Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 5926663, 17 pages
https://doi.org/10.1155/2022/5926663

Research Article

Incorporating Traffic Flow Model into A Deep Learning Method for Traffic State Estimation: A Hybrid Stepwise Modeling Framework

Yuyan Annie Pan,[1] Jifu Guo,[2] Yanyan Chen,[1] Siyang Li,[1] and Wenhao Li[1]

[1]*Beijing University of Technology, Beijing Key Laboratory of Traffic Engineering, Beijing, China*
[2]*Beijing Transport Institute, No. 9 LiuLiQiao South Road, Fengtai District, Beijing, China*

Correspondence should be addressed to Yanyan Chen; cdyan@bjut.edu.cn

Received 4 June 2022; Revised 19 September 2022; Accepted 11 October 2022; Published 4 November 2022

Academic Editor: Tomio Miwa

Traffic state estimation (TSE), which reconstructs the traffic variables (e.g., speed, flow) on road segments using partially observed data, plays an essential role in intelligent transportation systems. Generally, traffic estimation problems can be divided into two categories: model-driven approaches and data-driven approaches. The model-driven method is commonly used to solve TSE efficiently and calibrate the parameters of these models. The data-driven method requires a large amount of historical observed traffic data in order to improve performance accurately. In order to combine the advantages of model-driven and data-driven methods, this paper proposed a hybrid framework incorporating the traffic flow model into deep learning (TFMDL) modeling that contains both model-driven and data-driven components. This paper focuses on highway TSE with observed data from loop detectors. We build a hybrid cost function to adjust the weights of model-driven and data-driven proportions. We then evaluate the proposed framework using the open-access performance measurement system (PMS) dataset on a corridor of US I-405 in Los Angeles, California. The experimental results show the advantages of the proposed TFMDL approach in performing better than several benchmark models in terms of estimation accuracy and data efficiency.

1. Introduction

The intelligent transportation system (ITS) needs real-time traffic information and an accurate estimate of the future traffic state to make effective decisions so that users can obtain reliable information (e.g., travel time, route choice). The bottleneck is a critical topographic feature leading to traffic congestion because of its low capacity in the road network [1]. Accurate estimation and prediction of the traffic state could reduce traffic jams, pollution, and urban traffic pressure [2]. Therefore, traffic state estimation (TSE) plays a crucial role in urban traffic planning, control, operations, and other transportation services.

Traffic state estimation refers to inferring traffic state variables (i.e., flow, density, speed) on road segments using partially observed traffic data by traffic detectors [3]. The methods of estimating traffic conditions can be briefly divided into two main approaches: model-driven and data-driven. References [4, 5] have defined the data-driven method to infer traffic states based on the dependence learned from historical data using statistical or machine learning methods, such as convolutional neural networks (CNN), and long-term memory (LSTM) [6]. The model-driven approach is based on a priori knowledge of traffic dynamics, usually described by a physical model, e.g., the Lighthill-Whitham-Richards (LWR) models [7, 8], and the cell transmission model (CTM) [9].

Macroscopic traffic state variables denote the traffic conditions on road links in a traffic segment. Traffic flow models describe the dynamic behavior of vehicles and drivers. After calibrating the traffic flow model, the model can estimate and predict future traffic states more accurately

[10]. The model-driven approach uses real-time data as input to evaluate the traffic state in unobserved areas.Select traffic flow models could capture the relationship among traffic flow variables [11]. However, these traffic flow models were proposed based on ideal assumptions and conditions that may not fully capture real-world traffic phenomena.

With the rapid development of information processing technology, the neural network is widely used in estimation and prediction because of its advantages in modeling and optimizing nonlinear systems. It is necessary to choose a suitable neural network model according to the study problem and requirements [12]. The neural network trains the model using historical data and then estimates and predicts the unobserved traffic state. However, for the hidden anomalies and mutation points in the data sequence, due to the lack of training data, it is difficult for machine learning to make accurate estimations and predictions, which is especially obvious during the flow breakdown in the bottleneck. The traffic flow model could establish a stable fundamental diagram with some key parameters, especially since it is more accurate and can be interpreted for the mutation situation (e.g., breakdown in bottleneck). Therefore, combining the machine learning method with the traffic flow model is a practical way to build a hybrid model, improving the accuracy of estimation and prediction of traffic state.

There are some researchers focused on the hybrid method of modeling for estimation and prediction [5, 13, 14]. The latest representative study was published by [15, 16] and combined with LWR, CTM, and Deep Learning Neural Network to estimate traffic conditions accurately. Their study focused on the conditions with CAV and probe vehicle data in general conditions. However, it is more difficult and critical for traffic state estimation at the flow breakdown in the bottleneck. Therefore, this paper filled the gap in traffic state estimation and prediction under congestion conditions with a hybrid stepwise modeling framework, which can be further used for both modeling of the fundamental diagram and neural network intended to improve traffic state estimation and prediction accuracy. We hope to shed light on improving traffic state estimation accuracy across a hybrid modeling framework, especially suitable for the congestion situation. It makes possible the integration of the advantages of both model-driven and data-driven approaches while overcoming the weaknesses of either alone.

The rest of this paper is organized as follows. Section 2 reviews related work on traffic state estimation with a traffic flow model and a deep learning neural network. Section 3 builds a hybrid stepwise modeling framework for TSE by integrating the traffic flow model and deep learning to estimate and predict traffic states. Section 4 presents an experiment using the proposed method to verify the performance of the proposed method and details the evaluation of the experiment results and discussion. Section 5 concludes our work and suggests a research direction in the future.

2. Literature Review

2.1. Macroscopic Traffic Flow Model for Traffic State Estimation. Most traditional methods for traffic state estimation and prediction are model-driven. These models were proposed based on ideal assumptions and conditions that fully capture the traffic flow phenomena of the real world. The fundamental diagram is one of the most basic concepts in traffic flow theory, known as the relationship between flow-speed, density-flow, and density-speed. The fundamental diagram appears in almost all traffic flow theories related to the traffic state variables and contains remarkable information about traffic characteristics [17]. It needs to calibrate the critical parameters in the adopted traffic flow model to obtain traffic flow characteristics in a link, section, or corridor. Those key parameters include free-flow speed, critical speed, critical density, jam density, and ultimate capacity, among some other parameters [18]. Suppose the adopted traffic flow model could accurately capture the traffic state variables observed by the detector. The model-driven approach can precisely estimate and predict traffic states in unobserved areas [3]. Therefore, it is essential to use a suitable traffic flow model to systematically capture the fundamental relationships between traffic flow, speed, and density.

Many calibration methods have been proposed to identify the parameter values in the fundamental diagram. Some studies are based on sensor data [19, 20], and others are based on GPS data [21, 22]. In addition, some studies combine sensor data with probe vehicle data [23, 24]. The fundamental diagram relates traffic state variables to each other and is a core part of traffic flow theory.It plays a significant role in most traffic state estimation methods. Selected the Payne traffic flow model [25] to integrate into the Kalman filter's observation equation, [26] proposed a framework to estimate and predict travel time. Reference [27] highlighted a particle filter method to estimate traffic conditions in freeway networks based on a speed-extended CTM-based macroscopic model. Based on stochastic nonlinear macroscopic traffic flow modeling and an extended Kalman filtering model, [28] proposed a general approach to study traffic state estimation in freeways or networks.

Furthermore, the LWR model is the most widely used for model-driven TSE. Reference [29] modified the LWR partial differential equation (PDE) to incorporate a correction term that reduces the discrepancy between the observed and estimated traffic flow state. Reference [30] examined a dynamic first-order modeling approach to solving the traffic data-cleaning problem. Reference [31] developed a Bayesian recursive framework to estimate traffic flow states based on a freeway traffic model with aggregated states.

The model-driven TSE method could represent the physics of traffic flow and has high explanatory power for traffic flow characteristics. This means that even if the estimation is inaccurate, it would be possible to identify the reference value and confidence intervals. However, the performance of TSE might be poor if the model or calibrated models are inappropriate. Therefore, model-driven TSE requires careful selection of the traffic flow model and calibration of model parameters by measuring traffic flow data.

2.2. Data-Driven Method for Traffic State Estimation. Data-driven methods mainly rely on plentiful historical data, it predicts the traffic flow state by identifying the current traffic flow pattern from the collected historical traffic data, without detailed descriptions of intrinsic traffic dynamics based on traffic flow models [32]. Unlike the typical fundamental diagram, the data-driven methods are flexible in incorporating additional explanatory variables, such as time of day, day of week and weather [33].

The Neural Network (NN) and Deep neural network (DNN) have been widely and successfully applied to traffic state estimation and prediction because of their advantages in modeling and optimizing nonlinear systems [34, 35]. Two popular DNN modeling tools are most used in TSE, CNN method is good at dealing with visual data [36], while LSTM specializes in sequential data processing [37]. Many studies apply neural networks to enhance the accuracy and robustness of traffic state estimation. Reference [38] proposed a CNN-based method that learns traffic as images and predicts traffic speed in a large-scale network with high accuracy. Reference [39] introduced Bayesian learning to neural networks for the accurate prediction of transportation estimation. However, relying on historical data means these methods could have worse results if there are irregular data or cases. Some studies focused on the irregular situation with a data-driven approach. To name only a few, [40] researched the prediction accuracy for traffic flow by machine learning and the DNN approach during the disaster recovery period. Reference [41] developed a framework that combines a linear model with a deep learning model that captures the nonlinear spatial-temporal traffic flow phenomenon. It needs to be emphasized that the case study in this paper did not analyze the irregular data or long-term trends. In addition, [42] presented a new approach to predicting freeway travel time based on recurrent neural networks. Their model could be capable of dealing with spatial-temporal relationships implicitly. Furthermore, traditional neural network approaches for traffic flow estimation and prediction are usually single-task learning models, and there are some researchers focused on the multitask learning (MTL) method for traffic flow forecasting. Reference [43] highlight the multitask learning-based neural network method that improves generalization for traffic flow forecasting. Reference [44] incorporated multitask learning into its deep architecture and investigate homogeneous MTL and heterogeneous MTL for traffic flow prediction. Reference [45] proposed a deep learning-based multitask learning Gated Recurrent Units (MTL-GRU) method to improve the forecasting accuracy of traffic flow and speed.

