Data Science for Transportation (2023) 5:22
https://doi.org/10.1007/s42421-023-00085-8

RESEARCH

Probabilistic Traffic State Prediction Based on Vehicle Trajectory Data

Mohammadreza Khajeh Hosseini[1] · Alireza Talebpour[1]

Received: 28 April 2023 / Revised: 22 September 2023 / Accepted: 20 October 2023 / Published online: 13 November 2023

Abstract

Accurate prediction of traffic flow dynamics is a key step towards effective congestion mitigation strategies. The dynamic nature of traffic flow and lack of comprehensive data coverage (e.g., availability of data at loop detector locations), however, have historically prevented accurate traffic state prediction, leading to the widespread utilization of reactive congestion mitigation strategies. The introduction of connected automated vehicles provides an opportunity to address this challenge. These vehicles rely on trajectory-level prediction of their surrounding traffic environment to plan a safe and efficient path. This study proposes a methodology to utilize the outcome of such predictions to estimate the future traffic state. Moreover, the same approach can be applied to data from connected vehicles for traffic state prediction. Since in many driving scenarios, more than one maneuver is feasible, it is more logical to predict the location of the vehicles in a probabilistic manner based on the probability of different maneuvers. The key contribution of this study is to introduce a methodology to convert such probabilistic trajectory predictions to aggregate traffic state predictions (i.e., flow, space–mean speed, and density). The key advantage of this approach (over directly predicting traffic state based on aggregated traffic data) is its ability to capture the interactions among vehicles to increase the accuracy of the prediction. The down side of this approach, on the other hand, is that any increase in the prediction horizon reduces the accuracy of prediction (due to the uncertainty in the vehicles' interactions and the increase in the possibility of different maneuvers). At the microscopic level, this study proposes a probability-based version of the time–space diagram, and at the macroscopic level, this study proposes probabilistic estimates of flow, density, and space–mean speed using the trajectory-level predictions. To evaluate the effectiveness of the proposed approach in predicting traffic state, the mean absolute percentage error for each probabilistic macroscopic estimate is evaluated on multiple subsamples of the NGSIM US-101 and I-80 data sets. Moreover, while introducing this novel traffic state prediction approach, this study shows that the fundamental relation among the average traffic flow, density, and space–mean speed is still valid under the probabilistic formulations of this study.

Keywords Probabilistic traffic prediction · Fundamental diagram · Vehicle trajectory data

Introduction

Delaying the onset of traffic breakdown and congestion formation are the main objectives of any congestion prevention and mitigation strategy. While a number of reactive strategies have been proposed over time (e.g., see Papageorgiou et al. (1997) and Talebpour et al. (2013)), successful congestion management strategies rely on accurate prediction of traffic flow dynamics (Elfar et al. 2019). Conventionally, stationary vehicle detectors (e.g., inductive loops, piezoelectric sensors installed in the pavements, and cameras and radars installed on supporting structures) were utilized to monitor the traffic condition and provide the necessary data for traffic prediction. Accordingly, the available data only contained measurements along a very short segment of the road and have been mostly in the form of aggregated measures (e.g., flow, occupancy/density, and average speed). Unfortunately, such data cannot accurately capture the evolution of traffic flow throughout the roadway (e.g., cannot fully capture the shockwave formation and propagation). While introducing multiple sensors along the roadway can slightly mitigate this challenge, many key traffic flow

✉ Alireza Talebpour
ataleb@illinois.edu

Mohammadreza Khajeh Hosseini
mk49@illinois.edu

[1] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N Matthews Ave., Urbana, IL 61801, USA

features still cannot be captured due to the location-specific measurements by these sensors. For instance, Talebpour et al. (2013) showed that speed harmonization can prevent flow breakdown if the system can detect shockwave formation at its onset; which cannot be achieved with point measurements from conventional sensors. In other words, by the time the shockwave reaches the sensor location, it might be impossible to prevent its propagation using common congestion management strategies.

The dynamic nature of traffic flow combined with the aforementioned limitations of conventional data sources have historically prevented accurate traffic state prediction, limiting the usage of predictive congestion management strategies and leading to a widespread utilization of reactive methods. The introduction of connected automated vehicles provides an opportunity to address this challenge. Recent advances in vehicular communications have provided exceptional communications and data exchange opportunities between vehicles and their surroundings (Rahim et al. 2021). Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications enable monitoring the driving environment continuously over time and space and at individual vehicle level. Such information can be utilized to predict the future trajectory of individual vehicles. Moreover, since connected automated vehicles rely on trajectory-level prediction of their surrounding traffic environment to plan a safe and efficient path, by sharing such predictions, one can estimate the future traffic state in a distributed manner (without relying on a centralized prediction based on connected vehicles data).

Predicting the driving environment at the individual vehicle level is challenging due to the dynamic nature of the driving environment, the complex interactions among vehicles, and differences in driving behavior among different drivers. Structuring the driving task into multiple maneuvers based on the interactions among vehicles can improve vehicle trajectory prediction. These maneuvers represent the drivers' possible responses to their surrounding environment and interactions with other vehicles. Obviously, each vehicle's future trajectory changes depending on the driver's choice of maneuver. Since in many driving scenarios, more than one maneuver is feasible, focusing on a single driving maneuver cannot provide a comprehensive picture of potential future trajectories. As a result, this study adopts a probabilistic approach to predict the future movement of the vehicles conditioned on different maneuvers and by considering the interactions among the vehicles. Accordingly, the main objective of this study is to introduce a methodology to convert such probabilistic trajectory predictions to aggregate traffic state predictions (i.e., flow, space–mean speed, and density). The key advantage of this approach (over directly predicting traffic state based on aggregated traffic data) is its ability to capture the impacts of interactions among vehicles on traffic flow dynamics to increase the accuracy of the predictions.

The prediction horizon depends on the planning horizon, and in general, a longer but accurate prediction horizon is preferred for congestion management. Unfortunately, the prediction accuracy is expected to decrease with the increase in the prediction horizon due to the uncertainty in drivers' behavior and an increase in the possibility of various configurations and outcomes (Zhou et al. 2019). Capturing and characterizing such uncertainties is critical to better adjust the congestion management strategies. This is another key advantage of the proposed approach over the traditional traffic state prediction methods based on aggregated speed, flow, and density data. While aggregated level data-driven methodologies provide the means to predict for any prediction horizon, they mostly do not have the ability to explicitly capture the uncertainty in prediction as prediction horizon changes. In other words, their underlying models (that are well-trained for a particular prediction horizon) can be incapable of accurately predicting for any horizon (other than the one that they have been trained to predict). Our proposed approach, on the other hand, provides a probabilistic trajectory prediction for any time step between prediction time and the prediction horizon, accurately capturing and modeling the decrease in prediction accuracy as prediction horizon increases.

Accordingly, the main contributions of this study are threefold: (1) introducing a methodology to accurately predict the trajectory of the vehicles in the traffic stream while capturing all the interactions among vehicles; (2) introducing the concept of probabilistic time–space diagram and discuss its characteristics; and (3) proposing a set of methodologies to calculate probabilistic estimates of flow, density, and space–mean speed using the trajectory-level predictions. The traffic state prediction methodologies proposed in this study are based on the assumption of complete knowledge of the observed trajectory of all the vehicles in the study area considering a fully connected environment. Such information is either provided through a fully connected environment or based on the data measured and shared with connected and automated vehicles.

The remainder of this study is organized as follows: the following section provides a background on the existing macroscopic and vehicle trajectory prediction models. This section is followed by presenting the details on the adopted probabilistic trajectory prediction model at the microscopic level and introducing the proposed probabilistic traffic estimates at the macroscopic level. Next, the presented probabilistic traffic state prediction approach is evaluated using real-world vehicle trajectory data, and the results and a discussion on the findings are presented. Finally, the paper is concluded with some remarks and future research needs.

Background

Macroscopic-Level Traffic Prediction

Traffic prediction is one of the primary components of intelligent transportation systems. A traffic management system can implement appropriate traffic control measures based on the predicted traffic state. In addition, the energy management system in automated and connected vehicles can use traffic predictions to determine when to take control actions to optimize energy consumption. The process of traffic state prediction involves forecasting the state variables, such as flow, density, speed, and travel time, using observed data. Over the past two decades, a number of studies have proposed various methodologies for short-term (20 s up to 1hr) and long-term (1hr and more) traffic state prediction.

Different approaches to traffic state prediction use a variety of input data, traffic flow models, and estimation methodologies (Seo et al. 2017). Van Lint and Van Hinsbergen (2012) grouped traffic prediction methodologies into naive, parametric, and non-parametric categories. A naive approach assumes that future traffic states remain the same as the current ones (Huisken and Berkum 2003) or that future observations will be similar to historical ones (Eglese et al. 2006). The naive prediction methods are limited to situations, where the traffic conditions are stable over an extended period of time (notice that such situations are rare) or where the traffic patterns are nearly constant over time (e.g., daily, weekly). Parametric approaches refer to the methodologies that use a traffic flow model with parameters calibrated on historical data or in combination with new observations. The fundamental diagram represents the macroscopic characteristics of the traffic state and is a well-studied model of traffic flow. Accordingly, the fundamental diagram is mostly used in parametric traffic prediction and in conjunction with first- and second-order traffic flow models to estimate the state of the segments of roadways or networks based on noisy measurements collected from sensors or probe vehicles (Wang et al. 2009; Sun and Work 2014; Zheng and Su 2016; Wang et al. 2016; Seo et al. 2016). Another type of parametric model is mesoscopic traffic prediction models. The mesoscopic models predict the traffic state by tracking individual vehicles considering the macroscopic state of the traveling links (Mahmassani et al. 2009). One of the main challenges of the parametric approach is the trade-off between accuracy and complexity, especially when dealing with time-variant and abnormal traffic dynamics. Alternatively, non-parametric approaches are mostly based on simple data-driven methods that do not directly incorporate traffic flow models. To predict traffic state, a variety of data analysis and machine learning techniques are employed. Among the common approaches are linear regression (Smith et al. 2003; Wilby et al. 2014), various classes of neural networks (Chan et al. 2012; Lv et al. 2015; Duan et al. 2016; Polson and Sokolov 2017), support vector regression (Su et al. 2007; Castro-Neto et al. 2009), and time-series forecasting (Chen et al. 2011; Kumar and Vanajakshi 2015).

As data have become more widely available in the past decade, data-driven methods for traffic prediction have been gaining attention. Majority of non-parametric traffic state prediction models revolve around aggregated and macroscopic traffic data as input to the models, and there are only a few studies utilizing the vehicle trajectories as input to the macroscopic prediction models (Elfar et al. 2018; Khajeh Hosseini and Talebpour 2019a; Khajeh-Hosseini and Talebpour 2019b).As traffic flow and density increase (especially around the breakdown point), unexpected driving behaviors have an increasing impact on the traffic state and more complex dynamics can be observed in the traffic flow (Ossen 2008). As a result, better traffic predictions can be derived from capturing interactions among vehicles. Note that traffic prediction has a large and growing body of literature as well as comprehensive reviews, such as Seo et al. (2017), which are recommended for more in-depth study to interested readers.

Vehicle Trajectory Prediction

Although vehicle trajectory data are at the core of any traffic flow analysis, it has not been widely utilized for traffic state prediction. In fact, the majority of vehicle trajectory prediction methodologies has been developed to support the motion planning algorithms of automated vehicles. Predicting the movement of traffic agents in the surrounding driving environment is essential for safe path planning of the automated vehicles. Lefèvre et al. (2014) classifies the motion prediction methodologies based on their modeling hypotheses into three groups: (1) physics-based, (2) maneuver-based, and (3) interaction-aware.

