Metzger, Barbara; Loder, Allister; Kessler, Lisa; Bogenberger, Klaus

Article

Spatio-temporal prediction of freeway congestion patterns using discrete choice methods

EURO Journal on Transportation and Logistics (EJTL)

Provided in Cooperation with:
Association of European Operational Research Societies (EURO), Fribourg

Suggested Citation: Metzger, Barbara; Loder, Allister; Kessler, Lisa; Bogenberger, Klaus (2024) : Spatio-temporal prediction of freeway congestion patterns using discrete choice methods, EURO Journal on Transportation and Logistics (EJTL), ISSN 2192-4384, Elsevier, Amsterdam, Vol. 13, Iss. 1, pp. 1-19,
https://doi.org/10.1016/j.ejtl.2024.100144

This Version is available at:
https://hdl.handle.net/10419/325214

EURO Journal on Transportation and Logistics 13 (2024) 100144

Contents lists available at ScienceDirect

EURO Journal on Transportation and Logistics

journal homepage: www.elsevier.com/locate/ejtl

Spatio-temporal prediction of freeway congestion patterns using discrete choice methods

Barbara Metzger [a,*], Allister Loder [b], Lisa Kessler [a], Klaus Bogenberger [a]

[a] Chair of Traffic Engineering and Control, Technical University of Munich (TUM), Arcisstrasse 21, 80333 Munich, Germany
[b] Professorship of Mobility Policy, TUM School of Social Sciences and Technology, Technical University of Munich (TUM), Arcisstrasse 21, 80333 Munich, Germany

ARTICLE INFO

Keywords:
Traffic state prediction
Mixed logit
Congestion patterns
Freeway traffic

ABSTRACT

Predicting freeway traffic states is, so far, based on predicting speeds or traffic volumes with various methodological approaches ranging from statistical modeling to deep learning. Traffic on freeways, however, follows specific patterns in space–time, such as stop-and-go waves or mega jams. These patterns are informative because they propagate in space–time in different ways, e.g., stop and go waves exhibit a typical propagation that can range far ahead in time. If these patterns and their propagation become predictable, this information can improve and enrich traffic state prediction. In this paper, we use a rich data set of congestion patterns on the A9 freeway in Germany near Munich to develop a mixed logit model to predict the probability and then spatio-temporally map the congestion patterns by analyzing the results. As explanatory variables, we use variables characterizing the layout of the freeway and variables describing the presence of previous congestion patterns. We find that a mixed logit model significantly improves the prediction of congestion patterns compared to the prediction of congestion with the average presence of the patterns at a given location or time.

1. Motivation

Crawling, stop and go, or total stoppage — traffic jams are a phenomenon that occurs in road traffic all over the world. Speed in congestion ranges from slow rolling or stop and go to a complete stoppage in traffic. Congestion lengths vary from short stretches of roads to miles-long lines of vehicles. They can also widely vary in time: some congestion events dissipate in minutes, while others can paralyze traffic for hours.

A crucial element in the study of congestion is the analysis of historical data. In the data sets, a considerable amount of information on each congestion's spatial and temporal position can be found. This information is of great importance as it provides insights and can help to make data-based decisions on how to avoid, resolve, or predict congestion. Using historical data, statements can be made about the probabilities of 'congestion' or 'no congestion'. These statements and predictions about the traffic state help to increase traffic safety on the roads in general and the freeways in particular. Predicting traffic conditions is especially important for freeways because they have a high traffic density and are considered major arterials. On freeways, congestion and traffic delays can have a significant impact, not only because of delays, but they also increase the risk of accidents due to abrupt braking and the resulting rear-end collisions. We want to provide information on the probability of congestion, especially for freeway operators because it affects many vehicles or people in this context. The speed can be very high, especially without a general speed limit, and adequate prediction methods are not yet integrated.

With the information from our model, the freeway operators can adjust and prepare the traffic information systems and traffic management systems for the traffic for the following day or the following hours/minutes. According to Li et al. (2022), accurate and dependable short-term traffic forecasting holds a crucial significance in numerous key applications within the field of traffic and transportation. By providing trustworthy predictions of traffic quantities, short-term traffic forecasting enables traffic managers to promptly respond and make informed decisions to prevent congestion. Predicting congestion patterns and, thus, the duration, size, and impact of a congestion event contributes to improving traffic safety.

We focus on four congestion patterns mentioned by Karl et al. (2019). These are *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam*, whereas they range from a short speed breakdown to more distinctive congestion in time and space. The differentiation is conducted

* Corresponding author.
E-mail address: barbara.metzger@tum.de (B. Metzger).

https://doi.org/10.1016/j.ejtl.2024.100144
Received 14 March 2023; Received in revised form 10 September 2024; Accepted 13 September 2024
Available online 18 September 2024

via virtual trajectories to simulate driving vehicles through a congested space–time domain. We refer to the original paper for a more detailed description. Customized actions can be developed against individual congestion patterns to reduce the effects of congestion or keep congestion small. No method still predicts the traffic state, which is simultaneously simple, fast, and comprehensible for traffic managers.

The prediction of speeds and, thus, the prediction of travel times or travel time losses for entire routes are necessary for routing and route selection. For the rudimentary estimation of the danger of "traffic jam", the prediction of the exact speed is not necessary. Defining and predicting traffic conditions using specific thresholds rather than forecasting directly observed traffic variables has the advantage that the resulting condition classifications – such as *Jam Wave*, *Stop and Go*, *Wide Jam*, and *Mega Jam* – provide more meaningful information about traffic problems. As a result, more targeted measures can be initiated. These categorizations may be easier for discrete models or machine learning systems to predict because, unlike traffic variables such as speed and density, which are often volatile, they represent clearly defined and measurable criteria that can be captured more reliably in real-time. In the best case, the traffic flow can be optimized in advance by the model developed here, and a traffic jam does not occur at all or does not evolve to the same extent without prior knowledge and optimized traffic management.

This paper uses a discrete choice model: the multinomial logit model, to address the predictability of congestion patterns on freeways. This approach already existed in a comparable version in a 1998 paper. Still, it was not pursued further, and now we want to contribute and apply the latest computing technology to this methodology (Cottrell, 1998). By forecasting congestion events, freeway operators can control and interpret traffic optimally. When developing helpful tools for freeway operators, they must be easy to understand and logically interpretable. Also, the results are comprehensible and provide reliable and fast guidance. The proposed model is a statistical model that evaluates the historical traffic data of the freeway segment. Fig. 1 sketches the methodology applied in the paper: (1) Data preparation and processing; (2) Division of the study area into cells; (3) Spatio-temporal superimposition of historical data; (4) Prediction of the probability of a congestion pattern in a space–time cell.

In Kessler et al. (2020), an approach is described to identify congestion hot spots for these individual congestion patterns on a freeway stretch. The proposed algorithm first isolates coherent congested clusters from a spatio-temporally discretized speed matrix and then assigns one of the four congestion patterns to each cluster. Considering the spatial and temporal start and end points of each cluster, along with its assigned congestion pattern, accumulated occurrences of congestion can be determined. Regarding this analysis, the question arises about how the individual congestion events relate to each other spatially and temporally. The hot spot analysis of Kessler et al. (2020) showed that clusters of individual congestion patterns over time and road segments exist but lacks the question of whether individual congestion events are also predictable by historical data. Many authors have tried to predict various variables in traffic using many machine learning and artificial intelligence methods. However, we want to develop an easy-to-interpret model suitable for public authorities.

This paper is structured as follows. First, we describe the state of the art of traffic prediction models and the usage of statistical regression models in this field of research. Section 3 presents the data used for this study and explains four congestion patterns proposed in Karl et al. (2019). Thereafter, the prediction model is described in detail. Section 5 contains the model implementation results and the application of the methodology proposed in Molloy et al. (2021) to the data derived from the German autobahn A9. The last section gives a discussion, a summary, and an outlook on future research.

2. State of the art

This section contains the current state of research on different prediction capabilities of congestion, congestion patterns, and currently used statistical prediction models in this field of study.

2.1. Congestion prediction

In the literature, traffic prediction can be done using different techniques. These techniques can be grouped into several categories: statistical models such as Yildirimoglu and Geroliminis (2013), tree modeling such as Zhao et al. (2009), intelligence techniques such as Mahmuda et al. (2021) and Dia (2001), and mixed modeling.

The critical challenges in predicting traffic, particularly the importance of observability and uncertainty, have been identified in recent research (Li et al., 2022). It was noted that perfect observability, where all state variables can be fully reconstructed from available measurements, is a theoretical but not practical condition for perfect predictability of traffic systems. While strict determinism and perfect observability theoretically guarantee predictability, both conditions are difficult to meet in transportation networks due to unobservable variables such as demand, route choice patterns, and individual behavior. As a result, predicting traffic in large networks remains challenging, as highlighted by Li et al. (2022), and requires the development of interpretable models with traceable parameters.

Kerner's traffic flow theory (Kerner and Rehborn, 1996; Kerner, 2001, 2004; Kerner et al., 2004; Kerner, 2009; Palmer et al., 2011) describes the phenomenon of traffic congestion in detail. This theory distinguishes three phases of traffic: free-flow, synchronized flow, and wide moving jam. However, it should be emphasized that our study does not only divide congested traffic into two phases but we analyze four spatio-temporal patterns.

Many existing approaches deal with the prediction of travel times for specific routes. van Lint et al. (2002) presents a recurrent neural network approach for freeway travel time prediction. The proposed recurrent neural network addresses these limitations by implicitly capturing spatio-temporal relationships derived from a state-space formulation of the travel time prediction problem. van Lint (2006) showed that by using an ensemble of State-Space-Neural-Network (SSNN) models, a measure of the reliability of each prediction can also be generated.

A method combining stationary detector data and probe vehicle data to predict freeway congestion fronts was presented, as highlighted by Rempe et al. (2017). In Rempe and Bogenberger (2019), a forecast algorithm was applied to urban road networks with farther links taken into account. A clustering algorithm was used to analyze the level of congestion within clusters, which was integrated into a K-nearest neighbors travel time prediction algorithm. Rempe et al. (2021) proposed a physically informed deep learning method to estimate traffic conditions at locations without detection. The authors found that this method improved the estimation and understanding of traffic density in real freeway traffic data through data fusion.

Two methods for predicting the breakup point of traffic congestion were presented by Lee et al. (2014). The first method uses a model that records spatial and temporal changes in congestion on road networks with multiple junctions. In contrast, the second method uses an algorithm that finds similar historical congestion patterns and calculates drainage timing from these templates. Weather data were analyzed and used for congestion prediction in Lee et al. (2015), where big data processing technology and multiple linear regression analysis explored the relationship between weather and traffic congestion.

CPM-ConvLSTM, a novel deep learning-based model for traffic congestion prediction, was presented by Di et al. (2019). The model leverages the observed congestion patterns to predict the congestion level of road segments by using a spatial matrix to incorporate both the congestion patterns and the spatial relationships between road segments.

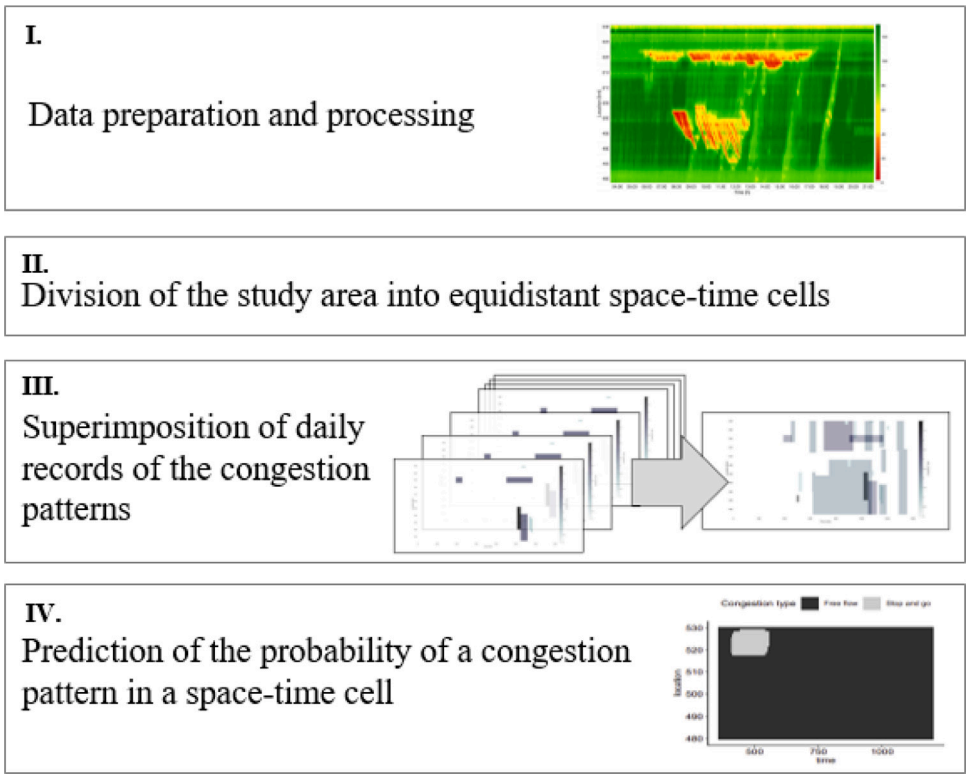


Fig. 1. Methodology described in the paper.

