c), the 50th hour of this total traffic volume results in the 45th hour on the main lane upstream and the 678th hour of the stream on the on-ramp - which results in LOS E. This example shows that the estimated LOS of a ramp junction can depend highly on the decision of which traffic stream is considered when determining the 50th hour or the DHV in general. This raises the question, which of these approaches provides the most appropriate result for the design of traffic facilities.

The concept for DHV estimation of the nth hour requests that the traffic facility may be congested at a maximum of n-1 hours in the observed year. This implies that only n-1 hours of the year may have a LOS of E or worse, as this is the threshold of acceptable traffic stream quality according to the definition of the German HCM (analogous to the HCM). As the LOS is a rather aggregated metric, it seems more appropriate to consider the metric from which the LOS is derived instead. According to the German HCM, the LOS is derived from the saturation rate. Hence, the DHV corresponds to the traffic volumes in the hour with the 50th highest saturation rate (d).

This paper examines how the nth highest saturated hour can be estimated by combining STC and PTC considering the topology of a node. For this purpose, we generalize the method proposed in the German HCM and analyze six concepts for DHV estimation at nodes. It is also examined how the number of available PTC affects the accuracy of an estimation. To achieve this, five data availability scenarios are distinguished in which the number and location of the available PTC are varied. Scenarios with and without PTC are considered. The combinations of a concept for DHV estimation and a data availability scenario lead to DHV estimation scenarios. These DHV estimation scenarios are calculated and analyzed to investigate the accuracy of estimating the nth highest saturated hour based on STC and the extrapolation with PTC with the method proposed in the German HCM.

LITERATURE REVIEW

The procedure for estimating the DHV on basic freeway segments based on STC can be divided into two sub-steps: The extrapolation of the results of an STC to an annual average daily traffic (AADT) and the determination of the DHV from the AADT. In the following literature review, studies are considered that deal with these two sub-steps.

To estimate the AADT from STC or short-period traffic counts (SPTC), the TMG (*3*) proposes the following method: First, grouping the available PTC stations into groups of road sections with similar temporal traffic variations. Then average seasonal adjustment factors are determined for each of these groups. The road section monitored with STC is assigned to one of the road section groups. The STC results are extrapolated to an AADT using the corresponding seasonal adjustment factor of the assigned road section group.

(*4*) and (*5*) further develop this method in the sense that the grouping of the PTC is performed using a fuzzy C-means algorithm, and Neural Networks are used for the assignment of the monitored road section. The authors select 25 PTC stations of a rural road network to validate the presented method. These PTC are grouped and used to derive 775 sample SPTC each covering one week. The study shows satisfactory results regarding the AADT estimates that could be obtained with SPTC of one week. It is found that the results depend on the period during which the SPTC are carried out, assuming that socio-economic and land-use characteristics influence the temporal pattern of the local travel demand. The authors recommend incorporating these characteristics in future models (*5*).

(*6*) apply machine-learning approaches such as Artificial Neural Networks and Support Vector Regression to develop models for estimating AADT from STC. According to the authors, Support Vector Regression is the most appropriate model, especially when incorporating hourly volume data and day-of-week and month-of-year as categorical features. It is found that including socio-economic factors into the model lowers the model performance. The authors conclude that the AADT of a location is primarily based on temporal traffic patterns like day-of-week or month-of-year.

(*7*) propose a method that improves the matching of STC to PTC or groups of PTC by including all historical counts collected for these sites in the model. The proposed two-pattern matching methods can significantly improve the performance of the TMG method (*3*). (*8*) compares the TMG method with a similar method used in Korea. The results show that it can be helpful to carry out two days of STC, one in each half of the year. (*9*) investigate the reliability of daily truck traffic estimates from STC and (*10*) the estimation of the AADT of non-motorized traffic.

(*11*) compare three different methods for determining the DHV. The first method is determining the DHV manually from the peak-hour of the design day during the week of STC. The remaining two methods estimate the DHV from the AADT by multiplying a K-factor, which represents the relative proportion of the DHV on the AADT. The methods differ regarding the approach to determine the K-factor. The performance of the methods is tested based on 74 PTC. For each PTC, 10 weekly counts are simulated by extracting the traffic volumes of the corresponding week and estimating the DHV with each of the three methods. The estimated DHV are then compared to the actual DHV (30^{th} hour), which are calculated based on all 8,760 hours of the year for the corresponding PTC. The simplest method of manually determining the DHV provides the best results. This result shows that the K-factor must be chosen very well to obtain accurate results.

(*12*) propose a model that also determines the DHV as a function of AADT but is considering the ratio of the daily volume of the design weekday to AADT and the ratio of the daily volume on Saturday to AADT. The model is developed using statistical analysis of 2016 PTC data and is then tested on the data of 2015 and 2017. The model provides accurate results with the advantage that classification of road sections based on the traffic flow characteristics is not required.

(*13*) investigate holiday peaking characteristics and the contributions of holiday travel to the yearly highest hourly volumes for rural highways in Alberta, Canada. Based on the analyzed data (PTC of 20 years), genetic algorithms are used to develop models for the prediction of DHV. The study demonstrates that models assisted by genetic algorithms are very robust and suitable for determining the DHV.

CONCEPTS FOR DHV ESTIMATION AT FREEWAY NODES: COMBINING PTC AND STC

The literature review indicates that STC may be a suitable data source for determining the DHV when data from PTC are not available. However, the studies all refer to count stations on basic freeway segments. An equivalent concept for DHV estimation in the context of a node has not yet been investigated in detail. The German HCM (*2*) shows an example of how to combine PTC and STC using topological relationships of a freeway interchange:

The example assumes that a cloverleaf interchange, as shown in Figure 2, has eight PTC, one for each inflow or outflow, and at least two STC for each ramp junction. Each PTC defines a specific demand-situation, which needs to be analyzed. In the following, these demand-situations are referred to as PTC demand-situation.