The physics-based model refers to the simple methodologies that consider the movements of vehicles with respect to dynamic or kinematic motion models without considering the vehicle's maneuvers or interaction among traffic agents. Dynamic modeling is usually adopted in control applications for an ego vehicle. However, measuring a target vehicle's position, speed, and acceleration is easier than measuring the forces resulting in the vehicle's motion. Consequently, kinematic models are simpler and more popular than dynamic models to predict a target vehicle's future state (Lefèvre et al. 2014). Some studies, such as Khajeh Hosseini et al. (2019), assume that the data on the vehicle's current state and the motion prediction model is adequate to estimate the future state of the target vehicle without considering any uncertainties. Another group of studies, such as Batz et al. (2009); Abbas et al. (2020), consider uncertainties in measurements and adopt Kalman Filter in motion modeling. Monte Carlo methods can also be used to sample from input data for the motion model to simulate different possibilities of the target vehicle's future state (Broadhurst et al. 2005). Unrealistic predictions can be removed in the post-processing of trajectories or by eliminating unrealistic input to the motion models. The physics-based models, however, do not consider the vehicle's maneuver or the interaction among vehicles, making them limited to short-term trajectory prediction (Lefèvre et al. 2014).

Maneuver-based models recognize the different vehicle maneuvers and predict the vehicle's movement considering the intended maneuver. In this approach, the intended maneuver is first detected by comparing the observed partial trajectory with a set of prototype trajectories for different maneuvers. Then, the trajectory of the vehicle is predicted based on the detected maneuver. In the prototype trajectory approach, it is assumed that the vehicles' trajectories can be grouped into different motion patterns represented by prototype trajectories learned from previous observations (i.e., from a training set). Gaussian Process has proven to be a good modeling approach for trajectory prototyping (Goli et al. 2018; Tran and Firl 2014; Joseph et al. 2011). The partially observed trajectory is compared to the prototype trajectories, and the rest of the trajectory is predicted based on the closest prototype trajectory. Some other studies identify the vehicle's maneuver by classifying the partially observed trajectory using machine learning approaches, such as support vector machine (Dou et al. 2016), multilayer perceptron (MLP) (Yoon and Kum 2016), logistic regression (Klingelschmitt et al. 2014) or even recurrent neural networks (Khosroshahi 2017). Kinematic motion models predict the rest of the vehicle's trajectory based on the detected maneuver. Maneuver-based trajectory prediction models are more accurate than the simple physics-based models; however, the maneuver-based models also do not consider the full interaction among the traffic agents that could impact the vehicle's trajectory (Lefèvre et al. 2014).

Interaction-aware methodologies consider the interactions among traffic agents when predicting the maneuver and trajectory of a target vehicle. A simple consideration of vehicles' interaction is using trajectory prototyping and dropping the conflicting trajectories between the pair of vehicles (Lawitzky et al. 2013). Another approach is to consider pairwise interaction among all the vehicles in the driving environment; however, the number of pairwise combinations increases quadratically with the number of vehicles in the scene. Some interaction-aware studies (Deo and Trivedi 2018a; Liebner et al. 2012) simplify the problem by assuming an asymmetric interaction between the target vehicle and environment, such that the environment impacts the target vehicle but the target vehicle does not impact the environment. The interaction-aware prediction methodologies are more accurate than the physics-based and maneuver-based methodologies as they consider the interdependency among the movement of traffic agents and consequently can be used for a longer prediction horizon (Lefèvre et al. 2014).

Recent trajectory prediction studies are mostly based on deep learning approaches considering the interaction among the vehicles. The interaction among vehicles is captured in the input representation of the driving environment to the deep learning prediction model (Mozaffari et al. 2020). The driving environment representation to these models could be in the form of trajectory history of the target vehicle and surrounding vehicles (Deo and Trivedi 2018b; Ma et al. 2019), bird's eye view representation of the processed surrounding environment (Deo and Trivedi 2018a; Cui et al. 2019; Sheng et al. 2022), or raw sensory data (Luo et al. 2018). The deep learning-based trajectory prediction models are mostly based on Recurrent neural networks (RNN) (Deo and Trivedi 2018b; Ma et al. 2019), convolutional neural networks (CNN) (Cui et al. 2019; Luo et al. 2018). or a combination of the two approaches (Deo and Trivedi 2018a). RNN architecture has a high capability to capture temporal dependencies in sequence data, and CNN architecture has a high capability to capture spatial dependencies in feature maps. Accordingly, the combination of two approaches seems suitable to capture both temporal and spatial dependencies in the traffic stream.

Methodology

The driving environment evolves as a result of interactions among individual vehicles. These interactions can be defined based on a combination of lateral and longitudinal maneuvers of vehicles in response to their driving environment. As discussed before, the accuracy of prediction at the individual vehicle level can increase by structuring the driving task into different maneuvers and considering the vehicles' interactions. In many driving scenarios, however, more than one maneuver is feasible. Therefore, to increase the realism of the predictions, the location of the vehicles should be estimated in a probabilistic manner based on different maneuvers. Obviously, with the increase in the prediction horizon, the accuracy of prediction decays due to the uncertainty in the vehicles' interactions and the increase in the possibility of different configurations and outcomes.

This issue can be partially resolved by aggregated level data-driven approaches. Those approaches can be trained to predict for any prediction horizon. The key downside of those approaches, however, is that they do not consider the interactions among agents that form the traffic flow. To address this shortcoming, this study proposes a methodology to capture the uncertainty in the prediction process. Our proposed approach provides a probabilistic trajectory prediction for any time step between the prediction time and the prediction horizon, accurately capturing and modeling the decrease in prediction accuracy as prediction horizon increases. Moreover, utilizing the trajectory level predictions, this study proposes estimating the macroscopic traffic states (e.g., flow, density, and space–mean speed) by aggregating probabilistic individual-level predictions. Accordingly, the prediction horizon can be extended at the microscopic level and interactions among vehicles can be captured at the macroscopic level. As discussed before, to capture the uncertainties, the macroscopic level predictions are also probabilistic in this approach.

Microscopic Level: Probabilistic Trajectory Prediction

The driving environment prediction at the individual vehicle level can be performed in the form of trajectory prediction. However, the dynamic nature of the driving environment, the interaction among vehicles, and the variation in driving behavior make predicting at the trajectory level challenging. The vehicle trajectory prediction can be improved by dividing the driving task into different maneuvers and considering the vehicle's interactions. In many driving scenarios, multiple maneuvers are feasible, and focusing on a single driving maneuver may not be sufficient. Thus, it is more realistic to adopt a probabilistic trajectory prediction model that includes the possibility of different maneuvers and considers the interactions among vehicles. This study utilizes the convolutional social pooling model proposed by Deo and Trivedi (2018a) as the core of individual vehicle trajectory prediction module. This model predicts the probabilistic trajectory of a target vehicle for different longitudinal and lateral maneuvers by considering the observed history of its surrounding vehicles, and captures the interactions among the vehicles considering a bird's eye view of the driving environment. At the time of prediction, it is assumed that a complete knowledge of the history of the surrounding driving environment, H, is available (either from a fully connected vehicle environment or based on the sensor data collected from connected and automated vehicles at a high penetration rate.) The maneuver-based probabilistic trajectory prediction can then be defined based on

$$P(T|H) = \sum_{i=1}^{M} P_{\Theta}(T|H, m_i)P(m_i|H) \tag{1}$$

where the future trajectory, T, is predicted based on the observed history of the surrounding environment, H, and the combination of M possible maneuvers (i.e., m_i for $i = 1 \dots M$). Θ is the parameters of the conditional probability distribution of the vehicle trajectory. The prediction model can be trained to estimate Θ and $P(m_i|H)$, the probability of individual maneuvers given the history.

The location of vehicle v over the roadway segment can be described with longitudinal and lateral coordinates of the vehicle ($X_v = [x_v, y_v]^T$). The probabilistic trajectory prediction model is trained to estimate the probability of every possible maneuver, as well as the future trajectory of vehicle v, in the form of the mean (μ) and covariance (Σ) of a bivariate normal distribution for every time step t in the future and for every maneuver m_i:

$$X_v^{t,m_i} \sim N(\mu_v^{t,m_i}, \sum_v^{t,m_i}) \tag{2}$$

This equation defines the probabilistic location of the vehicle considering one of the possible maneuvers. Accordingly, the probabilistic location of the vehicle can be defined based on a weighted summation of these bivariate Gaussian distribution models and the probability of each maneuver. Accordingly, the probability density function of the vehicle's location, $p(X^t)$ can be written as

$$p(X^t) = \sum_{i=1}^{M} P(m_i)\Phi(X^{t,m_i}|\mu^{t,m_i}, \sum^{t,m_i}) \tag{3}$$

In this bivariate Gaussian mixture model, $P(m_i)$ is the probability of taking maneuver m_i, and Φ is the probability density function of the vehicle location for maneuver m_i at time t. Utilizing these probabilistic trajectories, one can define probabilistic time–space diagram and utilize it to define the probabilistic macroscopic measures of traffic flow (i.e., flow, speed, density, and occupancy).

The convolutional social pooling model (Deo and Trivedi 2018a) uses an LSTM encoder–decoder network to predict the probability distribution of a target vehicle's future location. The model estimates parameters of the conditional distribution of the target vehicle's location, considering the observed trajectory histories. Moreover, this model expands each maneuver to different lateral and longitudinal maneuvers. The lateral maneuvers include three movements of maintaining the lane and lane-changing to the right and left. The longitudinal maneuvers include braking and not braking.

The convolutional social pooling model (Deo and Trivedi 2018a) is trained to predict the probability of different maneuvers (combination of lateral and longitudinal movements) of a target vehicle, as well as predicting the probabilistic location of the target in the form of the mean (μ) and covariance (σ) of a bivariate normal distribution for every time step t in the future and for every maneuver m. In other words, this model predicts a separate set of parameters (σ) for every time step in the prediction period and for every maneuver. Accordingly, the output of the probabilistic trajectory prediction model are

$$\text{Probability of maneuvers} = [P(m_1), P(m_2), P(m_3), P(m_4), P(m_5), P(m_6)] \tag{4}$$

$$\begin{aligned}
\theta_{m_1} &= [(\mu^{t,m_1}, \sigma^{t,m_1}), ..., (\mu^{t+t_f,m_1}, \sigma^{t+t_f,m_1})] \\
\theta_{m_2} &= [(\mu^{t,m_2}, \sigma^{t,m_2}), ..., (\mu^{t+t_f,m_2}, \sigma^{t+t_f,m_2})] \\
\theta_{m_3} &= [(\mu^{t,m_3}, \sigma^{t,m_3}), ..., (\mu^{t+t_f,m_3}, \sigma^{t+t_f,m_3})] \\
\theta_{m_4} &= [(\mu^{t,m_4}, \sigma^{t,m_4}), ..., (\mu^{t+t_f,m_4}, \sigma^{t+t_f,m_4})] \\
\theta_{m_5} &= [(\mu^{t,m_5}, \sigma^{t,m_5}), ..., (\mu^{t+t_f,m_5}, \sigma^{t+t_f,m_5})] \\
\theta_{m_6} &= [(\mu^{t,m_6}, \sigma^{t,m_6}), ..., (\mu^{t+t_f,m_6}, \sigma^{t+t_f,m_6})]
\end{aligned} \tag{5}$$

where m_i refers to maneuver i, and θ_{m_i} refers to the set of the parameters of the bivariate normal distribution describing the target vehicle's location for every time step in the future based on maneuver m_i.

The original deep learning model of Deo and Trivedi (2018a) predicts the probability of each of the lateral and longitudinal maneuvers separately and assumes independence between the lateral and longitudinal maneuvers. Accordingly, Deo and Trivedi (2018a) estimate the probability of the combination of maneuvers based on the following equation:

$$\begin{aligned}
P(m_{i,j}) &= P(lateral_i, longitudinal_j) \\
&= P(lateral_i)P(longitudinal_j)
\end{aligned} \tag{6}$$

In this equation, $P(m_{i,j})$ is the probability of taking the maneuver that is a combination of lateral movement i (e.g., lane changing to the right) and longitudinal movement j (e.g., braking). Predicting the probability of the lateral and longitudinal maneuvers separately results in the same probability of longitudinal maneuver independent of the lateral maneuver and vice versa. In other words, the probability of deceleration would be independent of the vehicle maintaining its current lane or moving to one of its neighboring lanes. However, the acceleration and deceleration due to a lateral maneuver depend on the new car following conditions (e.g., new leader) in the new lane. Accordingly, this study modifies the deep learning model to predict the joint probability of lateral and longitudinal maneuvers. The updated network is trained to predict the probability of six different maneuvers instead of predicting the probability of lateral and longitudinal movements separately. Each maneuver is a combination of lateral and longitudinal movements. In other words, the updated network is predicting $P(lateral_i, longitudinal_j)$ directly rather than predicting $P(lateral_i)$ and $P(longitudinal_j)$ separately.