The Congestion-based Traffic Prediction Model (CTPM), which improves predictions using congestion propagation patterns, was introduced, as demonstrated by Nagy and Simon (2021). Performance studies show a significant improvement in forecasts. The new model PCNN described in Chen et al. (2018) uses a deep convolutional neural network to analyze periodic traffic data and make short-term congestion predictions, incorporating techniques such as time series convolution and multi-grained learning to capture local time dependencies and multi-scale traffic patterns. Experiments with real traffic data show that PCNN performs significantly better than comparable methods.

The estimation and prediction of origin–destination (OD) flow is also a crucial issue and a possible solution in the fields of dynamic traffic management and traffic estimation and prediction systems, according to Antoniou et al. (2006). Advances in traffic data collection technologies have provided an abundance of untapped data that can be utilized in OD estimation and prediction. This study introduces a flexible and general methodology to estimate and predict OD flow, incorporating information from various conventional and innovative traffic data sources, such as automatic vehicle identification systems and probe vehicles.

Many of the methods mentioned specialize in forecasting travel times but not traffic conditions. The expected travel time is a statement for the individual driver and should be available before the start of a trip. Likewise, a constant update of the travel time prediction for the selected or possible alternative routes is handy. The prediction of the traffic condition and, if applicable, the size of congestion is mainly important for the freeway operators. The individual road user benefits from the generally safer and smoother traffic flow. In the literature, this has not yet been applied to freeways with a comprehensible modeling approach to the best of our knowledge.

2.2. Multinomial logit estimation

Logit models are frequently used for predictions in the transport sector, mostly for predicting accidents and their severities or the mode choice of road users.

Mixed logit models were developed to analyze driver injury severities and the differences between single-vehicle and multi-vehicle crashes on rural two-lane highways in New Mexico over two years, as demonstrated by Wu et al. (2014) and Dong et al. (2018). A series of significant contributing factors are considered, including driver behavior, weather conditions, environmental characteristics, roadway geometric features, and traffic compositions. The research findings indicate significant differences in the causal attributes determining driver injury severities between these two types of crashes.

The tendency of drivers to engage in crash avoidance maneuvers based on certain circumstances and characteristics of the accident is the focus of research by Kaplan and Prato (2012). The analysis is conducted utilizing a mixed logit model that represents the selection among five emergency lateral and speed control maneuvers. A methodology that addresses a driver's decision of a damaged car following a traffic accident is presented in Hamed and Al-Eideh (2020). The choice set includes three alternatives. A random parameter (mixed) logit model with heterogeneity in the means is specified and estimated to gain more insight into the driver's decision-making process following a traffic accident.

In Li et al. (2010), the discrete choice model is also used to predict the duration of an incident. There, a multinomial logit model is constructed to predict the duration of an incident. 62,941 recorded incidents of the Beijing Transportation Management Bureau were used for the development, and another 10,000 records for the validation. The average relative error of the model is 27.3%.

Route choice models were used for estimation approaches to obtain mode-specific values of travel time savings, based on data from Zurich, as demonstrated by Schmid et al. (2021). They use the estimates of the value of leisure and the values of benefits derived from the conditions experienced while traveling. Combining these two values at their individual level allows for a detailed analysis of the value of time assigned to travel distributions. The same code is used from Schatzmann and Axhausen (2021) for the study for the examination of the substitution effects of long-distance buses in Switzerland (50 km+) and how they

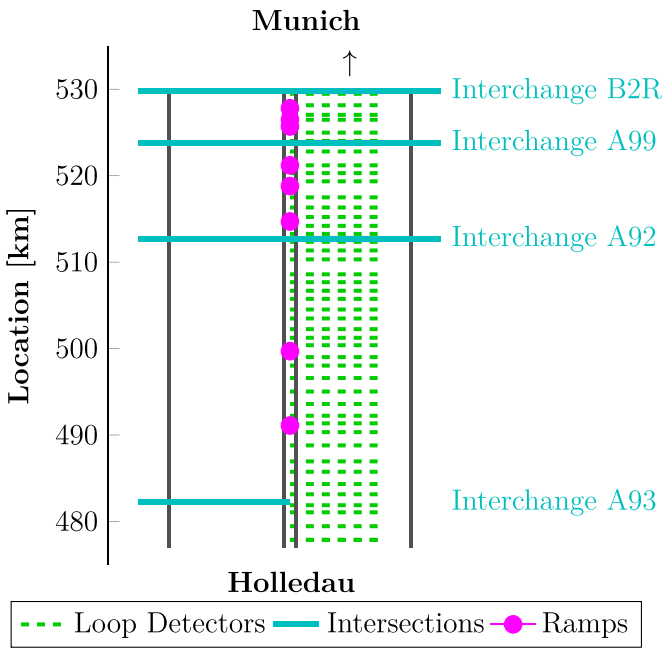


Fig. 2. Sketch of considered road stretch: stationary detectors (dashed green), interchanges (cyan), ramps (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

will affect the mode choice of trains and cars. They estimate standard Multinomial Logit and Mixed Multinomial Logit models to account for unobserved heterogeneity in cost and travel time sensitivities and also incorporate typical social-demographic variables.

2.3. Congestion prediction using mixed logit models

A discrete choice model that predicts recurrent congestion on freeways was developed, as described by Cottrell (1998). This paper from the year 1998 presents a model that estimates the probability of recurrent congestion at existing and potential freeway bottlenecks. The objectives of the model were to accept data that are commonly collected, evaluate existing conditions, predict future conditions, and evaluate the effects of mitigation measures. A logistic regression equation was developed from 163 observations. The model correctly classifies 83% of the data locations used in its development but should be tested on additional data. To the best of our knowledge, there is no other paper in this field of research, and the named model was not further pursued. We consider the potential of modeling with discrete choice methods to reach a very high level today, as there is now a higher and denser data availability and also the computational capacities have grown enormously.

3. Data

This section introduces the road stretch, presents all data used in the study and describes the data processing.

3.1. Freeway data

The proposed methodology is applied to data collected on the freeway A9 in Bavaria, Germany. As a test site, the freeway stretch from Holledau to Munich is chosen with a stretch length of 50 km. Holledau is located at km 480, Munich at km 530 (Fig. 2).

Speed data from inductive loops are available from 44 sensors with an average spacing of 1.2 km, which gathers speed data minute-by-minute. Data were recorded in eight months in 2019.

Table 1
ASM parameter values used for smoothing local speed data.

Parameter	Value
Spatial grid distance	500 m
Temporal grid distance	1 min
Speed in congestion	−18 km/h
Free-flow speed	80 km/h
Crossover from free to congested traffic	80 km/h
Width of the transition region	10 km/h

3.2. Data interpolation

All speed data are interpolated using the *Adaptive Smoothing Method (ASM)*, introduced by Treiber and Helbing (2002, 2003), Treiber et al. (2011) and Schreiter et al. (2010). Briefly summarized, raw data of a sparse input source are smoothed in two traffic-characteristic directions: v_{cong} denominating the wave speed in congested traffic conditions, and v_{free} denominating the wave speed in free-flow conditions. In a discrete time–space domain, the resulting complete speed matrices $V_{cong}(t,x)$ and $V_{free}(t,x)$ are combined cell-wise:

$$V_{ASM}(t,x) = w(t,x)V_{cong}(t,x) + (1 - w(t,x))V_{free}(t,x). \tag{1}$$

The weight $w(t,x)$ is adaptive and favors low speeds:

$$w(t,x) = \frac{1}{2}\left(1 + \tanh\left(\frac{V_{thr} - \min(V_{cong}(t,x), V_{free}(t,x))}{\Delta V}\right)\right) \tag{2}$$

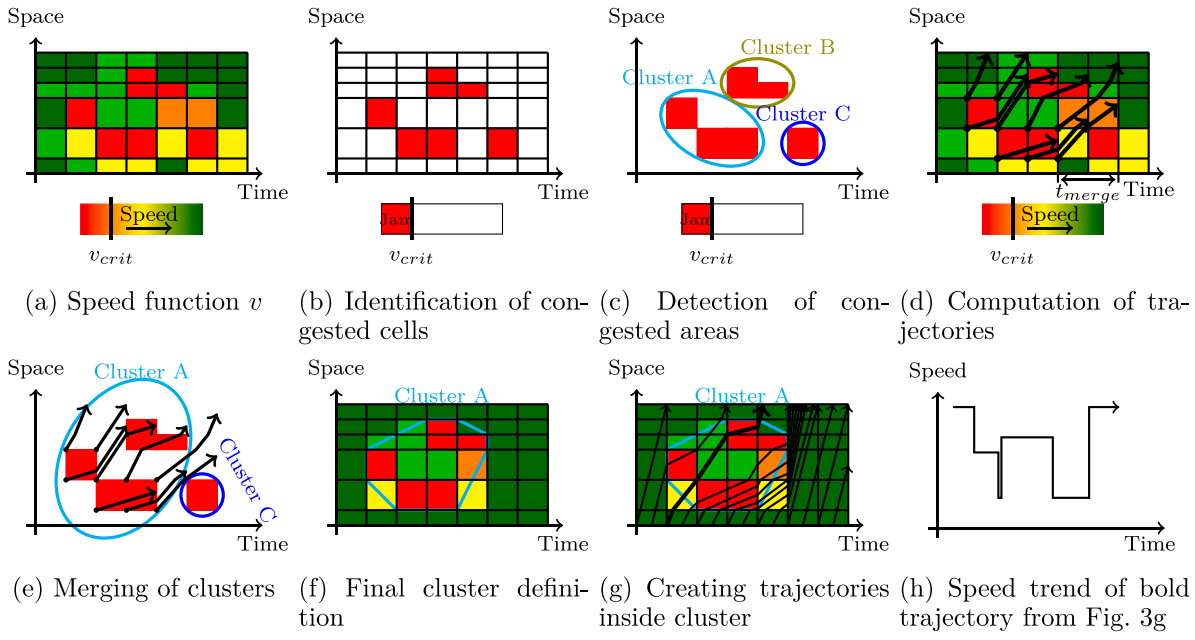
with V_{thr} a threshold where weight $w(t,x)$ equals to 0.5 and ΔV a parameter to control the steepness of the weight function.

As an example, the resulting interpolated speed distribution from Jul 13, 2019, is illustrated in Fig. 4. The x-axis shows the time running from early morning to late evening. The y-axis shows the locations in the direction of travel (from bottom to top). Each cell of the space–time domain is colored depending on the speed, which is indicated in the color bar on the right. To derive such a figure, the ASM parameter values given in Table 1 are applied. They refer only to the preparation – especially the interpolation – of the speed data out of the stationary detectors for the congestion pattern detection.

3.3. Congestion classification

The congestion classification introduced by Kessler et al. (2020) is then applied to the data set. The algorithm detects individual congestion elements based on the algorithm (Kessler et al., 2018) and assigns them to one of the congestion patterns defined by Karl et al. (2019). The schematic workflow of the methodology is sketched in Fig. 3. The parameter values of the congestion pattern definition can be found in Table 2. Starting from a discretized speed distribution, congested cells (speed below a threshold v_{crit}) are identified. A cluster is a set of neighbored congested cells. Using virtual trajectories, clusters that are located close to each other are merged. Close to each other means that the travel time in uncongested conditions between two clusters is not larger than a threshold t_{merge}. This way, the final congestion clusters are determined. In order to assign a congestion pattern to each cluster, the convex hull (to make the cluster unique) is embedded in free-flow conditions, again crossed by virtual trajectories, and the speed profile of each single trajectory traversing the congestion cluster is analyzed.

Depending on the frequency and the duration of each trajectory in congested conditions, one of the four congestion patterns is assigned to the cluster. The sensitivity of the methodology has been assessed in Karl et al. (2019). It is a robust algorithm that can identify and classify occurring congestion even beyond different data sources (Kessler and Bogenberger, 2023). It is also conceivable that the detection and classification work very well on the fused data. This is relevant for sections of freeways where more than one data source is available. A possible approach for the fusion would be Kessler et al. (2021). We refer



(a) Speed function v (b) Identification of congested cells (c) Detection of congested areas (d) Computation of trajectories

(e) Merging of clusters (f) Final cluster definition (g) Creating trajectories inside cluster (h) Speed trend of bold trajectory from Fig. 3g

Fig. 3. Computation of congestion clusters and classification of congestion type (Kessler et al., 2020, 2018; Karl et al., 2019). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Parameter values to identify and classify congestion clusters according to Karl et al. (2019).

Parameter	Value
Velocity threshold v_{crit}	40 km/h
Free-flow speed v_{freeflow}	120 km/h
Minimum free-flow time between congested areas t_{merge}	4 min
Minimum size of congested areas A_{min}	12 km min
Maximum duration of *Jam Wave* $t_{JamWave}$	3 min
Minimum duration of *Mega Jam* $t_{MegaJam}$	30 min
Minimum number of speed drops $n_{StopandGo}$	2
Temporal offset of starting trajectories t_r	5 min
Minimum share of congestion patterns (2 patterns) $n_{2\text{patterns}}$	0.51
Minimum share of congestion patterns (3 patterns) $n_{3\text{patterns}}$	0.41

to the original papers (Kessler et al., 2018; Karl et al., 2019; Kessler et al., 2020; Kessler, 2021) for further details. Applying this congestion classification to the speed distribution in Fig. 4, the result is depicted in Fig. 5.

In the literature mentioned above, the origin and justification of the four congestion patterns are explained in great detail. For the prediction with statistical models, mainly *recurrent* congestion is relevant. However, the data analysis shows that each of the four defined congestion patterns occurs more frequently in location and time. A list containing all identified congestion patterns and these interpolated speed distributions of the considered eight months in 2019 over the 50 km stretch forms the basis for further investigations.

