



Pan, Y., Guo, J., Chen, Y., Cheng, Q., Li, W., & Liu, Y. (2024). A fundamental diagram based interpretable framework for traffic flow estimation and prediction by combining a Markovian model with deep learning. *Expert Systems with Applications*, 238(E), [122219].
<https://doi.org/10.1016/j.eswa.2023.122219>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1016/j.eswa.2023.122219](https://doi.org/10.1016/j.eswa.2023.122219)

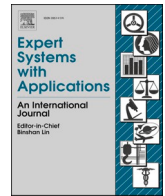
[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at <https://doi.org/10.1016/j.eswa.2023.122219>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



A fundamental diagram based hybrid framework for traffic flow estimation and prediction by combining a Markovian model with deep learning

Yuyan Annie Pan^a, Jifu Guo^b, Yanyan Chen^a, Qixiu Cheng^{c,*}, Wenhao Li^a, Yanyue Liu^d

^a Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing, China

^b Beijing Transport Institute, Beijing, China

^c University of Bristol Business School, University of Bristol, Bristol, UK

^d ITS Center, Research Institute of Highway Ministry of Transport, Beijing, China

ARTICLE INFO

Keywords:

Traffic state identification
Traffic flow prediction
Fundamental diagram
Hybrid interpretable model
Markov chain
LSTM

ABSTRACT

Accurate traffic congestion estimation and prediction are critical building blocks for smart trip planning and rerouting decisions in transportation systems. Over the decades, there have been many studies focusing on traffic congestion estimation and prediction with different statistical approaches (e.g., Markov chain) and machine learning models (e.g., clustering, Bayesian networks, and artificial neural networks). However, there is a lack of a unified framework to address the mechanisms of different models and integrate the advantages of different methods through combinations. This paper introduces the FD-Markov-LSTM model, a hybrid interpretable approach that combines the fundamental diagram (FD), Markov chain, and long short-term memory (LSTM). The aim is to estimate and predict traffic states by integrating statistical data in both congested and uncongested scenarios. The FD-Markov-LSTM model leverages the FD to identify hierarchical traffic states and utilizes the Markov process to capture the probabilistic transitions between these states. We employ the LSTM model to further capture the residual time series produced by the Markov chain model (assuming a memoryless property) to enhance the estimation and prediction performance. The proposed model's accuracy in estimating and predicting traffic flow is evaluated using empirical data from three case studies conducted in Beijing and Los Angeles. The results highlight a significant improvement in accuracy compared to classical benchmark models such as the Markov model, ARIMA model, k-Nearest Neighbor model, Random Forest model, and LSTM. Specifically, the FD-Markov-LSTM model achieves reductions of over 39% in mean absolute error, 35% in root mean squared error, and 7.4% in mean absolute percentage error. These results clearly demonstrate that the FD-Markov-LSTM model outperforms the benchmark models, enabling more precise predictions of traffic flow.

1. Introduction

Short-term traffic forecasting plays a crucial role in Intelligent Transportation Systems (ITS) (Zhang et al., 2023). The accurate and timely estimation and prediction of traffic conditions can aid ITS in promptly adjusting the relevant state, leading to a more precise and expedient alleviation of traffic congestion (Comert et al., 2021). Traffic state estimation and prediction can be categorized into two main approaches: analytical models, such as fundamental diagrams (FDs), and data-driven methods, including machine learning techniques. Traffic parameter variables, including flow, speed, and occupancy, are the key measurements of network performance, and their relationships are

usually described by FD models. FD models can analytically divide traffic states into uncongested and congested regimes; however, traffic dynamics over time cannot be captured. The increasing availability of transport big data and advancements in artificial intelligence models have led to a growing focus on data-driven methods, particularly deep learning approaches, for addressing traffic congestion estimation and prediction challenges (Gao et al., 2021). Data-driven methods usually need large amounts of historical data as training inputs, and they usually lack interpretability in model mechanisms and traffic flow phenomena. It is clear that both analytical model-based models and data-driven methods have their advantages and disadvantages, and it is natural to come up with an innovative approach by combining their advantages to

* Corresponding author.

E-mail addresses: yuyananniepan@gmail.com (Y.A. Pan), guojf@bjtrc.org.cn (J. Guo), cdyan@bjut.edu.cn (Y. Chen), qixiu.kevin.cheng@gmail.com (Q. Cheng), wenhaozi1999@126.com (W. Li), lyy@itsc.cn (Y. Liu).

<https://doi.org/10.1016/j.eswa.2023.122219>

Received 25 June 2023; Received in revised form 28 September 2023; Accepted 16 October 2023

Available online 17 October 2023

0957-4174/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

result in a hybrid model (Hou et al., 2022). Therefore, this paper develops a hybrid stepwise framework to estimate and predict traffic dynamics features with analytical state transition probabilities.

The proposed hybrid stepwise model presents a combination of a Markov chain model and a deep learning model for accurate traffic flow estimation and prediction. This method involves two main steps. Firstly, utilizing the FD, we identify uncongested and congested states for a single lane on the freeway. By employing the Markovian model, we determine traffic flow states and calculate the state transition probability matrix, capturing regular patterns and explicitly addressing temporal correlation errors between estimated and observed flows. In the second step, the long short-term memory (LSTM) model is utilized to capture additional traffic flow features within the residual series that are not entirely captured by the Markov chain model. The integration of deep learning methods within the Markov model allows for the extraction of important traffic flow characteristics from the residuals of the Markov prediction model. Ultimately, the predicted flows from the Markovian and LSTM models are combined to provide a comprehensive and accurate prediction result.

1.1. Literature review on traffic state identification

Accurate identification of traffic flow states within a transportation network is crucial in improving traffic flow prediction capabilities and enabling effective traffic management (Pan et al., 2023). Accurately identifying the traffic state serves as a fundamental component for implementing informed traffic management strategies. The efficacy of traffic state identification significantly impacts the precision of short-term traffic estimation and prediction.

Many previous studies concerning traffic state identification used fuzzy mathematical methods. Ban et al. (2007) used percentile speeds based on multiday data to identify and calibrate traffic flow in bottlenecks. Huang et al. (2011) introduced a novel real-time traffic state identification method utilizing the fuzzy c-means clustering method. Lu et al. (2015) presented an algorithm that leverages the fuzzy c-means model to improve the applicability of traffic state clustering for the identification and prediction of real-time traffic conditions. Moreover, researchers (Esfahani et al., 2018; Bao, 2019) have improved the accuracy of traffic flow state identification by introducing a modified fuzzy c-means clustering strategy.

Various data-driven intelligent approaches have been employed by researchers to detect changes in traffic states. Yang and Qiao (1998) utilized a self-organizing neural network technique, categorizing traffic states into seven distinct groups. Additionally, the identification of traffic flow states has seen extensive exploration through the lens of traffic flow theory. Long et al. (2008) introduced a congestion propagation model that utilizes the cell transmission model to accurately identify congestion bottlenecks in urban networks. Kerner (2013) proposed the three-phase traffic theory, which categorizes traffic flow into three distinct states: free flow, synchronized flow, and wide-moving jam. Kerner (2013) introduced the three-phase traffic theory, which classifies traffic flow into three distinct states, providing a comprehensive framework for understanding traffic dynamics. Building upon this, Lin (2019) developed an approach for traffic state identification that combines the Macroscopic Fundamental Diagram (MFD) with the support vector machine algorithm. This method allows for the identification of traffic states across the entire network.

Accurately determining the actual traffic conditions, especially during periods of congestion, presents a significant challenge due to the complex interplay of multiple factors that contribute to the overall traffic state (Chen et al., 2022). Previous studies have rarely utilized a comprehensive set of classification indicators, including critical parameters such as flow, speed, and density obtained from the FD, to identify distinct traffic states. It is widely known that FD is to establish the mathematical relationship between traffic flow, speed, and density. It classifies similar samples into one class (congested or non-congested)

by calibrating the FD with observed data, while different samples into another category. Therefore, the division of traffic states based on FD can not only conform to the characteristic of traffic flow, but also lead to a more accurate division of traffic states, especially for capturing congested and uncongested states. Thus, in this study, we will fill these knowledge gaps and introduce a classification indicator system based on the calibration of FD.

1.2. Literature review on traffic state estimation and prediction with Markov model

By examining the time series data of traffic flow, a strong relationship between the current and previous traffic flow states becomes apparent, which can be represented using conditional probabilities. The dynamic nature of traffic flow makes the Markov chain an appropriate model for describing this relationship. Consequently, several studies (Geroliminis and Skabardonis, 2005) have successfully utilized the Markov chain for traffic flow estimation and prediction. Shin and Sunwoo (2018) proposed a Markov-based model incorporating speed constraints to predict vehicle speed, particularly excelling in cornering sections. Evans et al. (2001) employed the Markov model for breakdown occurrences, demonstrating that higher arrival rates of vehicles increased the likelihood of breakdown. Kidando et al. (2018) developed a Markov-based structured model with random variations to describe the congestion, offering a valuable resource for detecting different traffic conditions. The Markov-based model has been widely utilized in travel time prediction in several studies (Kharoufeh and Gautam, 2004; Ramezani and Geroliminis, 2011, 2012; Tang et al., 2020). The experiments conducted have demonstrated the Markov chain approach's ability to accurately estimate the distribution of dependent link travel times and effectively capture correlations among link travel times.

Furthermore, there is a growing trend to integrate Markov models with other analytical or machine learning approaches to enhance traffic state prediction. Noroozi and Hellinga (2014) introduced a method that combines a short-memory time series model with a Markov model incorporating time-varying covariates, demonstrating its effectiveness in basic time series modeling. Li et al. (2020) devised a comprehensive Markov-based model that incorporates various external factors to predict traffic network conditions. By considering these factors, their model effectively captures the dynamic changes in traffic conditions, resulting in enhanced performance of the prediction system compared to single-model methods. The Markov model was utilized by Qi and Ishak (2014) and Sun et al. (2020) to predict congestion patterns during peak periods.

From the current literature, we can see that there is still a gap in the analysis of bottleneck with the Markov model and capturing the state transition dynamics of traffic flows during congested and uncongested periods in the urban transport system.

1.3. Literature review on estimation and prediction with machine learning

With the rapid progress of information processing technology, machine learning techniques have become widely utilized in the field of traffic state estimation and prediction. These methods bring notable benefits in modeling and optimizing nonlinear systems. Notable examples include LSTM networks, as utilized by Ma et al. (2015). Additionally, the Diffusion Convolutional Recurrent Neural Network, developed by Li et al. (2017), and the Spatial-Temporal Transformer Networks, proposed by Xu et al. (2020), have demonstrated effectiveness in this domain. Another notable model includes the Spatio-Temporal Graph Mixformer by Lablack and Shen (2023). These methods have significantly contributed to the advancement of traffic state estimation and prediction.

Antoniou et al. (2013) proposed a two-step dynamic framework for local traffic state prediction, utilizing available information. Fusco et al. (2016a, 2016b) conducted a comparative analysis of models such as

Bayesian Networks for short-term traffic estimation and prediction. In the context of urban transportation, Wang et al. (2019) presented a path-based deep learning framework for predicting city-wide traffic speeds, which provided logical and interpretable outcomes. In their recent work, Ma et al. (2023) introduced a novel multi-task learning approach for forecasting both traffic flow and speed, leveraging shared transferable features and task-specific correlations. Furthermore, data-driven methodologies have expanded to other domains, including demand estimation and prediction in network systems (Kim et al., 2020), traffic congestion forecasting for connected vehicles (Elfar et al., 2018), traffic speed interval prediction (Song et al., 2022), and the advancement of fusion techniques based on machine learning for precise predictions (Ma et al., 2020).

Despite the multitude of studies discussed above, many of them have overlooked the potential information contained in time series residuals during the estimation and prediction of traffic conditions. As residuals may still contain important traffic features, we have opted to utilize the LSTM model to further improve the accuracy of traffic estimation and prediction for congestion conditions in our study.