To learn time series with long time spans, long short-term memory (LSTM) neural networks have been effectively applied in short-term traffic flow prediction. Reference [46] applied LSTM to effectively capture nonlinear traffic dynamics using remote microwave sensor data collected for the Beijing network. Along this line, [47] proposed a deep and embedding learning approach to estimate and predict traffic flow characteristics. The model consists of an embedding component, a CNN component, and an LSTM component. Motivated by the success of CNNs and LSTMs, [48] proposed a spatiotemporal image-based convolutional networks approach to predict the network-wide traffic state. Reference [49] proposed a novel Spatio-Temporal Graph Convolutional Networks (STGCN) to enhance the time series prediction accuracy for multi-scale traffic networks. In addition, some other machine learning approaches have been developed for traffic flow prediction, such as Diffusion Convolutional Recurrent Neural Network (DCRNN) [50], Spatio-Temporal Attention-based Neural Network (STANN) [51], spatial-temporal graph convolutional network (ASTGCN) [52], spatiotemporal feature selection algorithm (STFSA) with CNN model [53].

Compared with model-driven methods, data-driven methods do not require explicit theoretical assumptions, such as the fundamental diagram and partial differential equation. They cannot guarantee that the results will be physically feasible. In addition, data-driven approaches typically rely on relatively simple models and heavily rely on the amount and variety of data used for training. However, reliance on historical data means these methods could have worse results if there is irregular data or long-term trends. The computational cost of training and learning can be very high [3]. In addition, the process can be considered a black box, meaning inductive insights are difficult to obtain. Therefore, some traffic experts constantly criticize the data-driven method for understanding the building blocks of the models [54].

2.3. Hybrid of Traffic Flow Models and Neural Networks. In a congestion bottleneck or adverse weather conditions, driver behavior becomes more complex and nonlinear, which leads to new challenges in traffic state estimation. However, neither model-driven nor data-driven methods alone can estimate such complex behavior with an allowable level of accuracy. Therefore, to mitigate the limitations of the previous TSE approaches, many researchers have studied a lot of research combining the advantages of data-driven and model-driven to compensate for their weaknesses.

A hybrid method could approximate the traffic states described by some traffic flow models in nonlinear differential equations and discover unknown model parameters using observed traffic data. It can improve the accuracy and explanation of model-driven or data-driven methods alone in traditional. However, many areas and questions still need to be explored in developing the hybrid method. The traffic flow model was first proposed in the deep learning method by [55, 56] to solve the TSE problem with the PDE feature. Moreover, [4] proposed a framework to solve nonlinear partial differential equations TSE problems. This problem has both a physics-informed neural network and a physics-uninformed neural network. Whereafter, [5] proposed a hybrid framework, a physics-informed deep learning model, to combine second-order traffic flow models and neural networks for the TSE. And they used experiments to demonstrate the proposed model in terms of data efficiency and estimation accuracy. Reference [15] presented a physics-informed deep learning method to improve the accuracy of the TSE problem. And they demonstrated the powerful capability of the model in utilizing limited data for real-time TSE. Reference [57] developed a hybrid model that integrates both linear regression models with a neural network layered by LSTM. They used this model to forecast the demand for taxi rides, and the results showed that this model could enhance the forecasting performance. Reference [58] proposed a kinematic wave-based deep CNN to estimate traffic speed dynamics. This method combines the advantages and mitigates the disadvantages of model-driven and data-driven approaches for TSE.

3. Methodology

3.1. Conceptual Illustration. It is critical for TSE to collect ample historical data. The recorded data is used to calibrate or identify traffic flow states in a model-driven method. Real-time information is collected at the moment when the traffic state is unknown. The conceptual illustration between traffic state estimation, real-time data, and historical data is shown in Figure 1. By combining the long-term historical and real-time data could estimate the unknown traffic state between the two locations

3.2. Model-Driven Method. The fundamental diagram (FD) appears in almost all traffic flow theories as it relates to the traffic state variables and contains remarkable information about traffic characteristics [59]. Some important traffic flow model parameters (free speed v_f, critical density k_c, critical speed v_c, jam density k_j, and capacity c) could influence the shape of the FD, especially in the undersaturated regime (A) and oversaturated regime (B). Figure 2 is the fundamental diagram with undersaturated and oversaturated conditions.

In this paper, the model-driven adopted of a new S-shaped three-parameter ($S3$) traffic flow model [60] to estimate the relationships among three fundamental variables (i.e., flow, speed, and density). This traffic flow model could capture the flow-speed-density relationship simultaneously under all possible densities, especially at high densities.

The S3 model has the following density-speed relationship, as shown in (1)

$$v = \frac{v_f}{\left[1 + \left(k/k_c\right)^m\right]^{2/m}}, \quad (1)$$

where, v_f is the free-flow speed and k_c is the critical density. The term m is introduced to control the smoothness or flatness of the curves for different planes across the feasible range of traffic congestion conditions. q, k, and v are flow, density, and speed, respectively. According to the conservative law $q = kv$, we could derive the density-flow relationship, as shown in (2)

$$q = \frac{k \cdot v_f}{\left[1 + \left(k/k_c\right)^m\right]^{2/m}}. \quad (2)$$

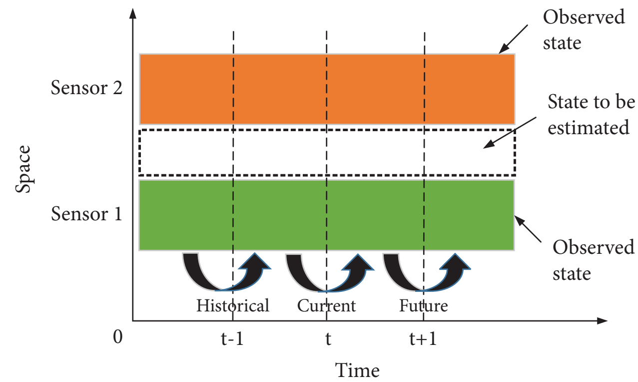


Figure 1: Illustration of traffic state estimation considers the spatial-temporal correlation.

Traffic flow models use empirical data from devices and develop hypotheses to calibrate and validate model parameters [34]. We could calibrate those key parameters v_f,k_c and m using observed multi-day speed and flow data.

For the results to be scale-independent, we normalize the average flow and average density as follows:

And then, we could derive the estimation traffic flow on a new day at the timeline, namely the estimated speed $\widehat{v}_{tf}(x,t)$ and estimated flow $\widehat{q}_{tf}(x,t)$, as shown in (3) and (4):

$$\widehat{v}_{tf}(x,t) = \frac{v_f}{\left[1 + \left(k_{obs}(x,t)/k_c\right)^m\right]^{2/m}}, \quad (3)$$

$$\widehat{q}_{tf}(x,t) = \frac{k_{obs}(x,t) \cdot v_f}{\left[1 + \left(k_{obs}(x,t)/k_c\right)^m\right]^{2/m}}. \quad (4)$$

For the results to be scale-independent, we normalize the flow data and speed data as follows

$$\widetilde{q}(x,t) = \frac{q(x,t) - q_{\min}(x,t)}{q_{\max}(x,t) - q_{\min}(x,t)}, \quad (5)$$

$$\widetilde{v}(x,t) = \frac{v(x,t) - v_{\min}(x,t)}{v_{\max}(x,t) - v_{\min}(x,t)}, \quad (6)$$

$q_{\max}(x,t)$ and $q_{\min}(x,t)$ are the global maximum and minimum flow over all time indexes. Similarly, $v_{\max}(x,t)$ and $v_{\min}(x,t)$ are the global maximum and minimum speed over all time index.

According to (5) and (6), we could normalize the observed flow data $\widetilde{q}_{obs}(x,t)$, the observed speed data $\widetilde{v}_{obs}(x,t)$, the estimated flow data $\widetilde{q}_{tf}(x,t)$ by model-driven method, and the estimated speed data $\widetilde{v}_{tf}(x,t)$ by model-driven method, respectively.

Then we could calculate the error between the normalized observed data and normalized estimated data by the traffic flow model.

$$err_1 = \left|\widetilde{q}_{obs}(x,t) - \widetilde{q}_{tf}(x,t)\right| + \left|\widetilde{v}_{obs}(x,t) - \widetilde{v}_{tf}(x,t)\right|. \quad (7)$$

The calibrated traffic flow model could accurately demonstrate the relationship among observed traffic state variables. The model-driven method could precisely estimate

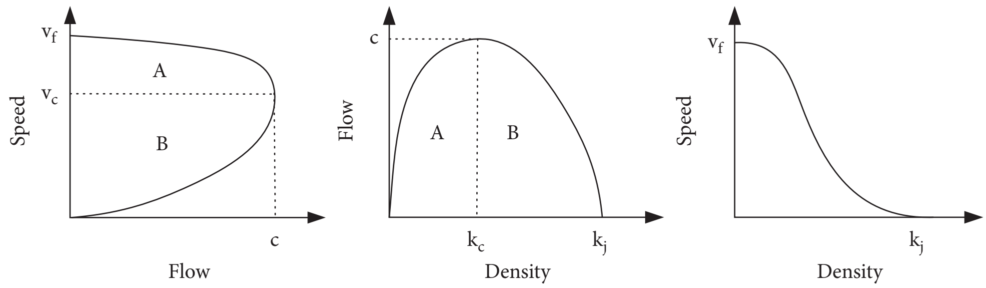


Figure 2: The curve of speed-flow, flow-density, and speed-density relationships.

and predict traffic states in unobserved areas and provide a higher resolution of the estimated traffic flow state. The traffic flow relationship curve can be obtained by calibrating the selected traffic flow model using observed actual data from a loop detector, which can obtain explainable parameters. However, the robustness of the calibrated traffic flow model needs to be improved in the event of traffic accidents or adverse weather conditions.