This study adopts the lateral and longitudinal movement classification of Deo and Trivedi (2018a). Lateral movements include lane changing to the right, lane changing to the left, and staying in the same lane, and longitudinal movements include braking or not braking. The lateral movement is defined based on the lane of the target vehicle at the time of prediction and its lane 5 s in the future. For example, if the target vehicle's lane at the time of prediction is the middle lane and the vehicle's lane 5 s in the future is one lane to the right of the middle lane, the lateral movement of the target vehicle is considered as lane changing to the right. The longitudinal movement is based on the target vehicle's speed at the time of the prediction and the vehicle's average speed during the prediction period. It is assumed that the target vehicle is performing a braking movement if its average speed during the prediction period is less than 80 percent of the vehicle's speed at the time of prediction. After updating the model to predict the lateral and longitudinal maneuvers jointly, the model's performance in predicting the maneuver is increased by more than 8 percent on the testing data set of this study. More details on the training and testing data set used for this model is provided in Sect. 4.

Probabilistic Occupancy Map

Occupancy map is a grid-based representation of the roadway segment. The roadway segment is divided into small cells (e.g., 1.2 by 1.2 ms in this study) represented by a matrix with values indicating the probability of that cell being occupied by a vehicle. The probabilistic occupancy map of the study area can be created considering the probabilistic location of all the vehicles on the study segment. Considering the probabilistic location of individual vehicles (Eq. 3), the occupancy probability of cell c with boundaries of $[x_1, x_2]$ and $[y_1, y_2]$ can be estimated based on

$$\begin{aligned}
P(O_c = 1) = &\sum_{v=1}^{\#vehicles} \int_{y_1}^{y_2} \int_{x_1}^{x_2} \sum_{i=1}^{M} P(m_{i,v}) \\
&\Phi(X_v^{t,m_i} | \mu_v^{t,m_i}, \Sigma_v^{t,m_i}) dxdy
\end{aligned} \tag{7}$$

This equation estimates the occupancy probability of each cell (e.g., c) in the occupancy matrix (O) by summing the probability of that cell being occupied by each of the vehicles on the roadway segment.

Probabilistic Time–Space Diagram

The time–space diagram is the plot of the trajectory of vehicles over time and space. To create the probabilistic time–space diagram, this study utilizes the representation proposed by Khajeh Hosseini and Talebpour (2019a). They proposed dividing the time and space domains into smaller cells and reconstructing the time–space diagram with a binary time–space matrix. In the binary time–space matrix of Khajeh Hosseini and Talebpour (2019a), a cell value of one indicates the presence of a vehicle at the respective time and location, and a value of zero is used for empty cells (Fig. 1). A probabilistic prediction of the time–space matrix can be created by breaking the binary constraint and replacing the value of each cell with the expected value of the time–space cell. The probability of vehicle v passing through a time–space cell ($[t_1, t_2]$,$[y_1, y_2]$) can be estimated by considering the probabilistic location of the vehicle at the beginning of the time period of the cell and the probabilistic location of the vehicle at the end on the time period of the cell. A vehicle can move from one time–space cell to another either by crossing the beginning of the space domain of the next cell during its time period or entering a new cell that shares the exact space boundaries similar to its previous cell. These two movements in time and space are depicted with red and yellow arrows in Fig. 2, respectively. The probability of a vehicle passing through a specific time–space cell is the summation of the probability of both possible ways to enter this cell. The expected value of each cell in the time–space matrix (TS) is the summation of the probability of every vehicle going through that cell multiplied by its contribution to the value of the cell, which is assumed to be one for all the vehicles in this study. Note that this contribution can vary based on the type of vehicles (e.g., trucks can have higher contribution than passenger cars).

$$E[TS_c] = \sum_{v=1}^{\#vehicles} [P(y_v^{t_1} \leq y_1, y_v^{t_2} \geq y_1) + P(y_v^{t_1} \geq y_1, y_v^{t_2} \leq y_2)] \times 1 \tag{8}$$

where v represents all the vehicles in the segment of interest. Considering the Gaussian mixture model of the location of the vehicles (Eq. 3), the expected value of each cell ($c = [[t_1, t_2], [y_1, y_2]]$) of the time–space matrix can be estimated based on the following equation:

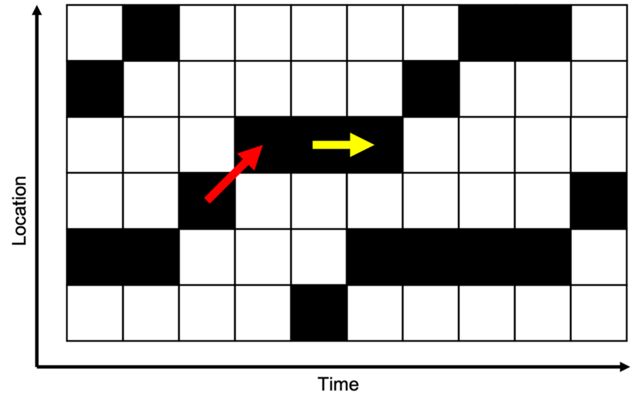


Fig. 2 Moving from one time–space cell to another

$$\begin{aligned} E[TS_c] = \sum_{v=1}^{\#vehicles} [&\int_{-\infty}^{y_1} \int_{-\infty}^{\infty} \sum_{i=1}^{M} P(m_{i,v}) \Phi(X_v^{t_1,m_i} | \mu_v^{t_1,m_i}, \Sigma_v^{t_1,m_i}) dxdy \\ &\times \int_{y_1}^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^{M} P(m_{i,v}) \Phi(X_v^{t_2,m_i} | \mu_v^{t_2,m_i}, \Sigma_v^{t_2,m_i}) dxdy \\ &+ \int_{y_1}^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^{M} P(m_{i,v}) \Phi(X_v^{t_1,m_i} | \mu_v^{t_1,m_i}, \Sigma_v^{t_1,m_i}) dxdy \\ &\times \int_{-\infty}^{y_2} \int_{-\infty}^{\infty} \sum_{i=1}^{M} P(m_{i,v}) \Phi(X_v^{t_2,m_i} | \mu_v^{t_2,m_i}, \Sigma_v^{t_2,m_i}) dxdy] \end{aligned} \tag{9}$$

It should be noted that in Eq. 9, the probability of passing the edge of the cell, $P(y_v^{t_1} \leq y_1, y_v^{t_2} \geq y_1)$ is approximated by multiplying the probability of the location of the vehicle being behind that point at the beginning of the given period multiplied by the probability of the vehicle being beyond that point at the end of the given period. This approximation is based on the consideration that the dependence between the

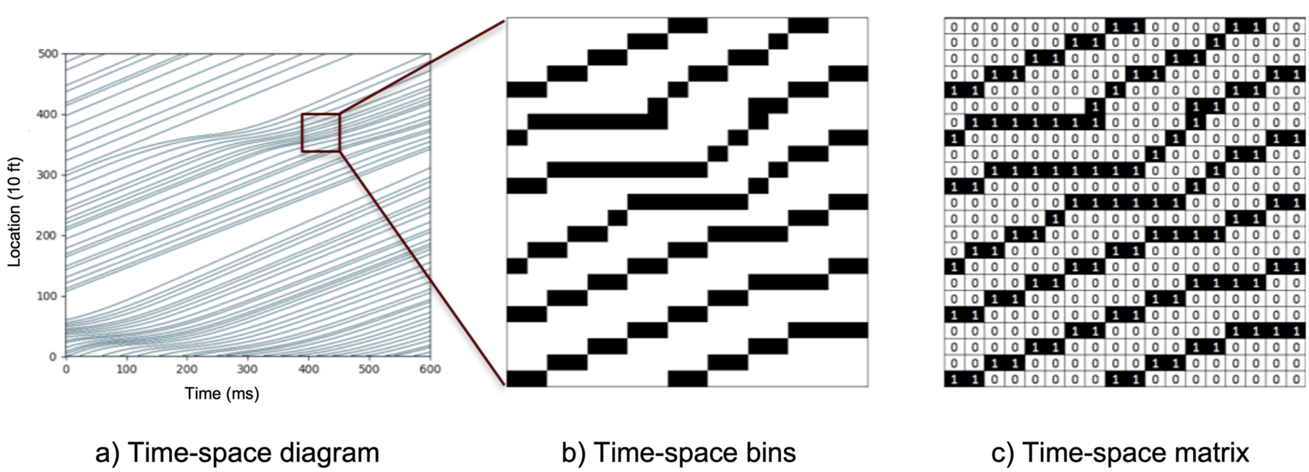


Fig. 1 Matrix representation of time–space diagram (Khajeh Hosseini and Talebpour 2019a)

mean (μ) and covariance (Σ) of the distributions for every time step t on the previous time steps' predictions is captured by the LSTM decoder component of the convolutional social pooling (Deo and Trivedi 2018a). Similar approximation is considered when estimating the probability remaining within the space boundaries ($P(y_v^{t_1} \geq y_1, y_v^{t_2} \leq y_2)$).

Macroscopic Level: Probabilistic Density and Flow Estimation

Probabilistic Density Estimation

Traffic density is defined by the number of vehicles occupying the unit length of the roadway. Based on this definition, the expected value of traffic density (K) can be estimated considering the expected number of vehicles on the roadway segment divided by the length of the roadway. The expected number of vehicles on a roadway segment with boundaries of $[x_1, x_2]$ and $[y_1, y_2]$ is the summation of the probability of every vehicle being on the segment multiplied by its contribution to the density, which is one for every passenger car. Note that this value can be adjusted to account for heavy vehicles and their impact on traffic flow dynamics (i.e., the concept of passenger car equivalent):

$$E[K(t)] = \frac{\sum_{v=1} P(y_1 < y_v^t < y_2, x_1 < x_v^t < x_2) \times 1}{y_2 - y_1} \tag{10}$$

Considering the probabilistic location of vehicles from Eq. 3 and the general definition of expected density in Eq. 10, the expected density can be estimated based on

$$E[K(t)] = \frac{\sum_v \int_{y_1}^{y_2} \int_{x_1}^{x_2} \sum_i P(m_{i,v}) \Phi(X_v^{t,m_i} | \mu_v^{t,m_i}, \Sigma_v^{t,m_i}) dxdy}{y_2 - y_1} \tag{11}$$

The numerator of Eq. 11 is comparable to the the occupancy probability in Eq. 7. The occupancy map is a matrix representation of the study area with values indicating the probability of that cell in space being occupied. Consequently, the expected density can also be estimated by summing all the values of the submatrix of the occupancy map that covers the target area and divided by the length of that segment:

$$E[K(t)] = \frac{\vec{1}^T O[[r_{y_1}, r_{y_2}];[c_{x_1}, c_{x_2}]]\vec{1}}{y_2 - y_1} \tag{12}$$

where O is the occupancy matrix and r_{y_1} and r_{y_2} are the row number of the occupancy map corresponding to y_1 and y_2, respectively. c_{x_1} and c_{x_2} denote the column number of the occupancy map corresponding to x_1 and x_2, respectively, and $\vec{1}$ is a vector of all ones. The vectors of all ones are used in the multiplication to facilitate summing all the members of the submatrix of the occupancy map.