3.4. Data processing

We divide the investigated stretch into equidistant space–time cells and store the known information per cell. The spatial cell size amounts to 500 m, the temporal to one minute. For each of the cells, spatial and temporal information is stored in addition to the congestion pattern and speed. Spatial information is whether the segment comprises a ramp (exit or parking lot), an interchange, and how far from Munich the cell is located. Temporal information is about the weekday and a binary value on the rush hour. This information is assigned to the individual cells of every investigated day. Thus, it is impossible that several congestion patterns are assigned to one cell since only one congestion pattern can be present at any given date, time, and location. In addition, each cell is assigned the current speed at that point in time.

Table 3
Sample characteristics.

Characteristics	Value	Percentage
Days	238	
Jams	836	
Location-sections (X)	100	
Time-sections (T)	840	
Total cells	20 088 269	1.0000
Free-flow cells	19 199 348	0.9558
Jam Wave cells	28 850	0.0014
Stop and Go cells	548 318	0.0273
Wide Jam cells	224 528	0.0112
Mega Jam cells	87 225	0.0043
Weekday cells	14 397 222	0.7167
Weekend cells	5 691 047	0.2833
Rush hour cells	7 165 836	0.3567
No rush hour cells	12 922 433	0.6433
Cells with a ramp	2 574 952	0.1282
Cells without a ramp	17 513 317	0.8718
Cells with an interchange	1 187 113	0.0591
Cells without an interchange	18 901 156	0.9409

Table 3 gives an overview of the cell properties. The entire data set (50 km and from 6 am to 8 pm) consists of 20, 088, 269 cells. From these, 95.58% are without congestion. 0.14% of the cells are assigned to *Jam Wave*, 2.73% to *Stop and Go*, 1.12% to *Wide Jam*, and 0.43% to *Mega Jam*. Fig. 6 shows the distribution of both the frequency and the location of the clusters of the congestion patterns, respectively. White means very rare or not at all, and black means frequent occurrences.

Regarding temporal characteristics, 72% of all cells are flagged as weekday and 36% as rush hour (defined from 6 to 9 am and from 4 to 7 pm). The proportion of the individual congestion patterns in the total number of cells can be seen in the graphs (a) to (e) in Fig. 6. The percentages of the respective congestion patterns are shown in black. In the graphs, the time is limited from 6 am til 8 pm (the time axis represents the minutes of a day) because congestion is hardly present

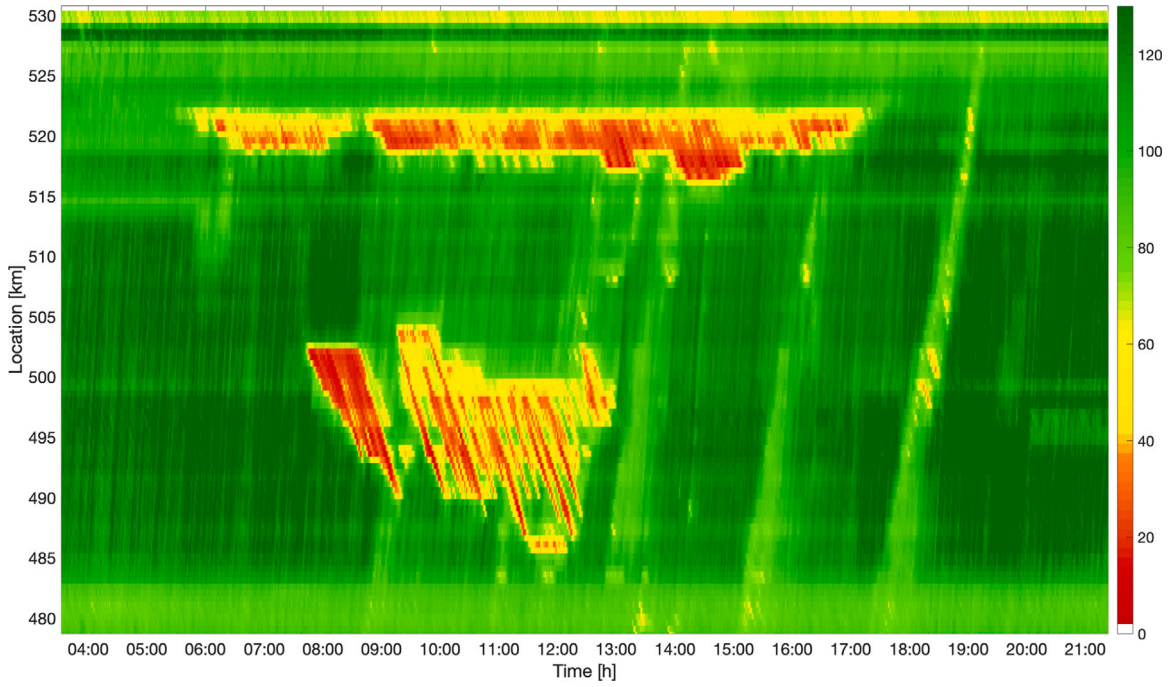


Fig. 4. Interpolated speed distribution. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

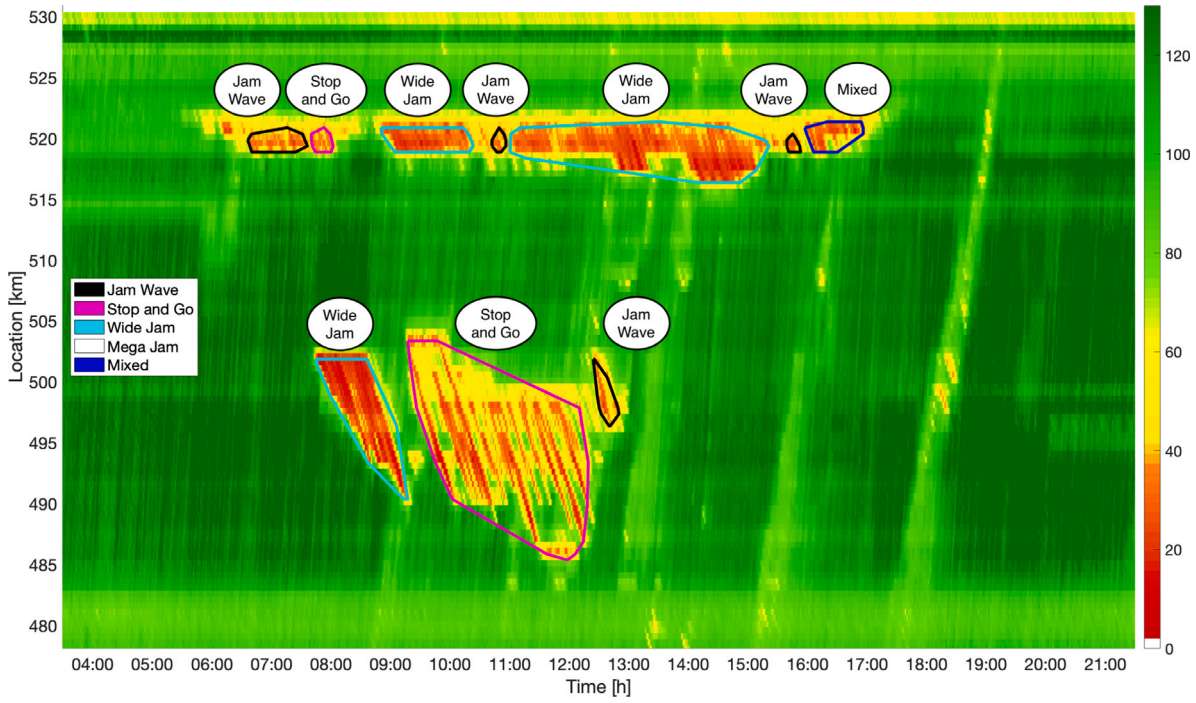


Fig. 5. Classified congestion clusters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

during night hours. In these graphs, it can also be seen that free flow is predominantly detected.

The spatial criteria conditions can be seen in Fig. 2. There are three interchanges in the interior of the entire stretch, excluding the intersection with B2R at the very end of the freeway. In addition, there are eight ramps and two parking lots along the stretch. The stretch contains several areas with a different number of lanes. It varies from two to four lanes. 13% of the cells have the information that there is a ramp for an exit or a parking lot, and 6% of the cells are flagged with the information that there is an interchange in the area.

Initial probabilities can already be drawn from the cells. For example, Fig. 7 shows the probability of Stop and Go congestion on weekdays. This diagram shows a hot spot on weekdays of Stop and Go traffic in the morning hours around km 520–530.

In the following, a mixed logit model is developed that can predict if there is a specific congestion pattern or free-flowing traffic for the

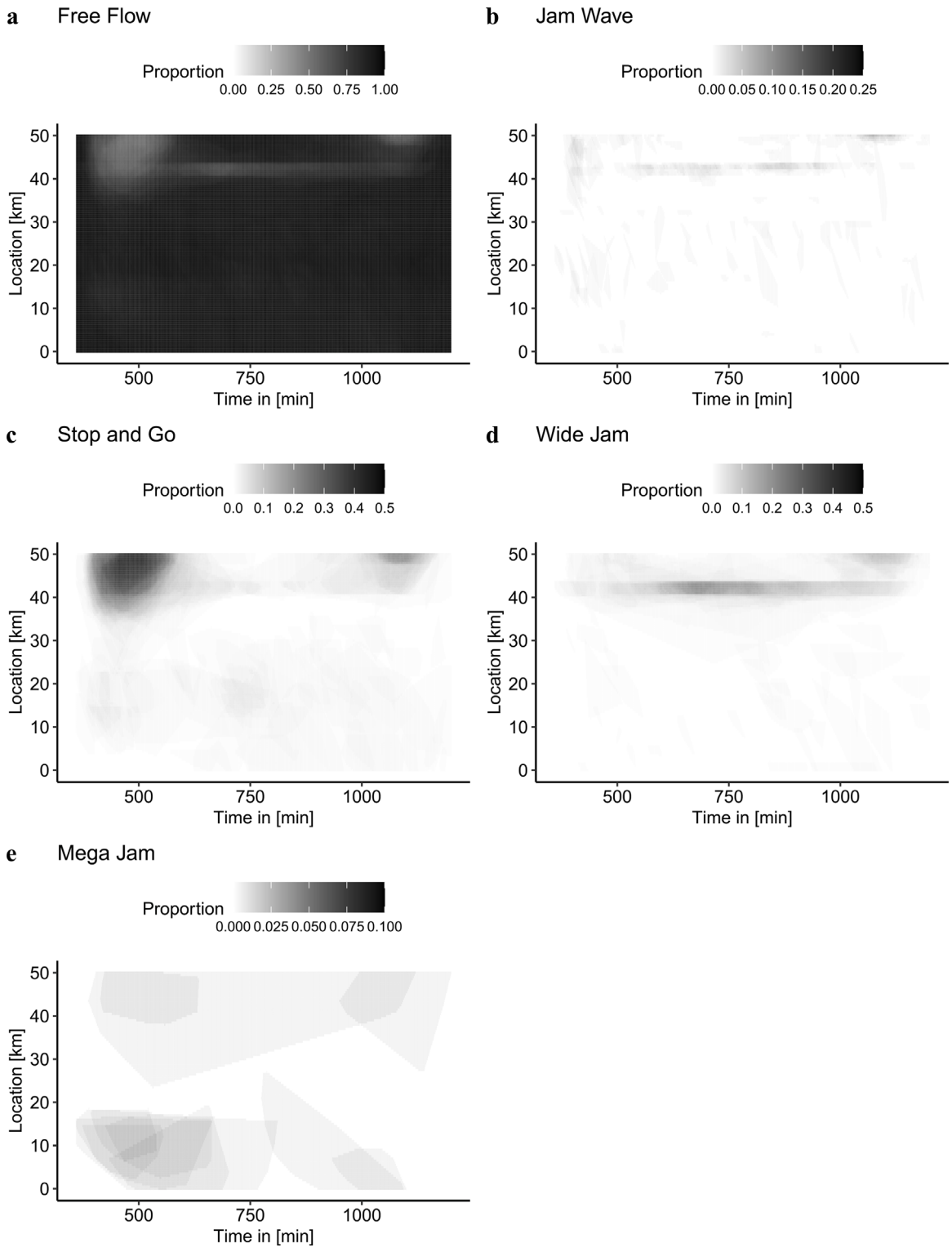


Fig. 6. Spatio-temporal distribution of the jam patterns.

next minutes. The model predicts the congestion pattern while considering both the speed and the congestion patterns in the temporally and spatially preceding cells and the infrastructural parameters. To compare the added value of labeling the prevailing cells with congestion patterns, we set a model that only includes the speed and the infrastructural parameters. The logit model predicts the probability of the presence of one of five traffic patterns: free-flow, Jam Wave, Stop and Go, Wide Jam, or Mega Jam. We select a mixed logit model to model the discrete outcomes or choices and a mixed model to correct for the panel data structure of our model (Train, 2009). In contrast to deep learning methods, explainable parameter values are retrieved that can be used for modeling and decision-making. This also enhances its applicability in traffic operations centers as the model's behavior is more transparent.

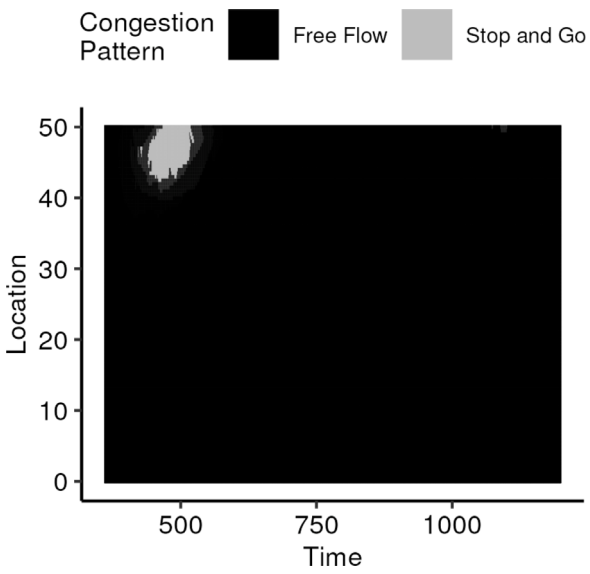


Fig. 7. Most likely weekday congestion pattern.