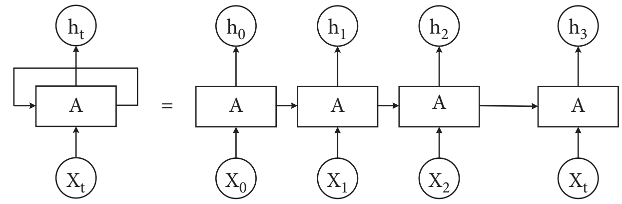


Figure 3: The structure of LSTM.

3.3. Data-Driven Method. Another solution to improve the accuracy of traffic state estimation is the data-driven method. Instead of relying on the calibrated traffic flow models based on empirical or historical observed data, it applies the rapidly emerging machine learning techniques to recognize the relationship between traffic state variables. Based on the real-time input data, the data-driven method could capture the traffic flow pattern in a specific environment, such as congestion and peak periods. Therefore, the data-driven approach is expected to be more reliable when the measured traffic data is abnormal or missing.

Deep learning is a subset learning method of machine learning techniques. Long short-term memory is an artificial recurrent neural network (RNN) architecture used in the deep learning area. It is a novel recurrent neural network architecture developed to capture the long-term temporal dependency for short-term travel speed or flow estimation and prediction. As the Long short-term memory recurrent neural network (LSTM) has been recommended by many existing works [6], especially for time series forecasting. Thus, LSTM was selected as the deep learning algorithm in the proposed model in the paper. A typical LSTM model consists of an input, output, forget, and external input gate. The structure of LSTM is shown in Figure 3.

The first step in the LSTM model is the forget gate f_t. It uses the sigmoid function δ to remove unnecessary information.

$$f_t = \delta\left(W_{ff} \cdot h_{t-1} + W_{xf} \cdot x_{t-1} + b_f\right). \tag{8}$$

The second step is the input gate i_t and external input gate g_t. It then updates and decides the new information.

$$\begin{aligned} i_t &= \delta\left(W_{gg} \cdot h_{t-1} + W_{xg} \cdot x_{t-1} + b_g\right), \\ g_t &= \tanh\left(W_{gg} \cdot h_{t-1} + W_{xg} \cdot x_{t-1} + b_g\right). \end{aligned} \tag{9}$$

The third step is mainly to update the old cell state s_t.

$$s_t = f_t \cdot f_t + i_t \cdot g_t. \tag{10}$$

The last step is the output gate o_t.

$$\begin{aligned} o_t &= \delta\left(W_{oo} \cdot h_{t-1} + W_{xo} \cdot x_{t-1} + b_o\right), \\ h_t &= o_t \cdot \tanh\left(s_t\right), \end{aligned} \tag{11}$$

δ denotes the standard logistics sigmoid function.

$$\delta(x) = \frac{1}{1 + e^{-x}}, \tag{12}$$

where $W_{ff}, W_{xf}, W_{gg}, W_{xg}, W_{oo}, W_{xo}$ is weight and b_f, b_g, b_o is bias. h_t, x_t is variable, $\otimes$ denotes the Hadamard product.

Before building the LSTM model, the parameters of the model should be determined. The parameters automatically adjusted by the model can be obtained through training and learning in model training without human intervention. The hyperparameters need to be set manually, including the number of input layers, the number of hidden layers in the network, the number of nodes in hidden layers, the number of output layers, activation function, loss function, optimization function, step size, iteration times, etc. The network structure of the LSTM model established in this paper includes one input layer, three hidden layers, and one output layer. In this study, the same data are used to test LSTM models with hidden layers 1, 2, and 3, respectively, and the model with the minimized error is selected by comparing the test results.

Similarly, we used (5) and (6) to normalize the observed data and estimated data by the data-driven method. Then we could calculate the error between the normalized observed data and normalized estimated data by the LSTM model, as shown in (13)

$$err_1 = \left|\tilde{q}_{obs}(x,t) - \tilde{q}_{dl}(x,t)\right| + \left|\tilde{v}_{obs}(x,t) - \tilde{v}_{dl}(x,t)\right|, \tag{13}$$

where, $\tilde{q}_{dl}(x,t)$ and $\tilde{v}_{dl}(x,t)$ are the normalized estimated flow and speed by the data-driven method. As previously mentioned, $\tilde{q}_{obs}(x,t)$ and $\tilde{v}_{obs}(x,t)$ are the normalized observed flow and observed speed data at location x in time t, respectively.

3.4. TFMDL-Based Traffic State Estimation Model. The traffic flow model could simulate the evolution of the traffic flow state, especially in particular circumstances like accidents or bad weather. We could obtain accurate and interpretable traffic state estimation results by calibrating the traffic flow model. In addition, considering the potential resource waste and privacy problems associated with the massive use of traffic data, traffic flow models can be used to minimize data utilization while achieving estimation accuracy. The neural network could improve the accuracy of data estimation, especially in the normal condition of good weather. The estimation and prediction results have high reliability [61]. Therefore, this paper proposes a hybrid model that could accurately estimate congestion duration on the bottleneck, the start and end times of congestion. It is suitable for specific weather and can also improve the real-time accuracy of traffic state estimation and prediction. When we derive the estimated data using different methods, we could use a hybrid stepwise modeling method to further improve the accuracy of TSE by incorporating the traffic flow model into the deep learning-based method (TFMDL). The flowchart of the hybrid method framework is shown in Figure 4. Algorithm 1 gives the algorithm framework of the proposed hybrid stepwise traffic state estimation framework.

The total cost function J of TFMDL for traffic state estimation is calculated by the following (14):

$$J = \alpha \cdot J_{tf} + (1-\alpha) \cdot J_{dl}, \tag{14}$$

where J is the total cost function, J_{tf} is cost from the model-driven method, and J_{dl} is cost from the data-driven method. The range of α is 0 or 1.

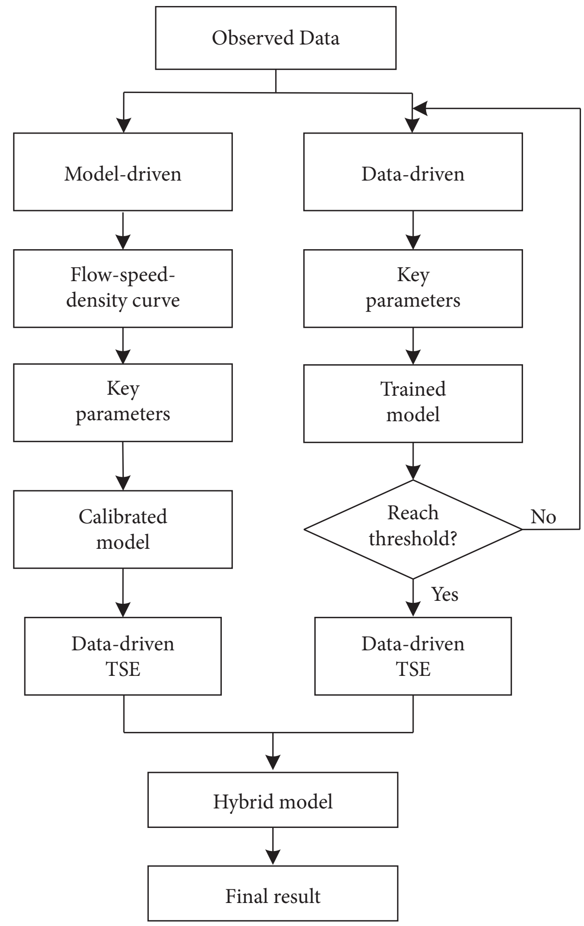


Figure 4: The framework of the proposed TFMDL structure.

4. Case Study

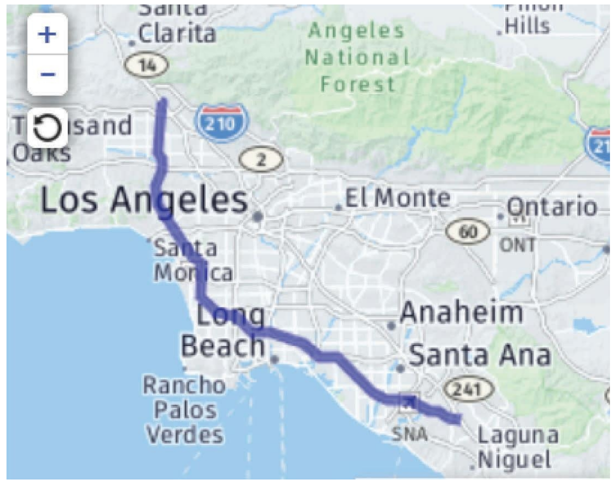
4.1. Data Description. The data collection site is the I-405 corridor in Los Angeles, the United States, with uninterrupted traffic flow, as shown in Figure 5. There are nine detectors with absolute mileage (9.87–13.74) in July 2021 (0: 00–24: 00). It includes 31 days of data, including 27 days of training data and 4 days of testing data. The observed data includes traffic flow, speed, and occupancy. The detailed detector information is shown in Table 1.