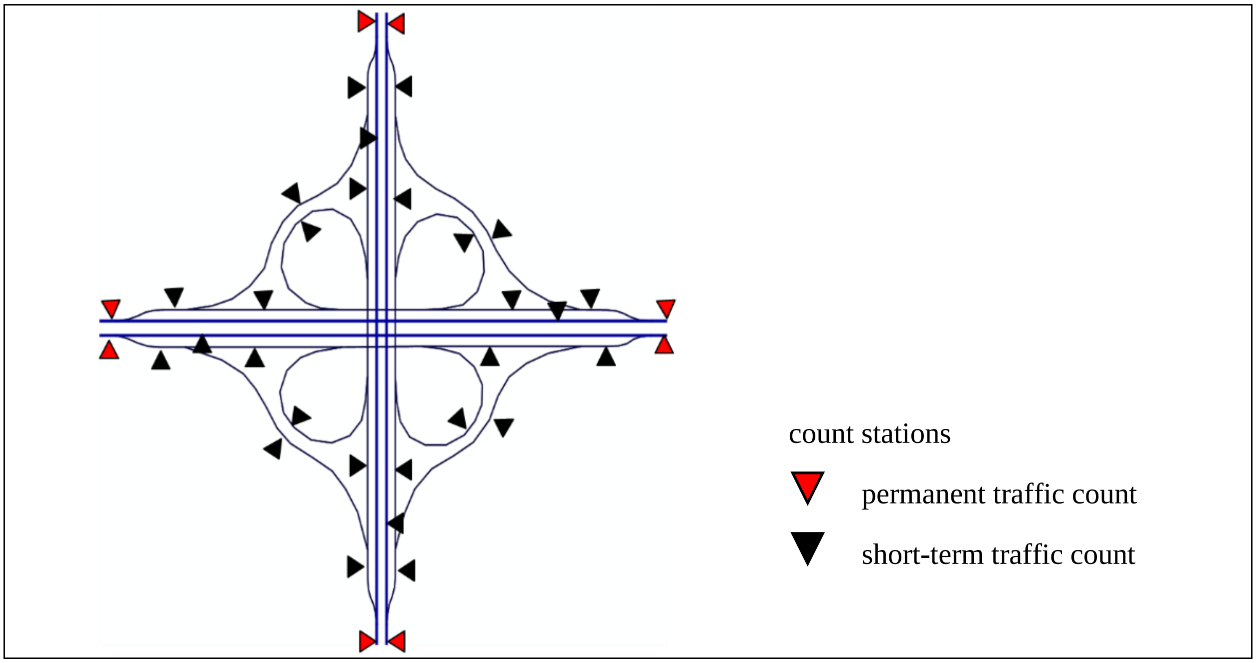


Figure 2: Count stations on a cloverleaf interchange based on an example in the German HCM (*2*).

Each PTC demand-situation describes a temporary state with consistent traffic volumes at the entire interchange, such that inflow equals outflow. In the German HCM (*2*), these demand-situations are defined using the method '50th hour of the PTC'. For the example in Figure 2 this leads to eight demand-situations, which may occur on different weekdays and times of day. Considering the exemplary merging segment in Figure 1, we can define three PTC demand-situations (a), (b), and (c), that is, one PTC demand-situation per PTC. Each PTC demand-situation has a different date and time, e.g., 12th April at 05:00pm (a), 24th April at 07:00am (b), and 8th September at 04:00pm (c).

STC are usually conducted at a different date. Therefore, the German HCM uses the day hour of the PTC demand-situation to derive a second demand-situation based on the STC (STC demand-situation). For the PTC demand-situation (a) in Figure 1, the STC demand-situation is 05:00pm. In the next step a matrix estimation method is applied using the PTC demand-situation as boundary conditions and the STC demand-situation as initial matrix to derive DHV for each count station.

This procedure is repeated for all eight PTC demand-situations of the cloverleaf interchange. After that, all eight demand-situations are evaluated. For each ramp junction, a separate saturation rate is estimated per demand-situation. The resulting saturation rate of a ramp junction is the worst-case saturation rate of all demand situations considered.

This approach can be generalized as follows. A concept for DHV estimation in the context of a node has four components:

- basic concept for DHV estimation for PTC, e.g., the 50th hour
- demand-situation(s) for the PTC (PTC demand-situation)
- demand-situation for the STC (STC demand-situation), which can depend on the PTC demand-situation
- data extrapolation procedure, e.g., matrix estimation

The basic concept for DHV estimation defines the traffic volumes of the PTC demand-situation. The combination of PTC demand-situation and STC demand-situation serves as input for the data extrapolation procedure. In this way, the PTC demand-situation defines boundary conditions for the extrapolation procedure and the STC demand-situation defines initial information of the traffic streams for all relations, i.e. origin-destination flows at the node. Using matrix estimation as a data extrapolation procedure, we get the general procedure summarized in Figure 3.

1. Define all PTC demand-situations, including traffic volumes for all PTC, regarding the basic concept for DHV estimation.
2. For each PTC demand-situation:
   a. Define STC demand-situations with volumes for all relations.
   b. For each STC demand-situation perform a matrix estimation with an initial matrix from the STC demand-situation and the boundary condition from the PTC demand-situation.

Figure 3: General procedure to calculate the DHV based on PTC and STC on freeway nodes.

For our analysis, we define four concepts for DHV estimation for the DHV in the context of a freeway node (Table 1) with PTC. They are applied to the general procedure described in Figure 3. Note: In this study, the 50th highest saturated hour is examined as a representative of the nth highest saturated hour.

1. 50th h (STC: day hour): This is the concept used in German HCM described in the example above. This concept applies PTC demand-situations with consistent traffic flows at the entire interchange (inflow equals outflow).
2. 50th h (STC: mPH or aPH): To analyze the influence of using a specific day hour for the STC demand-situation, which depends on the 50th hour of a PTC, we define a second concept by changing the way of deriving the STC demand-situation. Here the STC demand-situation does not depend on the day hour of the PTC demand-situation but is based on either the morning peak hour (mPH) or afternoon peak hour (aPH) of the STC considering the total traffic volume of all counting stations of the node. Which peak hour is chosen depends on whether the 50th hour occurs in the morning or in the afternoon. The PTC demand-situations are the same as in concept '50th h (STC: day hour)', thus providing consistent traffic flows.

3. 30^{th}-70^{th} h (STC: frequent day hour): This concept varies the way for the DHV estimation at the PTC. Instead of the 50^{th} hour, the mean traffic volumes of the 30^{th}-70^{th} hour are considered. To achieve consistent traffic flows at the node, each examined PTC demand-situation defines 41 hours of the year (calendar day and day hour). Traffic volumes at neighboring PTC stations are computed as the average of volume for these 41 time points. For the STC demand-situation this concept uses the most frequent day hour of all examined 41 hours.
4. 30^{th}-70^{th} h (STC: mPH+aPH): This concept also uses the 30^{th}-70^{th} hour for the DHV estimation at the PTC, but instead of considering one PTC demand-situation for each PTC, it only considers one PTC demand-situation with an average volume for each PTC station. This leads to a PTC demand-situation, where traffic flows are not consistent. As STC demand-situation the concept distinguishes two cases: one demand-situation regarding the morning peak hour of the STC considering all counting stations of the node and one demand-situation for the afternoon peak hour.

The STC demand-situation is consistent for all four concepts. If there is no PTC available, the German HCM suggests STC for a week, using the peak hour as DHV estimation for a single count station. To extend this approach to a node, we define two additional concepts covering only STC:

1. STC PH week (worst case): The concept defines one demand-situation per count station, which is derived from the weekly peak hour of the count station. For each ramp junction, the worst-case is considered. Each demand-situation shows consistent traffic flows.
2. STC PH week (one DS): This concept defines one demand-situation (DS) combining the peak hours of all count stations. This demand-situation has inconsistent traffic flows.

Furthermore, the number of PTC and STC can be varied, which leads to different data availability scenarios. They are defined and explained in the section ‘Data Availability Scenarios’.

Table 1: Overview of the six concepts for DHV estimation and their attributes.