1.4. Literature review on estimation and prediction with hybrid method

In the context of congestion bottlenecks or adverse weather, accurately estimating driver behavior poses challenges for traffic state estimation. Both traditional approaches struggle to handle the complexity of such behavior on their own. To overcome these limitations, researchers have proposed a combined approach that integrates the strengths of data-driven and model-driven methods. This approach aims to compensate for the weaknesses of each method and improve accuracy.

A promising approach to improve the accuracy and interpretability of traditional methods in traffic state estimation is through the integration of model-driven and data-driven techniques. This hybrid method combines the use of traffic flow models to approximate traffic states and the utilization of observed traffic data to discover unknown model parameters. By leveraging the strengths of both approaches, this hybrid approach offers enhanced performance. Nevertheless, there are still numerous areas and unanswered questions that necessitate further investigation in the advancement of hybrid methods for traffic state estimation.

Numerous studies have investigated hybrid approaches for traffic state estimation and prediction, aiming to combine the strengths of model-driven and data-driven techniques. For instance, Krakovna and Doshi-Velez (2016) combined LSTM and a hidden Markov model to enhance interpretability and prediction accuracy. Raissi (2018), Raissi and Karniadakis (2018) and introduced a deep learning method that incorporates partial differential equations (PDE) features for traffic prediction. Belezamo et al. (2019) combined the kinematic wave model with the Markov Chain model to estimate traffic state variables and develop an explainable framework. Huang and Agarwal (2020) introduced a physics-informed method that enhances traffic state estimation accuracy, particularly in real-time scenarios with limited data. Kim et al. (2020) proposed an interpretable model that combines linear regression models with a neural network, incorporating LSTM layers, to predict taxi ride demand. These studies highlight the ongoing exploration and development of hybrid methods for more effective and accurate traffic state estimation and prediction.

In different domains, hybrid approaches have been utilized to enhance model accuracy. For instance, Nguyen-Le et al. (2020) integrated LSTM and Markov models to forecast crack propagation in engineering. Similarly, Ma et al. (2021) introduced an LSTM-Markov model that leveraged Markov models to mitigate prediction errors in LSTM models, leading to improved accuracy in predicting COVID-19. The integration of data-driven and model-driven approaches through hybrid methods holds great promise for enhancing the accuracy and interpretability of traffic flow prediction. Ongoing research is actively

Table 1

Comparison of advantages and disadvantages of different short-time traffic flow forecasting methods.

Methods	References	Advantages	Disadvantages
Time series methods (e.g., Autoregressive integrated moving average, ARIMA; Markov chain)	Vlahogianni et al. (2004); Kumar and Vanajakshi (2015); Hou et al. (2022)	Efficient in capturing temporal patterns and seasonality in traffic data; relatively straightforward and easy to implement.	Constrained in capturing non-linear and intricate traffic patterns; susceptible to data omissions and outliers.
Kalman filtering methods	Guo et al. (2014); Cai et al. (2019); Emami et al. (2019)	Adaptable choice of predictive variables, robust real-time performance, and high precision; facilitate seamless online analysis.	The uncertainty and nonlinearity of traffic flow tend to escalate as the prediction time horizon shortens.
Regression methods (e.g., Lasso regression)	Chen et al. (2019); Wang et al. (2019); Lee et al. (2021)	Incorporate various influencing factors like weather, holidays, and events into the models; generally interpretable and can provide insights into the impact of different factors.	Limited modeling of complex nonlinear relationships; highly sensitive to data quality, noise or outliers in the data can significantly impact prediction results.
Machine learning methods (e.g., K-nearest neighbors, KNN; neural networks)	Cai et al. (2016); Sun et al. (2018); Khan et al. (2023)	Capable of representing intricate non-linear data relationships; can autonomously acquire pertinent features, diminishing the necessity for extensive feature engineering.	Might display overfitting tendencies when dealing with small datasets unless proper regularization is applied; refining parameters and selecting the model could entail a significant commitment of time and resources.
Deep learning methods (e.g., LSTM; Transformer)	Chen et al. (2019); Lu et al. (2022); Ren et al. (2023)	Efficient in capturing long-term dependencies and non-linear patterns; well-suited for capturing patterns and features in large-scale datasets and intricate traffic scenarios.	Requires a considerable volume of training data, which may not be readily available; and can entail substantial computational resources, especially during the training of larger models.
Hybrid models	Zheng et al. (2020); Sattarzadeh et al. (2023)	Combine multiple methods to leverage their respective strengths, potentially enhancing forecasting accuracy; adaptable and able to handle a variety of traffic situations.	Implementation can become more intricate, potentially necessitating meticulous integration and parameter fine-tuning.

dedicated to refining and exploring these hybrid techniques across various applications. Specifically, combining traffic flow models, statistical methods, and machine learning models not only improves the accuracy of the model but also enhances its interpretability. This comprehensive approach ensures that the resulting estimations and

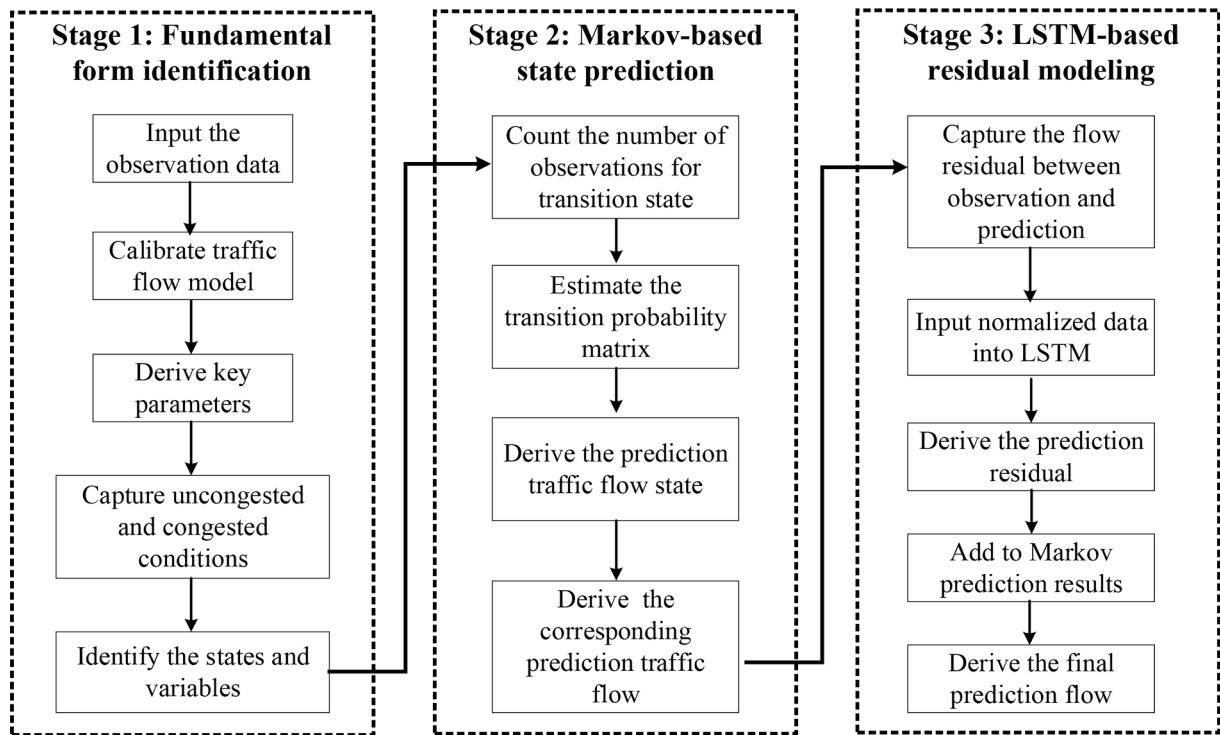


Fig. 1. Illustration of the proposed stepwise framework for traffic flow prediction.

predictions are both precise and readily understandable.

Various short-term traffic flow prediction methods possess distinct strengths and limitations. We compare the following commonly used short-time traffic flow forecasting methods, namely, time series methods, Kalman filtering methods, regression methods, machine learning methods, deep learning methods, and Hybrid models. The comprehensive analysis of the advantages and disadvantages of those models is shown in Table 1. In summary, selecting an appropriate prediction method hinges on careful consideration of factors such as data availability, model complexity, computational resources, and prediction time frame. Regardless of the chosen method, meticulous attention to data quality, preprocessing, and parameter tuning remains imperative for ensuring accurate and reliable short-term traffic flow predictions.

1.5. Objectives and contributions

Although numerous efforts have been made for different aspects of the traffic flow estimation and prediction problem, there are still some critical challenges that need to be further investigated. First, as for the traffic state identification, most existing methods apply the clustering method (), which lacks physical interpretability and cannot incorporate the traffic flow model into the traffic state identification method. The FD is recognized as a highly effective tool for capturing the intricate connection among traffic flow, speed, and density. Not only does it derive essential traffic flow parameters, but it also enables precise identification of traffic states. Nevertheless, numerous current methods for traffic flow estimation and prediction only utilize deep learning as a final stage in data processing, failing to fully exploit the valuable information present in the residuals. However, it is still possible to capture traffic features within the residuals sequence (Ma et al., 2020). This paper introduces a comprehensive hybrid framework, termed FD-Markov-LSTM, which combines the fundamental diagram, Markov model, and LSTM method. The primary goal of our research is to systematically address the interpretation and accuracy of prediction in bottleneck. The key contributions of this paper are summarized as follows:

(1) By embedding the FD model into the traffic state identification, the calibrated key parameter (e.g., critical speed and critical density) based on FD could accurately capture the change in traffic flow during the transition between different traffic states.

(2) We utilize the Markov-based model to determine the state-to-state transition probabilities, enabling precise prediction of future traffic states.

(3) A hybrid interpretable model framework is constructed to combine Markov and LSTM, which can further capture the data residual. The FD-Markov-LSTM model demonstrates significantly smaller prediction errors compared to benchmark models.

(4) The proposed FD-Markov-LSTM model enhances the accuracy of short- and medium-term trend prediction for bottleneck traffic flow.

The subsequent sections of this paper are structured as follows. Section 2 introduces an interpretable framework that combines FD, Markovian, and LSTM for traffic state estimation and prediction. In Section 3, we introduce the experimental description and evaluation performance of the proposed method. Section 4 analyzed and discussed the experimental results. Finally, Section 5 summarizes this work and provides insights into potential future research directions.

2. Methodology

Enhancing the precision of traffic state estimation and prediction stands as an effective approach in alleviating urban traffic congestion within traffic management systems. To this end, this paper introduces a stepwise hybrid model (FDM-LSTM) that amalgamates the strengths of the FD, Markov, and LSTM models. By leveraging the advantages of both traditional traffic flow prediction methods and deep learning techniques, this model enables accurate forecasting of future traffic flow states. The elaborate flowchart of the proposed FD-Markov-LSTM model in this paper is shown in Fig. 1. We will further explain the differences in each step in the following.

(i) In stage 1, this paper uses the calibrated FD to identify the traffic flow state, which could accurately capture the uncongested and congested conditions by critical speed v_c . Improving the accuracy of the

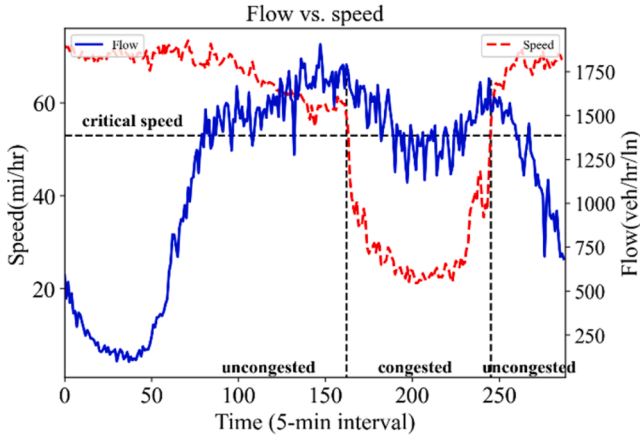


Fig. 2. Flow and speed evolution along with the time series.

Markov prediction model in the subsequent stage can be achieved by ensuring the precise input of traffic flow states.

(ii) In stage 2, we develop a Markov prediction model using the defined traffic flow state. The Markov model captures the transition between different traffic states, and the state transition probability matrix is used to forecast future traffic states. However, the Markov model only provides a general representation of traffic flow and cannot accurately capture the nonlinear characteristics of time series data.