4.2. Evaluation for Estimation Results. The performance of evaluation criteria focuses on traditional measures that are scalable and quickly explained: mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R-square (R^2). The definitions of them are shown as follows

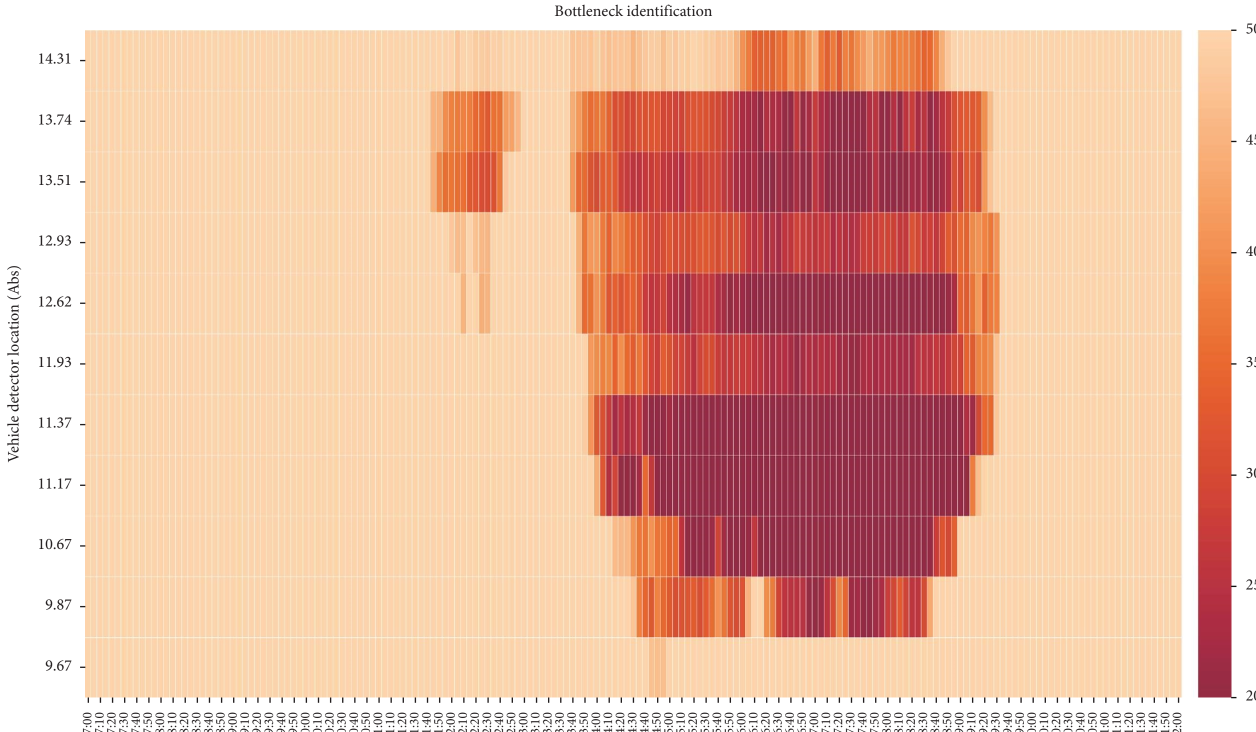
$$\begin{aligned}
MSE &= \frac{1}{N}\sum_{x=1}^{N}\left(\text{measured}_x - \text{estimated}_x\right)^2,\\
RMSE &= \sqrt{\frac{\sum_{x=1}^{N}\left(\text{measured}_x - \text{estimated}_x\right)^2}{N}},\\
MAE &= \frac{1}{N}\sum_{x=1}^{N}\left|\text{measured}_x - \text{estimated}_x\right|,\\
MAPE &= \frac{100\%}{N}\sum_{x=1}^{N}\left|\frac{\text{measured}_x - \text{estimated}_x}{\text{measured}_x}\right|,\\
R^2 &= 1 - \frac{\sum_{x=1}^{N}\left(\text{measured}_x - \text{estimated}_x\right)^2}{\sum_{x=1}^{N}\left(\text{measured}_x - \text{averaged}_x\right)^2},
\end{aligned} \tag{15}$$

where $measured_x$ is the measured traffic counts for x, $estimated_x$ is the estimated traffic counts for x, $averaged_x$ is the average traffic counts for x, N is the total number of counts.

4.3. Estimated Flow and Speed by Traffic Flow Model. We use the traffic flow model $S3$ to fit the measurement training data, and the fundamental diagram of the case data is shown



(a)



(b)

FIGURE 5: Data acquisition. (a) Freeway segment location. (b) Bottleneck identification.

Step 1: Select multi-day data to calibrate the key parameters in the traffic flow model, and we could derive the calibrated traffic flow model, including the flow-speed curve, flow-density curve, and speed-density curve.
Step 2: Using the calibrated model to estimate and predict the traffic flow state in an unobserved area (with partially observed data).
Step 3: Calculate the **error 1 (model-driven error)** between the observed and estimated value by the model-driven method.
Step 4: Build a machine learning model and divide the data set into training and testing data sets.
Step 5: Check if you have reached the maximum training threshold and return to Step 4 or step forward.
Step 6: Calculate **error 2 (data-driven error)** between the observed and estimated value by the data-driven method.
Step 7: Calculate the total cost function with the TFMDL model.

ALGORITHM 1: Implementation for building TFMDL-base TSE framework.

in Figure 6. The calibrated key parameters of the traffic flow are shown in Table 2.

Then we could derive the estimated speed and flow by calibrated traffic flow model, as shown in (16) and (17).

$$\widehat{v}_{tf}(x,t) = \frac{70}{\left[1 + \left(k_{obs}(x,t)/30\right)^{4.8}\right]^{0.42}}, \quad (16)$$

$$\widehat{q}_{tf}(x,t) = \frac{k_{obs}(x,t) \cdot 70}{\left[1 + \left(k_{obs}(x,t)/30\right)^{4.8}\right]^{0.42}}. \quad (17)$$

4.4. Estimated Flow and Speed by Machine Learning Model. The network structure of the LSTM model established in this paper includes one input layer, three hidden layers, and one output layer. The number of nodes of each hidden layer is 64, the optimization function is Adam, the excitation function is Sigmoid, and the number of iterations of each training is 100. This paper uses the same data to test LSTM models with the same parameters and hidden layers of 1, 2 and 3, respectively, for the selection of network layers. The comparison results are shown in Tables 3 and 4. It can be seen that the LSTM model with a three-layer hidden layer has the best performance, which has fewer RMSE, MAE, and MAPE than other tests. The estimated flow and speed by the data-driven method in the Los Angeles case are shown in Figures 7 and 8, respectively.

To determine the benefits of the proposed method, several benchmark models are also tested in the case study, and we provide more comparative analysis by referring to some literature [14, 62]. The benchmark algorithms include the LSTM model, k-Nearest Neighbor (KNN), ARIMA model, and Random Forest (RF). We further compare the observed data with estimated data by different methods, as shown in Tables 5 and 6. We could find that the proposed TFMDL framework is better than the other benchmark methods.

As shown in Table 5 and 6, we could find that the proposed TFMDL framework has better performance than the other models. The findings can be summarized as follows:

(1) On the whole, the precision of the proposed method is slightly higher than the other methods, especially in estimating the state of traffic speed. The proposed model shows the best performance in MAE, RMSE, and MAPE, with values of 1.20, 2.24, and 2.75%, respectively.

(2) In terms of flow estimation, the proposed model shows the best performance in MAPE, with a value of 4.26%, while the KNN method performed well in MAE and RMSE, respectively.

4.5. Spatiotemporal Analysis of TFMDL. We further evaluate the average performance and stability of the proposed TFMDL framework, and Figure 9 depicts the observed and estimated speed profiles on four days in spatiotemporal. The solid blue line indicates the prominent difference between the observation and the estimation. It shows the transition between the free-flow and the congested states. Reference [63] systematically analyze the spatiotemporal transition characteristics of traffic conditions near an on-ramp bottleneck. The accuracy of speed estimations on different days and for different absolute miles is shown in Table 7. Figure 10 shows the average comparison error in each location for four days. The findings can be summarized in Figures 9 and 10 Table 7 and by the following:

(1) The proposed framework could accurately estimate speed on all datasets. At absolute mile 12.93 on 07/31/2017, it has the lowest MAE (0.83 m/s), while at absolute mile 9.87 on 07/27/2017, it has the highest MAE (1.89 m/s). The average MAE of the dataset is all less than 1.8 m/s. In terms of MAPE, the average MAPE of the dataset are all less than 7%. Absolute mile 12.93 on 07/31/2017 has the lowest MAPE (1.87%), while absolute mile 10.67 on 07/26/2017 has the highest MAPE (10.54%).

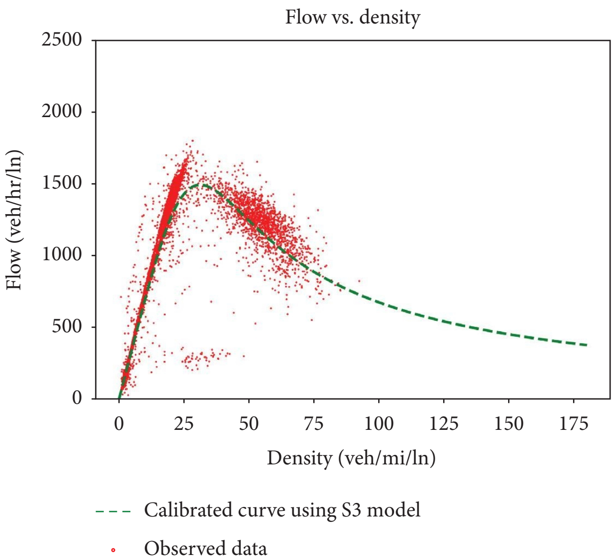
(2) The proposed framework could accurately estimate speed in oversaturated conditions. It may depend mainly on the traffic flow model ($S3$), which could better fit real-world observations, particularly under high traffic density ranges. However, compared to datasets with oversaturated conditions (i.e., absolute miles 9.87, 10.67, and 11.17), better estimation results are observed on datasets with light traffic conditions (i.e., absolute miles 12.93 and 13.74). We could further analyze whether model-driven or data-driven approaches occupy a more significant proportion during congestion period estimation in Section 4.6.

(3) This framework could capture the state of traffic congestion and noncongestion conditions, and we could obtain the start time of congestion and the end time of congestion. Obviously, the estimation effect is significantly better in the noncongestion period (i.e., green in the figure) and slightly worse in the congestion period (i.e., yellow and red in the figure).