Probabilistic Flow Estimation

The traffic flow rate is defined as the number of vehicles passing a point during a given period. Accordingly, the expected value of traffic flow (Q) at a given point can be estimated considering the expected number of vehicles passing the point over a given period. The probability of a vehicle passing a specific point of roadway over a period of time (e.g., t_1 to t_2) is the probability of the location of the vehicle being behind that point at the beginning of the given period and the vehicle being beyond that point at the end of the given period:

$$P_v^{crossing}(Y) = P(y_v^{t_1} < Y, y_v^{t_2} > Y) \tag{13}$$

where $P_v^{crossing}(Y)$ is the probability of vehicle v crossing point Y during the time period of $[t_1, t_2]$. The expected number of vehicles passing a specific point on the roadway over a fixed period is the summation of the probability of every vehicle passing that point during that period multiplied by its contribution to the flow rate, which is one for all passenger cars. Considering the Gaussian mixture model of the vehicle's location (Eq. 3), the expected flow rate, $E[Q]$, can be estimated based on the following equations:

$$E[Q(Y)] = \frac{\sum_v P(y_v^{t_1} < Y, y_v^{t_2} > Y) \times 1}{t_2 - t_1} \tag{14}$$

$$\begin{aligned} E[Q(Y)] = \sum_v [& \int_{-\infty}^{Y} \int_{-\infty}^{\infty} \sum_i P(m_{i,v}) \Phi(X_v^{t_1,m_i} | \mu_v^{t_1,m_i}, \Sigma_v^{t_1,m_i}) dxdy \\ & \times \int_{Y}^{\infty} \int_{-\infty}^{\infty} \sum_i P(m_{i,v}) \\ & \Phi(X_v^{t_2,m_i} | \mu_v^{t_2,m_i}, \Sigma_v^{t_2,m_i}) dxdy] / (t_2 - t_1) \end{aligned} \tag{15}$$

In Eq. 15, the probability of a vehicle passing a specific point of roadway over a period of $[t_1, t_2]$ is approximated by the probability of the location of the vehicle being behind that point at the beginning of the given period multiplied by the probability of the vehicle being beyond that point at the end of the given period. Note that similar to the density calculations, this approximation is based on the consideration that the dependence between the mean (μ) and covariance (Σ) of the distributions for every time step t on the previous time steps' predictions is captured by the LSTM decoder component of the convolutional social pooling (Deo and Trivedi 2018a).

Probabilistic Space–Mean Speed

There exists a variety of approaches to meaure space–mean speed ($\bar{U}_s$) in the literature (Hall 1996). A common definition of the space–mean speed is based on the average time taken by the vehicles to travel a specific segment of a roadway:

$$\bar{U}_s = \frac{D}{\frac{1}{N}\sum_i t_i} \tag{16}$$

where t_i is the time that took vehicle i to travel a specific roadway segment with the length D. One of the challenges with this definition is that it only averages the travel time of the vehicles that traveled the roadway segment completely. A more accurate definition of the space–mean speed, that is also more in line with the definition of the space–mean speed, considers all the vehicles traveling a specific segment of roadway over a given period of time (Hall 1996). In this definition, the space–mean speed is estimated by dividing the total distance traveled by all the vehicles by the total time spent by those vehicles on the specific segment of the roadway over a given period of time:

$$\bar{U}_s = \frac{d(A)}{t(A)} \tag{17}$$

where $d(A)$ and $t(A)$ denote the total distance traveled and total time spent by all the vehicles going through the time–space block A, respectively. This general definition of the space–mean speed is comparable to the ratio of Edie's (1963) generalized average flow and density of a time–space block A. Accordingly, the expected distance traveled by each vehicle in the time–space block of A with boundaries $[t_1, t_2]$ and $[y_1, y_2]$ can be estimated based on the following equations:

$$E[d_v] = \int_{y_1}^{y_2} P(y_v^{t_1} \leq y, y_v^{t_2} \geq y)dy \tag{18}$$

Equation 18 estimates the expected distance traveled by vehicle v in time–space block A, considering the probability of the vehicle going through every point of the segment of roadway in block A during the period of this time–space block. Similar to the concepts of the probabilistic time–space diagram, the time and space domains can be divided into smaller time–space bins. In this representation, the expected distance traveled by each vehicle can be estimated based on the following equation:

$$E[d_v] = \sum_i P(y_v^{t_1} \leq y_i, y_v^{t_2} \geq y_i)\Delta y \tag{19}$$

Equation 19 is similar to Eq. 18 except the integration is replaced with a summation. The expected time spent by each vehicle in the time–space block of A with boundaries $[t_1, t_2]$ and $[y_1, y_2]$ can be estimated based on the following equation:

$$E[t_v] = \int_{t_1}^{t_2} P(y_v^t \leq y_2, y_v^t \geq y_1)dt \tag{20}$$

Equation 20 estimates the expected time spent by vehicle v in the time–space block A, considering the probability of the vehicle being within the space boundaries of block A at every point of time within the period of time–space block. Similar to the concepts of the probabilistic time–space diagram, the time and space domains can be divided into smaller time–space bins. In this representation, the expected time spent by each vehicle can be estimated using summation instead of integration:

$$E[t_v] = \sum_t P(y_v^t \leq y_2, y_v^t \geq y_1)\Delta t \tag{21}$$

From Eq. 17, the space–mean speed ($\bar{U}_s$) can be estimated from the total expected distance traveled divided by the total time spent by all the vehicles on the study segment during the study period:

$$\bar{U}_s = \frac{E[d(A)]}{E[t(A)]} = \frac{\sum_v \sum_i P(y_v^{t_1} \leq y_i, y_v^{t_2} \geq y_i)\Delta y}{\sum_v \sum_t P(y_v^t \leq y_2, y_v^t \geq y_1)\Delta t} \tag{22}$$

Fundamental Equation of Traffic Flow

The traffic flow rate is defined as the number of vehicles passing a point of the roadway during a given period. The traffic flow rate is a function of location, $Q(y)$, and its average over the definite segment of roadway, $[y_1, y_2]$ can be estimated by integration of the traffic flow function:

$$\bar{Q} = \frac{\int_{y_1}^{y_2} q(y)dy}{(y_2 - y_1)} \tag{23}$$

The integration in Eq. 23 can be approximated by a summation over small roadway segments (Δy). Considering Eq. 14, the average traffic flow can be estimated based on the following equation:

$$\bar{Q} = \frac{\sum_i \sum_v P(y_v^{t_1} < Y, y_v^{t_2} > Y)\Delta y}{(t_2 - t_1)(y_2 - y_1)} \tag{24}$$

Traffic density is defined by the number of vehicles occupying the unit length of the roadway at a point of time. The traffic density is a function of time, and similar to the average flow, the average density over the definite time period, $[t_1, t_2]$ can be estimated by integrating the traffic density function:

$$\bar{K} = \frac{\int_{t_1}^{t_2} K(t)dt}{(t_2 - t_1)} \tag{25}$$

Taking into account Eq. 10 and replacing the integration of Eq. 25 with summation, the average density can be defined as

$$\bar{K} = \frac{\sum_t \sum_v P(y_1 < y_v^t < y_2)\Delta t}{(y_2 - y_1)(t_2 - t_1)} \tag{26}$$

From Eqs. 24 and 26, the ratio of the average flow and average density of a time–space block A with boundaries $[t_1, t_2]$ and $[y_1, y_2]$ is

$$\begin{aligned} \frac{\bar{Q}}{\bar{K}} &= \frac{\sum_i \sum_v P(y_v^{t_1} < Y, y_v^{t_2} > Y)\Delta y}{\sum_t \sum_v P(y_1 < y_v^t < y_2)\Delta t} \\ &= \frac{\sum_v \sum_i P(y_v^{t_1} < Y, y_v^{t_2} > Y)\Delta y}{\sum_v \sum_t P(y_1 < y_v^t < y_2)\Delta t} = \bar{U}_s \end{aligned} \tag{27}$$

The ratio of average flow and density of time–space block A in Eq. 27 is comparable to the space–mean speed in Eq. 22 when the order of summations are changed in both numerator and denominator. Accordingly, the fundamental relation among the average traffic flow, density and space–mean speed is preserved under the probabilistic formulations of this study:

$$\bar{Q} = \bar{K}\bar{U}_s \tag{28}$$

Data

One of the commonly adopted vehicle trajectory data sets is the FHWA Next generation Simulation Models (NGSIM) (U.S 2006). NGSIM is a well-known open-source trajectory data set collected in 2006 using digital cameras at different locations, including US Highway 101 and Interstate 80 freeway. The vehicle trajectories are extracted from the images of multiple cameras combined to create a single image that looks like an aerial shot. The NGSIM trajectory data contains the location of each vehicle at a frequency of 10 Hz over a 500 to 1000 ms stretch of roadway. The NGSIM data set contains three sets of 15 min trajectory data for each of the US Highway 101 and Interstate 80 freeway. This study uses the NGSIM data set for training the probablistice trajectory prediction model and evaluating the proposed probabilistic macroscopic estimates of this study. In this data set, every vehicle has a unique vehicle identification number that helps to identify the data points that belong to the same vehicle to store its trajectory. Each data point in the vehicle's trajectory includes multiple features, such as frame number (time), location (x and y), and velocity. For every point in the vehicle's trajectory, the lane feature is compared with the lane feature of the point 5 s later in the vehicle's trajectory (or the last point of the trajectory if the remaining trajectory is shorter than 5 s) to determine the lateral movement. Similarly, for every point in the vehicle's trajectory, the speed feature is compared with the average speed of the rest of the points in the trajectory up to the prediction period to determine the longitudinal movement.

Results and Discussion

The complete NGSIM data set is split into training (80%), validation (5%), and testing (15%) data sets. The updated probabilistic trajectory prediction model of Deo and Trivedi (2018a) is trained on the training portion of the data set to predict the trajectory of individual vehicles for different prediction periods, including 5, 10, 15, and 20 s. The trajectory prediction model takes as input 3 s of the track histories (locations over time) of a target vehicle and the vehicles within ± 27.4 ms in the longitudinal direction and within two adjacent lanes. The spatial configuration of surrounding vehicles of the target vehicle is summarized by a 13×3 matrix/grid. The output of the trajectory prediction model is the parameters of the conditional probability distribution of the location of the vehicle and the probability of individual maneuvers (based on Eq. 1) for time steps of 0.2 s for the next 5–20 s (depending on the model) in the future.

The trained trajectory prediction model is used to predict the future probabilistic trajectory of every vehicle in every tenth frame of a subsample of the NGSIM data set. The subsample is 5 min of the first set of trajectory data of US Highway 101. The predicted probabilistic trajectories are used to predict the proposed probabilistic microscopic and macroscopic traffic estimates (presented in the methodology section) for the middle 305 ms of the highway segment. The predicted macroscopic states are compared with the true traffic states calculated from the actual trajectories of the NGSIM data set. All the plots provided in this section are based on the selected subsample of the NGSIM data set.

Microscopic Level: Probabilistic Trajectory Prediction

The input to the maneuver-based model is 3 s of the past trajectory of a target vehicle as well as 3 s of the past trajectory of all its neighboring vehicles within the distance of ±27.4 meters. As it is mentioned in the methodology section, the deep learning model of Deo and Trivedi (2018a) is modified to predict the probability of lateral and longitudinal maneuvers jointly. Considering Eq. 1, the model is trained to predict Θ, the parameters of the conditional probability distribution of the location of the vehicle and $P(m_i|H)$, the probability of individual maneuvers for time steps of 0.2 s over the prediction period (e.g., 5, 10, 15 or 20 s). Table 1 presents the accuracy of four models trained for different

Table 1 Accuracy of the trajectory prediction model on the testing data set

Model	Location RMSE (m)	Lateral accuracy(%)	Longitudinal accuracy (%)
5 s	1.81	97.99	91.72
10 s	5.20	97.99	87.29
15 s	10.49	97.98	84.37
20 s	16.68	97.95	82.82

prediction periods. According to this table, the model's accuracy in predicting the vehicle's maneuver and location decreases with the increase in the prediction period. The trajectory prediction model predicts the probability of six different maneuvers (the combination of three lateral and two longitudinal movements) for every target vehicle. In Table 1, the lateral movement accuracy is estimated by counting the number of times that the maneuver with the highest probability matches the actual lateral movement of the vehicle. Similarly, the longitudinal movement accuracy is estimated by counting the number of times that the maneuver with the highest probability matches the actual longitudinal movement of the vehicle.

In this approach, the probabilistic location of the vehicle is described by a bivariate normal distribution (Eq. 2) for every maneuver, and the probability of each maneuver, $P(m_i)$, is predicted separately. Each maneuver is a combination of lateral and longitudinal movements. Figure 3 presents six plots of the predicted probabilistic density function of the location of a vehicle for 5 s in the future and for six different maneuvers. It should be noted that the color of points changes from dark blue to light yellow with the increase in the value of the probability density function of the points. According to this figure, the vehicle's location and its variance differ for various maneuvers.