Name of concept for DHV estimation	Basic concepts for DHV estimation for PTC	Number of DHV estimation scenarios	PTC demand-situation	STC demand-situation	Consistent traffic streams
50^{th} h (STC: day hour)	n^{th} hour	# PTC stations * # STC days	One for each PTC	Day hour of 50^{th} hour	yes
50^{th} h (STC: mPH or aPH)		# PTC stations * # STC days	One for each PTC	Morning peak hour or afternoon peak hour, regarding 50^{th} hour	yes
30^{th}-70^{th} h (STC: frequent day hour)	Mean value for a range of hours around the n^{th} hour	# PTC stations * # STC days	One for each PTC	Most frequent day hour in set of 30^{th} – 70^{th} hours	yes
30^{th}-70^{th} h (STC: mPH+aPH)		2 day times * # STC days	One including all PTC	Morning peak hour and afternoon peak hour	no
STC PH week (worst case)	Peak hour of week	# ramp junctions * # weeks of year	-	One demand-situation per ramp junction, no matrix estimation necessary	yes
STC PH week (one DS)		1 * # weeks of year	-	One demand-situation including peak hours of all ramp junctions	no

METHODICAL APPROACH

The intention of this study is to investigate how accurately the n^{th} highest saturated hour of a ramp junction can be determined using the method proposed by the German HCM - depending on the STC demand-situation and the number and location of the PTC stations.

For this purpose, a reference database is first created based on real-world PTC data from 2017, 2018, and 2019, which includes 72 nodes overall. The available traffic count data are expanded in such a way, that the traffic volume is known for each main lane and ramp for all 8,760 hours of the year. The German HCM evaluation procedures for the calculation of the saturation rate, as well as the proposed methods for DHV estimation based on STC, are implemented as a Python program. This program computes for each ramp junction the saturation rate for every hour of the year and derived the saturation rate at the 50^{th} highest saturated hour. This saturation rate serves as reference scenario to evaluate DHV estimation scenarios which combine a concept for DHV estimation and a data availability scenario.

To examine the concepts shown in Table 1, STC are simulated. For each year, all reasonable days (workdays outside of vacations) are determined. For each of these days, the STC is simulated, leading to different DHV estimation scenarios for every concept for DHV estimation combined with a data availability scenario. The traffic volumes of the corresponding day are extracted from the reference database for all count stations of the node which are not recorded by a PTC station. The volumes are used as the results of a simulated STC according to the method suggested by the German HCM. The set of count stations equipped with PTC and count stations requiring a STC depends on the data availability scenario. The tested data availability scenarios differ in the amount of available PTC at the node. Each simulated STC estimates the saturation rate of the 50^{th} highest saturated hour for each ramp junction. The estimated saturation rate is compared to the saturation rate of the reference scenario.

DATA PROCESSING

To perform the analysis described above, a reference database containing hourly traffic volumes for nodes and their ramp junctions is established. The database provides traffic volumes meeting the following requirements:

- For each count station, traffic volumes are available for all hours of the year.
- For each node, the available count stations represent a state of complete detection. In this state, count stations exist for all main lanes and ramps of the respective node.
- The data of the count station is consistent, meaning that within an hour, inflows correspond to outflows at the respective node and for all ramp junctions of this node.
- The reference database provides the input for computing the saturation rate and the LOS according to the German HCM for every ramp junction at every hour of the year.

The reference database is based on count station data provided by the (German) Federal Highway Research Institute and the states of Bavaria, Hesse, and North Rhine-Westphalia. The traffic volume data is aggregated to hourly intervals. To create a reference database meeting the requirements described above, the data is processed in three sequential steps:

1. Selection of nodes based on the data availability of the corresponding count station.
2. Processing on count station-level: Temporal data completion of missing hourly values for each count station.
3. Processing in the network context: Spatial data completion by inserting virtual count stations and ensuring consistency in the context of a node.

Since such a state of complete detection does not exist in reality, nodes with a good data availability are identified in the first step. Data availability is good if PTC stations exist not only on the main lane, but also on several ramps and if the PTC provides hourly volumes for all or almost all hours of a year. In the next step, missing time periods are filled in for all identified count stations. In case of short time periods missing, traffic volumes are scaled based on the context of the respective hour. In the case of missing hours, these are supplemented by the average traffic volume of the corresponding combination of hour of day and type of day (work day, vacation work day, Sunday, and public holiday). This serves as input values for the subsequent processing in the network context. For this purpose, the topology between the individual count stations has to be defined. Thus, the count stations must be combined to ramp junctions (e.g., for a merging segment, the count station representing the traffic volume of the on-ramp as well as the count station representing the traffic volume upstream of the segment on the main lane need to be combined). In addition, the meta information required for evaluating the ramp junction according to the German HCM, such as the type of the ramp junction or the longitudinal slope, is collected for each ramp junction. For the processing regarding the network context, the temporally completed traffic volumes of all existing count stations of the considered node are used as the boundary condition for a matrix estimation procedure, which is carried out for every hour of the year. In case the matrix estimation finds an admissible solution, each count station is assigned a traffic volume that is consistent in the network context. This step also determines the traffic volumes of the virtual count stations. Virtual count stations are count stations that do not exist in reality but are necessary to meet the state of complete detection.

Overall, a total of 72 nodes are processed with the described procedure. These nodes contain 701 ramp junctions. Traffic volumes come from 867 real and 519 virtual count stations. If necessary, missing data (e.g., due to malfunction of a PTC station) is derived based on the existing data to provide traffic volumes for all traffic streams of each node. These completed data represent traffic volumes in the context of the respective node and its surrounding PTC stations. It cannot be verified whether these are exactly the real traffic volumes which could not be measured. The traffic volumes of the reference database can be considered as realistic. All subsequent calculations are based on this database.

DATA AVAILABILITY SCENARIOS

To understand the impact of PTC and STC, five data availability scenarios are examined that differ in the number of available PTC and STC:

1. ‘PTC: all, STC: -‘: This data availability scenario assumes no STC but PTC for all count stations. This case represents a state with perfect data knowledge (Figure 4, first row).
2. ‘PTC: in-/outflow main lanes, STC: -‘: This data availability scenario again uses no STC data, but the numbers of PTC stations are reduced to one count station for each inflow or outflow on the main lanes of the node. This leads to eight PTC stations at a four-leg-interchange and to four PTC stations at a freeway exit (Figure 4, second row).
3. ’PTC: in-/outflow main lanes, STC: all’: This data availability scenario adds STC information for all counting stations (Figure 4, third row).
4. ‘PTC: inflow main lanes, STC: all’: This data availability scenario further reduces the number of available PTC to one PTC station for each inflow on main lanes. This leads to four PTC stations at interchanges and two stations at a freeway exit (Figure 4, fourth row).
5. ‘PTC: -, STC: all (week)’: This data availability scenario assumes only STC but extended to the period of one week.