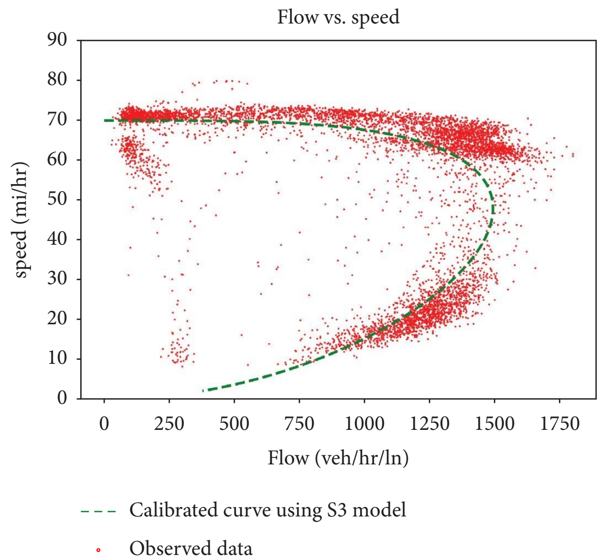
(4) It is observed that, compared to observed speed profiles, the estimated speed profiles look not smooth enough, and we can find some sharp speed changes. The main reason may be related to the parameter α in equation (14) only selecting integer values of 0 or 1. Considering the decimal value of the parameter α is expected to improve the estimation performance under congested traffic conditions, and it is also the further work of this study.

(5) As shown in Figure 10, the estimated average error on 07/31/2017 is the lowest, and the estimated average error on 07/28/2017 is the highest. Combining the last four Figures 9(e)–9(h) in Figure 9, we could find that this is probably related to the overall congestion range. It is obvious that the spatiotemporal congestion on 07/28/2017 is on a much larger scale than the other days.

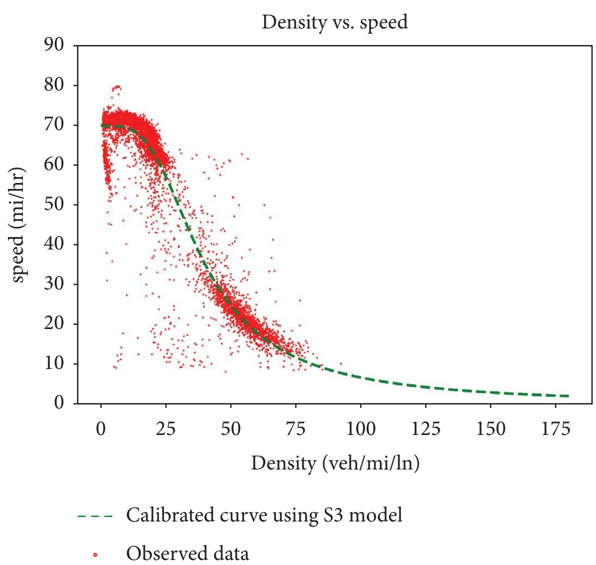
4.6. Model-Driven and Data-Driven Proportion Analysis. According to Figure 9, we could find that the congestion duration is 12:00 am 8:00 pm. Therefore, we further analyze the



(a)



(b)



(c)

Figure 6: Fundamental diagram in Los Angeles case. (a) flow-density relationship. (b) speed-flow relationship. (c) speed-density relationship.

Table 1: Descriptive information about detectors.

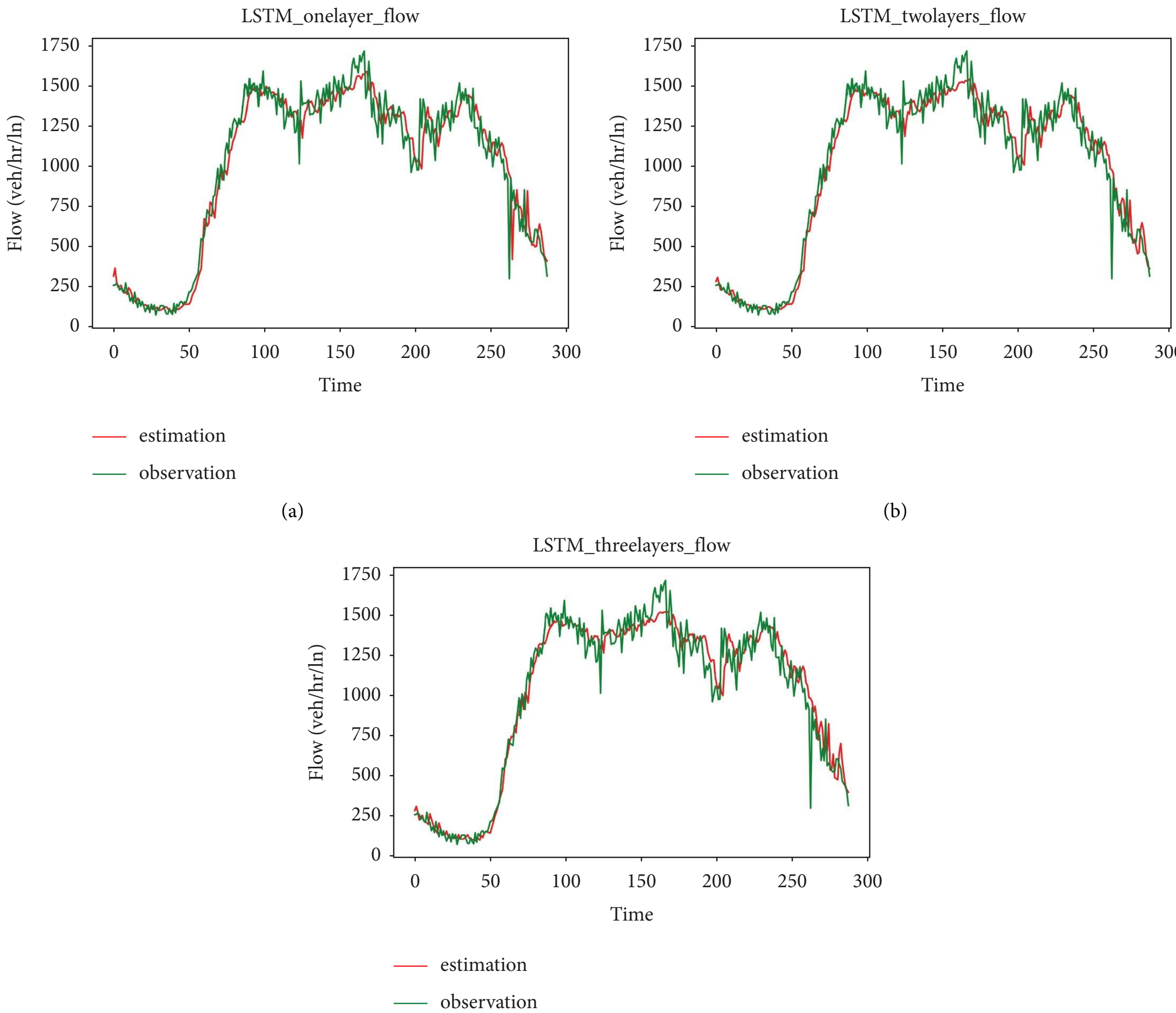
Detector ID	Absolute mile	Length (mile)	Number of lanes
1	13.74	0.57	4
2	13.51	0.23	4
3	12.93	0.58	5
4	12.62	0.31	5
5	11.93	0.69	6
6	11.37	0.56	6
7	11.17	0.2	6
8	10.67	0.5	6
9	9.87	0.8	4

Table 2: Key parameters in the Los Angeles case.

Free-flow speed (km/hr)	Critical speed (km/hr)	Critical density (veh/km/ln)	Ultimate capacity (veh/km/ln)	m
70	52.48	30	1574	4.8

Table 3: Compare evaluation indicators in the machine learning model.

Layers	RMSE	MAE	MAPE
One-layer	109.61	80.31	10.75
Two-layers	107.36	76.34	10.69
Three-layers	**105.87**	**76.19**	**10.38**



(a) (b) (c)

Figure 7: Estimated flow by the data-driven method in the Los Angeles case. (a) one-layer; (b) two-layer. (c) three-layer.

proportion of model-driven and data-driven in our proposed TFMDL framework all day (as shown in Figure 11) and during the peak period (as shown in Figure 12). The corresponding proportion analysis of model-driven and data-driven methods is shown in Tables 8 and 9, respectively. The following findings can be summarized in Figures 11 and 12, Tables 8 and 9.

(1) In all-day analysis, the estimated data by model-driven and data-driven proportions are almost identical, as shown in Figures 11(a)–11(c), while on 31/07/2017, as shown in Figure 11(d), the proportion of data-driven is more significant than the data estimated by model-driven.

(2) When we analyze the estimation data during the peak period (12: 00 am–8:00 pm), it is obvious that the data estimated by the model-driven method is larger than the data-driven method. However, one of the exceptions is the data on 31/07/2017. The estimated data by model-driven and data-driven proportions are almost identical.