(iii) In stage 3, we use the deep learning method to capture traffic flow features by residuals. We build the LSTM model and use the residual between the Markov prediction model and the observed data and the traffic state as input data in LSTM. Then we could further improve the prediction results for future traffic flow.

2.1. Traffic state identification based on fundamental diagram

Numerous researchers (Ma et al., 2020) have emphasized the significance of defining and categorizing traffic states across various levels, ranging from free-flow conditions to congested situations. Common approaches for achieving this include the utilization of k-means or fuzzy c-means clustering methods. However, a limitation of these clustering methods is the need to predefine the number of clusters. Additionally, the k-means method assumes spherical clusters and an equal distribution of observations among groups, which renders it less suitable for analyzing time-series data. The fuzzy c-means clustering method is not suitable for clustering data with uneven density or large numerical fluctuation in observed data. In a bottleneck or congested segment for a single lane, the time series data fluctuated significantly between congested and uncongested states, as shown in Fig. 2. Hence, employing the traditional clustering method to partition the traffic state would result in significant inaccuracies. This suggests that neither the k-means method

nor the fuzzy c-means clustering method is appropriate for handling time series data at the traffic bottleneck. Consequently, we categorize the traffic state by utilizing the FD, which is elaborated on in the following explanation.

Fig. 3 illustrates the limitations of the FD and time series plots in capturing the dynamic nature of traffic relationships over time. While the FD provides insights into traffic states and variable relationships, it does not fully capture the evolving nature of traffic. Conversely, time series plots effectively depict the temporal evolution of traffic but do not comprehensively capture traffic states through a combination of flow variables. In this study, we hope to capture the key parameters: capacity c , critical speed v_c , critical density k_c , free-flow speed v_f , which can be calibrated by the FD with observation data.

We could further use the critical speed or critical density to divide the traffic state into uncongested and congested conditions in the FD (as shown in Fig. 3). This FD-based method is more interpretable in terms of traffic flow characteristics. In addition, the number of clusters does not need to be estimated in advance. As a result, in this paper, we calibrated the FD to derive the critical parameters and further define traffic states accordingly into the bottleneck. The calibration method employed in this study involves collecting traffic flow data from observation equipment and using the least squares method to calibrate the speed-density function of the S3 model (Cheng et al., 2021, 2024). The S3 function is shown in Eq. (1).

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

The primary aim is to minimize the disparity between the observed speed data v_i and the estimated speed value \hat{v} calculated using the model. The unknown parameters in the objective function are the free-flow speed, critical density, and parameter m , as represented in Eq. (2).

$$\min Z = \sum_{i=1}^N [v_i - \hat{v}(k_i; v_f, k_c, m)]^2 \quad (2)$$

where v_f , k_c , and m are the parameters for calibration. N is the total amount of data, v_i and k_i are the observed speed data and density data.

By obtaining the critical density k_c from Eq. (2) and subsequently substituting it into Eq. (1), we can determine the critical speed v_c . This critical speed parameter serves as a means to differentiate between traffic congestion and non-congestion state.

2.2. State transition probability matrix in the Markovian model

The Markovian model operates on the assumption that the current state of the system is solely influenced by the preceding state, following the principle of memory lessness. In this model, the transition probability matrix is used to depict the stochastic process of traffic flow evolution.

The transition probability from state i to state j is represented as p_{ij} .

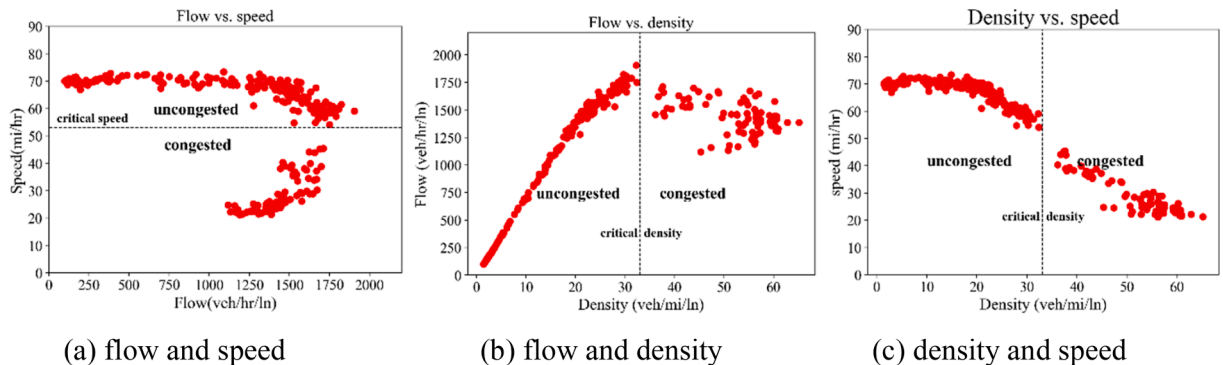


Fig. 3. Fundamental diagram of the bottleneck.

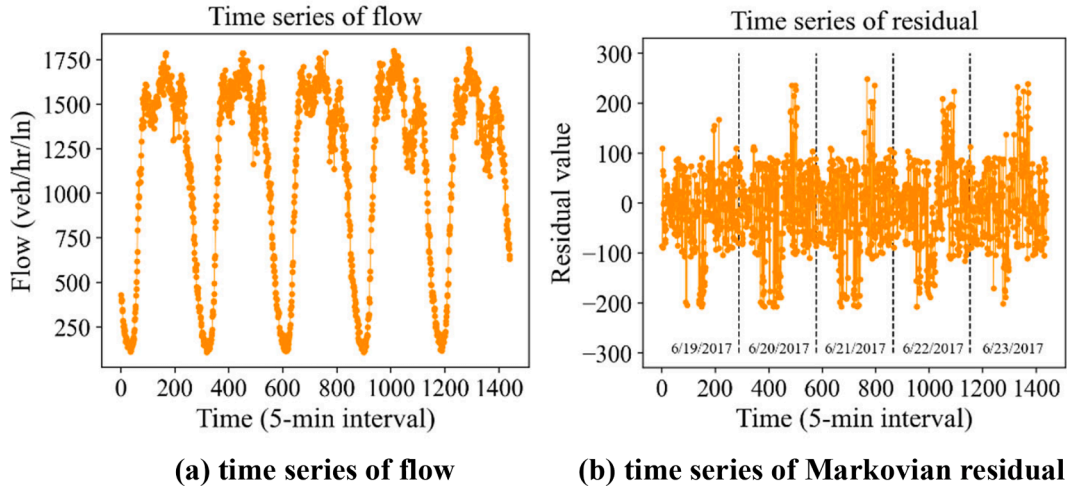


Fig. 4. Time series of data. (a) time series of original flow measurements; (b) time series of Markovian residuals.

The total number of transitions from state index S_i to state index S_j is denoted as m_{ij} , while m_i represents the total number of transitions from state index S_i to all other states, including state S_i itself. These definitions lead to the Eq. (3) as stated.

$$P(S_j|S_i) = p_{ij} = \frac{m_{ij}}{m_i}, i, j \in \{1, 2, \dots, n\} \quad (3)$$

The transition probability can be defined by the following Eq. (4).

$$P = [p_{ij}], i, j \in \{1, 2, \dots, n\} \quad (4)$$

In Eq. (4), p_{ij} is the transition probability from the state index S_i transition to state index S_j . For all $i, j \in \{1, 2, \dots, n\}$, $p_{ij} \geq 0$ and $\sum_{j \in \{1, 2, \dots, n\}} p_{ij} = 1$, $i = 1, 2, \dots, n$. The details of the notations used in this paper can be found in Table 9 in the Appendix.

P in Eq. (4) can be partitioned into four blocks for an FD in terms of the uncongested and congested period, as shown in Eq. (5).

$$P = \begin{bmatrix} P_{UU} & P_{UC} \\ P_{CU} & P_{CC} \end{bmatrix} \quad (5)$$

where P_{UU} represents the transfer from an uncongested state to another uncongested state. P_{UC} denotes the transition matrix from an uncongested state to a congested state. P_{CU} signifies the transition matrix from a congested state to an uncongested state. Finally, P_{CC} represents the transition matrix from a congested state to another congested state.

In the initial step of constructing the Markovian model, it is essential to systematically define different states. In this study, each state S_i within the Markov chain model is precisely characterized by a tuple comprising flow, density, and speed ranges. This representation encompasses a range of flow values, density values, and a specific speed level, enabling a comprehensive description of the traffic state. The FD describes the congestion level by the critical speed, which is used to distinguish the uncongested and congested conditions; namely, the traffic flow condition satisfies $\{(q, k, v) | v > v_c\}$ is classified as an uncongested condition, and it is classified as a congested situation if the traffic flow satisfies $\{(q, k, v) | v \leq v_c\}$. Furthermore, based on the uncongested and congested identification, we divide the flow and density data into intervals to obtain different traffic states.

2.3. Markovian-based model for capturing regular patterns

In this research, $X(t)$ is the observed link volume, S_i is the state of $X(t)$ after the classification. According to the transition probability matrix and the state of the initial predicted data, we could further predict the state of the subsequent time-series data by following Eq. (6), which

denotes how the current state S_i is most likely to transfer to state S_k .

$$p_{ik}(t) = \max(p_{ik}(t)), i, k \in \{1, 2, \dots, n\}, t \in \{1, 2, \dots, T\} \quad (6)$$

Then we could derive the predicted link flow for the state S_k , as shown in Eq. (7).

$$M(t) = \bar{q}_k(t) \bullet p_{ik}(t), i, k \in \{1, 2, \dots, n\}, t \in \{1, 2, \dots, T\} \quad (7)$$

where $M(t)$ is the estimated flow values by the Markov model at time t , $\bar{q}_k(t)$ is the average flow of all the traffic flow divided into states k at time t , $p_{ik}(t)$ is the transition probability from state i to another state k at time t .

The residual between the observed flow and the estimated flow could be derived as a new variable $r(t)$.

$$r(t) = X(t) - M(t), t \in \{1, 2, \dots, T\} \quad (8)$$

Fig. 4 shows a time series of flow plots using traffic flow values collected from a loop detector on the Los Angeles I-405 corridor, with a periodic pattern as shown in Fig. 4 (a). Fig. 4 (b) further depicts the residual time series of flow produced from the Markovian model with an interpretable state transition matrix. It indicates that there is still a significant repeating pattern within the time series. Therefore, the use of the LSTM model is motivated to capture additional traffic features present in the residual series of flow, which the Markov chain model might not capture comprehensively. This approach allows for a more thorough exploration of the traffic patterns and characteristics that may be overlooked by the Markovian model alone.

2.4. LSTM-based model for capturing the traffic features within the residual time series

LSTM is extensively employed in traffic flow prediction due to its unique capabilities and consistent outperformance of traditional time series models in various benchmarks (Chen et al., 2019; Lu et al., 2022). LSTM was chosen for several reasons: firstly, LSTM's aptitude for handling sequential data and capturing temporal dynamics, both short and long-term, aligns perfectly with the inherent sequential nature of traffic flow data. Secondly, its proficiency in capturing complex and nonlinear traffic patterns, which traditional linear models struggle with, is attributed to the recursive neural network (RNN) architecture's ability to decipher intricate data relationships. Thirdly, LSTM excels in integrating diverse data sources like traffic sensors, weather data, and historical patterns, significantly enhancing prediction accuracy. Moreover, LSTM's automatic feature learning reduces the need for extensive manual feature engineering, which proves valuable when dealing with intricate and time-varying relationships between factors such as weather

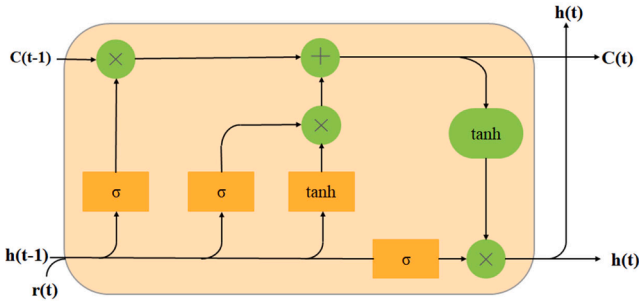


Fig. 5. The structure of LSTM.

conditions, holidays, and accidents.