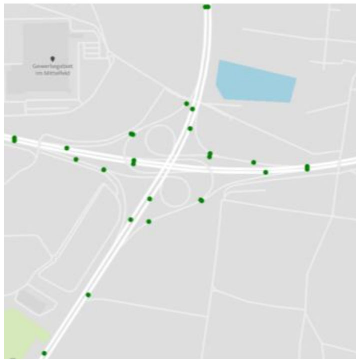
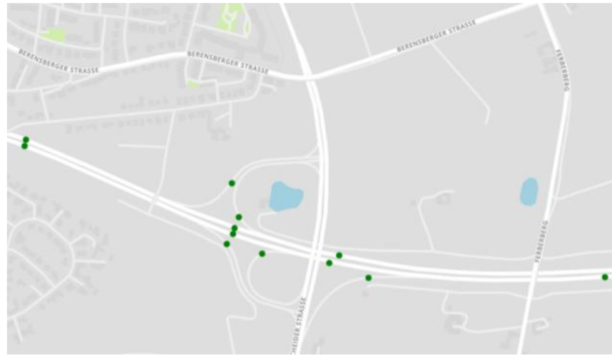
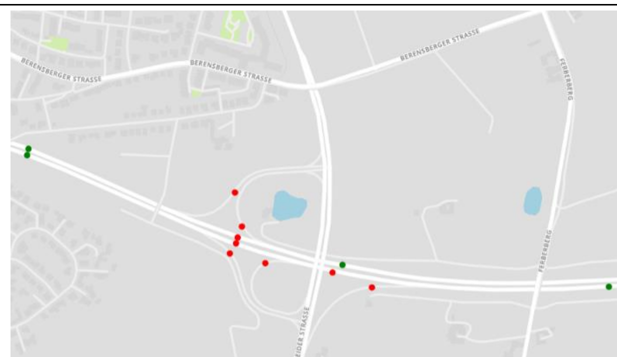
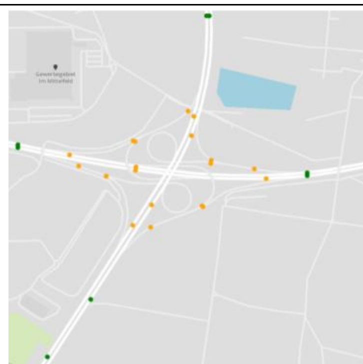
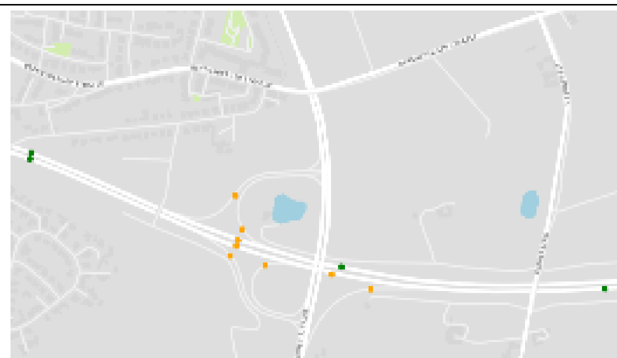
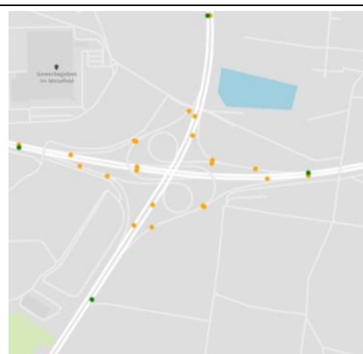
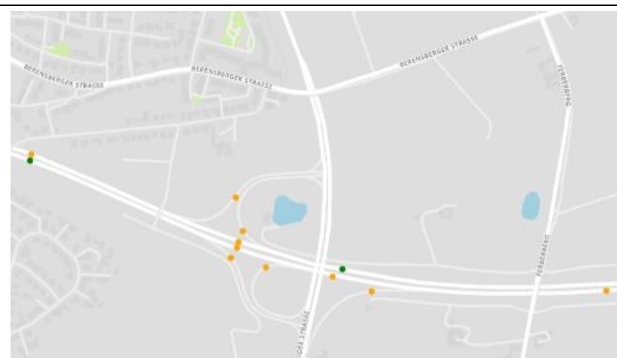
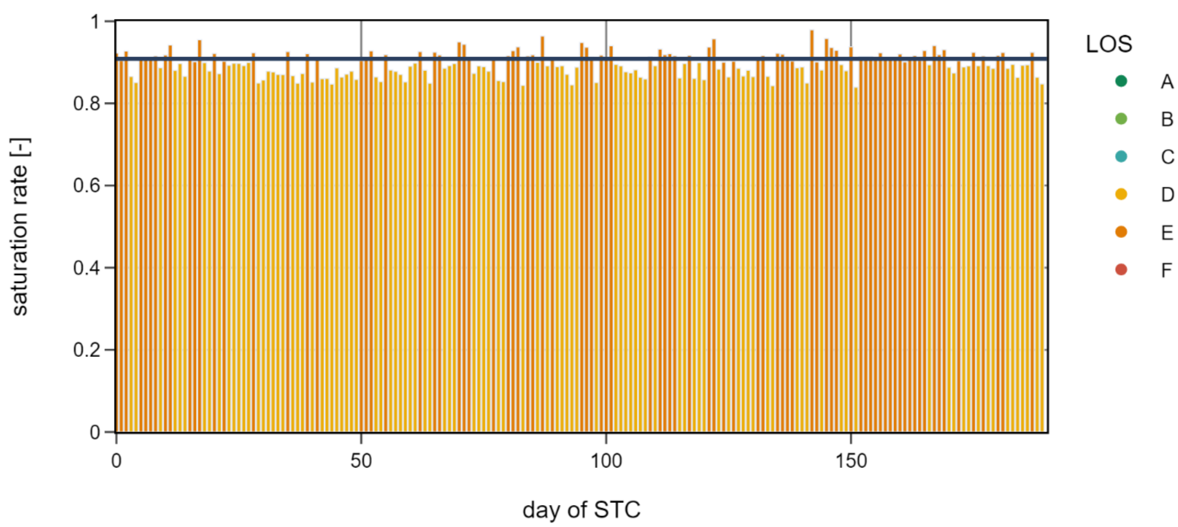
Data availability scenario	Example four-leg-interchange	Example freeway exit
PTC: all, STC: -		
PTC: in-/outflow main lanes, STC: -		
PTC: in-/outflow main lanes, STC: all		
PTC: inflow main lanes, STC: all		
	● PTC - permanent traffic count station ● STC – short-term traffic count station ● no data	

Figure 4: Data availability scenarios, applied to an interchange and a freeway exit (map: ©Mapbox (*14*), ©OpenStreetMap contributors (*15*)).

EVALUATION

For each ramp junction, the saturation rate is calculated for each DHV estimation scenario. For data availability scenarios with STC every potential count day is simulated. As the resulting saturation rate differs depending on the day of the STC an estimation scenario contains multiple saturation rates. Figure 5 shows exemplary results for one DHV estimation scenario, where each bar shows the result for one potential day of the STC. The saturation rate is represented by the height of the bar, whereas the color of the bar represents the LOS corresponding to the saturation rate. The target saturation rate of the reference scenario (50th highest saturated hour) is marked by the horizontal black line.



Horizontal black line: saturation rate of the reference scenario (50th highest saturated hour).

Figure 5: Exemplary results of one ramp junction for a DHV estimation scenario with several STC in one year.

The correlation between the saturation rate of the simulated STC and that of the reference scenario is shown in Figure 6. For each ramp junction of a node, the results of the simulated STC are plotted against the target saturation rate of the reference scenario (each bar of Figure 5 has a corresponding point in Figure 6). The results of one specific ramp junction are plotted on the same x-coordinate since the reference scenario does not depend on the day of the STC. The colored squares illustrate the corresponding LOS. If a point is located in one of the squares, the saturation rate of the simulated STC results in the same LOS as that of the reference scenario. Otherwise, the simulated LOS differs from the target LOS of the reference scenario.