Figure 4 presents five plots of the predicted probabilistic density function of the location of a single vehicle (Eq. 2) for different prediction periods. The predicted location of the vehicle in these plots is depicted based on the density function of the predicted bivariate normal distribution for the vehicle's location in the future for the second maneuver, and the color of points changes from dark blue to light yellow proportion to the density function value. The uncertainty in the vehicle's location increases with the increase in the prediction period. This increase in the uncertainty of the vehicle's location is also evident in the plots of Fig. 4 with the increase in the spread of the density function along both axes.

Figure 5 presents the plots of the bivariate Gaussian mixture models (Eq. 3) of the location of a single vehicle for different prediction periods. In most cases, the predicted probability of a single maneuver is significantly higher than the probability of other maneuvers. Consequently, the shape of the resulting Gaussian mixture model is closer to the shape of the density function of the maneuver with the highest probability.

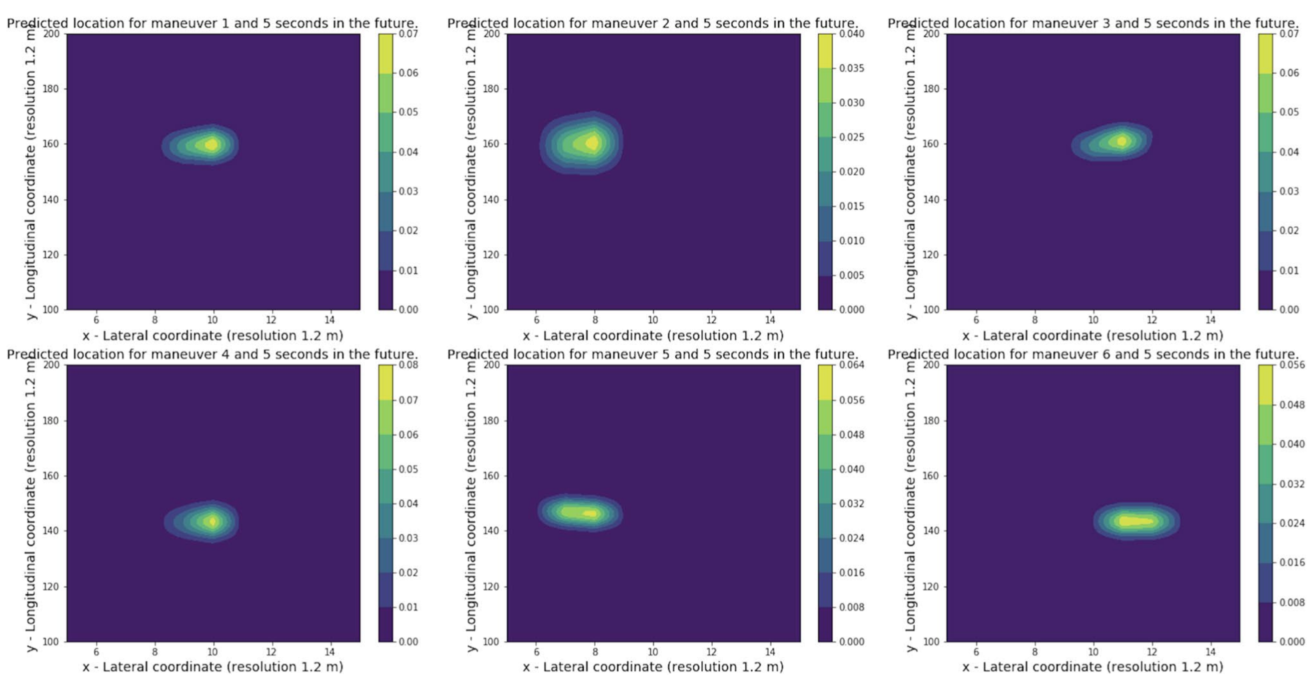


Fig. 3 Predicted location of a single vehicle for 5 s in the future and for different maneuvers

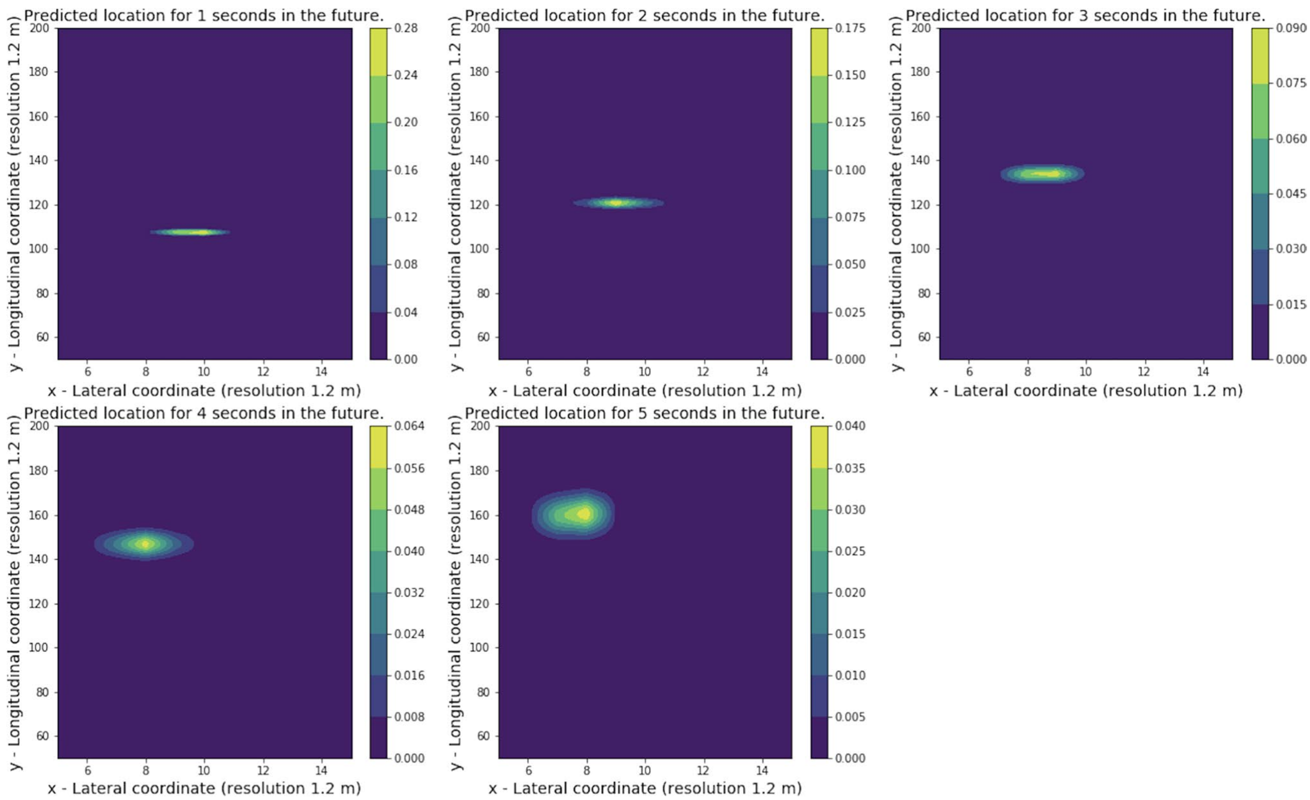


Fig. 4 Predicted location of a single vehicle for 5 s in the future and for the second maneuver

Probabilistic Occupancy Map

In this study, the space domain is divided into small cells of 1.2 by 1.2 ms, and the occupancy map of the roadway segment is represented by a matrix with values estimated using Eq. 7 indicating the probability of that cell being occupied. Figure 6 presents the probabilistic occupancy map of the segment of roadway for up to 5 s in the future. In these plots, the color of each cell changes from dark blue to light yellow proportional to the probability of the cell being occupied. The lighter the color of the cell indicates a higher probability of being occupied. The uncertainty in the vehicle's location increases with the increase in the prediction period. The increase in the uncertainty of the location of the vehicle is also evident in the plots of Fig. 6 with the spread of cells with low occupancy probability.

Probabilistic Time–Space Diagram

This study proposes the use of probabilistic time–space matrix representation of the time–space diagram. In this matrix, the time and space domains are divided into smaller cells, and the value of each cell is the expected value of that cell being occupied by a vehicle. Equation 8 estimates the probability of a vehicle passing through a time–space cell based on two probabilities either by crossing the beginning of the space domain of the cell during its time period or remaining within the exact space boundaries similar to its previous cell. The probability of each of these two movements in time and space as well as their summation are depicted in Fig. 7 for a single vehicle. For the specific vehicle presented in this figure, the probability of the vehicle remaining within the exact space boundaries similar to its previous cell (Fig. 7b) is relatively low compared to the probability of entering new space domains (Fig. 7a) particularly at the beginning of the time domain. With the increase in the prediction time, the uncertainty in the location of the vehicle increases, and the probability of the two possible movements from one time–space cell to another gets slightly closer to each other. Figure 7c presents the summation of the probability of these two possible movements in the time–space diagram. In addition, Figs. 8 and 9 present predicted probabilistic time–space matrix and corresponding actual time–space matrix of the study area for the next 5 s. Please note that these figures are provided to present an example of the probabilistic time–space matrix proposed in this study; a quantitative evaluation of the macroscopic traffic state measures proposed in this study is provided in the next section. The predicted probabilistic time–space matrix is comparable to the true time–space matrix for most lanes

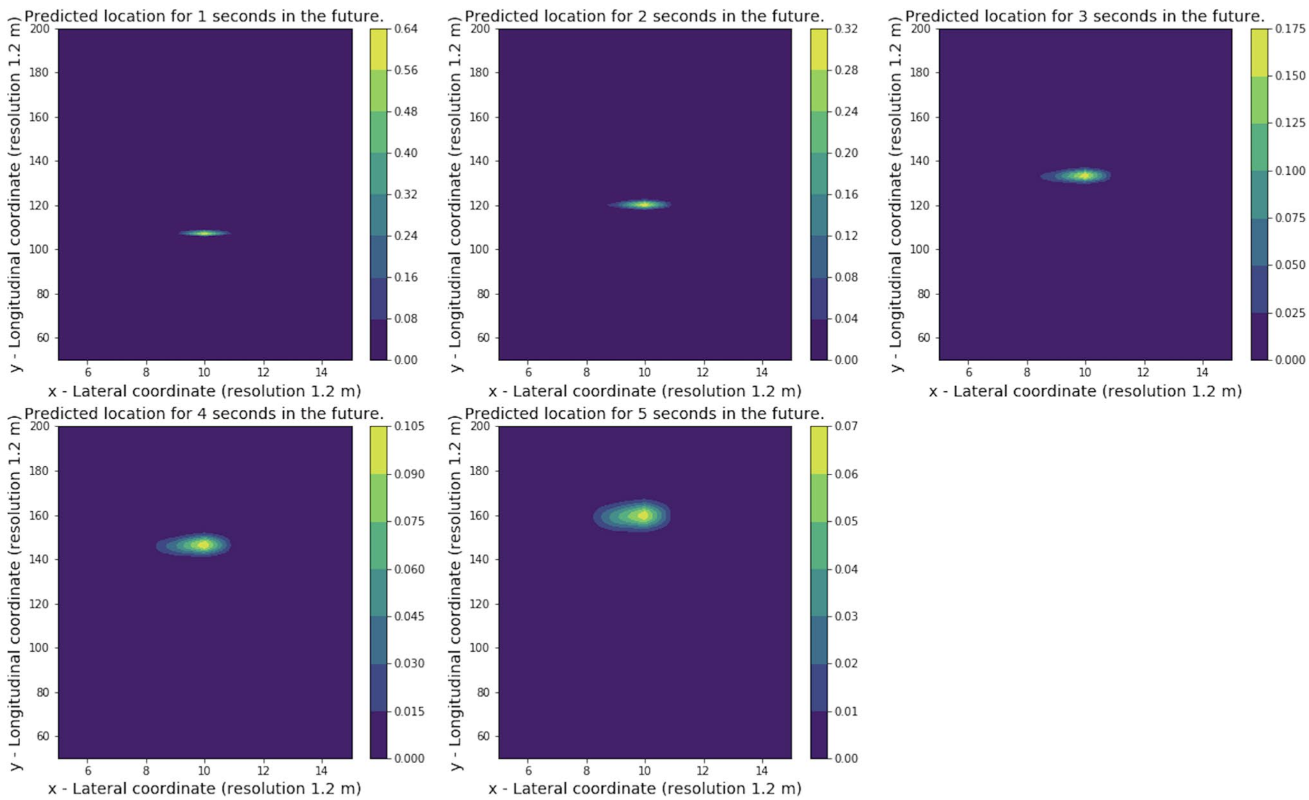


Fig. 5 Predicted location of a single vehicle for 5 s in the future considering all maneuvers