Traffic flow data represents a non-stationary random sequence, yet it displays discernible regularities and clear trends within a continuous time series. The proposed approach selects LSTM networks, a specific type of recurrent neural network model, for short-term traffic prediction. LSTM is well-suited for this task because it has the ability to (i) handle time series data (Abdoos and Bazzan, 2021) and (ii) capture the pronounced fluctuations and nonlinearity present in the data (Ma et al., 2015). As a result, this paper utilizes the LSTM model to effectively capture the flow characteristics within the residual time series.

We use the residual derived by the Markovian model as the input variable in the LSTM model. The model input is denoted as $r(t)$, and the output sequence is denoted as $L(t)$, where T is the maximum time index, and $L(t)$ is the residual time series predicted by the LSTM. The structure of LSTM is shown in Fig. 5. The LSTM model employed in this paper is composed of one input layer, three hidden layers, and one output layer. It is worth noting that when working with time series data using LSTM, increasing the number of hidden layers can lead to a substantial increase in computation time and memory usage (Hu and Chen, 2018; Hua et al., 2019). Determining the various parameters used in the LSTM models is a crucial aspect of the modeling process, ensuring fair comparisons and optimal performance (Pan et al., 2022). While some parameters are automatically adjusted during model training through learning, hyperparameters must be set manually. These hyperparameters encompass critical elements such as the number of input layers, hidden layer configurations (including the number of hidden layers and nodes within them), output layer specifications, activation functions, loss functions, optimization algorithms, learning rates, and iteration counts, among others. In the context of our study, we adopted LSTM models with distinct network structures to explore their impact on predictive performance. Specifically, our models included one input layer, followed by multiple hidden layers. To ensure fairness and rigor in our comparisons,

we conducted experiments with LSTM models employing one, two, and three hidden layers, respectively, all trained on the same dataset. The selection of the optimal model was based on minimizing error metrics obtained through rigorous testing and evaluation, ensuring a comprehensive and unbiased assessment of model performance.

The first step in the LSTM model is the utilization of the forget gate, denoted as $f(t)$. It employs the sigmoid function δ to attenuate irrelevant noise, ensuring that the values of $f(t)$ are confined within the range of 0 to 1.

$$f(t) = \delta[W_{ff} \bullet h(t-1) + W_{rf} \bullet r(t-1) + b_f] \quad (9)$$

The second step in the LSTM model involves the input gate, denoted as $i(t)$, and the external input gate, denoted as $g(t)$. These gates play a crucial role in updating and determining the incorporation of new information into the LSTM model.

$$i(t) = \delta[W_{ii} \bullet h(t-1) + W_{ri} \bullet r(t-1) + b_i] \quad (10)$$

$$g(t) = \tanh[W_{gg} \bullet h(t-1) + W_{rg} \bullet r(t-1) + b_g] \quad (11)$$

The third step in the LSTM model focuses on updating the previous cell state, denoted as $s(t)$. This step ensures that the LSTM model retains relevant information from previous time steps while discarding unnecessary information.

$$s(t) = f(t) \otimes f(t) + i(t) \otimes g(t) \quad (12)$$

The last step in the LSTM model is the output gate $o(t)$.

$$o(t) = \delta[W_{oo} \bullet h(t-1) + W_{ro} \bullet r(t-1) + b_o] \quad (13)$$

$$h(t) = o(t) \otimes \tanh[s(t)] \quad (14)$$

δ denotes the sigmoid function.

$$\delta(r) = \frac{1}{1 + e^{-r}} \quad (15)$$

where W_{ff} , W_{rf} , W_{gg} , W_{rg} , W_{oo} , W_{ro} are weight and b_f , b_g , b_o are bias. h_t , r_t are variable, \otimes denotes the Hadamard product, $t \in \{1, 2, \dots, T\}$. The details of the notations used in this paper can be found in Table 10 in the Appendix.

2.5. Integrated modeling framework: FD-Markov-LSTM

In this study, we present a hybrid model that seamlessly integrates three distinct components: the FD, Markov chain, and LSTM models. By combining the strengths of traffic-theoretic statistical models with machine learning approaches, our proposed method is expected to syner-

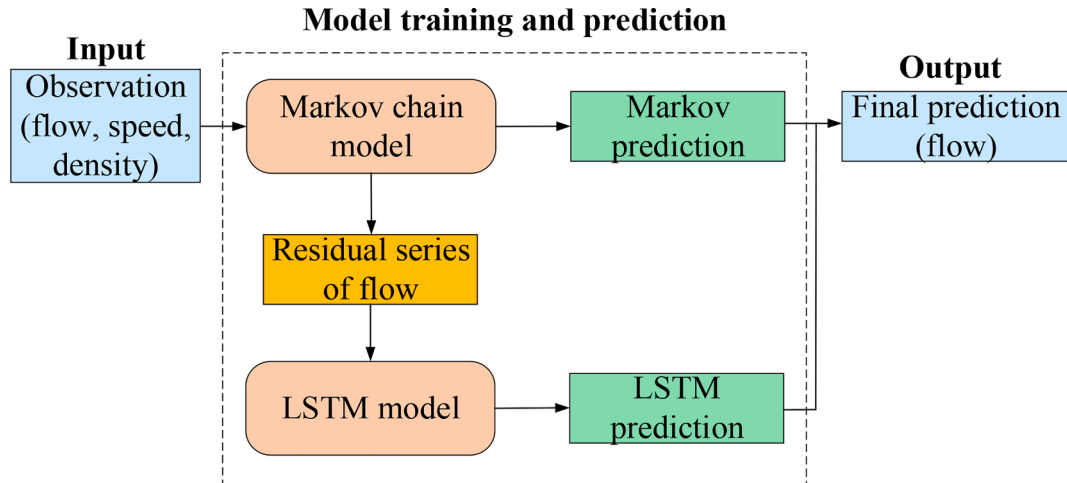


Fig. 6. The structure of the proposed model.



Fig. 7. Areas of the selected corridors used in the datasets.

Table 2
Description of empirical datasets.

Data set	Data type and information	Temporal resolution	Time period of data collection	Bottleneck location	Corridor location
Dataset 1: Los Angeles I-405	Loop detectors with traffic flow, speed, and occupancy data	5-min interval	(00:00 a.m.-12:00p.m.) Training data: June 19 to 22, 2017 Test data: June 23, 2017	Absolute Postmile (Abs) 13.51	I-405 in Los Angeles, California, US
Dataset 2: Beijing West Third Ring	Remote traffic microwave sensor with traffic flow, speed data	2-min interval	(06:00 a.m.-10:00p.m.) Training data: June 25 to 28, 2018 Test data: June 29, 2018	Zizhu Bridge HI2090c	West Third Ring in Beijing, China
Dataset 3: Beijing North Third Ring	Remote traffic microwave sensor with traffic flow, speed, and occupancy data	10-min interval	(00:00 a.m.-12:00p.m.) Training data: June 18 to 21, 2013 Test data: June 27, 2013	Jimen Bridge ID3059	West Third Ring in Beijing, China

Table 3
Calibrated parameters in the three datasets.

Datasets	Free-flow speed	Critical speed	Critical density	Capacity
Los Angeles I-405	65 (mi/hr)	49 (mi/hr)	33 (veh/mi/ln)	1617 (veh/hr/ln)
Beijing West Third Ring	65 (km/hr)	33(km/hr)	37 (veh/km/ln)	1221 (veh/hr/ln)
Beijing North Third Ring	64 (km/hr)	34 (km/hr)	50 (veh/km/ln)	1700 (veh/hr/ln)

gistically complement one another and effectively capture complex patterns in traffic flows, thereby improving the accuracy of estimation and prediction. The simplified structure of the proposed FD-Markov-LSTM model is shown in Fig. 6. The comprehensive model can be written in Eq. (16).

$$X(t) = M(t) + L(t), t \in \{1, 2, \dots, T\} \quad (16)$$

where $X(t)$ represents the time series observation flow, $M(t)$ is the prediction of the Markov model, $L(t)$ is the residual time series predicted by the LSTM.

3. Case study

3.1. Data description

In order to demonstrate the subtle but essential difference between the various traffic flow prediction models, we use three different datasets for the examples shown in Fig. 7 with the descriptions in Table 2. It should be emphasized that the three selected corridors are one of busiest

freeways in California and the city of Beijing.

Traffic data for the busiest freeways in California is sourced from the Performance Measurement System (PEMS), a widely utilized repository for transportation research and analysis. PEMS gathers real-time traffic data, encompassing traffic flow, speed, and occupancy, from sensors and detectors on major California highways. This publicly available data can be accessed via the California Department of Transportation's website (<https://pems.dot.ca.gov/>) or data portal. Meanwhile, in Beijing, traffic data is collected through strategically placed loop detectors managed by local transportation authorities. These detectors capture real-time information on flow, speed, and occupancy, which is then aggregated and processed. Accessing this Beijing data typically necessitates coordination with relevant transportation agencies due to its proprietary sensor networks and data management systems (Fig. 8).

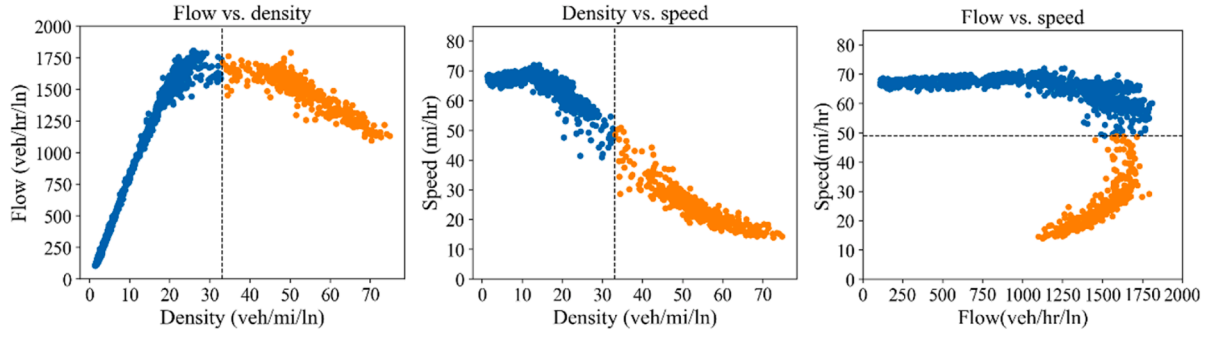
The raw data comprised aggregated flows, occupancies, and spot speeds for each mainline lane. To transform the occupancy data into density data, the conversion equation Eq. (17) introduced by May (1990) was employed.

$$k = \frac{5280 \cdot occ}{l + d} \quad (17)$$

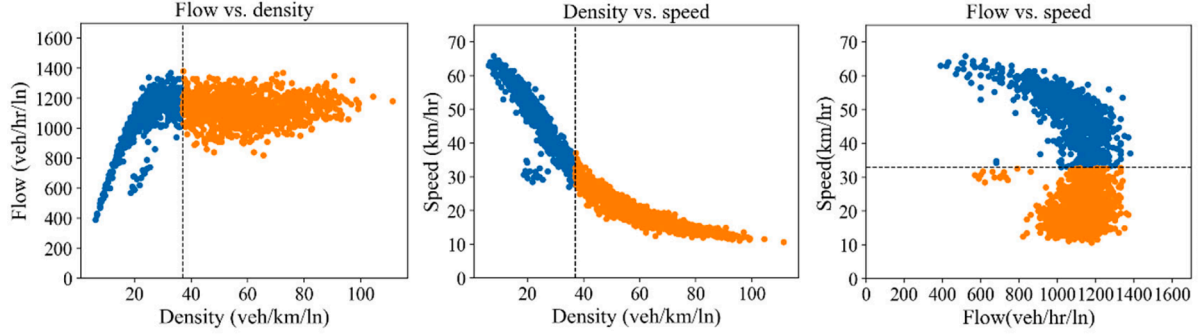
where occ is occupancy, l is the average length of the vehicle, and d is the average area of influence for a detector. In this paper, $l + d$ is 25 feet.