For a better prediction of the current congestion pattern, the cells that are spatially upstream and temporally ahead of the current cell are considered. The triangular area of the cells, which are zero to five steps in space and in time before the current cell builds a funnel, are considered; see Fig. 8. The properties of these cells in the funnel and the cells on the diagonal at the funnel (marked in bold in Fig. 8) are added to the already given information of each individual cell.

4. Modeling the congestion prediction

This section describes the setup of the prediction model and its estimation. In Section 4.1, two very simple naive forecasting methods from data analysis are used. The results are in Table 7. Section 4.2 shows the structure of two variants of a mixed logit model. The estimation model formulas are presented in the last section of this chapter. We used the package *mixl* in R (Molloy et al., 2021). Predicting congestion patterns using a mixed logit model is not based on random utility theory, and no human utility function is associated with each congestion pattern. Here, we use the model structure of the mixed logit model to predict congestion patterns as they are in the data in an explanatory way. Our result is the latent propensity score of congestion patterns.

4.1. Base model specification

When analyzing the basic strategies for predicting traffic patterns, it becomes clear that simplicity, especially at the beginning of an analysis, often takes precedence over complexity when selecting methods. Initial forecasts benefit from naive methods due to their simplicity, their economy of calculation, and their usually easy possibility of interpretation. There is a variety of naive methods: two of these methods are characterized by their practicality and applicability to our problem: the persistence model, which assumes that the future value reflects the immediate past, and the mean value method, which makes forecasts based on the average of past data.

The persistence model **Base model P** is remarkable as it assumes short-term continuity of traffic flows and is therefore suitable for stable conditions. However, it is not sufficient during the beginning and termination of congestion. This has safety implications and emphasizes the need for accurate predictions at critical moments. In contrast, the averaging approach uses historical frequency to anticipate future conditions and provides a statistical prediction: **Base model A**. The very simple basis variation that analyzes the data used to determine the probability of a traffic condition in a cell. An example of a result of the basic variant for weekdays is shown in Fig. 7.

These naive forecasting models serve as a valuable benchmark. If sophisticated algorithms fail to outperform the baseline established by these methods, their application can be reconsidered. However, the results must also be treated with caution and scrutinized carefully. A simple method can perform well, but it does not necessarily reflect complexity. This illustrates the importance of such elementary prediction techniques in the early stages of traffic pattern analysis.

4.2. Mixed logit model specification

The data set with approx. 20 million cells is used as a basis for the logit model (see Table 3). The data set is divided into a training set of 80% of random days and a test set consisting of the remaining 20% of the days. This way, we get more than six months of training data and almost two months of test data. Then, two variants (models I and II) and a basic investigation of the model are created. The simple basis variation that analyzes the data visualized in Fig. 1 determines the probability of a traffic condition in a cell. An example of a result of the basic variant for weekdays is shown in Fig. 7.

Model I calculates the probability of a traffic condition from historical data and spatial and temporal information. The input data are the historical data of the traffic jams and the parameters of the individual cells: the information on ramps, parking lots, or interchanges, and weekday or rush hour. To estimate the probability of the current traffic pattern, the speed is analyzed and included in the determination of the likelihood of the current cell.

The development of **model II** has the same parameters as model I but with the labeling of the congestion patterns. More precisely, the model I only has the infrastructural information and speed information of the current and the previous funnel cells. In contrast, model II additionally uses the information on the congestion patterns in the funnel cells. This aims to demonstrate the benefit of historical knowledge of the congestion pattern in the previous cells and allows the analysis of the impact of congestion patterns for the prediction. Model II also considers the speed in *cell −3* (see Fig. 8). The speed of this cell allows us to infer the congestion pattern by the different frequencies and values of the speed. *Cell −3* has a value frequently lower than 50 km/h for Stop and Go, whereas the speed for Wide Jam is significantly more often above 75 km/h (see Fig. 9). Additionally to the speed in *cell −3*, the difference of the speed of *cell −5* to *cell −1* is considered in model II. Fig. 10 shows that the speed difference is often zero for Stop and Go as the general speed value is low. For Wide Jam, the speed drop is often more recognizable. It should be noted that the number of observations per congestion pattern is different. The plots in Figs. 9 and 10 are normalized. To include the characteristic spatio-temporal propagation of congestion, we use a co-moving coordinate system in the modeling through the funnel. The explanatory variables are summarized in the following list, and their hypothesis on the effects of prediction is explained:

- **Average speed in funnel** $speed_{funnel}$: This parameter is used for the first sequence of models I and II. We expect that if the average speed in the funnel is lower, the probability of congestion is higher.
- **Location to Munich** $distance_{Munich}$: The nearer the city and thus the metropolitan area, the higher the probability of congestion.
- **Speed cell −3** v_{cell-3}: The speed in *cell −3* of the funnel is shown in detail in Fig. 9. The higher the speed in this cell, the more likely are the congestion patterns Wide Jam or Mega Jam.
- **Rush hour** rh: The cells have the binary information on rush hour. The assumption is that at rush hour times, smaller congestion patterns such as Jam Wave and Stop and Go will occur more likely.

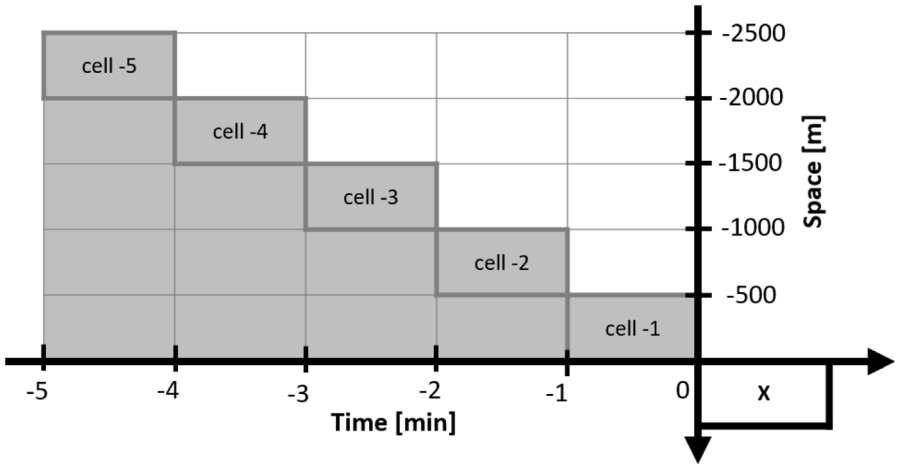


Fig. 8. Look-ahead funnel for a cell x.

- **Weekday** we: Whether it is a weekday and, therefore, a working day, a weekend day, or a public holiday is an important aspect for the probability of emergence of congestion.
- **Section with ramps or parking lot** $weavingsection$: On sections with ramps or entrances to junctions or car parks, the likelihood of congestion may be increased compared to the mainline.
- **Section with interchange** $interchange$: On sections with a freeway junction and many interconnecting processes, the likelihood of minor traffic jams is increased.
- **Speed difference of cell −5 to cell −1** $v_{cell-5to1}$: A larger positive difference from the speed in *cell −5* to *cell −1* and thus a larger speed drop is decisive for a higher probability of a rather sudden congestion object (congestion pattern Stop and Go or Mega Jam). Congestion pattern Jam Wave or Wide Jam are more likely to occur in areas with a generally low-speed level. The distribution of the speed difference is presented in Fig. 10.
- **Proportion of cells with congestion pattern Jam Wave in the funnel** $indic_{JamWave}$: If in the previous cells of the funnel, the proportion of cells with Jam Wave pattern was already high, the probability that this congestion pattern is still present is high.
- **Proportion of cells with congestion pattern Stop and Go in the funnel** $indic_{StopGo}$: If in the previous cells of the funnel the proportion of cells with Stop and Go pattern was already high, the probability that this congestion pattern is still present is high.
- **Proportion of cells with congestion pattern Wide Jam in the funnel** $indic_{WideJam}$: If in the previous cells of the funnel, the proportion of cells with Wide Jam pattern was already high, the probability that this congestion pattern is still present is high.
- **Proportion of cells with congestion pattern Mega Jam in the funnel** $indic_{MegaJam}$: If in the previous cells of the funnel the proportion of cells with Mega Jam pattern was already high, the probability that this congestion pattern is still present is high.
- **Existence of congestion pattern Jam Wave in cell −1** $JamWave_{cell-1}$: The probability of the Jam Wave pattern is higher if the Jam Wave pattern was also present at the location at the previous time interval.
- **Existence of congestion pattern Jam Wave in cell −3** $JamWave_{cell-3}$: On the other hand, the probability of the jam pattern Jam Wave to occur decreases if the jam pattern was already present in *cell −3*. This is because a Jam Wave is more likely to stand alone. If there were several in succession, it would be of the Stop and Go pattern.

We know that not all possibilities of data processing are being used up. Information on demand clusters of the examined days and months can be integrated so that, in addition to normal weekdays, days before public holidays or vacations will also be considered. These influences are clearly noticeable on Bavarian freeways. In addition, exogenous factors influencing traffic, such as the weather, are missing in the current model.

4.3. Model estimation

The model is formulated as a mixed logit model (Train, 2009) as estimated using the high-performance computing package *mixl* in R (Molloy et al., 2021). A mixed logit model is a discrete choice model for several categorical outcomes, such as congestion patterns, where a single decision maker, here a cell in the space–time diagram, is observed several times. In our case, each cell is observed 240 times. In our model, we use error components and no random coefficients. A mixed logit model can be used without a random-coefficients interpretation by representing the correlation among the utilities for the different alternatives by error components,

$$p(y = i|x, \theta) = \frac{e^{U_i}}{\sum_j e^{U_i}} \tag{3}$$

with the context of the following form of a utility function, according to Train (2009):

$$U_i = x'\beta_i + \epsilon_i \tag{4}$$

The model is estimated using simulation techniques (Train, 2009) that approximate the normal distribution of the random intercepts utilizing a sequence of Halton draws (20 draws). The utility functions for each alternative are specified linearly. The starting values are set to 0.1. In the estimation, we experienced issues of co-linearity, as in our observed data, where several categorical attributes are perfectly co-linear and are removed consequently.

We specified five utility functions, one for each congestion pattern, plus free flow. The free-flow utility function is set to zero. The other four functions, for the congestion patterns Jam Wave, Stop and Go, Wide Jam, and Mega Jam, are defined with several parameters. The parameters are presented in an itemized list in Section 4.2. The model is built up in two levels: in the first level, a preliminary analysis is calculated, in which the probability of one of the congestion types is calculated compared to free flow. For this purpose, the average speed in the funnel before the current cell and the distance in kilometers to the city (Munich) are used as parameters:

$$V_J = @C_{Jam} + @\sigma_J \cdot draw_1 + @speed_{funnel} \cdot \$meanspeed_{funnel} + distance_{Munich} \cdot \$x \tag{5}$$

In the second step, the individual utility functions

$$U_1 = 0 \tag{6a}$$

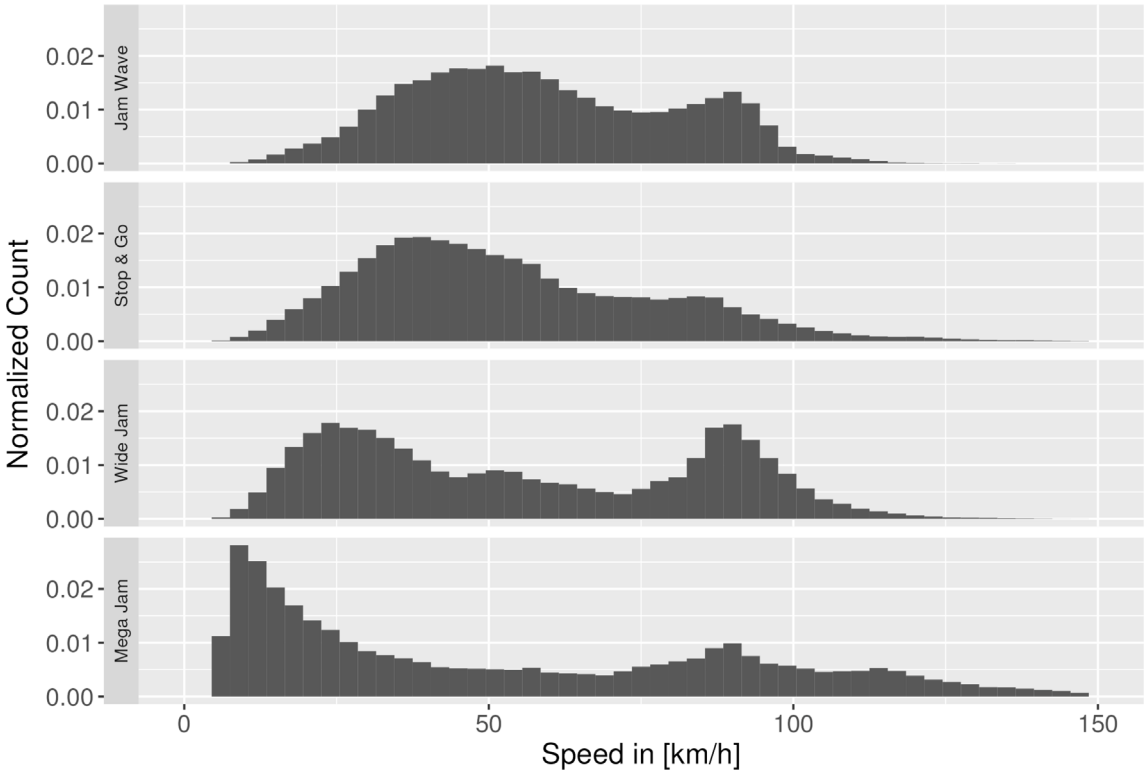


Fig. 9. Normalized histogram of speed distribution in cell −3 (Fig. 8) before current cell with jam pattern.