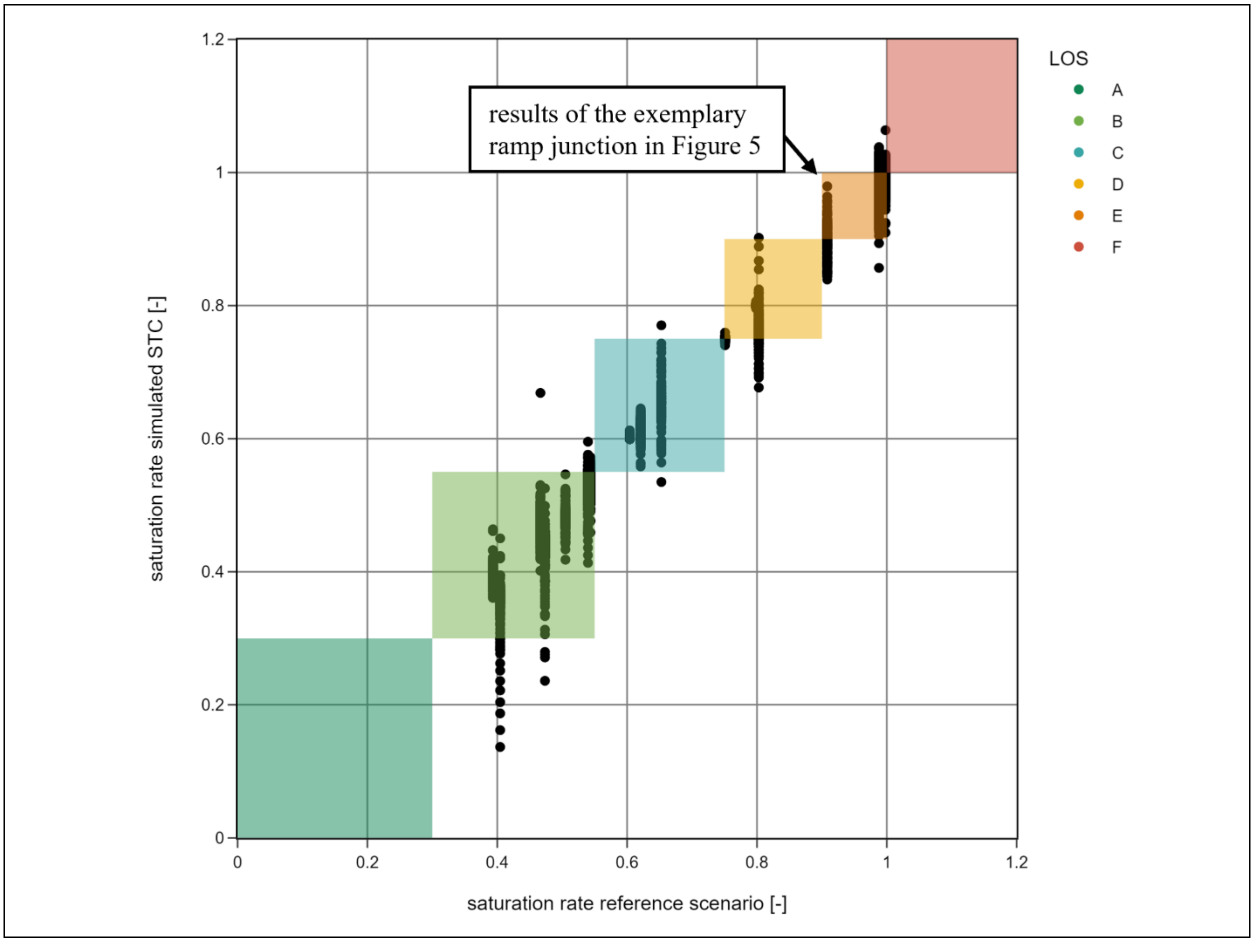


Figure 6: Exemplary results of a node (with 16 ramp junctions) for a DHV estimation scenario for several STC in 2017.

Figure 6 shows the results for one node with 16 ramp junctions. To be able to evaluate the DHV estimation scenarios for all nodes, further aggregation of the results is required. For this reason, the metric of the 'average LOS-accuracy' is introduced. Based on the results of the simulated STC, this metric describes the relative share of the simulated STC, that achieve the target LOS of the reference scenario. Regarding the visualization in Figure 5, this corresponds to the proportion of points located within one of the colored LOS-squares.

$$average\ LOS\text{-accuracy} = \frac{n(LOS_{es} = LOS_{rs})}{n}$$

with

$n = number\ of\ estimation\ scenarios$

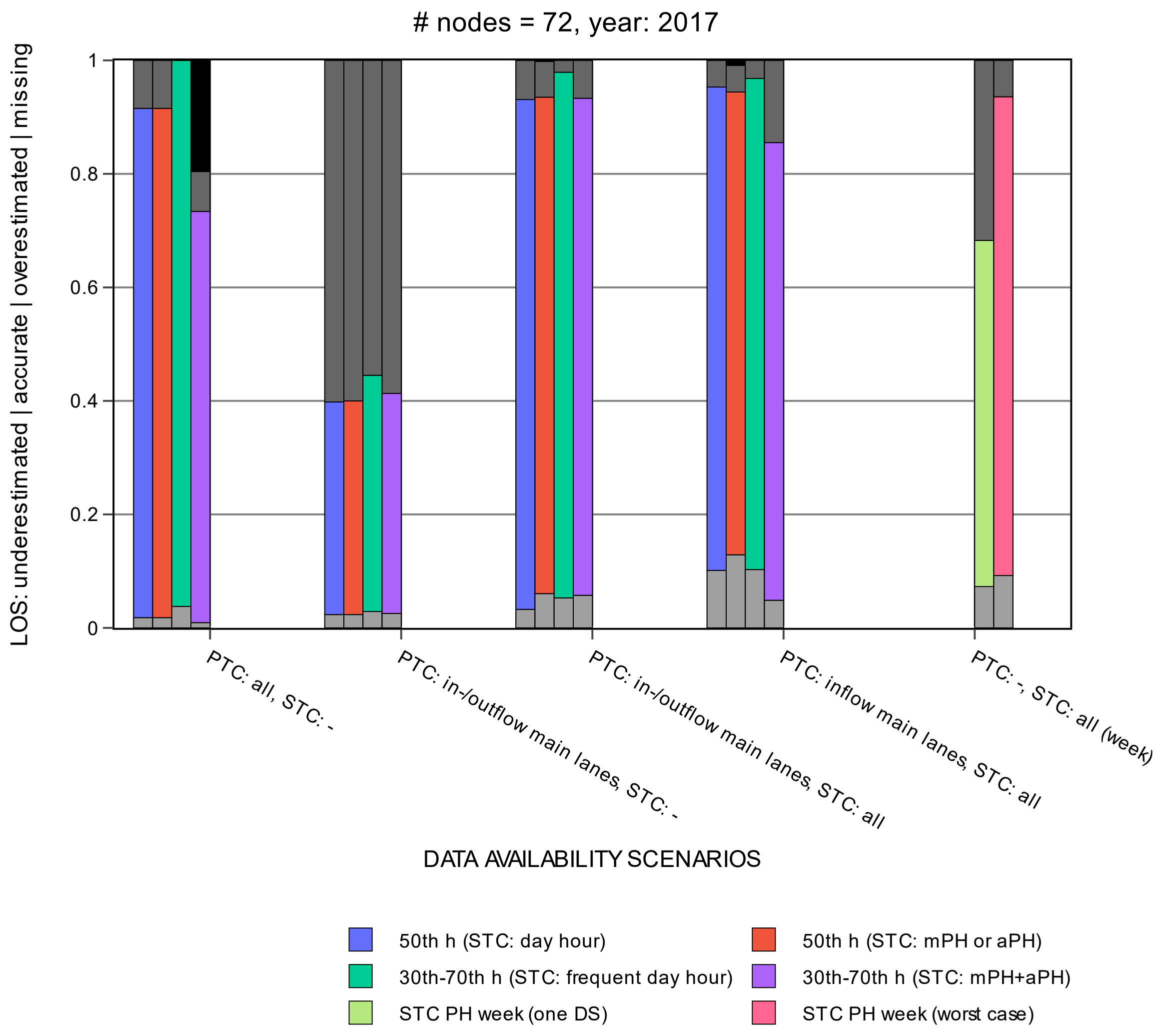
$LOS_{es} = calculated\ LOS\ of\ estimation\ scenario$