3.2. Evaluation metrics for prediction results

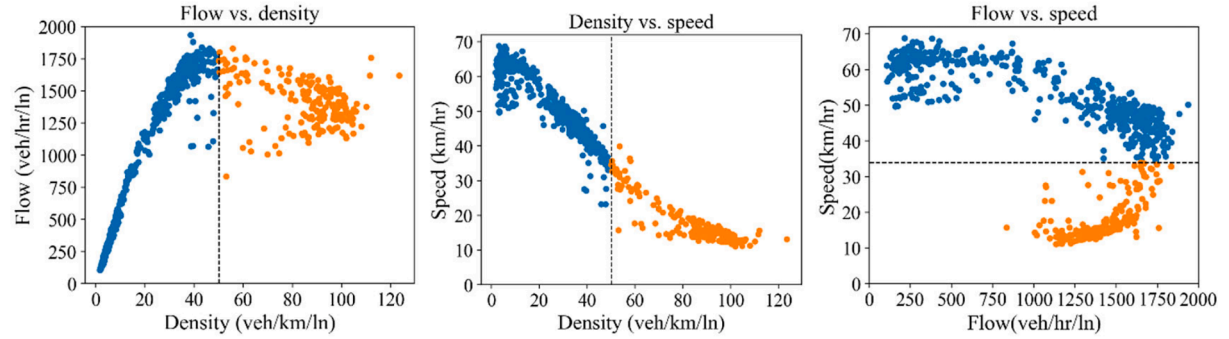
To assess the performance of different traffic flow prediction models, it is essential to establish evaluation criteria that can gauge prediction quality and facilitate method enhancement. This evaluation process primarily involves comparing predicted traffic outcomes with real-time traffic conditions. Three key metrics we will delve into are Mean



(a) Los Angeles I-405



(b) Beijing West Third Ring



(c) Beijing North Third Ring

Fig. 8. Uncongested and congested state identification in the three datasets.

Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

MAE is computed as the mean of the discrepancies between the actual values (ground truth) and the predicted values. In mathematical notation, it is represented as:

$$MAE = \frac{1}{M} \sum_{m=1}^M |measured_m - estimated_m| \quad (18)$$

where $measured_m$ is the measured traffic flow value for observation m , $estimated_m$ is the estimated traffic flow value for observation m , M is the total number of flow counts.

RMSE is the square root of the mean of the squared disparities between the predicted values generated by the regression model and the actual target values. This can be expressed mathematically as:

$$RMSE = \sqrt{\frac{\sum_{m=1}^M (measured_m - estimated_m)^2}{M}} \quad (19)$$

MAPE is commonly employed in traffic forecasting to mitigate the issue of a small number of instances with low traffic flow having a disproportionate impact on error measurement. It takes into account not only the absolute error between the predicted and actual values but also the relative error as a ratio to the actual value. The definition of MAPE is expressed as:

$$MAPE = \frac{100\%}{M} \sum_{m=1}^M \frac{|measured_m - estimated_m|}{measured_m} \quad (20)$$

4. Results and discussion

4.1. Traffic state identification

The FD establishes a relationship between flow, speed, and density, which captures the characteristics of traffic flow and road segments. We could use the critical speed v_c to distinguish the congested and uncongested conditions. In Section 2.1, we provide a detailed description of how the least squares methodology is applied to calibrate the observed

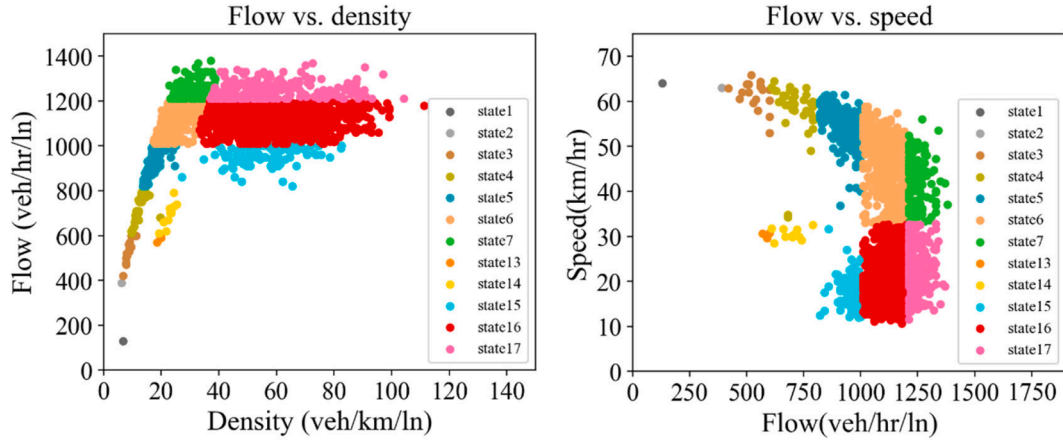
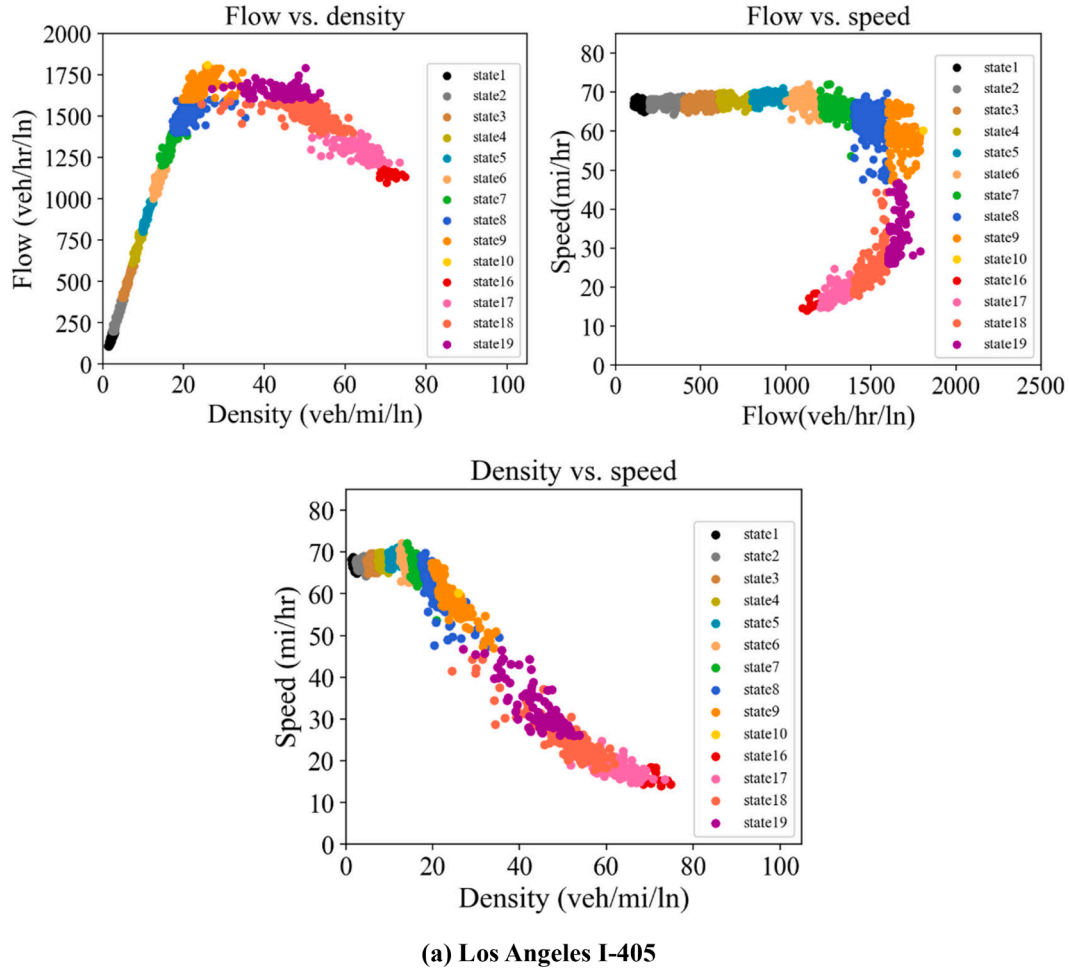


Fig. 9. Different traffic flow state identification in the Markovian model.

data and derive the key parameters of the FD.

Defining different states systematically is a crucial step in establishing the Markovian model. In this study, each state in the Markov chain model is precisely defined using a tuple that includes ranges for flow, density, and speed. This approach effectively captures the variations in traffic conditions by incorporating specific flow and density intervals along with speed levels. First, we divide traffic flow into the uncongested and congested state by the critical speed, which is calibrated by FD. Then, we divide traffic flow and density into different states by equal intervals, and eventually, we can obtain 20 different

traffic states in the Markovian model. We plot the fundamental diagram with different states defined in the Markovian model, as shown in Fig. 9. It is obvious that some states are unavailable in the observation data, such as states 8–12 and 18–20 in the Beijing West Third Ring dataset. These states correspond to a flow of more than 1400 veh/hr/ln and less than 400 veh/hr/ln, which is consistent with the FD. In addition, it is feasible and visible for transportation planners and analysts to examine the state transition from uncongested to congested conditions.

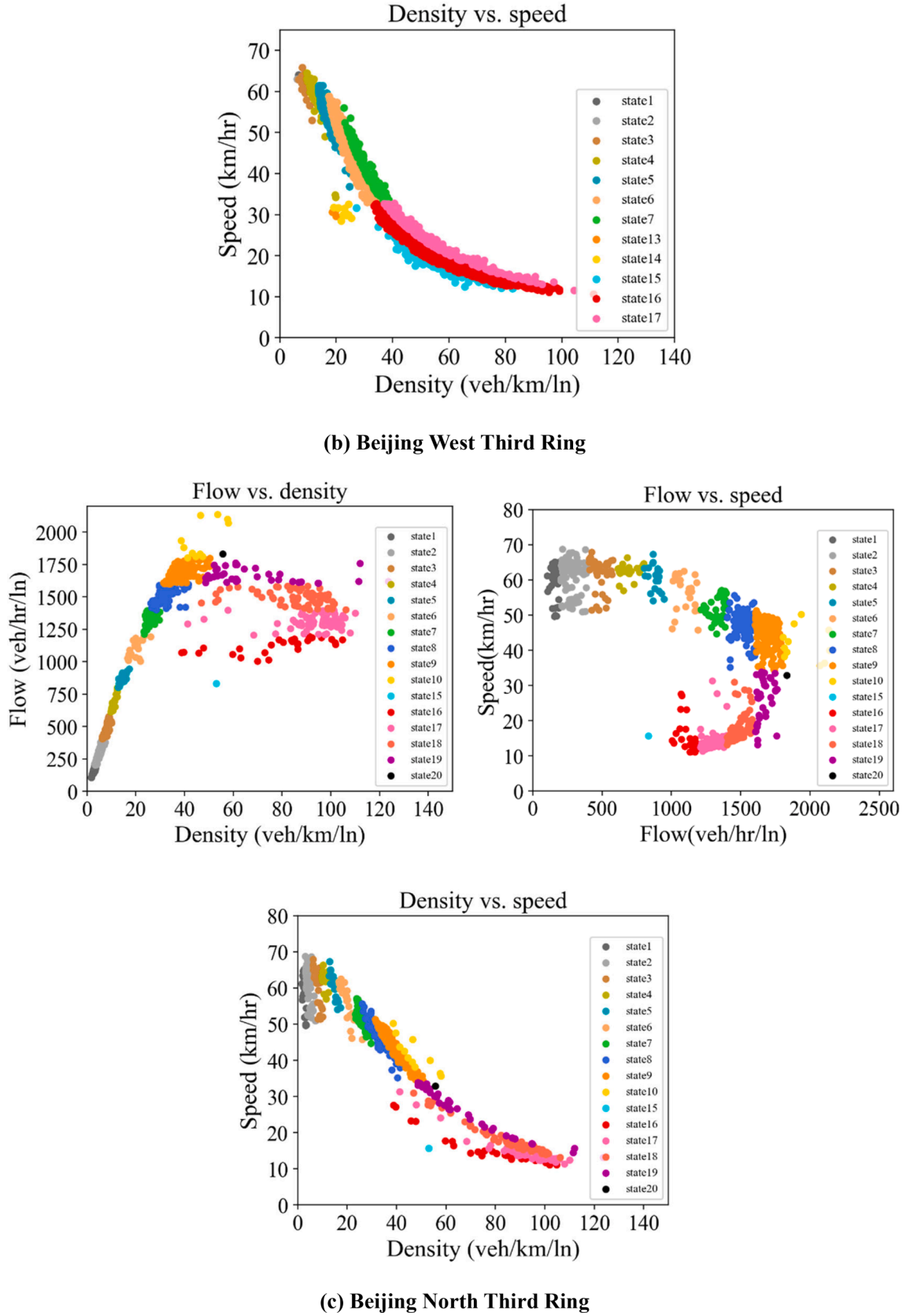


Fig. 9. (continued).