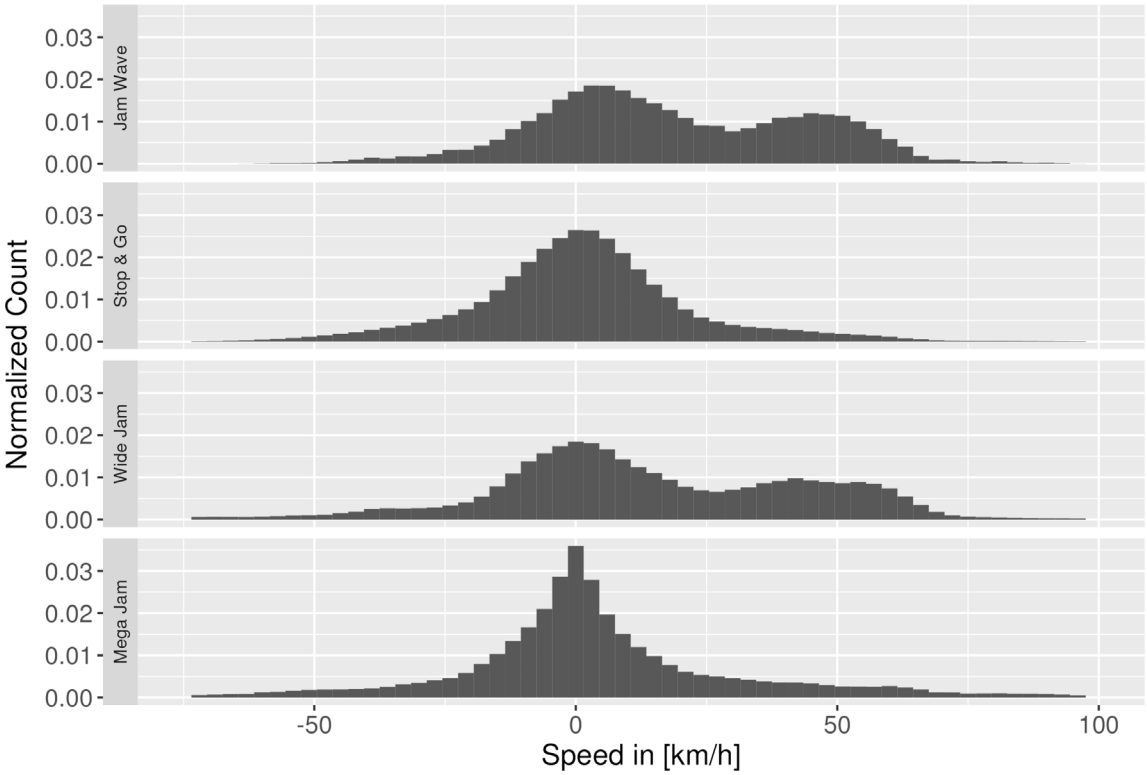


Fig. 10. Normalized histogram of speed difference of cell −5 to cell −1 (Fig. 8) before current cell with jam pattern.

$$U_2 = @C_2 + EC_2 + @speed_{U2-cell-3} \cdot \$v_{cell-3} + @indic_{JamWave} \cdot \$funnelshare_{JamWave} + @rushhour_{JamWave} \cdot \$rh + @weekday_{JamWave} \cdot \$wd + @weavingsection_{JamWave} \cdot ws + @interchange_{JamWave} \cdot \$ic + @JamWave_{cell-1} \cdot \$JamWave_{cell-1} + @JamWave_{cell-3} \cdot \$JamWave_{cell-3} \quad (6b)$$

$$U_3 = @C_3 + EC_3 + @speed_{U3-cell-3} \cdot \$v_{cell-3} + @indic_{StopGo} \cdot \$funnelshare_{StopGo} + @rushhour_{StopGo} \cdot \$rh + @weekday_{StopGo} \cdot \$wd + @weavingsection_{StopGo} \cdot ws + @interchange_{StopGo} \cdot \$ic \quad (6c)$$

$$U_4 = @C_4 + EC_4 + @speed_{U4-cell-3} \cdot \$v_{cell-3} + @indic_{WideJam} \cdot \$funnelshare_{WideJam} + @rushhour_{WideJam} \cdot \$rh + @weekday_{WideJam} \cdot \$wd + @speedWJ_{cell-5to1} \cdot \$v_{cell-5to1} \quad (6d)$$

$$U_5 = @C_5 + EC_5 + @speed_{U5-cell-3} \cdot \$v_{cell-3} + @indic_{MegaJam} \cdot \$funnelshare_{MegaJam} + @rushhour_{MegaJam} \cdot \$rh + @speedMJ_{cell-5to1} \cdot \$v_{cell-5to1} \quad (6e)$$

are defined to predict the probability of the individual congestion patterns. Each alternative has its own utility function, referred to as U_1 to U_5 in this case. Each utility function consists of a constant term $(C_{Jam}, C_2, C_3, C_4, C_5)$, fixed effects marked with the @ sign multiplied by the parameters out of the data set with the $ sign, and random effects (EC_2, EC_3, EC_4, EC_5) that come from a normal distribution and are calculated as follows:

$$EC_2 = @\sigma_{12} \cdot draw_1 + @\sigma_2 \cdot draw_5 \quad (7a)$$

$$EC_3 = @\sigma_{13} \cdot draw_1 + @\sigma_3 \cdot draw_2 \quad (7b)$$

$$EC_4 = @\sigma_{14} \cdot draw_1 + @\sigma_4 \cdot draw_3 \quad (7c)$$

$$EC_5 = @\sigma_{15} \cdot draw_1 + @\sigma_5 \cdot draw_4 \quad (7d)$$

In the combined probability function, the preselection V_J is put ahead by multiplication:

$$P_{freeflow} = \frac{e^{V_J} \cdot \$selChoice}{1 + e^{V_J}} + \frac{1 - \$selChoice}{1 + e^{V_J}} \quad (8)$$

The variable V_J represents the utility of the first alternative and consists of a constant term, a random effect, fixed effects, and a linear function of the characteristics. $P_{freeflow}$ is the probability that the first alternative will be chosen, which is calculated by taking the exponent of the utility of the first alternative and then normalizing it. Based on this probability and the utility functions of the congestion patterns, the probabilities of all five possible outcome functions are calculated.

5. Results

This section examines the prediction of the probability of the congestion patterns by the discrete choice model. It describes a validation to show how well the three proposed models predict the used data set. In addition, we show the application of the model to one congestion pattern as an example.

5.1. Estimation results

The models are based on the congestion patterns of the historical database described in Section 3 in Table 3 and include actual speed. The basis model contains no statistical variables and is evaluated in the next section. The results of the statistical variants model I and model II are shown in Table 4, Table 5, and Table 6.

Table 4
Model estimation results for the two model variants: random effects and error components.

Parameter	Model I	Model II
C Jam indicator C_{Jam}	6.36 (0.02)***	6.07 (0.02)***
C Jam Wave C_2	1.01 (0.03)***	−7.11 (0.12)***
C Stop and Go C_3	2.50 (0.02)***	−8.43 (0.07)***
C Wide Jam C_4	3.45 (0.02)***	−14.22 (0.17)***
C Mega Jam C_5	2.31 (0.02)***	−21.08 (1.70)***
SIGMA Stop and Go σ_3	−0.01 (0.07)	−1.71 (0.04)***
SIGMA link of free-flow and Stop and Go σ_{13}	−1.34 (0.01)***	1.23 (0.05)***
SIGMA jam indicator σ_{Jam}	−0.89 (0.01)***	0.22 (0.01)***
SIGMA Mega Jam σ_5	1.33 (0.01)***	−0.02 (0.09)
SIGMA Wide Jam σ_4	−0.93 (0.01)***	−0.07 (0.05)
SIGMA link of free-flow and Wide Jam σ_{14}	−0.40 (0.01)***	3.45 (0.09)***
SIGMA link of free-flow Mega Jam σ_{15}	−0.39 (0.02)***	5.28 (0.68)***
SIGMA link of free-flow and Jam Wave σ_{12}	−1.05 (0.02)***	1.61 (0.08)***
SIGMA Jam Wave σ_2	0.05 (0.03)	−1.48 (0.08)***

Standard errors in parentheses. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

After finding a model that predicts the occurrence of an event – in this paper, the congestion pattern of Jam Wave, Stop and Go, Wide Jam, Mega Jam, or free-flow – it is appropriate to check the prediction quality or model quality. In principle, for many classification models, such as logistic regression, it is possible to specify a pseudo-R^2. For the logit model, however, there are three popular variants of the pseudo-R^2: McFadden R^2, Cox&Snell R^2, and the Nagelkerke R^2. For their calculation, the likelihood of the zero model LL is used. In our case, we consider the McFadden R^2, which can take a value between 0-very poor and 1-excellent. The McFadden R^2 in Table 6 is 0.93 for model I and 0.99 for model II. These two values are relatively high for a logit model. The reason for this is the immensely high correct prediction of free flow. The speed pre-selection model also improves the quality of the prediction. However, the increase of the value from model I to model II shows that it improves the prediction by including the information on the traffic patterns before the currently considered cell. The improvement in describing the data from model I to model II is also supported by Akaike's Information Criteria (AIC), which is lower in the second variant (Brewer et al., 2016; Groß, 2015).

From Tables 4 and 5, statements can also be made about the probability of the congestion patterns and the function of the model can be checked. We will limit ourselves here to the best version of the models examined: model II. On the selected stretch, the probability of congestion, in general, increases when the mean speed in the funnel decreases. A falling speed in *cell −3* of the funnel indicates a higher probability of Jam Wave and Stop and Go. In contrast, it means a lower likelihood of Wide and Mega Jam. The time, especially the defined rush hour, indicates a higher probability for all congestion patterns, especially for Wide Jam. According to the simulation, the impact of the day of the week on the prediction would mean that the probability of congestion is higher on weekends than on weekdays. This does not seem directly plausible to us, but the parameters increase the general correctness of the model.

The probability of Jam Wave and Stop and Go decreases in the areas where merging operations (weaving section in Table 5) occur and where a freeway interchange exists. This is contrary to the first consideration for the probability of congestion at these locations. Concerning the input data, the presence of a weaving section or an interchange is always only the 500 m where the connection point or the exit onto the other freeway at the interchange is located. The real merging processes, however, take place before or after these areas. Therefore, the statement of the model is plausible at this point. Splitting the areas even further in a subsequent study would be conceivable. The frequency of the respective congestion pattern in the funnel strongly indicates the respective congestion pattern in the forecast. Again, the

Table 5
Model estimation results for the two model variants: values for the data dependent parameters in Eq. (4) with @ sign.

Parameter	Model I	Model II
Average speed in funnel $speed_{funnel}$	−0.11 (0.00)***	−0.11 (0.00)***
Kilometers to Munich $distance_{munich}$	−0.01 (0.00)***	−0.00 (0.00)***
Speed in cell −3 by Jam Wave $speed_{U2-cell-3}$	−0.08 (0.00)***	−0.03 (0.00)***
Speed in cell −3 by Stop and Go $speed_{U3-cell-3}$	−0.09 (0.00)***	−0.01 (0.00)***
Speed in cell −3 by Wide Jam $speed_{U4-cell-3}$	−0.09 (0.00)***	0.02 (0.00)***
Speed in cell −3 by Mega Jam $speed_{U5-cell-3}$	−0.09 (0.00)***	0.03 (0.00)***
Jam Wave while rush hour $rushhour_{JamWave}$	−0.15 (0.03)***	0.11 (0.06)
Stop and Go while rush hour $rushhour_{StopGo}$	0.67 (0.02)***	0.75 (0.03)***
Wide Jam while rush hour $rushhour_{WideJam}$	−0.53 (0.02)***	0.62 (0.07)***
Mega Jam while rush hour $rushhour_{MegaJam}$	−0.09 (0.02)***	2.28 (0.24)***
Jam Wave on weekdays $weekday_{JamWave}$	−0.60 (0.02)***	−0.57 (0.04)***
Stop and Go on weekdays $weekday_{StopGo}$	0.86 (0.01)***	−0.11 (0.02)***
Wide Jam on weekdays $weekday_{WideJam}$	−0.81 (0.01)***	0.04 (0.03)
Jam Wave at weaving sections $weavingsecion_{JamWave}$	0.00 (0.04)	0.22 (0.07)***
Stop and Go at weaving sections $weavingsecion_{StopGo}$	−0.18 (0.03)***	−0.23 (0.05)***
Jam Wave at interchanges $interchange_{JamWave}$	−1.10 (0.06)***	−1.17 (0.11)***
Stop and Go at interchanges $interchange_{StopGo}$	−0.34 (0.04)***	−0.40 (0.07)***
Speed difference of cell −5 to −1 by Wide Jam $speedWJ_{cell-5to1}$	0.05 (0.00)***	0.09 (0.00)***
Speed difference of cell −5 to −1 by Mega Jam $speedMJ_{cell-5to1}$	0.04 (0.00)***	0.10 (0.00)***
Proportion Jam Wave in front of Jam Wave $indic_{JamWave}$		18.18 (0.33)***
Proportion Stop and Go in front of Stop and Go $indic_{StopGo}$		14.18 (0.08)***
Proportion Wide Jam in front of Wide Jam $indic_{WideJam}$		23.10 (0.27)***
Proportion Mega Jam in front of Mega Jam $indic_{MegaJam}$		33.40 (2.78)***
Pattern in cell −1 by Jam Wave $JamWave_{cell-1}$		4.32 (0.12)***
Pattern in cell −3 by Jam Wave $JamWave_{cell-3}$		−9.11 (0.20)***

Standard errors in parentheses. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table 6
Model estimation results for the two model variants: statistical measures.