$LOS_{rs} = LOS\ of\ corresponding\ reference\ scenario$

$n(LOS_{es} = LOS_{rs}) = number\ of\ estimation\ scenarios\ , which\ hit\ the\ LOS\ of\ their\ reference\ scenario$

RESULTS

The evaluation of the results for the years 2017, 2018, and 2019 shows similar values of LOS-accuracy per year. The following discussion considers the results of all years (Table 2). Figure 7 shows the results for 2017 as an example. We aggregate the average LOS-accuracy for each concept for DHV estimation (colored bars in Figure 7) with respect to the year and data availability scenario (x-axis). Figure 8 shows the results of Figure 7 split into results for interchanges and results for exits.



light gray bar:	share underestimating the LOS of the reference scenario (LOS underestimated)
coloured bar:	share meeting the LOS of the reference scenario (LOS accurate)
dark grey bar:	share overestimating the LOS of the reference scenario (LOS overestimated)
black bar:	share for which the LOS could not be computed (LOS missing)

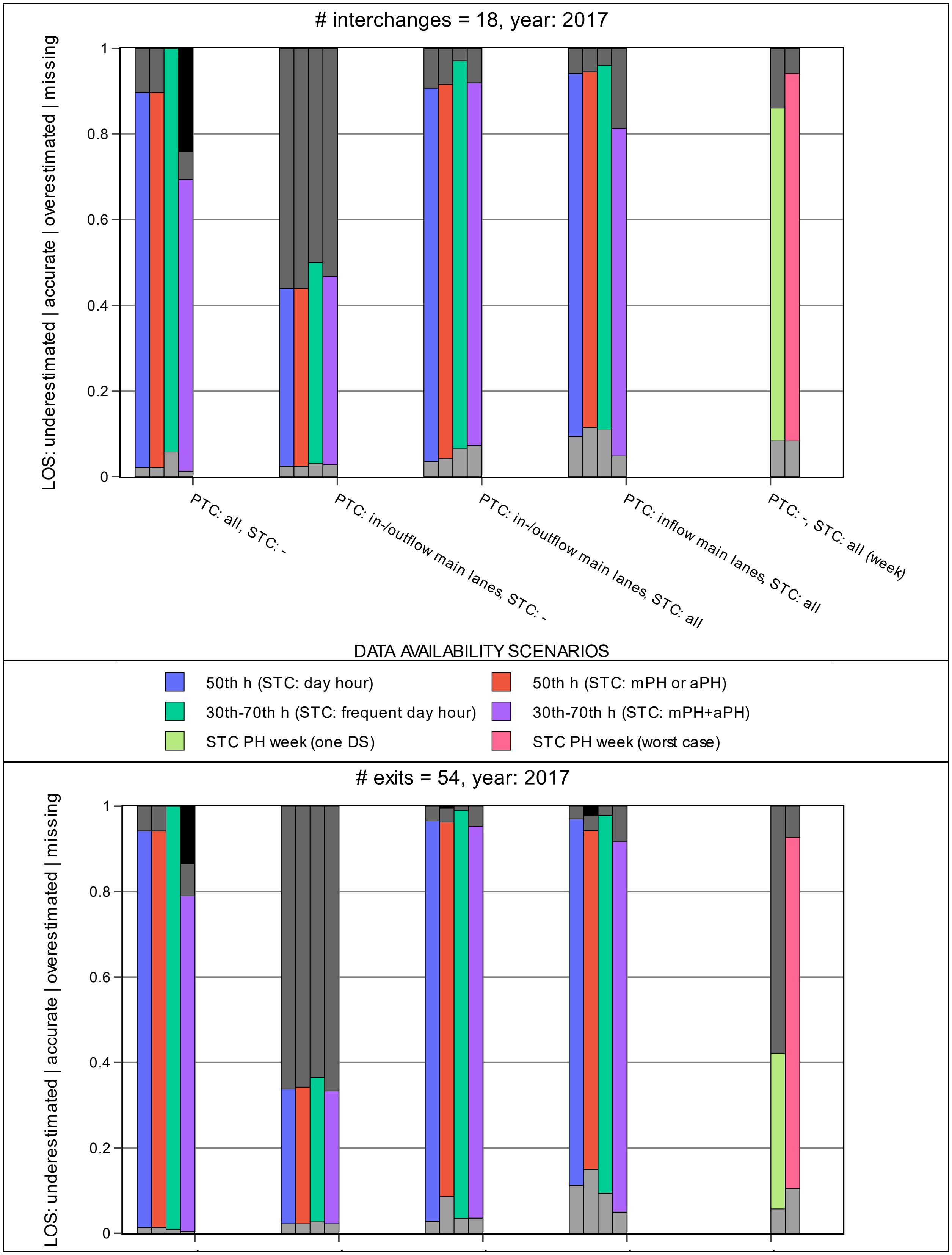
LOS underestimated:	LOS is lower than reference scenario, e.g. C < D
LOS accurate:	LOS is equal to reference scenario (= average LOS-accuracy)
LOS overestimated	LOS is higher than reference scenario, e.g. E > D
LOS missing	LOS could not be computed.

Figure 7: Aggregated evaluation for all estimation scenarios.