(3) According to Tables 8 and 9, we can find that the proportion of model-driven estimated data during the peak period is significantly higher than the data-driven method. Therefore, we could conclude that the model-driven could better capture the congested

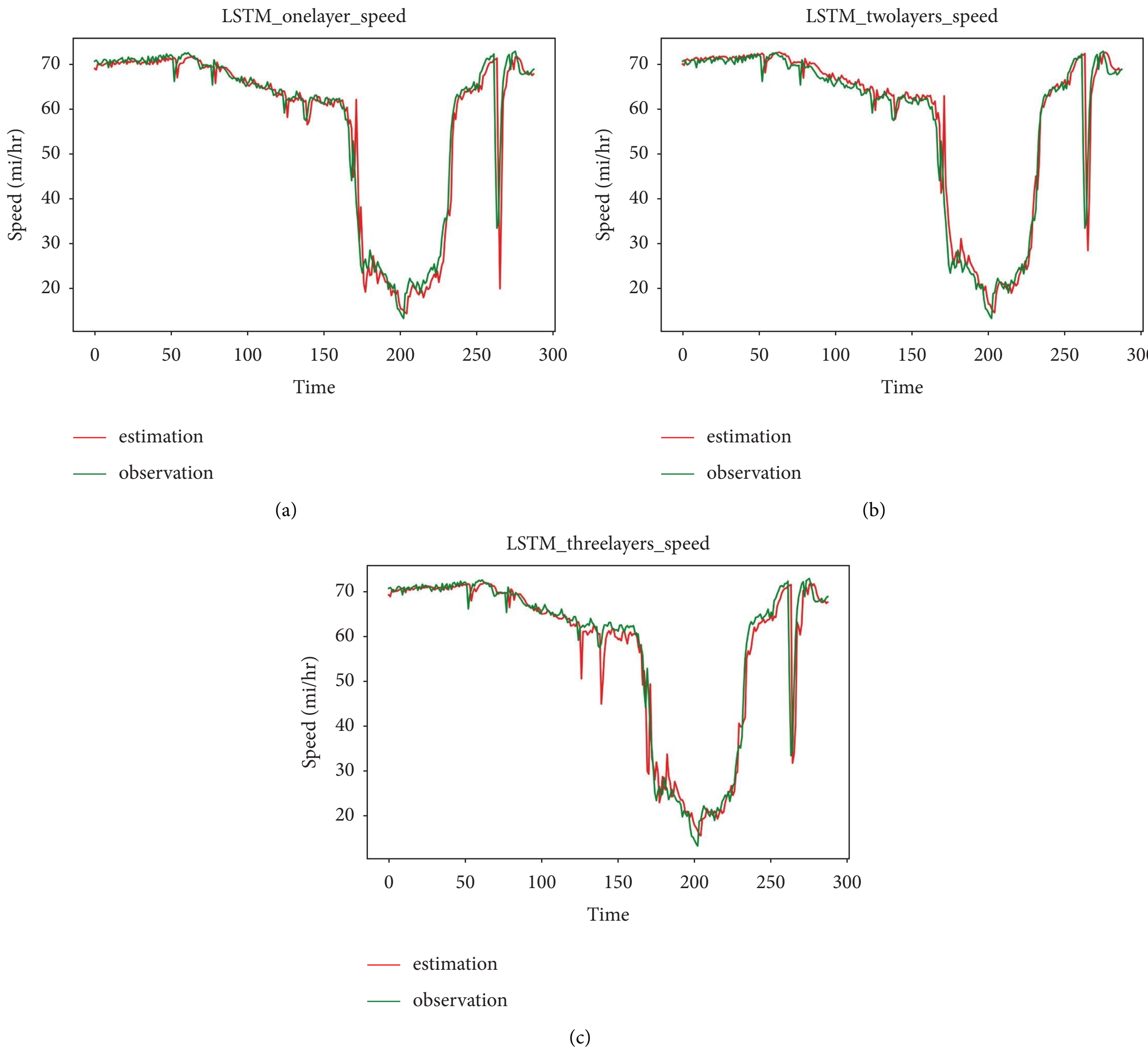


Figure 8: Estimated speed by the data-driven method in the Los Angeles case. (a) one-layer. (b) two-layer. (c) three-layer.

Table 4: Compare evaluation indicators in the machine learning model.

Layers	RMSE	MAE	MAPE
One-layer	4.28	2.11	5.16
Two-layers	4.06	1.91	4.92
Three-layers	**3.85**	**1.85**	**4.51**

Table 5: Compare evaluation indicators with different models using flow data.

Methods	MAE	RMSE	MAPE (%)
LSTM	76.19	105.87	10.38
KNN	**31.19**	**44.69**	8.94
ARIMA	32.84	47.52	10.85
Random forest	112.70	156.25	9.91
This paper	41.94	69.38	**4.26**
Average	58.97	84.74	8.87

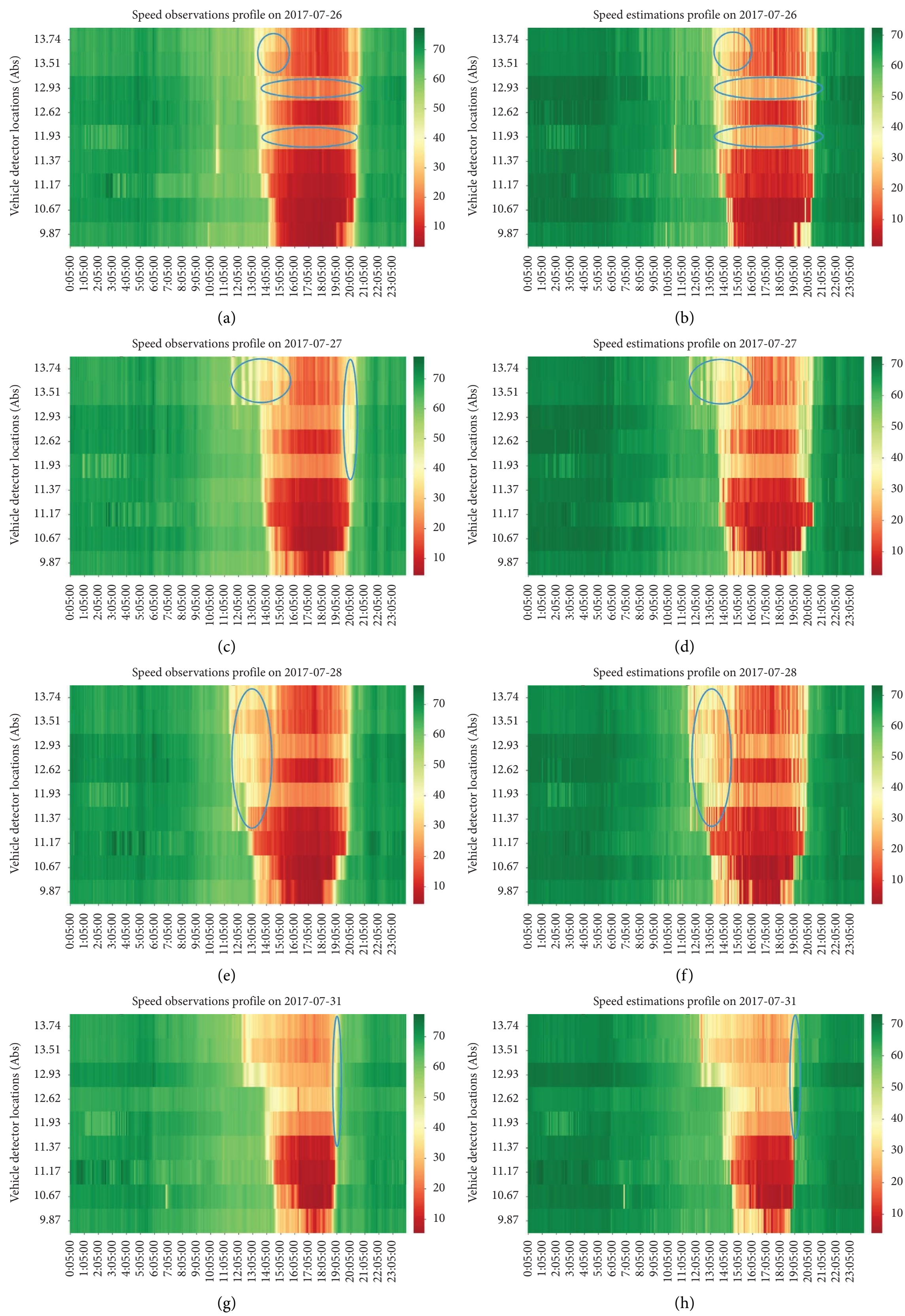


Figure 9: Comparison speed between observations and estimations on different days. (a) Speed observations on 07/26/2017, (b) Speed estimations on 07/26/2017, (c) Speed observations on 07/27/2017, (d) Speed estimations on 07/27/2017, (e) Speed observations on 07/28/2017, (f) Speed estimations on 07/28/2017, (g) Speed observations on 07/31/2017, and (h) Speed estimations on 07/31/2017.

Table 6: Compare evaluation indicators with different models using speed data.

Methods	MAE	RMSE	MAPE (%)
LSTM	1.85	3.85	4.51
KNN	4.35	6.86	8.83
ARIMA	1.42	2.71	3.41
Random forest	13.13	19.09	4.09
This paper	**1.20**	**2.24**	**2.75**
Average	4.39	6.95	4.72

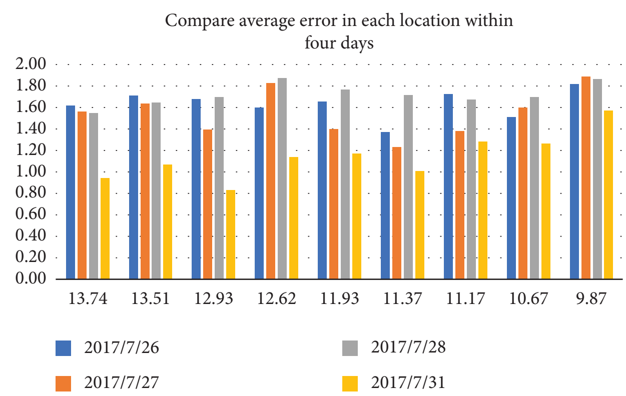


Figure 10: Comparison of average error in each location within four days in all day.

Figure 11: Comparison of speed estimation between model-driven and data-driven in all days. (a) Comparison estimation on 07/26/2017, (b) Comparison estimations on 07/27/2017, (c) Comparison estimation on 07/28/2017, and (d) Comparison estimations on 07/31/2017.