Parameter	Model I	Model II
Number of parameters	33.00	39.00
Number of respondents	81 582.00	81 582.00
Number of choice observations	15 407 162.00	15 407 162.00
Number of draws	20.00	20.00
LL(null)	−24 853 418.98	−24 853 418.98
LL(final)	−1 781 441.15	−252 278.68
McFadden R2	0.93	0.99
AIC	3 562 948.30	504 635.35

Standard errors in parentheses. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

influence of the real congestion pattern in *cell −3* of the funnel was not so clearly predictable. The probability of Jam Wave increases if Jam Wave exists in *cell −1*. This is logical, but if Jam Wave exists in *cell −3* of the funnel, the probability of Jam Wave decreases in the current cell of the prediction. This phenomenon could be explained by the fact that a Jam Wave occurs singularly: if it happened frequently, it would be declared as Stop and Go.

5.2. Validation

The first estimation results from the error and combination terms are shown in Table 4, and the results on the influences of the individual parameters can be read in Table 5. Table 7 compares the models' prediction from the training data set to the test data set. The five traffic conditions are compared: free-flow, Jam Wave, Stop and Go, Wide Jam, and Mega Jam. The table compares the traffic conditions of each of the four models' predictions with the correct traffic conditions.

To evaluate the classification results comprehensively, we selected three key metrics that are particularly suitable for imbalanced datasets like ours, since we have many more cells with free flow than cells with congestion, see Table 3: Balanced Accuracy (*BAcc*), Macro F1 (*MacF1*), and Cohen's Kappa (κ). These metrics provide valuable insights into model performance by addressing class imbalance and prediction quality across all classes.

BAcc improves upon traditional accuracy by accounting for class imbalance. It calculates each class's average recall (true positive rate), ensuring that both frequent and rare classes are equally represented in the evaluation. *BAcc* is defined as:

$$BAcc = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \tag{9}$$

where N is the number of classes, TP_i represents the true positives for class i, and FN_i represents the false negatives for class i. By averaging the recall across all classes, *BAcc* offers a fairer evaluation of a classifier's performance, particularly for imbalanced datasets (Brodersen et al., 2010).

The Macro F1 Score is another essential metric for evaluating models on imbalanced datasets. It computes the F1 score (the harmonic mean of precision and recall) for each class individually and then averages these scores across all classes. Unlike micro F1, which weights the overall performance by class frequency, Macro F1 ensures that each class contributes equally to the final score, regardless of size. This is critical when rare courses like Mega Jam in our dataset are just as important to predict as frequent ones like free-flow. The formula for Macro F1 is:

$$MacroF1 = \frac{1}{N}\sum_{i=1}^{N} F1_i \tag{10}$$

where $F1_i$ is the F1 score for class i, and N is the total number of classes. This metric helps ensure the model performs well across all classes, making it ideal for datasets where the correct prediction of all traffic states is equally important (Sokolova and Lapalme, 2009).

Cohen's Kappa is a robust metric that measures the agreement between predicted and actual classifications, considering the possibility that it could occur by chance. This makes it a more reliable performance indicator than simple accuracy, particularly in multiclass classification problems. κ is beneficial in imbalanced datasets because it adjusts for the expected agreement due to the class distribution:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{11}$$

where p_0 is the observed proportional agreement between actual and predicted values, and p_e is the expected proportional agreement under random classification. Kappa values range from 0 (agreement by chance) to 1 (perfect agreement), with negative values indicating worse-than-random agreement. In the context of this study, κ helps

us differentiate between models that perform well across all traffic conditions and those that may appear accurate but are biased towards the majority class (Towards Data Science, 2020).

Given the imbalanced nature of the traffic conditions in our dataset, these three metrics are particularly suitable for our evaluation. *BAcc* ensures that the model's performance is fairly evaluated across frequent and rare classes, making it a more robust metric than simple accuracy, which could be misleading in highly imbalanced scenarios. Macro F1 allows us to evaluate the model's performance across all classes equally, emphasizing precision and recall, which is crucial when rare traffic conditions are just as important as common ones. Cohen's κ adds an additional layer of understanding by adjusting for the agreement expected by chance, making it particularly valuable for imbalanced datasets where simple accuracy might overestimate model performance. Together, these metrics provide a comprehensive and fair assessment of the model's ability to predict the various traffic conditions in the dataset, ensuring that no single class dominates the evaluation.

Table 7 shows all the results for a one-minute prediction of the models listed above: **Model I** predicts 99.4% free-flow, 0% Jam Wave, 36.1% Stop and Go, 36.2% Wide Jam, and 0% of Mega Jam correctly, resulting in a κ of 0.49 and an *MacF1* of 0.25 and *BAcc* of 0.34. Model II gives much more useful and better results than the base model. 99.7% of free-flow, 73.6% of Jam Wave, 93.9% of Stop and Go, 88.9% of Wide Jam, and 96.4% of Mega Jam are predicted correctly. This leads to better classification rates than the previous models with a κ of 0.93, a *MacF1*, and *BAcc* of 0.91. **Model II** is needed to improve the base model because model I, without prior information, does not provide a much better model than the base A variant. For the naive models, Model P also has outstanding results, in contrast to Model A. The exact classification rates and forecast results for one minute can be found in Table 7.

We know that these are basic model variants and that the prediction may include overfitting by the high proportion of free flow. However, the high percentages of the correct classifications in the diagonal of the third row in Table 7 and the relatively high correlation coefficients show that the principle of the logit model works quite well and can be applied to this use case. The results will fit even more by further input parameters or investigations of the congestion pattern parameters and properties.

In order to make more precise statements about the quality of the models, and also to analyze beyond the one-minute forecast, a 30-min forecast was created and the results calculated for 1 min, 5 min, 10 min and 30 min. Here, the results are limited to the two best models from Table 7, i.e. only Base Model P and Model II. Similarly, only the classification rates are calculated for each of the forecast periods and not the entire cross-correlation matrix for reasons of clarity. Table 8 presents the performance metrics for Base Model P and Model II across different prediction horizons. The table is divided into two sections: full-class predictions and binary classifications. The three key metrics are reported: *BAcc*, *macF1*, κ.

For the Full-Class predictions: The Base Model P generally shows high values across all metrics, with particularly strong performance in short-term predictions (1 min) and gradually decreasing performance as the prediction horizon extends to 30 min. Model II, has similar values, but for all metrics substantially lower. As the prediction interval increases, the model's performance still remains below the performance of Base Model P. For the binary Classifications, just the differentiation of free flow or jam (i.e. all four congestion patterns) is made. Here, Base Model P continues to perform consistently well, with only a small decline in performance as the prediction horizon increases. Model II shows stronger binary classification performance, especially at the longer prediction horizons (5, 10, and 30 min), but still falls short compared to Base Model P.

While the performance metrics in Tables 7 and 8 show Base Model P with slightly better results in some areas, these figures should be viewed in context. Base Model P's higher metrics likely stem from its simplistic assumption that the current traffic state will continue, which may not reflect the complexities of real-world traffic. In contrast, Model II's lower scores reflect its proactive and risk-averse nature, as it actively predicts potential congestion even when conditions appear stable. This makes Model II more suitable for use in preventive traffic management, where the goal is to anticipate and prevent problems before they arise, rather than merely reacting to them.

A key advantage of Model II is its focus on minimizing false negatives—situations where traffic congestion occurs but is not predicted. By prioritizing congestion forecasting, Model II reduces the risk of dangerous, unexpected traffic jams, even if this sometimes results in false positives (predicting congestion when none occurs). Base Model P, on the other hand, is more conservative and less likely to predict congestion unless it has already started, which could lead to missed opportunities for early intervention and safety precautions. In addition to this, Model II incorporates contextual data, such as time of day and location, which enhances its ability to detect recurring traffic patterns and make more informed predictions. This is particularly useful in scenarios where traffic patterns change rapidly or are influenced by external factors, something Base Model P cannot account for with its limited focus on the most recent traffic state.

Moreover, Model II bases its predictions on the class with the highest probability, which means that in some cases, the chosen class may only have a marginally higher probability than other classes. This could lead to a situation where a slightly wrong class is predicted, though the difference is minimal. However, this probabilistic approach also allows traffic managers to use the predicted probabilities to inform more nuanced decision-making, something Base Model P cannot offer, as it provides only a single, deterministic class prediction.

In summary, while Base Model P shows better results, Model II offers a more sophisticated approach to traffic forecasting by proactively predicting congestion and integrating contextual data into its decision-making process. This allows traffic operators to take preemptive action before congestion becomes critical, making Model II more effective for real-world traffic management, where safety and prevention are key priorities.

5.3. Application

We use the model estimates for model II from Table 4 to show how the proposed congestion pattern prediction model can be used in practice. In this application example, we focus on February 10, 2019, when a Wide Jam occurred around noon. The jam was detected at 11:37 am (t = 697). Fig. 11 shows a more detailed progress of the congestion.

In this prediction application scenario, we set the prediction horizon to 30 min and predict congestion patterns at one-minute intervals. As shown in Fig. 8, information from the previous five minutes is used to predict congestion patterns. The information in the funnel is used so that all reliably available information is known up to the time of the start of the forecast, including the information on the main congestion pattern in the funnel cells directly before the forecast time. Although this information is actually only reliably available after a congestion pattern has been resolved, it is possible to determine a trend in the congestion pattern by applying an optimized automated algorithm to identify the congestion pattern even before a congestion pattern has been completely resolved. This information is taken as given for the application in this model.

The 30-min prediction works as follows. The first predicted time interval uses measured information from the previous five minutes. The output is a probability for each congestion pattern. In the second predicted time interval, four minutes of measured data and one minute of predicted data are used as input. Note that measured data for each congestion pattern is in a binary format, i.e., the pattern is present or not, while the predicted data is a probability. The third predicted time

Table 7
Classification table of observed versus predicted congestion patterns with classification rates. Each prediction is made for a 1-min forecast horizon, starting from every time step and location in the test dataset. All data up to the start time is assumed to be known, and predictions are made from that point forward. The values in each row of the observed congestion patterns sum to one.

Predicted congestion pattern	Observed congestion pattern					Classification rate		
	Free-flow	Jam Wave	Stop and Go	Wide Jam	Mega Jam	κ	*MacF1*	*BAcc*
Base model A								
Free-flow	0.995	0	0.005	0	0			
Jam Wave	0.923	0	0.077	0	0			
Stop and Go	0.654	0	0.346	0	0	0.32	0.16	0.27
Wide Jam	0.966	0	0.034	0	0			
Mega Jam	0.942	0	0.058	0	0			
Base model P								
Free-flow	0.996	0.001	0.001	0.001	0.001			
Jam Wave	0.060	0.940	0	0	0			
Stop and Go	0.010	0.001	0.990	0	0	0.98	0.98	0.98
Wide Jam	0.009	0	0	0.991	0			
Mega Jam	0.004	0	0	0	0.996			
Model I								
Free-flow	0.994	0	0.006	0.001	0			
Jam Wave	0.639	0	0.163	0.198	0			
Stop and Go	0.557	0	0.361	0.082	0	0.49	0.25	0.34
Wide Jam	0.448	0	0.190	0.362	0			
Mega Jam	0.524	0	0.401	0.075	0			
Model II								
Free-flow	0.997	0	0.002	0.001	0			
Jam Wave	0.264	0.736	0	0	0			
Stop and Go	0.061	0	0.939	0	0	0.93	0.91	0.91
Wide Jam	0.111	0	0	0.889	0			
Mega Jam	0.036	0	0	0	0.964			

Table 8
Performance metrics for Base Model P and Model II, with prediction horizons of one, five, ten, and thirty mins. Forecasts are generated every 30 min (on the hour and half-hour).

	Full-class predictions											
	Balanced Accuracy *BAcc*				Macro F1 *macF1*				Cohen's κ			
Prediction of horizon	1 min	5 min	10 min	30 min	1 min	5 min	10 min	30 min	1 min	5 min	10 min	30 min
Base Model P	0.982	0.944	0.906	0.824	0.983	0.945	0.904	0.814	0.988	0.966	0.940	0.856
Model II	0.791	0.774	0.764	0.724	0.751	0.626	0.614	0.504	0.612	0.659	0.625	0.390
	Binary classifications											
	Balanced Accuracy *BAcc*				Macro F1 *macF1*				Cohen's κ			
Prediction of horizon	1 min	5 min	10 min	30 min	1 min	5 min	10 min	30 min	1 min	5 min	10 min	30 min
Base Model P	0.994	0.983	0.970	0.930	0.994	0.983	0.970	0.927	0.988	0.966	0.940	0.854
Model II	0.981	0.968	0.962	0.926	0.901	0.827	0.810	0.690	0.724	0.656	0.622	0.387

interval then uses three minutes of measured data and two minutes of predicted data. The input from measured and predicted data continues until the fifth interval (five-minute prediction). After that, only predicted values are used as input until the 30-min prediction horizon is reached.