1 **Table 2: Aggregated evaluation for all DHV estimation scenarios.**

data availability scenario	concept for DHV estimation	LOS-accuracy			LOS-under-estimation			LOS-over-estimation			Without result		
		2017	2018	2019	2017	2018	2019	2017	2018	2019	2017	2018	2019
PTC: all, STC: -	30th-70th h (STC: frequent day hour)	96%	92%	94%	4%	8%	6%	0%	0%	0%	0%	0%	0%
PTC: all, STC: -	30th-70th h (STC: mPH+aPH)	72%	78%	63%	1%	3%	1%	7%	5%	5%	20%	15%	31%
PTC: all, STC: -	50th h (STC: day hour)	90%	92%	92%	2%	2%	2%	8%	6%	6%	0%	0%	0%
PTC: all, STC: -	50th h (STC: mPH or aPH)	90%	92%	92%	2%	2%	2%	8%	6%	6%	0%	0%	0%
PTC: in-/outflow main lanes, STC: -	30th-70th h (STC: frequent day hour)	42%	42%	41%	3%	4%	3%	55%	55%	55%	0%	0%	0%
PTC: in-/outflow main lanes, STC: -	30th-70th h (STC: mPH+aPH)	39%	39%	38%	3%	4%	3%	59%	57%	59%	0%	0%	0%
PTC: in-/outflow main lanes, STC: -	50th h (STC: day hour)	37%	38%	39%	2%	3%	3%	60%	58%	58%	0%	0%	0%
PTC: in-/outflow main lanes, STC: -	50th h (STC: mPH or aPH)	38%	38%	40%	2%	3%	3%	60%	58%	58%	0%	0%	0%
PTC: in-/outflow main lanes, STC: all	30th-70th h (STC: frequent day hour)	93%	90%	92%	5%	9%	6%	2%	1%	2%	0%	0%	0%
PTC: in-/outflow main lanes, STC: all	30th-70th h (STC: mPH+aPH)	88%	87%	88%	6%	8%	6%	7%	6%	5%	0%	0%	0%
PTC: in-/outflow main lanes, STC: all	50th h (STC: day hour)	90%	89%	90%	3%	6%	5%	7%	5%	6%	0%	0%	0%
PTC: in-/outflow main lanes, STC: all	50th h (STC: mPH or aPH)	87%	87%	87%	6%	8%	9%	6%	5%	5%	0%	0%	0%
PTC: inflow main lanes, STC: all	30th-70th h (STC: frequent day hour)	87%	83%	83%	10%	15%	14%	3%	2%	3%	0%	0%	0%
PTC: inflow main lanes, STC: all	30th-70th h (STC: mPH+aPH)	81%	81%	82%	5%	6%	6%	14%	12%	12%	0%	0%	0%
PTC: inflow main lanes, STC: all	50th h (STC: day hour)	85%	82%	85%	10%	15%	12%	5%	3%	3%	0%	0%	0%
PTC: inflow main lanes, STC: all	50th h (STC: mPH or aPH)	82%	79%	81%	13%	18%	14%	5%	3%	3%	1%	0%	1%
PTC: -, STC: all (week)	STC PH week (one DS)	61%	57%	56%	7%	9%	9%	32%	33%	35%	0%	0%	0%

PTC: -, STC: all (week)	STC PH week (worst case)	84%	81%	82%	9%	12%	12%	6%	6%	6%	0%	0%	0%



1 **Figure 8: Aggregated evaluation for all DHV estimation scenarios by interchanges and exits.**

PTC: all, STC: -

This data availability scenario shows the deviation of a concept for DHV estimation compared to the reference with perfect data availability, i.e. PTC for all count stations. All concepts for DHV estimation perform quite similar with an average LOS-accuracy between 90-96%. An exception is the estimation concept ‘30^{th}-70^{th} h (STC: mPH+aPH)’. The latter cannot compute a LOS for about 20% of the scenarios due to an inconsistent PTC demand-situation. That happens more often on nodes of type interchange than on type freeway exit nodes. An interchange has far more traffic streams and count stations, which lead to more boundary conditions for matrix estimation.
The results for the estimation concepts using the 50^{th} hour as basic concept for DHV estimation, are exactly the same as the data availability scenario does not rely on STC. This is also a reason for the estimation concept ‘30^{th}-70^{th} h (STC: frequent day hour)’ having the highest average LOS-accuracy. Using an average volume of multiple highly loaded hours is more robust compared to 50^{th} highest saturation than just using the 50^{th} hours as DHV. Especially the number of scenarios that overestimated the LOS are reduced using the estimation concept ‘30^{th}-70^{th} h (STC: frequent day hour)’.

PTC: in-/outflow main lanes, STC: -

The reduction of available PTC in this data availability scenario leads to drastically decreasing average LOS-accuracies for all concepts for DHV estimation. Most scenarios (about 60%) overestimate the LOS compared to the references, even more for nodes of type freeway exit. Without additional STC, there is no information about traffic stream distribution on ramps. Interchanges deal better with this availability scenario because there are PTC on each inflow or outflow. For freeway exits, the traffic stream on ramps is overestimated.

PTC: in-/outflow main lanes, STC: all

In this data availability scenario STC provide the missing information on the distribution of flows on the ramps. This leads to an improvement of the average LOS-accuracy up to 93%, which is near the result of the first data availability scenario ‘PTC: all, STC:-‘. This approach works for interchanges and freeway exits, with exits having minor advantages due to their less complex topology. The estimation concept ‘50^{th} h (STC: day hour)’, which derives the STC demand-situation based on the DHV of the PTC, has the fewest cases underestimating LOS compared to the reference. In contrast the estimation concept ‘30^{th}-70^{th} h (STC: frequent day hour)’ shows the lowest share of overestimates. In this data availability scenario, there is little difference between the concepts for DHV estimation considering PTC demand-situations for each PTC station and the estimation concept ‘30^{th}-70^{th} h (STC: mPH+aPH)’ that creates artificial demand-situations for the morning and afternoon peak hour. Comparing this data situation to the previous availability scenario ‘PTC: in-/outflow main lanes, STC: -‘, it is reasonable to conclude that additional information about traffic flows provided by STC are crucial for a precise estimation of the LOS reference.

PTC: inflow main lanes, STC: all

Due to the reduction of PTC stations, a decrease of the average LOS-accuracy of about 5% to 9% can be observed for all concepts for DHV estimation. The three concepts that define multiple PTC demand-situations to maintain consistency tend to underestimate the LOS. In contrast, the estimation concept ‘30^{th}-70^{th} h (STC: mPH+aPH)’ tends to overestimate the LOS reference, especially for interchanges.

PTC: - , STC: all (week)

The evaluation of the data availability scenario without PTC but with STC over a whole week shows that considering a demand-situation, including peak hours for all ramp junctions, overestimates the LOS in many cases, especially for freeway exits. The concept 'STC PH week (worst case)', which evaluates STC demand-situations for each ramp junction and considers a worst case per ramp junction, has an average LOS-accuracy of almost 80%. This is similar to the result for the data availability scenario 'PTC: inflow main lanes, STC: all'. This good LOS-accuracy occurs because the idea of a weekly peak hour is quite equivalent to the 50^{th} hour, considering that a year has 52 weeks.

CONCLUSIONS AND FURTHER RESEARCH

For the design of a traffic facility, the German HCM provides a concept for DHV estimation that determines the design hourly volume from the n^{th} hour (usually n=50) of a year. To determine this value exactly, the volumes of all traffic flows of a traffic facility must be recorded for all 8,760 hours of a year. Permanent traffic count stations required for this are usually only available on the main lanes in the freeway network, but not for ramps. For this case, the German HCM recommends determining the unrecorded ramp traffic volumes by means of short-term traffic counts with a subsequent projection using the available permanent traffic counts.