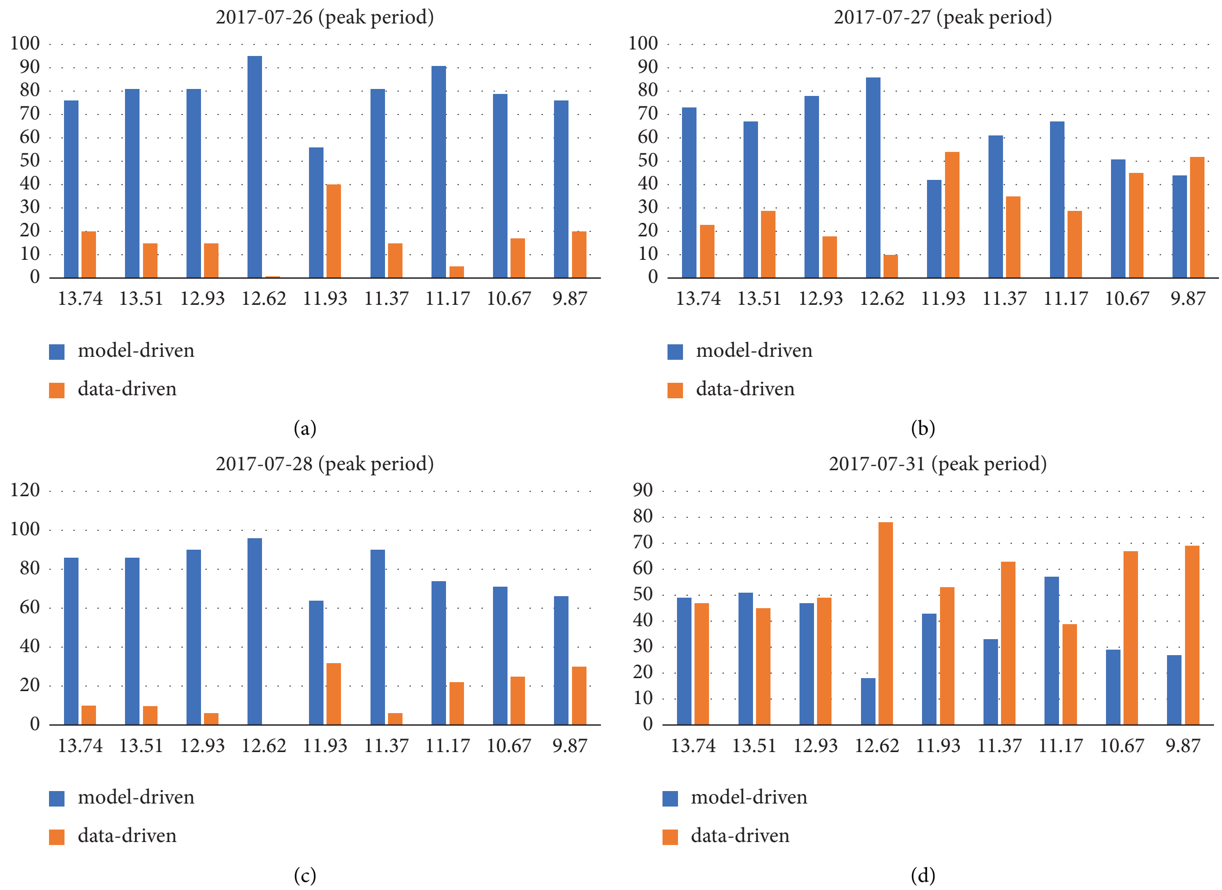


Figure 12: Comparison of speed estimation between model-driven and data-driven in the peak period. (a) Comparison estimation on 07/26/2017, (b) Comparison estimations on 07/27/2017, (c) Comparison estimation on 07/28/2017, and (d) Comparison estimations on 07/31/2017.

Table 7: Comparison error between observations and estimations on different days on different absolute mile.

	07/26/2017	07/27/2017	07/28/2017	07/31/2017
Absolute mile	MAE/MAPE	MAE/MAPE	MAE/MAPE	MAE/MAPE
13.74	1.62/5.24	1.56/4.23	1.55/4.78	0.95/2.19
13.51	1.71/5.73	1.64/4.61	1.65/5.68	1.07/2.58
12.93	1.68/5.13	1.40/3.36	1.70/4.82	0.83/1.87
12.62	1.60/5.59	1.83/5.39	1.88/5.72	1.14/2.52
11.93	1.66/4.70	1.40/3.41	1.77/4.72	1.17/2.41
11.37	1.37/5.81	1.24/4.59	1.72/6.93	1.01/3.26
11.17	1.73/7.98	1.39/4.71	1.67/5.65	1.29/3.63
10.67	1.52/10.54	1.60/8.33	1.70/8.88	1.26/4.59
9.87	1.82/9.19	1.89/7.25	1.87/8.25	1.58/4.35
Average	1.63/6.66	1.55/5.10	1.72/6.16	1.14/3.04

traffic flow with fewer data points. Furthermore, the mode-driven method could better explain the traffic flow dynamics, especially during the congestion period. On the contrary, the data-driven approach infers traffic states based on historical data and a lack of accurate estimates of congestion periods.

Table 8: Compare account for estimation between model-driven and data-driven in all-day.

	07/26/2017		07/27/2017		07/28/2017		07/31/2017	
Absolute mile	Model-driven	Data-driven	Model-driven	Data-driven	Model-driven	Data-driven	Model-driven	Data-driven
13.74	**0.57**	0.43	0.47	**0.53**	**0.51**	0.49	0.40	**0.60**
13.51	**0.60**	0.40	0.48	**0.52**	**0.51**	0.49	0.48	**0.52**
12.93	**0.66**	0.34	**0.51**	0.49	**0.55**	0.45	0.38	**0.62**
12.62	**0.73**	0.27	**0.74**	0.26	**0.76**	0.24	0.09	**0.91**
11.93	0.47	**0.53**	0.36	**0.64**	0.41	**0.59**	0.36	**0.64**
11.37	**0.63**	0.37	**0.53**	0.47	**0.60**	0.40	0.36	**0.64**
11.17	**0.78**	0.22	**0.61**	0.39	**0.61**	0.39	0.50	**0.50**
10.67	**0.63**	0.38	0.45	**0.55**	**0.53**	0.47	0.34	**0.66**
9.87	**0.51**	0.49	0.28	**0.72**	0.44	**0.56**	0.28	**0.72**
Average	**0.62**	0.38	0.49	**0.51**	**0.55**	0.45	0.35	**0.65**

Table 9: Compare accounts for estimation between model-driven and data-driven in the peak period.

	07/26/2017		07/27/2017		07/28/2017		07/31/2017	
Absolute mile	Model-driven	Data-driven	Model-driven	Data-driven	Model-driven	Data-driven	Model-driven	Data-driven
13.74	**0.79**	0.21	**0.76**	0.24	**0.90**	0.10	**0.51**	0.49
13.51	**0.84**	0.16	**0.70**	0.30	**0.90**	0.10	**0.53**	0.47
12.93	**0.84**	0.16	**0.81**	0.19	**0.94**	0.06	0.49	**0.51**
12.62	**0.99**	0.01	**0.90**	0.10	**1.00**	0.00	0.19	**0.81**
11.93	**0.58**	0.42	0.44	**0.56**	**0.67**	0.33	0.45	**0.55**
11.37	**0.84**	0.16	**0.64**	0.36	**0.94**	0.06	0.34	**0.66**
11.17	**0.95**	0.05	**0.70**	0.30	**0.77**	0.23	**0.59**	0.41
10.67	**0.82**	0.18	**0.53**	0.47	**0.74**	0.26	0.30	**0.70**
9.87	**0.79**	0.21	0.46	**0.54**	**0.69**	0.31	0.28	**0.72**
Average	**0.83**	0.17	**0.66**	0.34	**0.84**	0.16	0.41	**0.59**

5. Conclusion and Future Work

Traffic flow estimation and forecasting is a critical problem in transportation planning. In this paper, we proposed a hybrid framework incorporating model-driven and data-driven methods for traffic state estimation. The results of numerical experiments based on real-world datasets demonstrate the effectiveness of the proposed framework. Experiments in the real-world show that the proposed TFMDL framework could improve performance better than model-driven or data-driven methods alone in terms of estimation accuracy and data efficiency. Meanwhile, this hybrid stepwise modeling framework could better capture dynamic traffic flow states during congestion conditions.

The study focuses on traffic speed and flow estimation and prediction, but a comprehensive traffic state estimation and prediction, which includes travel time and queue length, would have more significance for passengers. Therefore, the future work of this study includes: (1) trying to consider the relations among different formats of traffic data and then building a multiple input/output traffic framework to output a comprehensive traffic state estimation and prediction result. (2) embedding queue models in the proposed framework to improve the performance under oversaturated traffic conditions. (3) considering the dynamic value of a parameter α is expected to improve the estimation performance under different traffic conditions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Dong and H. S. Mahmassani, "Flow breakdown and travel time reliability," *Transportation Research Record*, vol. 2124, no. 1, pp. 203–212, 2009.

[2] J. Long, Z. Gao, H. Ren, and A. Lian, "Urban traffic congestion propagation and bottleneck identification," *Science in China - Series F: Information Sciences*, vol. 51, no. 7, pp. 948–964, 2008.

[3] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, "Traffic state estimation on highway: a comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017.

[4] R. Shi, Z. Mo, K. Huang, X. Di, and Q. Du, "Physics-informed deep learning for traffic state estimation," 2021a, Available: http://arxiv.org/abs/2101.06580.

[5] R. Shi, Z. Mo, and X. Di, "Physics-informed deep learning for traffic state estimation: a hybrid paradigm informed by second-order traffic models," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 1, pp. 540–547, 2021b.

[6] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep

learning theory," *PLoS One*, vol. 10, no. 3, Article ID e0119044, 2015a.

[7] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences*, vol. 229, no. 1178, pp. 317–345, 1955.

[8] P. I. Richards, "Shock waves on the highway," *Operations Research*, vol. 4, no. 1, pp. 42–51, 1956.

[9] C. F. Daganzo, "The cell transmission model, part II: network traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995.

[10] A. Spiliopoulou, M. Kontorinaki, M. Papageorgiou, and P. Kopelias, "Macroscopic traffic flow model validation at congested freeway off-ramp areas," *Transportation Research Part C: Emerging Technologies*, vol. 41, pp. 18–29, 2014.