In Fig. 12, we show the prediction at 11:30 am (t = 690), 11:38 am (t = 698), and 12:00 pm (t = 720). The vertical red line shows the current time interval from which on-wards congestion patterns are predicted. The observations to the left of this line are the observed states used for the prediction. Fig. 12 shows the probability of each congestion pattern in each space–time cell. At t = 690, no information on congestion is present in the five minutes before the start of the prediction interval, and thus, the model predicts that free-flow conditions prevail in the next 30 min. At t = 698, the jam has just been detected, as can be seen in the tiny black space–time area for congestion type *Wide Jam* in the second row of Fig. 12. It can be seen that, for the next time intervals, the prediction model increases the probability that this congestion type prevails. However, as only a small area in the space–time funnel (c.f. Fig. 8) is affected, the prediction is not very distinctive but can indicate an emerging trend. Moving the prediction point further into the future to 12:00 pm (t = 720), we see a distinctive prediction of the Wide Jam congestion pattern prevailing for more than 15 min. The prediction further shows a familiar pattern of congestion moving backward. This is partly encoded in the used funnel structure and the estimated model parameters. It can also be concluded from the lower part of Fig. 12 that the model learned to some degree to predict the end of congestion as it does not predict that this pattern will prevail forever.

Fig. 13 shows the forecast behavior at the end of the congestion event. Again, the probability of the congestion pattern decreases as the time step progresses. The upper left sub-image in Figs. 12 and 13 represent the ground truth, respectively. It shows when the congestion pattern should be predicted and when it should not be. The model takes less than five minutes to recognize and correctly forecast a traffic jam beginning and ending. However, it should be noted that the predicted end can be updated when using new measurements or by including additional parameters specifically targeted at the end of a congestion pattern.

Table 9 demonstrates the high accuracy of our model in predicting congestion patterns for one minute, five minutes, and the entire 30 min. The cells that correctly predicted the traffic pattern were compared with the cells at the time step over all mispredicted locations. The resulting probability value from the model was rounded to 0 or 1, with all results showing over 90% correctness for all time steps, as depicted in Figs. 12 and 13. In summary, Model II is well-suited for this example of a *Wide Jam*. The next model design will include the duration parameter to predict the more precise duration of a congestion pattern.

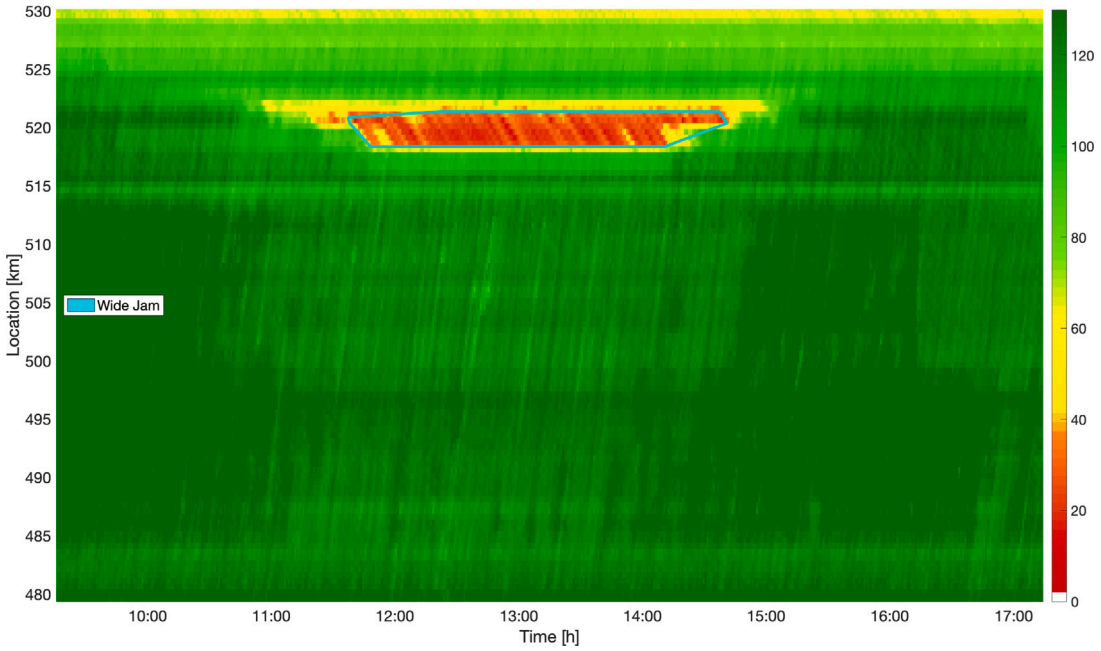


Fig. 11. Spatio-temporal distribution of the application example. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 9
Hit rate of application example for prediction horizon of one, five and 30 min.

Prediction at time	Hit rate of 1 min prediction	Hit rate of 5 min prediction	Hit rate of 30 min prediction
690	1	1	0.965
698	0.980	0.976	0.949
700	0.980	0.974	0.947
870	0.980	0.976	0.989
875	0.970	0.974	0.994
880	0.980	0.988	0.998

6. Discussion, conclusion, and outlook

In this paper, we applied a mixed logit model, a discrete choice model, to predict congestion patterns using empirical data from a freeway in Germany. We make use of the congestion patterns introduced by Karl et al. (2019). The data set comprises the daytime hours from 6 am to 8 pm and a stretch length of 50 km. Four different models were built for prediction: (i) a base model (A) that predicts congestion patterns by the average occurrence of a congestion pattern at a given location; (ii) a second base model (P) which uses the actual observed pattern from the last period as a forecast for the next period; (iii) a mixed logit model only with infrastructure effects and speed (Model I); (iv) a mixed logit model with infrastructure effects, speed and information on the existence of previous congestion patterns (Model II). However, while Base Model P shows slightly better results in terms of certain metrics, it is a reactive model that simply extends the current traffic state without considering broader contextual factors. In contrast, Model II, despite showing lower scores, offers a more advanced and proactive approach by incorporating infrastructure, speed, and previous congestion patterns. This makes Model II better suited for real-world traffic management, where anticipating and preventing congestion is crucial for improving safety and traffic flow.

The model is a complex statistical tool that provides detailed insights into decision-making processes. The model can choose between the five traffic patterns (Free Flow, Jam Wave, Stop&Go, Wide Jam, and Mega Jam). The biggest influencing parameters are the presence and distribution of the congestion patterns in the funnel. Here, the funnel mainly takes into account the congestion front movement. It is conceivable that the vehicle movement, i.e., almost perpendicular to the congestion front, also influences the forecast — similar to the anisotropic smoothing kernel.

We conclude that the approach to predicting congestion pattern probabilities with discrete choice methods and then analyzing the result for the most likely pattern is applicable and a promising avenue to improve the prediction of traffic patterns. However, the free-flow pattern is a dominant outcome in the data, which can result in overfitting issues. We will further extend the parameter set by adding some dynamic information about the start or end of each congestion pattern. In addition, more attention will be paid to how artificial intelligence, e.g., neural networks, can support the prediction of congestion patterns – like van Cranenburgh et al. (2022) presented in their work – but without losing the comprehensibility as in the logit model presented here.

In conclusion, the practical implications of this novel and innovative approach are at least twofold: First, the prediction of congestion times comprises the spatio-temporal extent of the congestion patterns in the space–time diagram. This can exceed the prediction of traffic patterns in a single segment. Therefore, additional information is available for traffic management centers and map operators to improve their predictions. This prediction can be conceptualized in two steps, firstly for a long-term forecast using only a simple static accumulation analysis of the congestion patterns and including the information about the location and time or date. The second step is to use the results of the newly developed multinomial logit model to produce a short-term forecast of up to 30 min. There are advantages here regarding traffic safety, optimization of traffic flow, and knowledge of possible diversions in the event of major and atypical traffic jams. This is particularly important for freeways, as these are (still) the main traffic arteries —

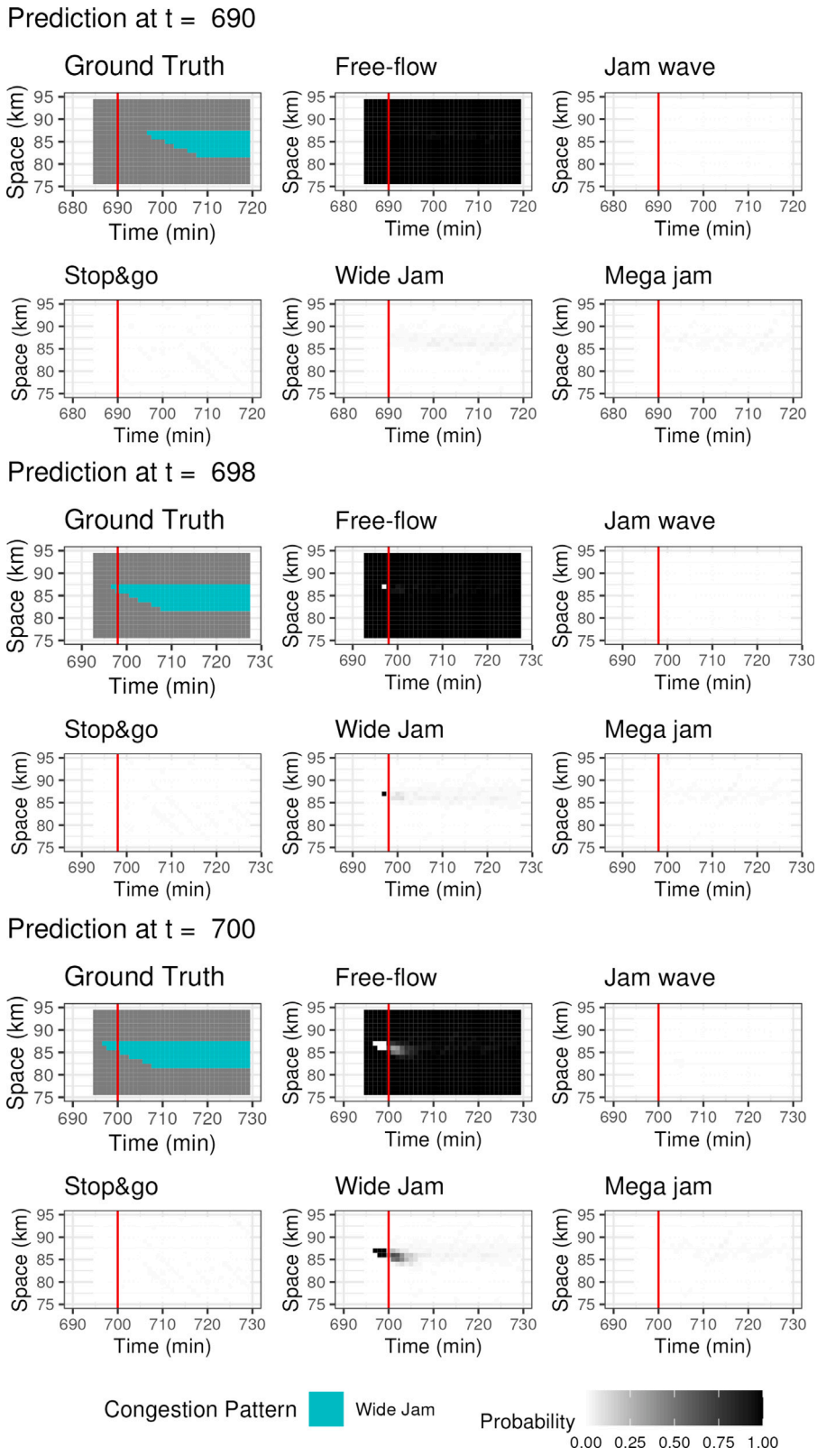


Fig. 12. Predicting congestion patterns; Start of congestion. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