and vehicle trajectories. However, there are some instances of dissimilarity between predicted and actual trajectories in the future. The trajectory prediction model predicts the probability of different maneuvers at the beginning of the prediction period and the corresponding probabilistic trajectories for each maneuver. The trajectory prediction model predicts six sets of probabilistic trajectories (one trajectory per maneuver) based on the maneuvers initiated at the beginning of the prediction period. However, the trajectory prediction model is incapable of predicting a combination of maneuvers over the prediction period, for example, predicting maintaining the same lane and changing lanes further in the future. The accuracy of the maneuver prediction decays with the increase in the prediction period due to the possibility of the occurrence of new maneuvers further in the future. The majority of the discrepancies observed between the predicted trajectories and the actual trajectories in Figs. 8 and 9 are due to the occurrence of more than one maneuver over the prediction period, such as maintaining the lane and then lane changing (additional maneuvers) that happen further in the future, especially in the right-most lane (lane 5) and the auxiliary lane (lane 6). This drawback is due to one of the shortcomings of the trajectory prediction model adopted in this study that cannot capture a combination of maneuvers (more than one maneuver) as a result of future interactions or lane-changing maneuvers further in the future. The current trajectory prediction model predicts the trajectory of the vehicles individually considering the current observed environment, without taking into account the future trajectory of the surrounding vehicles and their interactions and the possibility of a combination of maneuvers. Accordingly, developing a probabilistic trajectory prediction model that could capture the future interactions among all the vehicles and predicts the trajectory of all the vehicles in the scene simultaneously could potentially address this shortcoming. However, developing such a holistic prediction approach is left as the future research direction.

Macroscopic-Level Traffic State Prediction

This study proposes probabilistic estimates of the traffic flow, density, and space–mean speed considering uncertainities in location of the vehicles. First, the trajectory prediction model is used to predict the probabilistic trajectory of every vehicle in every tenth frame of multiple subsamples of the NGSIM data set. Then, the probabilistic traffic states are estimated for the middle 305 ms of the highway segment considering the probabilistic trajectories of the vehicles.

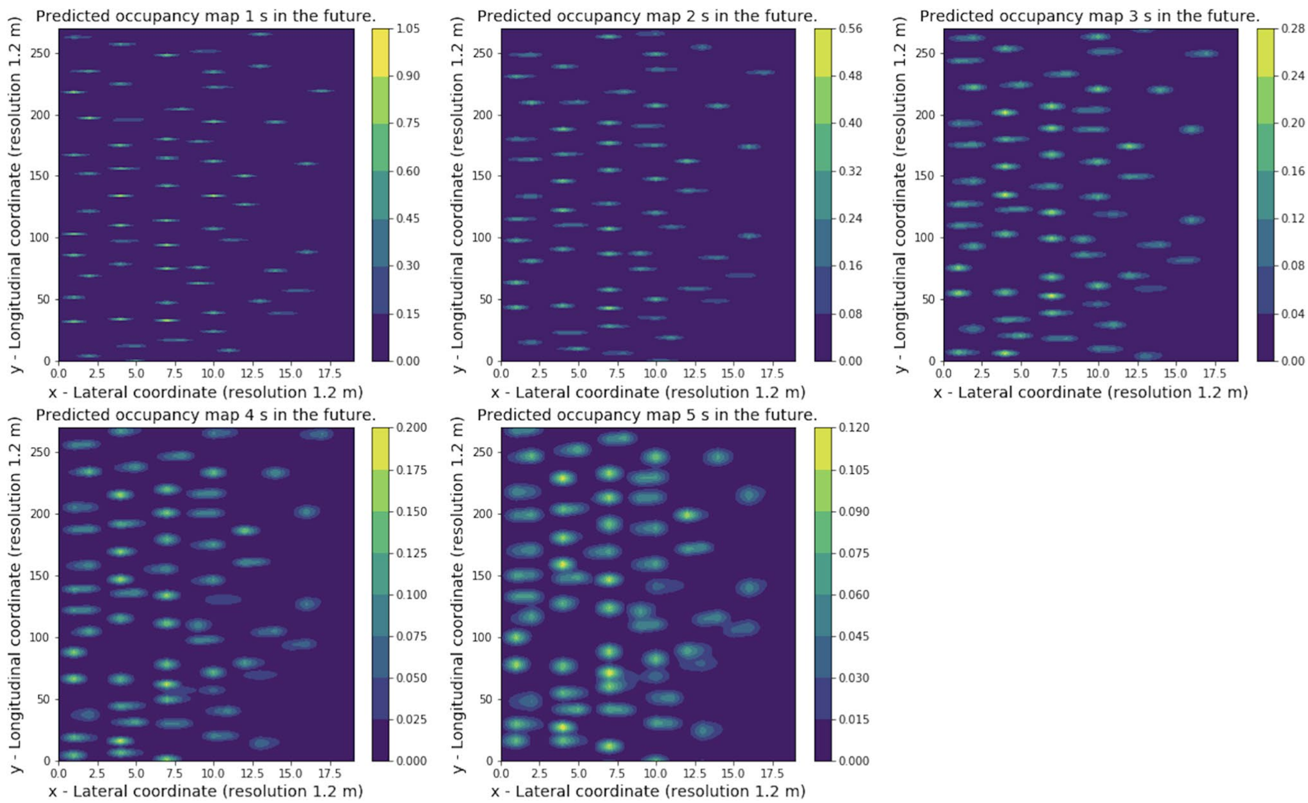
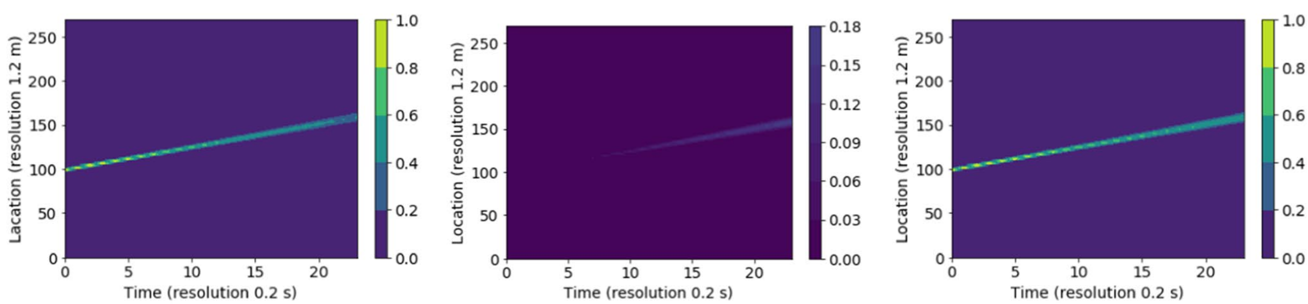


Fig. 6 Probabilistic occupancy map for 5 s in the future



(a) Probability of a vehicle crossing the beginning of space domain

(b) Probability of a vehicle remaining within the exact space boundaries of its previous cell

(c) Probability of a vehicle occupying cells of time-space matrix

Fig. 7 Probabilistic time–space matrix for a single vehicle

In this section, the expected traffic flow is estimated based on Eq. 24. In this equation, the duration of time $(t_2 - t_1)$ is considered equal to the prediction period (e.g., 20 s), and the expected traffic flow is estimated for every 3.05 ms of the middle 305 ms of the study segment (i.e., at 100 locations). The mean absolute percentage error for traffic flow rate ($MAPE_Q$) is estimated based on the following equation:

$$MAPE_Q = \frac{100}{n \times f} \sum_{f=1}^{F} \sum_{i=1}^{n} \frac{|E[Q_{i,f}] - Q_{i,f}|}{|Q_{i,f}|} \tag{29}$$

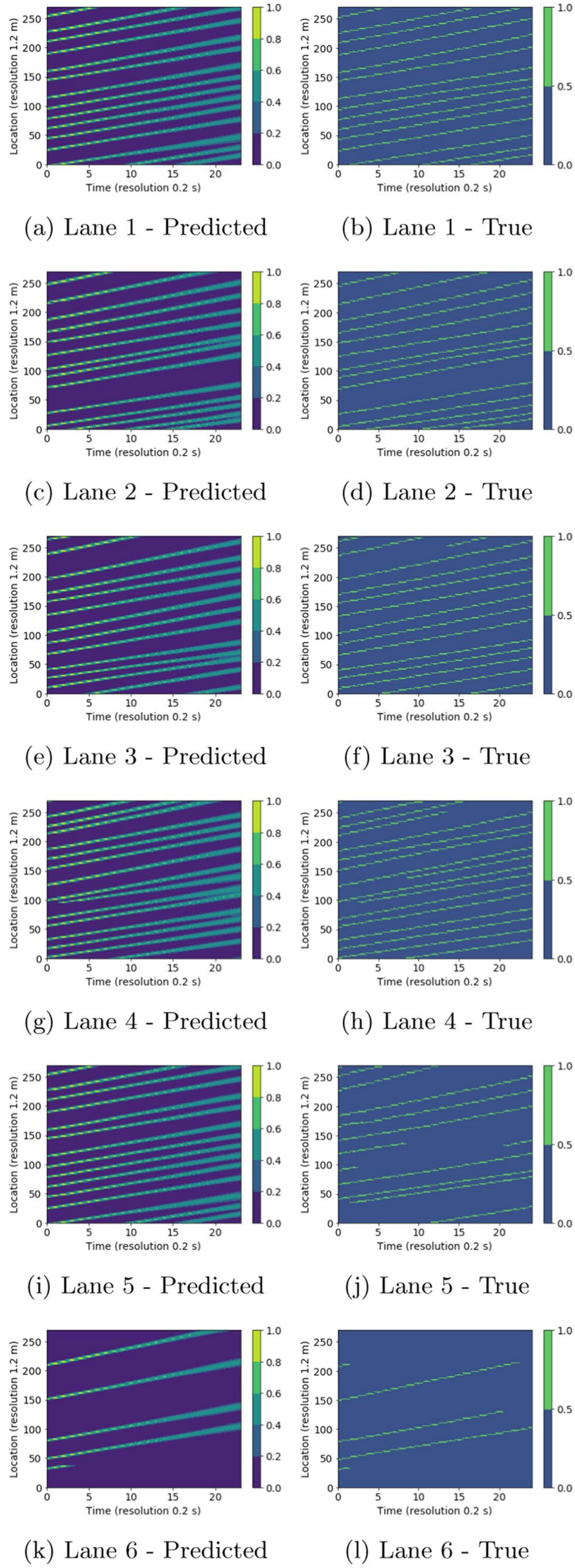


Fig. 8 Predicted and True time–space matrix for individual lanes and all lanes

In this equation, $E[Q_{i,f}]$ is the expected traffic flow rate at the ith on the study area estimated using the probabilistic trajectories predicted based on frame f and the probabilistic traffic flow estimate proposed in this study (Eq. 15), and $Q_{i,f}$ is the actual traffic flow rate at the ith location on the study area estimated based on the true trajectories of the vehicles. The traffic flow is estimated every 3.05 ms at 100 locations ($n = 100$).

The expected traffic density is estimated using Eq. 26 for the middle 305 ms of the highway segment and for every 0.2 s time step in the future up until the prediction period (e.g., 20 s). For example, if the prediction period is 20 s, the density of the middle 305 ms of the highway segment is estimated for 0.2, 0.4, 0.6,..., 20 s in the future. The mean absolute percentage error for traffic density ($MAPE_K$) is estimated using the following equation:

$$MAPE_K = \frac{100}{m \times F} \sum_{f=1}^{F} \sum_{j=1}^{m} \frac{|E[K_{j,f}] - K_{j,f}|}{|K_{j,f}|} \tag{30}$$

In this equation, $E[K_{j,f}]$ is the expected traffic density of the middle 305 ms of the highway segment at the jth time step in the future estimated using the probabilistic trajectories predicted based on frame f and the probabilistic traffic density estimate proposed in this study (Eq. 26). $K_{j,f}$ is the actual density at the jth time step based on the true trajectory of the vehicles. m is the number of time steps in the prediction period that density is estimated for.