4.2. State transition probability matrices

The transition probability matrix is partitioned into four blocks: P_{UU} ,

P_{UC} , P_{CU} and P_{CC} . The transition probability matrices derived from the Markovian model revealed the probability of traffic flow transitioning to various possible states.

Table 4The 5-min state transition probability matrix in Los Angeles I-405: Transition matrix P_{UU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_1	0.908	0.092	0	0	0	0	0	0	0	0
S_2	0.102	0.820	0.078	0	0	0	0	0	0	0
S_3	0	0.137	0.767	0.096	0	0	0	0	0	0
S_4	0	0	0.128	0.723	0.149	0	0	0	0	0
S_5	0	0	0	0.115	0.672	0.213	0	0	0	0
S_6	0	0	0	0	0.153	0.671	0.176	0	0	0
S_7	0	0	0	0	0	0.174	0.593	0.233	0	0
S_8	0	0	0	0	0	0	0.068	0.760	0.172	0
S_9	0	0	0	0	0	0	0	0.286	0.677	0.006
S_{10}	0	0	0	0	0	0	0	0	1	0

Table 5The 5-min state transition probability in Los Angeles I-405: Transition matrix P_{UC} .

	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_1	0	0	0	0	0	0	0	0	0	0
S_2	0	0	0	0	0	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	0
S_4	0	0	0	0	0	0	0	0	0	0
S_5	0	0	0	0	0	0	0	0	0	0
S_6	0	0	0	0	0	0	0	0	0	0
S_7	0	0	0	0	0	0	0	0	0	0
S_8	0	0	0	0	0	0	0	0	0	0
S_9	0	0	0	0	0	0	0	0	0.031	0
S_{10}	0	0	0	0	0	0	0	0	0	0

Table 6The 5-min state transition probability in Los Angeles I-405: Transition matrix P_{CU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0	0	0	0	0	0	0
S_{14}	0	0	0	0	0	0	0	0	0	0
S_{15}	0	0	0	0	0	0	0	0	0	0
S_{16}	0	0	0	0	0	0	0	0	0	0
S_{17}	0	0	0	0	0	0	0	0	0	0
S_{18}	0	0	0	0	0	0	0	0.030	0	0
S_{19}	0	0	0	0	0	0	0	0	0	0
S_{20}	0	0	0	0	0	0	0	0	0	0

Taking the Los Angeles I-405 corridor as an example, Table 4 is the transition probability matrix of transferring from an uncongested state to another uncongested state P_{UU} . It shows that the uncongested state tends to evolve to its original state with a high probability. In the uncongested circumstances, the orange diagonals in the figure with the high value illustrate the pattern in the free-flow traffic period that will generally remain for a period of time before transferring. Besides, except for the possibility of maintaining the current state, there are some

transitions with the largest value from one state evolving to the before/after state, such as state 10 is most likely to evolve to state 9 with a probability of 100 %.

The transition matrix P_{UC} of transferring from an uncongested state to a congested state is shown in Table 5. It is obvious that most of the probability value is 0 %, indicating that the transition from an uncongested to a congested state is low frequency. This outcome can be attributed to the breakdown phenomenon, which is characterized by an

Table 7The 5-min state transition probability in Los Angeles I-405: Transition matrix P_{CC} .

	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0	0	0	0	0	0	0
S_{14}	0	0	0	0	0	0	0	0	0	0
S_{15}	0	0	0	0	0	0	0	0	0	0
S_{16}	0	0	0	0	0	0.125	0.750	0.125	0	0
S_{17}	0	0	0	0	0	0.110	0.560	0.330	0	0
S_{18}	0	0	0	0	0	0.024	0.161	0.613	0.173	0
S_{19}	0	0	0	0	0	0	0.012	0.388	0.600	0
S_{20}	0	0	0	0	0	0	0	0	0	0

Table 8

Comparisons of prediction performance with different models in datasets.

Datasets	Models	MAE	RMSE	MAPE
Dataset 1: Los Angeles I-405	Markov chain	92.47	122.96	10.42 %
	ARIMA	83.51	99.25	9.39 %
	KNN	88.04	110.28	9.43 %
	LSTM	68.09	88.73	8.19 %
	RF	224.71	342.03	57.34 %
	FD-Markov-LSTM	63.81	83.27	7.95 %
Dataset 2: Beijing West Third Ring	Markov chain	74.32	99.28	7.07 %
	ARIMA	104.34	74.98	7.12 %
	KNN	90.92	131.47	8.10 %
	LSTM	64.81	88.56	6.17 %
	RF	74.42	101.82	7.24 %
	FD-Markov-LSTM	37.22	45.87	3.32 %
Dataset 3: Beijing North Third Ring	Markov chain	126.90	189.02	9.96 %
	ARIMA	151.76	157.12	11.62 %
	KNN	124.86	180.63	11.78 %
	LSTM	123.57	165.40	12.81 %
	RF	199.29	311.63	38.78 %
	FD-Markov-LSTM	94.17	146.68	9.63 %

Table 9

Comparisons of prediction performance with the uncongested and congested period.

Datasets	Period	Models	MAE	RMSE	MAPE
Dataset 1: Los Angeles I-405	Uncongested period	Markov	82.67	110.10	11.31 %
		LSTM	55.08	70.69	8.01 %
		FD-Markov-LSTM	55.44	72.01	8.71 %
	Congested period	Markov	115.78	149.13	8.31 %
		LSTM	83.68	105.31	6.14 %
		FD-Markov-LSTM	71.99	95.14	5.22 %
Dataset 2: Beijing West Third Ring	Uncongested period	Markov	78.28	114.16	8.29 %
		LSTM	64.28	96.37	6.88 %
		FD-Markov-LSTM	47.91	68.63	5.22 %
	Congested period	Markov	72.60	92.10	6.55 %
		LSTM	65.04	84.96	5.86 %
		FD-Markov-LSTM	38.79	47.08	3.51 %
Dataset 3: Beijing North Third Ring	Uncongested period	Markov	121.32	181.40	15.15 %
		LSTM	119.12	161.37	13.67 %
		FD-Markov-LSTM	63.47	89.58	8.44 %
	Congested period	Markov	179.01	242.46	12.93 %
		LSTM	135.87	189.02	10.41 %
		FD-Markov-LSTM	126.90	176.05	9.56 %

instantaneous and abrupt reduction in traffic flow or speed.

In addition, there is a similar phenomenon that the transition probability matrix P_{CU} has a small transition probability value, as shown in Table 6. With the P_{CC} for congested states, each element in the matrix represents the probability of transitioning from one congested state to another. Table 7 displays the probabilities of traffic states transitioning either within the congested state or to a different state. Notably, the orange data in the matrices represents the highest probabilities of remaining within the same state, while the green data indicates the highest probabilities of transitioning from one state to another. Tables 10–17 in the Appendix provide the state transition probabilities for the remaining two cases.

Table 10

Nomenclature used in this paper.

Symbols	Definitions
i, j, k	State number, $i, j, k \in \{1, 2, \dots, n\}$
t	Time index, $t \in \{1, 2, \dots, T\}$
$X(t)$	The sequence of observed flow measurements
$M(t)$	The sequence of estimated flow values by the Markov model
$L(t)$	The residual time series predicted by the LSTM
$r(t)$	The residual flow between observed data and estimated data by the Markov model
S_i	State index of state i
m_i	The total number of transitions from the state index S_i to all other states
m_{ij}	The total number of transitions from the state index S_i to state index S_j
p_{ij}	The transition probability from one state i to another state j
P	State transition matrix
P_{UU}	Transition matrix representing the transfer from one uncongested state to another uncongested state
P_{UC}	Transition matrix of transferring from an uncongested state to a congested state
P_{CU}	Transition matrix of transferring from a congested state to an uncongested state
P_{CC}	Transition matrix of transferring from a congested state to another congested state
q_i	Flow range of state i
k_i	Density range of state i
$\bar{q}_k(t)$	The average flow of all the traffic flow divided into state k at time t
v_f	Free-flow speed
v_c	Critical speed
k_c	Critical density
c	Capacity
$f(t)$	Forget gate
δ	Denotes the standard logistics sigmoid function
W_{ff}	Weight from previous forget gate to current forget gate
W_{rf}	Weight from the input layer to forget gate
$h(t-1)$	Hidden variable from the previous time step $t-1$
$h(t)$	Hidden variable at the current time step t
$r(t-1)$	Input variable (Residuals from Markov model)
b_f	Bias at forget gate
$i(t)$	Input gate
W_{ii}	Weight from the previous input gate to the current input gate
W_{ri}	Weight from input-to-input layer gate
b_i	Bias at the input gate
$g(t)$	External input gate
W_{gg}	Weight from previous external gate to current external gate
W_{rg}	Weight from input to external input gate
$s(t)$	Variable at current time step t
$o(t)$	Output gate
W_{oo}	Weight from the previous output gate to the current output gate
W_{ro}	Weight from input to output gate
b_o	Bias at output gate

4.3. Performance comparison of the proposed model with other benchmark models

In addition to the hybrid interpretable models proposed in this study, several benchmark models were tested in the experiments. The benchmark algorithms consisted of two traditional time-series forecasting models, namely the Markov chain model and the ARIMA model, as well as several machine learning approaches, such as k-Nearest Neighbor (KNN), Random Forest (RF), and LSTM. The performance comparison of the compared model is shown in Table 8. Three case studies were conducted using empirical data collected in Beijing and Los Angeles to evaluate the accuracy of traffic flow estimation and prediction achieved by the proposed FD-Markov-LSTM model. The results demonstrate superior performance compared to classical benchmark models, indicating higher prediction accuracy. Specifically, the MAE of the FD-Markov-LSTM model is reduced by an average of 39 % compared to the benchmark models, while the RMSE is reduced by an average of 35 %. Furthermore, the proposed FD-Markov-LSTM model shows an average reduction of 7.4 % in MAPE. These improvements across all performance indicators clearly indicate that the FD-Markov-LSTM model outperforms other models, resulting in more accurate traffic flow predictions, particularly in bottleneck scenarios (Table 18).

Table 11The 2-min state transition probability in Beijing West Third Ring: Transition matrix P_{UU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_1	0	0	1	0	0	0	0	0	0	0
S_2	0	0	1	0	0	0	0	0	0	0
S_3	0	0	0.500	0.444	0	0.056	0	0	0	0
S_4	0	0	0.108	0.432	0.351	0.027	0	0	0	0
S_5	0	0.005	0.005	0.048	0.527	0.396	0.019	0	0	0
S_6	0	0	0.004	0.002	0.163	0.626	0.130	0	0	0
S_7	0	0	0	0	0.026	0.436	0.308	0	0	0
S_8	0	0	0	0	0	0	0	0	0	0
S_9	0	0	0	0	0	0	0	0	0	0
S_{10}	0	0	0	0	0	0	0	0	0	0

Table 12The 2-min state transition probability in Beijing West Third Ring: Transition matrix P_{UC} .