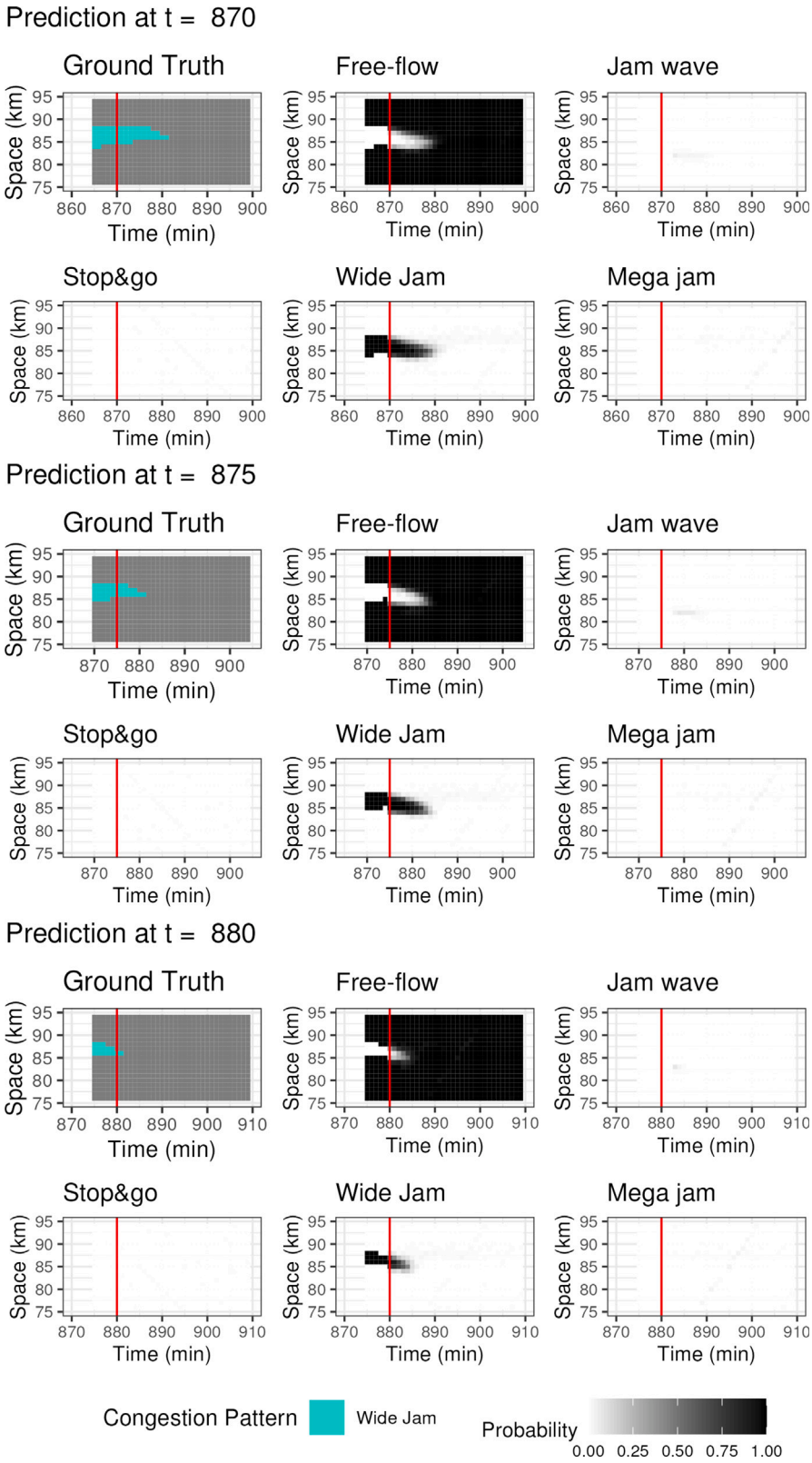


Fig. 13. Predicting congestion patterns; End of congestion. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

especially in Germany. Second, the applicability of a prediction based on a mixed logit or related model in traffic management centers is feasible. In particular, as the model parameters can be re-estimated on a rolling horizon basis and these parameters are explainable, it is a feature that is difficult to achieve in novel deep learning approaches. This is evidently an advantage to getting the support of traffic managers and decision-makers to implement such a prediction model.

CRediT authorship contribution statement

Barbara Metzger: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Allister Loder:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Lisa Kessler:** Conceptualization, Data curation, Formal analysis, Supervision, Writing – review & editing. **Klaus Bogenberger:** Conceptualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank *Landesbaudirektion Bayern* for providing the data.

References

Antoniou, Constantinos, Ben-Akiva, Moshe, Koutsopoulos, Haris N., 2006. Dynamic traffic demand prediction using conventional and emerging data sources. IEE Proc. Intell. Transp. Syst. 153, 97–104.

Brewer, Mark J., Butler, Adam, Cooksley, Susan L., 2016. The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. Methods Ecol. Evol. 679–692.

Brodersen, Kay H, Ong, Cheng Soon, Stephan, Klaas E, Buhmann, Joachim M, 2010. The balanced accuracy and its posterior distribution. In: 2010 20th International Conference on Pattern Recognition. IEEE, pp. 3121–3124.

Chen, Meng, Yu, Xiaohui, Liu, Yang, 2018. PCNN: Deep convolutional networks for short-term traffic congestion prediction. IEEE Trans. Intell. Transp. Syst. 19, 3550–3559.

Cottrell, Wayne D., 1998. Estimating the probability of freeway congestion recurrence. In: Transportation Research Record 1634, No. 1. TRR.

Di, Xiaolei, Xiao, Yu, Zhu, Chao, Deng, Yang, Zhao, Qinpei, Rao, Weixiong, 2019. Traffic congestion prediction by spatiotemporal propagation patterns. In: 2019 20th IEEE International Conference on Mobile Data Management. MDM, pp. 298–303.

Dia, Hussein, 2001. An object-oriented neural network approach to short-term traffic forecasting. European J. Oper. Res. 131 (2), 253–261, URL https://www.sciencedirect.com/science/article/pii/S0377221700001259. Artificial Intelligence on Transportation Systems and Science.

Dong, Bowen, Ma, Xiaoxiang, Chen, Feng, Chen, Suren, 2018. Investigating the differences of single-vehicle and multivehicle accident probability using mixed logit model. J. Adv. Transp. 2018, 9, URL https://www.hindawi.com/journals/jat/2018/2702360/.

Groß, Marcus, 2015. Logistische regression - beurteilung der klassifikationsgüte. URL https://www.inwt-statistics.de/blog-artikel-lesen/Logistische_Regression_Klassifikationsguete.html. [Online; Viewed 29 July 2021].

Hamed, Mohammad M., Al-Eideh, Basel M., 2020. An exploratory analysis of traffic accidents and vehicle ownership decisions using a random parameters logit model with heterogeneity in means. Anal. Methods Accid. Res. 25, 100116, URL https://www.sciencedirect.com/science/article/pii/S2213665720300063.

Kaplan, Sigal, Prato, Carlo Giacomo, 2012. Associating crash avoidance maneuvers with driver attributes and accident characteristics: A mixed logit model approach. Traffic Inj. Prev. 13 (3), 315–326, PMID: 22607255.

Karl, Barbara, Kessler, Lisa, Bogenberger, Klaus, 2019. Automated classification of different congestion types. In: 2019 IEEE Intelligent Transportation Systems Conference. ITSC, IEEE.

Kerner, Boris S., 2001. Tracing and forecasting of congested patterns for highway traffic management. In: 2001 IEEE Intelligent Transportation Systems Conference Proceedings. ITSC, pp. 88–93.

Kerner, Boris S., 2004. The Physics of Traffic. Springer Berlin Heidelberg.

Kerner, Boris S., 2009. Introduction to Modern Traffic Flow Theory and Control. Springer Berlin Heidelberg.

Kerner, Boris S., Rehborn, Hubert, 1996. Experimental properties of complexity in traffic flow. Phys. Rev. E 53, R4275–R4278, URL https://link.aps.org/doi/10.1103/PhysRevE.53.R4275.

Kerner, Boris S., Rehborn, Hubert, Aleksic, Mario, Haug, Andreas, 2004. Recognition and tracking of spatial–temporal congested traffic patterns on freeways. Transp. Res. C 12, 369–400.

Kessler, Lisa, 2021. Strategies for Detection of Congestion Patterns Using Multiple Sensor Technologies (Ph.D. thesis). Technische Universität München, p. 148.

Kessler, Lisa, Bogenberger, Klaus, 2023. Detection rate of congestion patterns comparing multiple traffic sensor technologies. IEEE Open J. Intell. Transp. Syst. PP, 1–1.

Kessler, Lisa, Huber, Gerhard, Kesting, Arne, Bogenberger, Klaus, 2018. Comparing speed data from stationary detectors against floating-car data. IFAC-PapersOnLine 51 (9), 299–304, 15th IFAC Symposium on Control in Transportation Systems CTS 2018.

Kessler, Lisa, Karl, Barbara, Bogenberger, Klaus, 2020. Congestion hot spot identification using automated pattern recognition. In: 2020 IEEE Intelligent Transportation Systems Conference. ITSC, IEEE.

Kessler, Lisa, Rempe, Felix, Bogenberger, Klaus, 2021. Multi-sensor data fusion for accurate traffic speed and travel time reconstruction. Front. Future Transp. 2.

Lee, Kyungmin, Hong, Bonghee, Jeong, Doseong, Lee, Jiwan, 2014. Congestion pattern model for predicting short-term traffic decongestion times. In: 17th International IEEE Conference on Intelligent Transportation Systems. ITSC, pp. 2828–2833.

Lee, Jiwan, Hong, Bonghee, Lee, Kyungmin, Jang, Yang-Ja, 2015. A prediction model of traffic congestion using weather data. In: 2015 IEEE International Conference on Data Science and Data Intensive Systems. pp. 81–88.

Li, Guopeng, Knoop, Victor L., van Lint, Hans, 2022. Estimate the limit of predictability in short-term traffic forecasting: An entropy-based approach. Transp. Res. C 138, 103607.

Li, Ruimin, Zhao, Xiaoqiang, Yu, Xinxin, Cheng, Nan, Zhu, Jianan, 2010. Incident duration model on urban freeways based on discrete choice model. In: 2010 International Conference on Electrical and Control Engineering. pp. 3826–3829.

van Lint, Hans, 2006. Reliable real-time framework for short-term freeway travel time prediction. In: J. Transp. Eng.. ASCE.

van Lint, Hans, Hoogendoorn, Serge Paul, van Zuylen, Henk J., 2002. Freeway travel time prediction with state-space neural networks: Modeling state-space dynamics with recurrent neural networks. Transp. Res. Rec. 1811 (1), 30–39.

Mahmuda, Akhtar, Sara, Moridpour, Michael, Bazant, 2021. A review of traffic congestion prediction using artificial intelligence. J. Adv. Transp. 2021.

Molloy, Joseph, Becker, Felix, Schmid, Basil, Axhausen, Kay W., 2021. mixl: An open-source r package for estimating complex choice models on large datasets. J. Choice Model. 39, 100284, URL https://www.sciencedirect.com/science/article/pii/S1755534521000178.

Nagy, Attila M., Simon, Vilmos, 2021. Improving traffic prediction using congestion propagation patterns in smart cities. Adv. Eng. Inform. 50, 101343.

Palmer, Jochen, Rehborn, Hubert, Gruttadauria, Ivan, 2011. Reconstruction quality of congested freeway traffic patterns based on Kerner's three-phase traffic theory. Int. J. Adv. Syst. Meas. 168–181.

Rempe, Felix, Bogenberger, Klaus, 2019. Feature engineering for data-driven traffic state forecast in urban road networks. In: Transportation Research Board 98th Annual Meeting Transportation Research Board.

Rempe, Felix, Kessler, Lisa, Bogenberger, Klaus, 2017. Fusing probe speed and flow data for robust short-term congestion front forecasts. In: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems. MT-ITS, pp. 31–36.

Rempe, Felix, Loder, Allister, Bogenberger, Klaus, 2021. Estimating motorway traffic states with data fusion and physics-informed deep learning. In: IEEE Intelligent Transportation Systems Conference. ITSC, IEEE.

Schatzmann, Thomas, Axhausen, Kay W., 2021. Long-Distance Buses in Switzerland - an Examination of Their Substitution Effects for Long-Distance Travel, Project Report. IVT, ETH Zurich, URL https://www.research-collection.ethz.ch/handle/20.500.11850/472308.

Schmid, Basil, Molloy, Joseph, Peer, Stefanie, Jokubauskaite, Simona, Aschauer, Florian, Hössinger, Reinhard, Gerike, Regine, Jara-Diaz, Sergio R., Axhausen, Kay W., 2021. The value of travel time savings and the value of leisure in zurich: Estimation, decomposition and policy implications. Transp. Res. A 150, 186–215, URL https://www.sciencedirect.com/science/article/pii/S0965856421001658.

Schreiter, Thomas, van Lint, Hans, Treiber, Martin, Hoogendoorn, Serge, 2010. Two fast implementations of the adaptive smoothing method used in highway traffic state estimation. In: 13th International IEEE Conference on Intelligent Transportation Systems. ITSC, pp. 1202–1208.

Sokolova, Marina, Lapalme, Guy, 2009. A systematic analysis of performance measures for classification tasks. Inf. Process. Manage. 45 (4), 427–437.

Towards Data Science, 2020. Cohen's kappa: A robust measure for classification agreement. URL https://towardsdatascience.com/cohens-kappa-how-to-measure-agreement-between-two-raters-558d6e8bee3a.

Train, Kenneth E., 2009. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge.

Treiber, Martin, Helbing, Dirk, 2002. Reconstructing the spatio-temporal traffic dynamics from stationary detector data. Coop. Transp. Dyn. 1, 3.1–3.24, (Internet Journal, www.TrafficForum.org/journal).

Treiber, Martin, Helbing, Dirk, 2003. An adaptive smoothing method for traffic state identification from incomplete information. In: Interface and Transport Dynamics. Springer, Heidelberg, pp. 343–360.

Treiber, Martin, Kesting, Arne, Wilson, R. Eddie, 2011. Reconstructing the traffic state by fusion of heterogeneous data. Comput.-Aided Civ. Infrastruct. Eng. 26, 408–419.

van Cranenburgh, Sander, Wang, Shenhao, Vij, Akshay, Pereira, Francisco, Walker, Joan, 2022. Choice modelling in the age of machine learning - discussion paper. J. Choice Model. 42, 100340.

Wu, Qiong, Chen, Feng, Zhang, Guohui, Liu, Xiaoyue Cathy, Wang, Hua, Bogus, Susan M., 2014. Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways. Accid. Anal. Prev. 72, 105–115, URL https://www.sciencedirect.com/science/article/pii/S0001457514001870.

Yildirimoglu, Mehmet, Geroliminis, Nikolas, 2013. Experienced travel time prediction for congested freeways. Transp. Res. B 53, 45–63, URL https://www.sciencedirect.com/science/article/pii/S0191261513000465.

Zhao, Xiaoqiang, Li, Ruimin, Yu, Xinxin, 2009. Incident duration model on urban freeways based on classification and regression tree. In: 2009 Second International Conference on Intelligent Computation Technology and Automation. Vol. 3, pp. 625–628.