The space–mean speed is estimated considering the expected total distance traveled (Eq. 19) and expected total time spent (Eq. 21) by all the vehicles traveling the middle 305 ms during the prediction period (e.g., 20 s). Consequently, one space–mean speed is predicted for every frame. Then, the mean absolute percentage error for space–mean speed ($MAPE_K$) is estimated using the following equation:

$$MAPE_{U_s} = \frac{100}{F} \sum_{f=1}^{F} \frac{|\bar{U}_{s,f} - U_{s,f}|}{|U_{s,f}|} \tag{31}$$

In this equation, $U_{s,f}$ is the space–mean speed predicted based on the expected total distance traveled and total time spent using the probabilistic trajectories predicted based on frame f. $U_{s,f}$ is the space–mean speed estimated based on the actual total distance and actual total time spent by all the vehicles based on the true trajectories.

Figure 10 presents the mean absolute percentage error (MAPE) of the traffic estimates for different prediction periods and for six different 5-min subsamples of the NGSIM data set. Table 2 presents the average traffic flow and density of the 5-min subsamples adopted from the six trajectory data sets. According to this table, the traffic conditions in the second and third data sets of the US 101 are slightly more congested (higher density and lower flow) compared to the

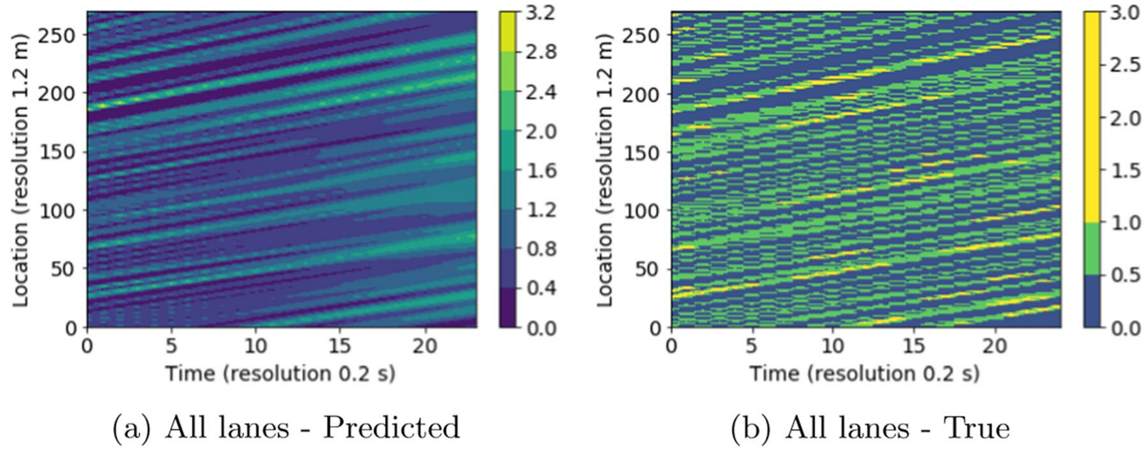


Fig. 9 Predicted and True time–space matrix for all lanes

Table 2 Traffic state of the subsamples of NGSIM data set

Data set	Average flow (vehicles per hour per lane)	Average density (vehicles per kilometer per lane)
NGSIM US101, first 15 min	1373.5	35.0
NGSIM US101, second 15 min	1191.1	44.6
NGSIM US101, third 15 min	1272.5	42.5
NGSIM I80, first 15 min	1238.4	36.1
NGSIM I80, second 15 min	1225.7	50.7
NGSIM I80, third 15 min	1308.4	52.3

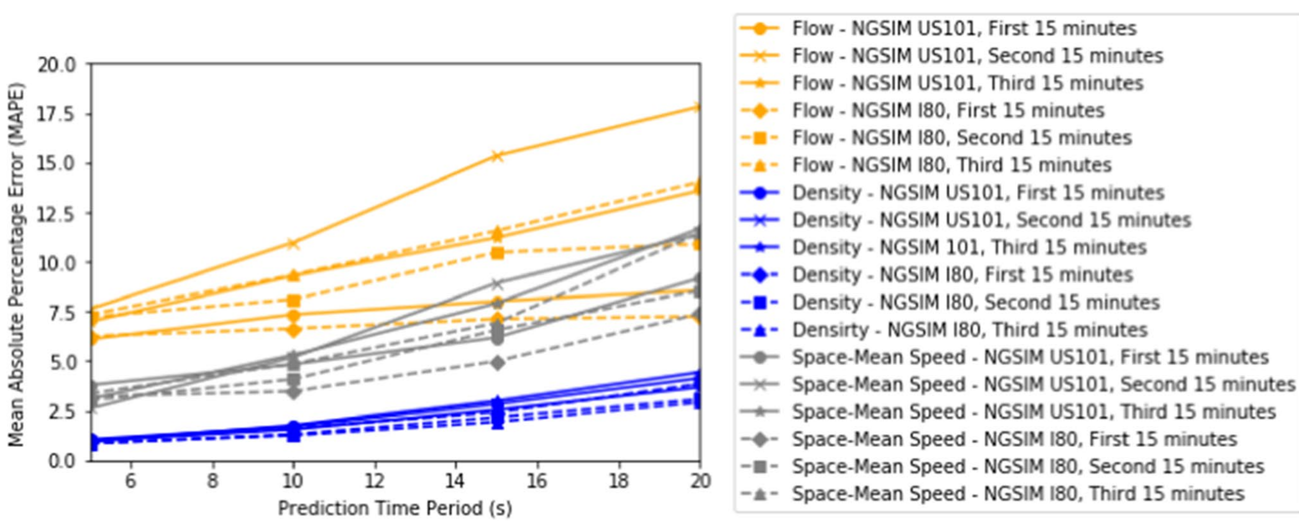


Fig. 10 Mean absolute percentage error (MAPE) of the traffic estimates for different prediction period

first data set of US 101. The second data set of the US 101 has the highest average density and the lowest traffic flow compared to other two data sets of US 101. According to this table, the traffic density of the subsamples of the I80 have a relatively higher average density. The second and third data sets of the I80 are more congested compared to the first data set of the I80.

According to Fig. 10, the mean absolute percentage error for traffic flow increases from 6.1 to 8.6 on the first data set of US101, 7.6 to 17.8 percent for the second data set of US101, and 7.0 to 13.6 percent for the third data set of US101 with the increase in the prediction period. The mean absolute percentage error for traffic flow changes between 6.3 to 7.3 percent for the first data set of I80, 7.2 to 10.9 percent for the second data set of I80, and 7.3 to 14.0 percent for the third data set of I80, with the increase in the prediction period. The MAPE reported for density in Fig. 10 is the average of the mean absolute percentage error over all the time-steps of the prediction period. According to this figure, the mean absolute percentage error for traffic density increases from 1.0 to 4.1 percent for the first data set of US101, 1.0 to 3.7 percent for the second data set of US101, and from 0.9 to 4.4 percent for the third data set of US101, with the increase in the prediction period. The mean absolute percentage error for traffic density increases from 1.0 to 3.9 percent for the first data set of I80, 0.8 to 3.1 percent for the second data set of I80, and 0.9 to 2.9 percent for the third data set of I80 with the increase in the prediction period. Moreover, the mean absolute percentage error for the space-mean speed increases from 3.8 to 9.2 percent for the first data set of US101, 2.6 to 11.4 percent for the second data set of US101, and from 3.1 to 11.7 percent for the third data set of US 101 with the increase in the prediction period. The mean absolute percentage error for space-mean speed increases from 3.2 to 7.4 percent for the first data set of I80, from 3.0 to 8.6 percent for the second data set of I80, and from 3.3 to 11.0 percent for the third data set of I80. In general, the model performance decays in predicting the traffic state for all three measures with the increase in the prediction period and traffic congestion. The performance decay is more significant for the traffic flow compared to density and space-mean speed.

These MAPE values are comparable to or even lower than most of the existing studies, showing the importance of capturing the interactions among individual vehicles in the prediction process. A recent example of such studies is a study by Khajeh-Hosseini and Talebpour (2019b). They proposed a deep learning traffic state prediction model based on the observed time-space diagram of the roadway to capture the interaction among the vehicles when predicting the traffic state. The performance of their deep neural network model in terms of MAPE on predicting the flow and density for the next 20 s on the NGSIM data set is 22.23 and 22.50, respectively. Moreover, they also compared the performance of their proposed model with other non-parametric models, including multilayer perceptron (MLP), support vector regression (SVR), and autoregressive integrated moving average (ARIMA). In their study, the ARIMA model trained on the NGSIM data set performed relatively better than the other non-parametric models with MAPE of 10.55 and 10.04 for predicting the average flow and density of the segment in the next 20 s. In this study, the average MAPE is 9.54 percent for measuring flow and 2.16 percent for measuring density. The traffic density and flow estimates top the performance of the models evaluated in Khajeh Hosseini and Talebpour (2019a). However, it should be noted that the performances of the probabilistic traffic state estimates are directly dependent on the performance of the probabilistic trajectory prediction model at the individual level. Therefore, the performances reported in this study are just proof of concept investigating the opportunity of probabilistic estimates of the traffic state. Accordingly, a more accurate trajectory prediction can even further improve the accuracy of the macroscopic predictions.

Figure 11 presents the MAPE for density estimate over time steps on the 5-min subsample of the first 15 min of the US Highway 101 data set of NGSIM, and for different trajectory prediction models trained for different prediction periods. The general trend of the error in the density estimation is also increasing with the increase in the time steps, due to the decrease in the accuracy of the trajectory prediction with the increase in the prediction period. The MAPE in density gets to 13.45 percent when predicting the density of the segment at 20 s in the future. According to Fig. 11, the MAPE in density estimates for different trajectory prediction models (trained for different prediction periods) performs relatively similar on their overlapping steps, specifically for models with prediction periods of 5, 10, and 15. However, the model with a prediction period of 20 s performs slightly better than the other models, indicating that the model trained to predict for a more extended period learns a better generalization. It should be noted that all the four trajectory prediction models are trained with five training epochs and random initialization of the parameters. The random initialization could also result in slightly different performances among the models. Please note that the MAPE reported in Fig. 10 is the average of the error when predicting density for all those time steps and for all the frames. However, Fig. 11 presents the MAPE for density estimates

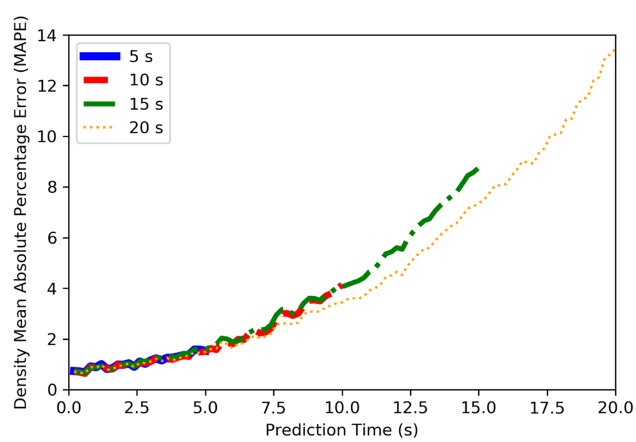


Fig. 11 Mean absolute percentage error (MAPE) of the density estimates for different prediction period

for every time step in the future separately, and the accuracy of the prediction decays with time. In Fig. 10, the MAPE is the average of the density errors over all the time steps; consequently, the MAPE over the whole prediction period (Fig. 10) is less than the MAPE of the density prediction at the end of the prediction period (i.e., at the time step for 20 s in the future).

Conclusion

This study proposes a novel methodology for probabilistic estimation and prediction of the traffic state based on probabilistic predictions of vehicle trajectories. At the microscopic level, this study develops a probability-based version of the time-space diagram of the vehicles. In this representation, the time and space are divided into smaller cells, where the value of each cell is the expected value of that cell being occupied by a vehicle. With the increase in the prediction period, the uncertainty in the location of the vehicles increases. As a result, more cells are expected to have values larger than zero further along the end of the prediction horizon but with smaller values compare to the cell values at the start of the prediction period. The predicted probabilistic time-space matrix is comparable to the actual time-space matrix for most lanes and vehicle trajectories. There are instances, however, where predicted and actual future trajectories diverge. In most cases, the discrepancies are due to lane-changing maneuvers in the future. This shortcoming is due to a limitation in the trajectory prediction model used in this study that does not capture future interactions that can lead to lane-changing maneuvers.