[11] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: a general approach," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 141–167, 2005.

[12] B. L. Smith, W. T. Scherer, and J. H. Conklin, "Exploring imputation techniques for missing data in transportation management systems," *Transportation Research Record*, vol. 1836, no. 1, pp. 132–142, 2003.

[13] B. Yu, Z. Z. Yang, K. Chen, and B. Yu, "Hybrid model for prediction of bus arrival times at next station," *Journal of Advanced Transportation*, vol. 44, no. 3, pp. 193–204, 2010.

[14] T. Ma, C. Antoniou, and T. Toledo, "Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast," *Transportation Research Part C: Emerging Technologies*, vol. 111, pp. 352–372, 2020.

[15] J. Huang and S. Agarwal, "Physics informed deep learning for traffic state estimation," in *Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, Rhodes, Greece, 2020, September.

[16] A. J. Huang and S. Agarwal, "Physics-informed deep learning for traffic state estimation: illustrations with LWR and CTM Models," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 503–518, 2022.

[17] J. Zhang, X. Qu, and S. Wang, "Reproducible generation of experimental data sample for calibrating traffic flow fundamental diagram," *Transportation Research Part A: Policy and Practice*, vol. 111, pp. 41–52, 2018.

[18] X. Qu, S. Wang, and J. Zhang, "On the fundamental diagram for freeway traffic: a novel calibration approach for single-regime models," *Transportation Research Part B: Methodological*, vol. 73, pp. 91–102, 2015.

[19] G. Dervisoglu, G. Gomes, J. Kwon, R. Horowitz, and P. Varaiya, "Automatic calibration of the fundamental diagram and empirical observations on capacity," *Transportation Research Board 88th Annual Meeting*, vol. 15, pp. 31–59, 2009.

[20] R. Zhong, C. Chen, A. H. F. Chow, T. Pan, F. Yuan, and Z. He, "Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models," *Journal of Advanced Transportation*, vol. 50, no. 3, pp. 363–385, 2016.

[21] R. Herring, A. Hofleitner, P. Abbeel, and A. Bayen, "Estimating arterial traffic conditions using sparse probe data," in *Proceedings of the 13th international IEEE conference on intelligent transportation systems*, pp. 929–936, IEEE, Madeira Island, Portugal, 2010, September.

[22] A. S. Nagle and V. V. Gayah, "Accuracy of networkwide traffic states estimated from mobile probe data," *Transportation Research Record*, vol. 2421, no. 1, pp. 1–11, 2014.

[23] T. Seo, T. Kusakabe, and Y. Asakura, "Estimation of flow and density using probe vehicles with spacing measurement equipment," *Transportation Research Part C: Emerging Technologies*, vol. 53, pp. 134–150, 2015.

[24] W. Deng, H. Lei, and X. Zhou, "Traffic state estimation and uncertainty quantification based on heterogeneous data sources: a three detector approach," *Transportation Research Part B: Methodological*, vol. 57, pp. 132–157, 2013.

[25] H. J. Payne, "FREFLO: a macroscopic simulation model of freeway traffic," *Transportation Research Record*, vol. 722, 1979.

[26] C. Nanthawichit, T. Nakatsuji, and H. Suzuki, "Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway," *Transportation Research Record*, vol. 1855, no. 1, pp. 49–59, 2003.

[27] L. Mihaylova, R. Boel, and A. Hegyi, "Freeway traffic estimation within particle filtering framework," *Automatica*, vol. 43, no. 2, pp. 290–300, 2007.

[28] Y. Wang, M. Papageorgiou, A. Messmer, P. Coppola, A. Tzimitsi, and A. Nuzzolo, "An adaptive freeway traffic state estimator," *Automatica*, vol. 45, no. 1, pp. 10–24, 2009.

[29] J. C. Herrera and A. M. Bayen, "Incorporation of Lagrangian measurements in freeway traffic state estimation," *Transportation Research Part B: Methodological*, vol. 44, no. 4, pp. 460–481, 2010.

[30] H. Haj-Salem and J. P. Lebacque, "Reconstruction of false and missing data with first-order traffic flow model," *Transportation Research Record*, vol. 1802, no. 1, pp. 155–165, 2002.

[31] L. Mihaylova and R. Boel, "A particle filter for freeway traffic estimation,"vol. 2, pp. 2106–2111, in *Proceedings of the 2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No. 04CH37601)*, vol. 2, pp. 2106–2111, IEEE, Los Alamitos, CA, 2004, December.

[32] S. Oh, Y. J. Byon, K. Jang, and H. Yeo, "Short-term travel-time prediction on highway: a review of the data-driven approach," *Transport Reviews*, vol. 35, no. 1, pp. 4–32, 2015.

[33] C. Antoniou, H. N. Koutsopoulos, and G. Yannis, "Dynamic data-driven local traffic state estimation and prediction," *Transportation Research Part C: Emerging Technologies*, vol. 34, pp. 89–107, 2013.

[34] D. B. Work, O. P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton, "An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices," in *Proceedings of the 2008 47th IEEE Conference on Decision and Control*, pp. 5062–5068, IEEE, Cancun, Mexico, 2008, December.

[35] I. Lana, J. L. Lobo, E. Capecci, J. Del Ser, and N. Kasabov, "Adaptive long-term traffic state estimation with evolving spiking neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 126–144, 2019.

[36] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[37] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[38] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, p. 818, 2017.

[39] H. Park, A. Haghani, and X. Zhang, "Interpretation of Bayesian neural networks for predicting the duration of detected incidents," *Journal of Intelligent Transportation Systems*, vol. 20, no. 4, pp. 385–400, 2016.

[40] M. Chikaraishi, P. Garg, V. Varghese et al., "On the possibility of short-term traffic prediction during disaster with machine

learning approaches: an exploratory analysis," *Transport Policy*, vol. 98, pp. 91–104, 2020.
[41] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017.
[42] J. W. C. Van Lint, S. P. Hoogendoorn, and H. J. van Zuylen, "Freeway travel time prediction with state-space neural networks: modeling state-space dynamics with recurrent neural networks," *Transportation Research Record*, vol. 1811, no. 1, pp. 30–39, 2002.
[43] F. Jin and S. Sun, "Neural network multitask learning for traffic flow forecasting," in *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp. 1897–1901, IEEE, Hong Kong, China, 2008, June.
[44] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014.
[45] K. Zhang, L. Wu, Z. Zhu, and J. Deng, "A multitask learning model for traffic flow and speed forecasting," *IEEE Access*, vol. 8, pp. 80707–80715, 2020.
[46] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015b.
[47] Z. Zheng, Y. Yang, J. Liu, H. N. Dai, and Y. Zhang, "Deep and embedded learning approach for traffic flow prediction in urban informatics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3927–3939, 2019.
[48] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting," in *International Joint Conference on Artificial Intelligence*, pp. 3634–3640, 2018.
[49] H. Yu, Z. Wu, S. Wang, Y. Wang, and X. Ma, "Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks," *Sensors*, vol. 17, no. 7, p. 1501, 2017b.
[50] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: data-driven traffic forecasting," in *Proceedings of the ICLR*, Scottsdale, AZ, USA, 2017.
[51] L. N. Do, H. L. Vu, B. Q. Vo, Z. Liu, and D. Phung, "An effective spatial-temporal attention based neural network for traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 12–28, 2019.
[52] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 922–929, 2019.
[53] W. Zhang, Y. Yu, Y. Qi, F. Shu, and Y. Wang, "Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning," *Transportmetrica: Transportation Science*, vol. 15, no. 2, pp. 1688–1711, 2019.
[54] M. G. Karlaftis and E. I. Vlahogianni, "Statistical methods versus neural networks in transportation research: differences, similarities and some insights," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 3, pp. 387–399, 2011.
[55] M. Raissi, "Deep hidden physics models: deep learning of nonlinear partial differential equations," *Journal of Machine Learning Research*, vol. 19, no. 1, pp. 932–955, 2018.
[56] M. Raissi and G. E. Karniadakis, "Hidden physics models: machine learning of nonlinear partial differential equations," *Journal of Computational Physics*, vol. 357, pp. 125–141, 2018.
[57] T. Kim, S. Sharda, X. Zhou, and R. M. Pendyala, "A stepwise interpretable machine learning framework using linear regression (LR) and long short-term memory (LSTM): city-wide demand-side prediction of yellow taxi and for-hire vehicle (FHV) service," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102786, 2020.
[58] B. T. Thodi, Z. S. Khan, S. E. Jabari, and M. Menéndez, "Incorporating kinematic wave theory into a deep learning method for high-resolution traffic speed estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 17849–17862, 2022.
[59] B. D. Greenshields, J. R. Bibbins, W. S. Channing, and H. H. Miller, "A study of traffic capacity," in *Highway Research Board Proceedings*Vol. 1935, National Research Council (USA), Highway Research Board, Washington, DC, USA, 1935.
[60] Q. Cheng, Z. Liu, Y. Lin, and X. S. Zhou, "An s-shaped three-parameter (S3) traffic stream model with consistent car following relationship," *Transportation Research Part B: Methodological*, vol. 153, pp. 246–271, 2021.
[61] X. B. Jin, R. J. RobertJeremiah, T. L. Su, Y. T. Bai, and J. L. Kong, "The new trend of state estimation: from model-driven to hybrid-driven methods," *Sensors*, vol. 21, no. 6, p. 2085, 2021.
[62] L. Qu, W. Li, W. Li, D. Ma, and Y. Wang, "Daily long-term traffic flow forecasting based on a deep neural network," *Expert Systems with Applications*, vol. 121, pp. 304–312, 2019.
[63] Q. Wan, G. Peng, Z. Li, and F. H. T. Inomata, "Spatiotemporal trajectory characteristic analysis for traffic state transition prediction near expressway merge bottleneck," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102682, 2020.