	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_1	0	0	0	0	0	0	0	0	0	0
S_2	0	0	0	0	0	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	0
S_4	0	0	0.027	0.027	0	0.027	0	0	0	0
S_5	0	0	0	0	0	0	0	0	0	0
S_6	0	0	0	0	0	0.061	0.015	0	0	0
S_7	0	0	0	0	0	0.111	0.120	0	0	0
S_8	0	0	0	0	0	0	0	0	0	0
S_9	0	0	0	0	0	0	0	0	0	0
S_{10}	0	0	0	0	0	0	0	0	0	0

Table 13The 2-min state transition probability in Beijing West Third Ring: Transition matrix P_{CU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0	0	0	0	0	0	0
S_{14}	0	0	0	0.182	0	0	0	0	0	0
S_{15}	0	0	0	0	0.010	0.029	0	0	0	0
S_{16}	0	0	0	0	0.003	0.028	0.010	0	0	0
S_{17}	0	0	0	0	0	0.035	0.012	0	0	0
S_{18}	0	0	0	0	0	0	0	0	0	0
S_{19}	0	0	0	0	0	0	0	0	0	0
S_{20}	0	0	0	0	0	0	0	0	0	0

Table 14The 2-min state transition probability in Beijing West Third Ring: Transition matrix P_{CC} .

	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0.667	0	0.333	0	0	0	0
S_{14}	0	0	0.182	0.545	0.091	0	0	0	0	0
S_{15}	0	0	0	0.010	0.343	0.533	0.076	0	0	0
S_{16}	0	0	0	0.001	0.055	0.767	0.136	0	0	0
S_{17}	0	0	0	0	0.015	0.660	0.278	0	0	0
S_{18}	0	0	0	0	0	0	0	0	0	0
S_{19}	0	0	0	0	0	0	0	0	0	0
S_{20}	0	0	0	0	0	0	0	0	0	0

This study employs a hybrid approach that integrates multiple models to enhance prediction accuracy and adaptability to diverse traffic flow conditions, particularly in complex congestion scenarios. When compared to both the basic Markov model and the LSTM model, the hybrid model in this study demonstrates superior performance. Additionally, hybrid models mitigate overfitting risk by amalgamating multiple methods, as each model may exhibit overfitting in different

aspects, thereby enhancing the model's generalization capacity. However, creating a hybrid model through multiple stages can elevate model complexity and computational demands, necessitating increased computational resources and training time. Consequently, this model may not be suitable for large-scale traffic management but proves advantageous for precise traffic control and congestion management. It is important to note that this study solely evaluated prediction accuracy,

Table 15The 10-min state transition probability in Beijing North Third Ring: Transition matrix P_{UU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_1	0.681	0.319	0	0	0	0	0	0	0	0
S_2	0.156	0.771	0.073	0	0	0	0	0	0	0
S_3	0	0.200	0.629	0.143	0.029	0	0	0	0	0
S_4	0	0	0.227	0.591	0.182	0	0	0	0	0
S_5	0	0	0	0.250	0.450	0.300	0	0	0	0
S_6	0	0	0	0.208	0.333	0.250	0.125	0.083	0	0
S_7	0	0	0	0	0.027	0.189	0.324	0.405	0.054	0
S_8	0	0	0	0	0	0.011	0.163	0.467	0.283	0.033
S_9	0	0	0	0	0	0.015	0.015	0.185	0.593	0.044
S_{10}	0	0	0	0	0	0	0	0.182	0.545	0

Table 16The 10-min state transition probability in Beijing North Third Ring: Transition matrix P_{UC} .

	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_1	0	0	0	0	0	0	0	0	0	0
S_2	0	0	0	0	0	0	0	0	0	0
S_3	0	0	0	0	0	0	0	0	0	0
S_4	0	0	0	0	0	0	0	0	0	0
S_5	0	0	0	0	0	0	0	0	0	0
S_6	0	0	0	0	0	0	0	0	0	0
S_7	0	0	0	0	0	0	0	0	0	0
S_8	0	0	0	0	0	0.011	0	0.022	0.011	0
S_9	0	0	0	0	0	0	0	0.037	0.104	0.007
S_{10}	0	0	0	0	0	0	0	0	0.273	0

Table 17The 10-min state transition probability in Beijing North Third Ring: Transition matrix P_{CU} .

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0	0	0	0	0	0	0
S_{14}	0	0	0	0	0	0	0	0	0	0
S_{15}	0	0	0	0	0	0	0	0	0	0
S_{16}	0	0	0	0	0	0	0.036	0	0	0
S_{17}	0	0	0	0	0	0	0.015	0.015	0	0
S_{18}	0	0	0	0	0	0	0	0	0.129	0.029
S_{19}	0	0	0	0	0	0	0	0.086	0.286	0
S_{20}	0	0	0	0	0	0	0	0	0	0

Table 18The 10-min state transition probability in Beijing North Third Ring: Transition matrix P_{CC} .

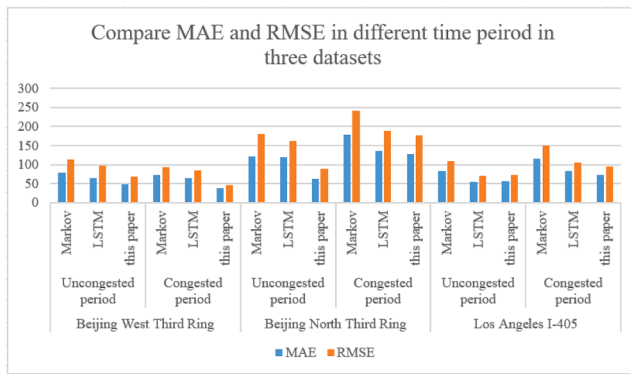
	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}	S_{20}
S_{11}	0	0	0	0	0	0	0	0	0	0
S_{12}	0	0	0	0	0	0	0	0	0	0
S_{13}	0	0	0	0	0	0	0	0	0	0
S_{14}	0	0	0	0	0	0	0	0	0	0
S_{15}	0	0	0	0	0	1	0	0	0	0
S_{16}	0	0	0	0	0	0.536	0.357	0.071	0	0
S_{17}	0	0	0	0	0	0.123	0.585	0.200	0.062	0
S_{18}	0	0	0	0	0	0.029	0.229	0.486	0.100	0
S_{19}	0	0	0	0	0.029	0.029	0.029	0.371	0.171	0
S_{20}	0	0	0	0	0	0	0	1	0	0

omitting considerations of model training speed and time.

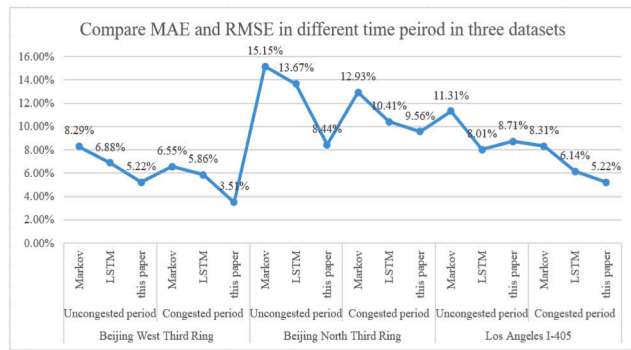
4.4. Performance comparison of congested and uncongested period

Furthermore, the objective of this paper is to assess the accuracy of the proposed method across different traffic flow conditions. To accomplish this, we partitioned the data from a single day into two

categories: congestion periods and non-congestion periods based on the prevailing congestion conditions. Subsequently, we conducted separate tests on the identified study cases. According to traffic state identification in [Section 4.1](#), we could use the calibrated critical speed v_c in [Table 3](#) to distinguish the uncongested and congested periods. Then we could provide an analysis of prediction performances with different traffic conditions, and the results are shown in [Table 9](#). It is obvious that



(a) comparison of MAE and RMSE



(b) comparison of MAPE

Fig. 10. Performance comparison of the dataset during the different periods.

the method proposed in this article has small errors, indicating that the proposed hybrid interpretable model could reasonably capture the congestion and uncongested patterns.

Furthermore, we conducted a comparison of model performance in different time periods, including all-day, congested, and uncongested periods, as depicted in Fig. 10. The superiority of the proposed model is evident, particularly during congested periods, demonstrating its suitability for accurate traffic flow estimation and prediction in the presence of congestion, despite the increased complexity associated with congested traffic conditions.

Our study presents several key advantages in the realm of traffic flow estimation and prediction. Firstly, we introduce the FD-Markov-LSTM hybrid model, combining the strengths of the fundamental diagram, Markov chain, and LSTM to enhance accuracy and interpretability. This innovative approach addresses limitations inherent in conventional models and consistently outperforms classical benchmark models, establishing its superior accuracy in traffic flow prediction. However, it is important to note some limitations. Our model's complexity may pose implementation and resource challenges compared to simpler, traditional statistical approaches. Moreover, while excelling in short- and medium-term trend prediction, further research is required to bolster its long-term forecasting capabilities, particularly for extended time intervals. Additionally, the method relies on historical data, and its performance could be influenced by specific dataset characteristics in the case studies.

5. Conclusions

In summary, this study introduces the FD-Markov-LSTM model, a pioneering hybrid framework for traffic flow estimation and prediction. This unique model amalgamates the strengths of FD, Markov chain, and LSTM techniques to elevate the interpretability and accuracy of traffic forecasting. The key contributions of this research encompass improved

traffic state identification through FD integration, precise state-to-state transition prediction using the Markov-based model, the introduction of an interpretable hybrid framework combining Markov and LSTM, and superior short- and medium-term trend prediction. Empirical findings consistently validate the FD-Markov-LSTM model's superior performance compared to traditional benchmark models.

The study's results highlight the outstanding predictive capabilities of the FD-Markov-LSTM model compared to classical benchmark models. When pitted against established benchmarks like Markov, ARIMA, k-Nearest Neighbor, Random Forest, and LSTM, the FD-Markov-LSTM model exhibits substantial improvement, with a remarkable 39 % reduction in MAE, over 35 % decrease in RMSE, and a notable 7.4 % drop in MAPE. This model signifies a significant leap forward in traffic flow forecasting, successfully bridging the gap between analytical and data-driven approaches. Its practical utility is underscored by the substantial reduction in prediction errors. These results underscore the FD-Markov-LSTM model's exceptional accuracy in traffic flow prediction, positioning it as a valuable choice for real-world applications within intelligent transportation systems.

Future research can focus on real-world validation, long-term forecasting refinements, and the incorporation of multi-source data fusion techniques, promising further improvements in dynamic traffic flow estimation and prediction for intelligent transportation systems. It can be developed in the following aspects: (1) To ensure the robustness and practicality of the proposed method, it is essential to validate its performance in real-world scenarios that involve various sources of nonrecurring delays, including adverse weather conditions, incidents, and work zones. (2) Further refinement of the proposed model is required to enhance its accuracy in making long-term forecasts, specifically for time intervals such as 30 min or 1 h. (3) To achieve more accurate estimation and prediction of dynamic traffic flow, it is recommended to employ multi-source data fusion techniques.

CRedit authorship contribution statement

Yuyan Annie Pan: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Jifu Guo:** Conceptualization, Methodology, Resources. **Yanyan Chen:** Conceptualization, Methodology, Supervision. **Qixiu Cheng:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – review & editing, Supervision. **Wenhao Li:** Conceptualization, Methodology, Validation. **Yanyue Liu:** Conceptualization, Methodology, Validation.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

References

- Abdoos, M., & Bazzan, A. L. (2021). Hierarchical traffic signal optimization using reinforcement learning and traffic prediction with long-short term memory. *Expert Systems with Applications*, 171, Article 114580.
- Antoniou, C., Koutsopoulos, H. N., & Yannis, G. (2013). Dynamic data-driven local traffic state estimation and prediction. *Transportation Research Part C: Emerging Technologies*, 34, 89–107.
- Ban, X., Chu, L., & Benouar, H. (2007). Bottleneck identification and calibration for corridor management planning. *Transportation Research Record*, 1999(1), 40–53.
- Bao, D. (2019). A multi-index fusion clustering strategy for traffic flow state identification. *IEEE Access*, 7, 166404–166409.
- Belezamo, B., Wu, X., Avci, C., & Zhou, X. (2019). Kinematic wave-oriented Markov Chain model to capture the spatiotemporal correlations of coupled traffic states. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (pp. 2343–2348). IEEE.

- Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., & Sun, J. (2016). A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. *Transportation Research Part C: Emerging Technologies*, 62, 21–34.
- Cai, L., Zhang, Z., Yang, J., Yu, Y., Zhou, T., & Qin, J. (2019). A noise-immune Kalman filter for short-term traffic flow forecasting. *Physica A: Statistical Mechanics and its Applications*, 536, Article 122601.
- Chen, M. R., Zeng, G. Q., Lu, K. D., & Weng, J. (2019). A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM. *IEEE Internet of Things Journal*, 6(4), 6997–7010.
- Chen, X., Zhang, S., & Li, L. (2019). Multi-model ensemble for short-term traffic flow prediction under normal and abnormal conditions. *IET Intelligent Transport Systems*, 13(2), 260–268.
- Chen, Y., Li, S., Pan, Y., & Zhang, J. (2022). Urban expressway congestion forewarning based on slope change of traffic flow fundamental diagram. *Journal of Transportation Engineering, Part A: Systems*, 148(6), 04022030.
- Cheng, Q., Lin, Y., Zhou, X. S., & Liu, Z. (2024). Analytical formulation for explaining the variations in traffic states: A fundamental diagram modeling perspective with stochastic parameters. *European Journal of Operational Research*, 312(1), 182–197.
- Cheng, Q., Liu, Z., Lin, Y., & Zhou, X. S. (2021). An s-shaped three-parameter (S3) traffic stream model with consistent car following relationship. *Transportation Research Part B: Methodological*, 153, 246–271.
- Comert, G., Khan, Z., Rahman, M., & Chowdhury, M. (2021). Grey models for short-term queue length predictions for adaptive traffic signal control. *Expert Systems with Applications*, 185, Article 115618.
- Elfar, A., Talebpoor, A., & Mahmassani, H. S. (2018). Machine learning approach to short-term traffic congestion prediction in a connected environment. *Transportation Research Record*, 2672(45), 185–195.
- Emami, A., Sarvi, M., & Asadi Bagloee, S. (2019). Using Kalman filter algorithm for short-term traffic flow prediction in a connected vehicle environment. *Journal of Modern Transportation*, 27, 222–232.
- Esfahani, R. K., Shahbazi, F., & Akbarzadeh, M. (2018). Three-phase classification of an uninterrupted traffic flow: A k-means clustering study. *Transportmetrica B: Transport Dynamics*, 7(1), 546–558.
- Evans, J. L., Eleftheriadou, L., & Gautam, N. (2001). Probability of breakdown at freeway merges using Markov chains. *Transportation Research Part B: Methodological*, 35(3), 237–254.
- Fusco, G., Colombaroni, C., & Isaenko, N. (2016a). Comparative analysis of implicit models for real-time short-term traffic predictions. *IET Intelligent Transport Systems*, 10(4), 270–278.
- Fusco, G., Colombaroni, C., & Isaenko, N. (2016b). Short-term speed predictions exploiting big data on large urban road networks. *Transportation Research Part C: Emerging Technologies*, 73, 183–201.
- Gao, Y., Li, J., Xu, Z., Liu, Z., Zhao, X., & Chen, J. (2021). A novel image-based convolutional neural network approach for traffic congestion estimation. *Expert Systems with Applications*, 180, Article 115037.
- Geroliminis, N., & Skabardonis, A. (2005). Prediction of arrival profiles and queue lengths along signalized arterials by using a Markov decision process. *Transportation Research Record*, 1934(1), 116–124.
- Guo, J., Huang, W., & Williams, B. M. (2014). Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. *Transportation Research Part C: Emerging Technologies*, 43, 50–64.
- Hou, Y., Zheng, X., Han, C., Wei, W., Scherer, R., & Polap, D. (2022). Deep learning methods in short-term traffic prediction: A survey. *Information Technology and Control*, 51(1), 139–157.
- Hu, Y. L., & Chen, L. (2018). A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. *Energy Conversion and Management*, 173, 123–142.
- Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019). Deep learning with long short-term memory for time series prediction. *IEEE Communications Magazine*, 57(6), 114–119.
- Huang, A. J., & Agarwal, S. (2020, September). *Physics informed deep learning for traffic state estimation* (pp. 1–6). IEEE.
- Huang, Y., Xu, L., & Guo, Y. (2011, December). The urban road traffic state identification method based on FCM clustering. In *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEEE)* (pp. 596–600).
- Kerner, B. S. (2013). Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory: A brief review. *Physica A: Statistical Mechanics and its Applications*, 392(21), 5261–5282.
- Khan, A., Fouda, M. M., Do, D. T., Almaleh, A., & Rahman, A. U. (2023). Short-term traffic prediction using deep learning long short-term memory: taxonomy, applications, challenges, and future trends. *IEEE Access*.
- Kharoufeh, J. P., & Gautam, N. (2004). Deriving link travel-time distributions via stochastic speed processes. *Transportation Science*, 38(1), 97–106.
- Kidando, E., Moses, R., Sando, T., & Ozguven, E. E. (2018). Evaluating recurring traffic congestion using change point regression and random variation Markov structured model. *Transportation Research Record*, 2672(20), 63–74.
- Kim, T., Sharda, S., Zhou, X., & Pendyala, R. M. (2020). 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*, 120, Article 102786.
- Krakovna, V., & Doshi-Velez, F. (2016). Increasing the interpretability of recurrent neural networks using hidden markov models. *arXiv preprint arXiv:1606.05320*.
- Kumar, S. V., & Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review*, 7(3), 1–9.
- Lablack, M., & Shen, Y. (2023). Spatio-temporal graph mixformer for traffic forecasting. *Expert Systems with Applications*, 228, Article 120281.
- Lee, K., Eo, M., Jung, E., Yoon, Y., & Rhee, W. (2021). Short-term traffic prediction with deep neural networks: A survey. *IEEE Access*, 9, 54739–54756.
- Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*.
- Li, T., Ma, J., & Lee, C. (2020). Markov-based time series modeling framework for traffic-network state prediction under various external conditions. *Journal of Transportation Engineering, Part A: Systems*, 146(6), 04020042.
- Lin, X. (2019). A road network traffic state identification method based on macroscopic fundamental diagram and spectral clustering and support vector machine. *Mathematical Problems in Engineering*, 2019, 6571237.
- Long, J., Gao, Z., Ren, H., & Lian, A. (2008). Urban traffic congestion propagation and bottleneck identification. *Science in China Series F: Information Sciences*, 51(7), 948–964.
- Lu, H. P., Sun, Z. Y., & Qu, W. C. (2015). Big data-driven based real-time traffic flow state identification and prediction. *Discrete Dynamics in Nature and Society*, 2015, Article 284906.
- Lu, K. D., Wu, Z. G., & Huang, T. (2022). Differential evolution-based three stage dynamic cyber-attack of cyber-physical power systems. *IEEE/ASME Transactions on Mechatronics*, 28(2), 1137–1148.
- Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54, 187–197.
- Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. *Transportation Research Part C: Emerging Technologies*, 111, 352–372.
- Ma, R., Zheng, X., Wang, P., Liu, H., & Zhang, C. (2021). The prediction and analysis of COVID-19 epidemic trend by combining LSTM and Markov method. *Scientific Reports*, 11(1), 17421.
- Ma, D., Zhu, J., Song, X. B., & Wang, X. (2023). Traffic flow and speed forecasting through a Bayesian deep multi-linear relationship network. *Expert Systems with Applications*, 213, Article 119161.
- May, A. D. (1990). *Traffic flow fundamentals*. Englewood Cliffs, NJ: Prentice-Hall.
- Nguyen-Le, D. H., Tao, Q. B., Nguyen, V. H., Abdel-Wahab, M., & Nguyen-Xuan, H. (2020). A data-driven approach based on long short-term memory and hidden Markov model for crack propagation prediction. *Engineering Fracture Mechanics*, 235, Article 107085.
- Noroozi, R., & Hellenga, B. (2014). Real-time prediction of near-future traffic states on freeways using a Markov model. *Transportation Research Record*, 2421(1), 115–124.
- Pan, Y. A., Guo, J., Chen, Y., Li, S., & Li, W. (2022). Incorporating traffic flow model into a deep learning method for traffic state estimation: A hybrid stepwise modeling framework. *Journal of Advanced Transportation*, 2022.
- Pan, Y. A., Zheng, H., Guo, J., & Chen, Y. (2023). Modified volume-delay function based on traffic fundamental diagram: A practical calibration framework for estimating congested and uncongested conditions. *Journal of Transportation Engineering, Part A: Systems*, 149(11), 04023112.
- Qi, Y., & Ishak, S. (2014). A Hidden Markov Model for short term prediction of traffic conditions on freeways. *Transportation Research Part C: Emerging Technologies*, 43, 95–111.
- Raissi, M. (2018). Deep hidden physics models: Deep learning of nonlinear partial differential equations. *The Journal of Machine Learning Research*, 19(1), 932–955.
- Raissi, M., & Karniadakis, G. E. (2018). Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357, 125–141.
- Ramezani, M., & Geroliminis, N. (2011). Arterial route travel time distribution estimation with a Markov chain procedure. *11th Swiss Transport Research Conference*.
- Ramezani, M., & Geroliminis, N. (2012). On the estimation of arterial route travel time distribution with Markov chains. *Transportation Research Part B: Methodological*, 46(10), 1576–1590.
- Ren, Q., Li, Y., & Liu, Y. (2023). Transformer-enhanced periodic temporal convolution network for long short-term traffic flow forecasting. *Expert Systems with Applications*, 227, Article 120203.
- Sattarzadeh, A. R., Kutadinata, R. J., Pathirana, P. N., & Huynh, V. T. (2023). A novel hybrid deep learning model with ARIMA Conv-LSTM networks and shuffle attention layer for short-term traffic flow prediction. *Transportmetrica A Transport Science*, 1–23.
- Shin, J., & Sunwoo, M. (2018). Vehicle speed prediction using a Markov chain with speed constraints. *IEEE Transactions on Intelligent Transportation Systems*, 20(9), 3201–3211.
- Song, Z., Feng, W., & Liu, W. (2022). Interval prediction of short-term traffic speed with limited data input: Application of fuzzy-grey combined prediction model. *Expert Systems with Applications*, 187, Article 115878.
- Sun, B., Cheng, W., Goswami, P., & Bai, G. (2018). Short-term traffic forecasting using self-adjusting k-nearest neighbours. *IET Intelligent Transport Systems*, 12(1), 41–48.
- Sun, T., Huang, Z., Zhu, H., Huang, Y., & Zheng, P. (2020). Congestion pattern prediction for a busy traffic zone based on the Hidden Markov Model. *IEEE Access*, 9, 2390–2400.
- Tang, J., Hu, J., Hao, W., Chen, X., & Qi, Y. (2020). Markov Chains based route travel time estimation considering link spatio-temporal correlation. *Physica A: Statistical Mechanics and its Applications*, 545, Article 123759.
- Vlahogianni, E. I., Golias, J. C., & Karlaftis, M. G. (2004). Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24(5), 533–557.
- Wang, D., Wang, C., Xiao, J., Xiao, Z., Chen, W., & Havyarimana, V. (2019). Bayesian optimization of support vector machine for regression prediction of short-term traffic flow. *Intelligent Data Analysis*, 23(2), 481–497.

- Wang, J., Chen, R., & He, Z. (2019). Traffic speed prediction for urban transportation network: A path based deep learning approach. *Transportation Research Part C: Emerging Technologies*, 100, 372–385.
- Xu, M., Dai, W., Liu, C., Gao, X., Lin, W., Qi, G. J., & Xiong, H. (2020). Spatial-temporal transformer networks for traffic flow forecasting. arXiv preprint arXiv:2001.02908.
- Yang, H., & Qiao, F. (1998). Neural network approach to classification of traffic flow states. *Journal of Transportation Engineering*, 124(6), 521–525.
- Zhang, R., Mao, S., & Kang, Y. (2023). A novel traffic flow prediction model: Variable order fractional grey model based on an improved grey evolution algorithm. *Expert Systems with Applications*, 224, Article 119943.
- Zheng, H., Lin, F., Feng, X., & Chen, Y. (2020). A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22(11), 6910–6920.