This study also proposes probabilistic estimates of flow, density, and space-mean speed at the macroscopic level. Moreover, this study proves that the fundamental relation among the average traffic flow, density, and space-mean speed is preserved under the probabilistic formulations of this study. The presented approach was tested using NGSIM US-101 and I-80 data sets. The mean absolute percentage of error (MAPE) for each of the probabilistic estimates at the macroscopic level is estimated for multiple subsamples of the NGSIM data set for different prediction periods. With an increase in prediction period time, the MAPE increases for all three traffic state estimates. As the uncertainty in probabilistic trajectories increases, the accuracy of the prediction decreases. The average MAPE of flow rises from 6.92 to 12.03 percent, with increasing the prediction period from 5 to 20 s. Moreover, the average MAPE of the space-mean speed rises from 3.17 to 9.89 percent, with an increase in the prediction period from 5 to 20 s. The average MAPE reported for density is the average of the MAPE over all the time-steps of the prediction period and changes from 0.93 to 3.52 percent from 5 to 20 s, respectively. It should be noted that the performances of the probabilistic traffic state estimates proposed in this study are directly dependent on the performance of the probabilistic trajectory prediction model at the individual level. Therefore, despite being more accurate than many existing approaches, this study should be treated as a proof-of-concept for probabilistic estimation of the traffic state.

The current trajectory prediction model adopted for this study predicts the trajectory of the vehicles individually considering the current observed environment, without taking into account the future trajectory of the surrounding vehicles and their interactions and the possibility of a combination of maneuvers. Accordingly, developing a probabilistic trajectory prediction model that could capture the future interaction among all the vehicles and predict the trajectory of all the vehicles in the scene simultaneously could potentially help the trajectory prediction for a more extended period. However, developing such a holistic prediction approach is left as the future research direction. Moreover, the traffic state prediction methodologies proposed in this study are based on the assumption of complete knowledge of the observed trajectory of all the vehicles in the study area considering a fully connected environment. Such information is either provided through a fully connected environment or based on the data measured and shared with connected and automated vehicles. In practice, the future traffic stream could be a mix of conventional, connected, and connected automated vehicles. While it is feasible to capture and monitor the majority of the vehicles in the traffic stream based on the sensory data from the connected and automated vehicles when their market penetration rate is above a minimum level, Talebpour et al. (2016) showed that due to signal interference, many information packets would not reach their destinations, even in a fully connected driving environment. Accordingly, it is critical to investigate and improve the proposed models in this study to predict the traffic state based on partial or incomplete data from the traffic stream in future research studies.

Author Contributions Conceptualization: MKH and AT; methodology: MKH and AT; formal analysis and investigation: MKH; writing-original draft preparation: MKH; writing-review and editing: AT; funding acquisition: AT; resources: AT; supervision: AT.

Funding This material is based upon work supported by the National Science Foundation under Grant No. 1826410 and Grant No. 2112650.

Availability of Data and Materials All the data utilized in this study is available upon request.

Declarations

Conflict of Interest The authors do not have any financial or non-financial conflicts of interest to declare.

Ethics Approval and Consent to Participate Not applicable.

References

Abbas MT, Jibran MA, Afaq M, Song W-C (2020) An adaptive approach to vehicle trajectory prediction using multimodel Kalman filter. Trans Emerg Telecommun Technol 31(5):3734

Batz T, Watson K, Beyerer J (2009) Recognition of dangerous situations within a cooperative group of vehicles. In: 2009 IEEE Intelligent Vehicles Symposium, IEEE pp 907–912

Broadhurst A, Baker S, Kanade T (2005) Monte carlo road safety reasoning. In: IEEE Proceedings. Intelligent Vehicles Symposium, 2005, IEEE, pp 319–324

Castro-Neto M, Jeong Y-S, Jeong M-K, Han LD (2009) Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. Expert Syst Appl 36(3):6164–6173

Chan KY, Dillon TS, Singh J, Chang E (2012) Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm. IEEE Trans Intell Transp Syst 13(2):644–654

Chen C, Hu J, Meng Q, Zhang Y (2011) Short-time traffic flow prediction with arima-garch model. In: Intelligent Vehicles Symposium (IV), 2011 IEEE, pp 607–612

Cui H, Radosavljevic V, Chou F-C, Lin T-H, Nguyen T, Huang T-K, Schneider J, Djuric N (2019) Multimodal trajectory predictions for autonomous driving using deep convolutional networks. In: 2019 International Conference on Robotics and Automation (ICRA) IEEE, pp 2090–2096

Deo N, Trivedi MM (2018a) Convolutional social pooling for vehicle trajectory prediction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp 1468–1476

Deo N, Trivedi MM (2018b) Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms. In: 2018 IEEE Intelligent Vehicles Symposium (IV) IEEE, pp 1179–1184

Dou Y, Yan F, Feng D (2016) Lane changing prediction at highway lane drops using support vector machine and artificial neural network classifiers. In: 2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), IEEE, pp 901–906

Duan Y, Lv Y, Liu Y-L, Wang F-Y (2016) An efficient realization of deep learning for traffic data imputation. Transp Res Part C 72:168–181

Edie LC (1963) Discussion of traffic stream measurements and definitions. Port of New York Authority

Eglese R, Maden W, Slater A (2006) A road timetabletm to aid vehicle routing and scheduling. Comput Oper Res 33(12):3508–3519

Elfar A, Xavier C, Talebpour A, Mahmassani HS (2018) Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles. Technical report

Elfar A, Talebpour A, Mahmassani HS (2019) Predictive speed harmonization in a connected environment: A machine learning approach. Technical report

Goli SA, Far BH, Fapojuwo AO (2018) Vehicle trajectory prediction with gaussian process regression in connected vehicle environment. In: 2018 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp 550–555

Hall FL (1996) Traffic stream characteristics. Traffic Flow Theory. US Federal Highway Administration, p 36

Huisken G, Berkum EC (2003) A comparative analysis of short-range travel time prediction methods. In: 82nd Annual Meeting of the Transportation Research Board, Washington, DC

Joseph J, Doshi-Velez F, Huang AS, Roy N (2011) A Bayesian nonparametric approach to modeling motion patterns. Auton Robot 31(4):383

Khajeh Hosseini M, Talebpour A (2019a) Traffic prediction using time-space diagram: a convolutional neural network approach. Transp Res Rec 2673(7):425–435

Khajeh-Hosseini M, Talebpour A (2019b) Back to the future: Predicting traffic shockwave formation and propagation using a convolutional encoder-decoder network. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, pp 1367–1372

Khajeh Hosseini M, Talebpour A, Shakkottai S (2019) Privacy risk of connected vehicles in relation to vehicle tracking when transmitting basic safety message type 1 data. Transp Res Rec 2673(12):636–643

Khosroshahi A (2017) Learning, classification and prediction of maneuvers of surround vehicles at intersections using lstms. PhD Thesis, UC San Diego

Klingelschmitt S, Platho M, Groß H-M, Willert V, Eggert J (2014) Combining behavior and situation information for reliably estimating multiple intentions. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings, IEEE, pp 388–393

Kumar SV, Vanajakshi L (2015) Short-term traffic flow prediction using seasonal Arima model with limited input data. Eur Transp Res Rev 7(3):21

Lawitzky A, Althoff D, Passenberg CF, Tanzmeister G, Wollherr D, Buss M (2013) Interactive scene prediction for automotive applications. In: 2013 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp 1028–1033

Lefèvre S, Vasquez D, Laugier C (2014) A survey on motion prediction and risk assessment for intelligent vehicles. ROBOMECH J 1(1):1–14

Liebner M, Baumann M, Klanner F, Stiller C (2012) Driver intent inference at urban intersections using the intelligent driver model. In: 2012 IEEE Intelligent Vehicles Symposium IEEE, pp 1162–1167

Luo W, Yang B, Urtasun R (2018) Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3569–3577

Lv Y, Duan Y, Kang W, Li Z, Wang F-Y et al (2015) Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst 16(2):865–873

Ma Y, Zhu X, Zhang S, Yang R, Wang W, Manocha D (2019) Trafficpredict: trajectory prediction for heterogeneous traffic-agents. Proc AAAI Conf Artif Intell 33:6120–6127

Mahmassani HS, Dong J, Kim J (2009) Dynasmart-p 1.5 user's guide and programmer's guide

Mozaffari S, Al-Jarrah OY, Dianati M, Jennings P, Mouzakitis A (2020) Deep learning-based vehicle behavior prediction for autonomous driving applications: a review. IEEE Trans Intell Transp Syst 23:33–47

Ossen SJL (2008) Longitudinal driving behavior: theory and empirics

Papageorgiou M, Hadj-Salem H, Middelham F (1997) Alinea local ramp metering: summary of field results. Transp Res Rec 1603(1):90–98

Polson NG, Sokolov VO (2017) Deep learning for short-term traffic flow prediction. Transp Res Part C 79:1–17

Rahim MA, Rahman MA, Rahman MM, Asyhari AT, Bhuiyan MZA, Ramasamy D (2021) Evolution of IoT-enabled connectivity and applications in automotive industry: a review. Veh Commun 27:100285

Seo T, Tchrakian TT, Zhuk S, Bayen AM (2016) Filter comparison for estimation on discretized pdes modeling traffic: Ensemble Kalman filter and minimax filter. In: 2016 IEEE 55th Conference on Decision and Control (CDC), IEEE, pp 3979–3984

Seo T, Bayen AM, Kusakabe T, Asakura Y (2017) Traffic state estimation on highway: a comprehensive survey. Annu Rev Control 43:128–151

Sheng Z, Xu Y, Xue S, Li D (2022) Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving. IEEE Trans Intell Transp Syst 23(10):17654–17665

Smith B, Scherer W, Conklin J (2003) Exploring imputation techniques for missing data in transportation management systems. Transp Res Rec 1836:132–142

Su H, Zhang L, Yu S (2007) Short-term traffic flow prediction based on incremental support vector regression. ICNC Third Int Conf Natl Comput 1:640–645

Sun Y, Work DB (2014) A distributed local Kalman consensus filter for traffic estimation. In: 2014 IEEE 53rd Annual Conference on, Decision and Control (CDC), IEEE, pp 6484–6491

Talebpour A, Mahmassani HS, Hamdar SH (2013) Speed harmonization: evaluation of effectiveness under congested conditions. Transp Res Rec 2391(1):69–79

Talebpour A, Mahmassani HS, Bustamante FE (2016) Modeling driver behavior in a connected environment: integrated microscopic simulation of traffic and mobile wireless telecommunication systems. Transp Res Rec 2560(1):75–86

Tran Q, Firl J (2014) Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings, IEEE, pp 918–923

U.S. Federal Highway Administration (2006) Next generation simulation (ngsim). https://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm

Van Lint J, Van Hinsbergen C (2012) Short-term traffic and travel time prediction models. Artif Intell Appl Crit Transp Issues 22(1):22–41

Wang Y, Papageorgiou M, Messmer A, Coppola P, Tzimitsi A, Nuzzolo A (2009) An adaptive freeway traffic state estimator. Automatica 45(1):10–24

Wang R, Work DB, Sowers R (2016) Multiple model particle filter for traffic estimation and incident detection. IEEE Trans Intell Transp Syst 17(12):3461–3470

Wilby MR, Díaz JJV, Rodríguez Gonźlez AB, Sotelo MÁ (2014) Lightweight occupancy estimation on freeways using extended floating car data. J Intell Transp Syst 18(2):149–163

Yoon S, Kum D (2016) The multilayer perceptron approach to lateral motion prediction of surrounding vehicles for autonomous vehicles. In: 2016 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp 1307–1312

Zheng Z, Su D (2016) Traffic state estimation through compressed sensing and Markov random field. Transp Res Part B 91:525–554

Zhou Y, Ravey A, Péra M-C (2019) A survey on driving prediction techniques for predictive energy management of plug-in hybrid electric vehicles. J Power Sources 412:480–495