The aim of the study is to develop concepts for DHV estimation at freeway nodes and to compare them with a reference concept based on the n^{th} highest saturated hour of a year. When comparing the concepts, the available databases (permanent traffic count stations, short-term traffic counts) are varied to examine how this affects the accuracy of an estimation.

Using a reference database, combinations of six concepts for DHV estimation and five data availability scenarios are tested. The study provides the following results on the databases:

- The data availability scenario has a much stronger influence on the quality of the results than the choice of the concept for DHV estimation.
- It is sufficient to have one permanent traffic count on each inflowing main lane of a freeway node.
- Short term traffic counts on all ramps of the node are crucial for an accurate estimation of the DHV because this is the only way to collect information for all origin-destination flows at the node.
- The results indicate that the method proposed in the German HCM for determining the DHV at nodes using short-term traffic counts and extrapolating at permanent traffic counts precisely meets the required n^{th} highest saturated hour.

The results for the concepts for DHV estimation can be summarized as follows:

- The estimation concepts provide rather similar results for identical data availability scenarios.
- Estimation concepts defining STC demand-situations with regard to the PTC demand-situation ('50th h (STC: day hour) ' and '30th-70th h (STC: frequent day hour) ') lead to slightly better results.
- Estimation concepts using peak-hours for the STC demand-situation ('50th h (STC: mPH or aPH)' and '30th-70th h (STC: mPH+aPH)') have the disadvantage that the choice of the corresponding peak-hour is dominated by the traffic volumes on the main lanes, since these lanes contribute most to the total traffic volume of the node. The traffic volumes of the ramps are underrepresented concerning their influence on the selection of the

demand-situation, although they usually have a larger influence on the traffic flow quality at a ramp junction.

- Estimation concepts which only consider short-term traffic counts of a week ('STC PH week (worst case)') can provide a good estimate for the n^{th} highest saturated hour, if the counts are carried out in representative weeks (e.g. outside of holidays).

The results of this study are based on a reference database. This reference database is the result of a data processing approach, in which missing data/traffic volumes were complemented by realistic values based on available permanent traffic count data in the vicinity. The concepts presented were tested on different traffic facilities (merging, diverging and weaving segments of freeway exits and interchanges). The concepts and the influence of the availability scenarios (short-term traffic counts and/or permanent traffic counts) were evaluated aggregately for these facilities. In future research, the influence of the concepts and data availability scenarios on the individual facilities should be investigated in more detail since the effects might differ for clusters of similar facilities. The same applies to the temporal aggregation: Figure 5 shows that the estimation accuracy differs depending on the day the short-term traffic counts is simulated. A careful identification of suitable days could further improve the method. In addition, the question arises of how other data sources, such as Floating Car Data, would perform if these were used instead of short-term traffic counts. The procedures used to calculate the LOS in this study and the use of the saturation rate are taken from the German HCM. The traffic facilities evaluated are also located in Germany. In the future, it would be interesting to evaluate the method in other countries in combination with the guidelines applicable there (e.g., the HCM, which determines the LOS of a ramp junction based on traffic density).

ACKNOWLEDGMENTS

This paper was based on research sponsored by the German Federal Ministry for Digital and Transport, represented by the Federal Highway Research Institute.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: M.V. Baumann, M. Friedrich, S. Reichert, M. Schilling, P. Vortisch, V. Wassmuth; literature review: M.V. Baumann, M. Schilling; data processing: M.V. Baumann, M. Schilling; analysis and interpretation of results M.V. Baumann, M. Friedrich, S. Reichert, M. Schilling, P. Vortisch, V. Wassmuth; draft manuscript preparation: M.V. Baumann, M. Friedrich, M. Schilling, P. Vortisch. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. National Academies of Sciences (US). *Highway Capacity Manual*. HCM, Washington, D.C., 2016.
2. FGSV. *Handbuch Für Die Bemessung Von Straßenverkehrsanlagen*. HBS. FGSV-Verl., 299 B, 2015.
3. United States. Federal Highway Administration. *Traffic Monitoring Guide* FHWA-PL-01-021, United States. Federal Highway Administration. Office of Highway Policy Information, 01.01.2001. https://rosap.ntl.bts.gov/view/dot/41607.
4. Gecchele, G., R. Rossi, M. Gastaldi, and S. Kikuchi. Advances in Uncertainty Treatment in FHWA Procedure for Estimating Annual Average Daily Traffic Volume. *Transportation Research Record*, Vol. 2308, No. 1, 2012, pp. 148–156.
5. Gastaldi, M., G. Gecchele, and R. Rossi. Estimation of Annual Average Daily Traffic from One-Week Traffic Counts. A Combined ANN-Fuzzy Approach. *Transportation Research Part C: Emerging Technologies*, Vol. 47, 2014, pp. 86–99.
6. Khan, S. M., S. Islam, M. Z. Khan, K. Dey, M. Chowdhury, N. Huynh, and M. Torkjazi. Development of Statewide Annual Average Daily Traffic Estimation Model from Short-Term Counts: A Comparative Study for South Carolina. *Transportation Research Record*, Vol. 2672, No. 43, 2018, pp. 55–64.
7. Bagheri, E., M. Zhong, and J. Christie. Improving AADT Estimation Accuracy of Short-Term Traffic Counts Using Pattern Matching and Bayesian Statistics. *Journal of Transportation Engineering*, Vol. 141, No. 6, 2015, A4014001.
8. Ha, J. A. A Comparison of Short Duration Traffic Counts Methodologies for Estimating Annual Average Daily Traffic. *Machine Design and Manufacturing Engineering III*, Vol. 607, 2014, pp. 657–663.
9. Sharma, S. C., Z. Luo, and G. X. Liu. Short-Period Counts with a Focus on Truck Traffic Estimation. *ITE Journal*, November, 2002, pp. 42–45.
10. Figliozzi, M., P. Johnson, C. Monsere, and K. Nordback. Methodology to Characterize Ideal Short-Term Counting Conditions and Improve AADT Estimation Accuracy Using a Regression-Based Correcting Function. *Journal of Transportation Engineering*, Vol. 140, No. 5, 2014.
11. Capparuccini, D. M., A. Faghri, A. Polus, and R. E. Suarez. Fluctuation and Seasonality of Hourly Traffic and Accuracy of Design Hourly Volume Estimates. *Transportation Research Record*, Vol. 2049, No. 1, 2008, pp. 63–70.
12. Petković, M., V. Tubić, and N. Stepanović. Prediction of Design Hourly Volume on Rural Roads. *Transportation Research Record*, Vol. 2675, No. 3, 2021, pp. 112–121.
13. Liu, Z., and S. Sharma. Predicting Directional Design Hourly Volume from Statutory Holiday Traffic. *Transportation Research Record*, Vol. 1968, No. 1, 2006, pp. 30–39.
14. Mapbox-Contributors. Mapbox. Available from Https://www.Mapbox.Com, 2022.
15. OpenStreetMap-Contributers. OpenStreetMap. Map Data Copyrighted OpenStreetMap Contributors and Available from Https://www.Openstreetmap.Org, 2021.