Journal of Intelligent Transportation Systems
Technology, Planning, and Operations

ISSN: 1547-2450 (Print) 1547-2442 (Online) Journal homepage: www.tandfonline.com/journals/gits20

Road traffic attributes prediction using deep learning hybridization by the traffic fundamental diagram

Sajjad Shafiei, Qiyuan Zhu, A. K. Qin, Simona Adriana Mihaita, Hanna Grzybowska & Hussein Dia

JOURNAL OF INTELLIGENT TRANSPORTATION SYSTEMS
https://doi.org/10.1080/15472450.2025.2577735

Road traffic attributes prediction using deep learning hybridization by the traffic fundamental diagram

Sajjad Shafiei[a], Qiyuan Zhu[a], A. K. Qin[a], Simona Adriana Mihaita[b], Hanna Grzybowska[c], and Hussein Dia[d]

[a]Department of Computing Technologies, Swinburne University of Technology, Melbourne, Australia; [b]Data Science Institute, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia; [c]Data 61, CSIRO, Sydney, Australia; [d]Department of Civil and Construction Engineering, Swinburne University of Technology, Melbourne, Australia

ABSTRACT

The prediction of road traffic attributes, such as flow, speed, and density, plays a crucial role in traffic management systems. Deep learning (DL) techniques provide an effective way to predict the future based on historical data. However, it is often impractical to measure all road traffic attributes for DL-based predictions. Given that traffic flow is the most frequently measured traffic attribute, there is a large body of research on DL-based traffic flow prediction. Nevertheless, using only traffic flow data is insufficient to comprehensively depict road traffic conditions. Traffic fundamental diagrams (TFDs) offer a way to estimate other traffic attributes, based on traffic flow. However, the inherent non-monotonic relationship between flow and density poses a significant challenge because a specific flow may correspond to two different densities, reflecting uncongested and congested road conditions. To address this issue, we propose a novel framework for predicting road traffic attributes by hybridizing DL with TFDs. The proposed framework comprises two streams. In the supplementary stream, traffic data (flow, density, and speed) are used to calibrate TFDs and generate congestion labels, which are then used to train a congestion predictor. The main stream relies solely on traffic flow data, aligning with real-world scenarios. One DL model predicts future flow, while another predicts congestion labels based on historical flow data. These labels, combined with the predicted flow, enable the calibrated TFDs to determine density and speed values. Experiments based on a case study of a freeway in Melbourne, Australia, demonstrate the effectiveness of the proposed framework.

ARTICLE HISTORY
Received 28 March 2023
Revised 12 October 2025
Accepted 16 October 2025

KEYWORDS
deep learning; road traffic fundamental diagram; traffic flow prediction

I. Introduction

The effective management of road traffic, particularly in reducing congestion and optimizing traffic flows, is a critical challenge for modern Intelligent Transport Systems (ITS). Accurate prediction of road traffic attributes, including flow, speed, and density, plays a pivotal role in operational decisions of these systems. The advent of machine learning (ML) and DL techniques has ushered in a new era of data-driven traffic management, allowing for more nuanced and reliable predictions based on historical data.

Among the various traffic attributes, traffic flow (volume) is the most commonly measured and analyzed due to its direct relevance and ease of acquisition. Traffic flow data is frequently used in emissions models, fuel consumption models, and road surface maintenance (Kumar & Raubal, 2021). Traffic flow can be measured using traditional traffic sensors like loop detectors or traffic counter tubes. In contrast, other traffic attributes often require more advanced sensors or extensive data analysis, such as using number plate recognition technology to measure road segment travel time.

Given its accessibility and applications, traffic flow prediction has traditionally been the focal point of numerous research studies and practical applications (Panayiotou et al., 2023; Tedjopurnomo et al., 2020). However, relying solely on traffic flow to represent the full spectrum of traffic conditions has significant limitations. Traffic flow values alone do not represent the traffic congestion condition due to the non-

CONTACT Sajjad Shafiei sshafiei@swin.edu.au

monotonic relationship between road traffic flow and density. For example, low traffic flow on a freeway in off peak hours may indicate that the freeway is uncongested, while the same flow in peak hours can indicate congestion and downstream queuing. This issue makes traffic flow prediction alone less informative for network operators and traffic management centers.

A practical approach to addressing this issue is by incorporating additional types of traffic data alongside traffic flow, which can provide a more comprehensive understanding of traffic conditions (Zhuang et al., 2025). Predicting multiple traffic attributes can enhance this understanding, yet a key challenge lies in the availability of diverse traffic data for the same time intervals, which may be insufficient in some applications. While many studies have focused on enhancing traffic flow prediction algorithms to increase accuracy, less attention has been given to determining which traffic congestion conditions the traffic flow corresponds to.

On a parallel research track, there is comprehensive literature on macroscopic traffic models that describe bivariate relationships between traffic attributes, such as flow-speed, flow-density, and speed-density (Bai et al., 2021; Qu et al., 2017). These relationships are often presented in TFDs and are widely used in analytical traffic modeling analyses. Accordingly, in view of the advances in both model-based and data-driven approaches, there is a critical gap in the disconnection between these two approaches. Incorporating these approaches into the one modeling framework leverages the benefits of both and results in more comprehensive outcomes. The data-driven model can capture the spatiotemporal correlations that exist in historical traffic flow data to predict traffic flow more accurately. Additionally, employing TFDs attempts to map traffic flow to other traffic attributes, reducing traffic condition ambiguity.

In line with this, this study seeks to investigate the possibility of predicting additional traffic information beyond the traffic flow values when the traffic flow information is only available. This study investigates the performance of a proposed framework that can be applied to different existing data-driven traffic flow prediction models (TFPMs) to provide a clearer understanding of traffic conditions on the road network.

To clarify the framework of this study, we refer to Greenshields' model, one of the simplest and most fundamental in traffic flow theory. It illustrates that a single traffic flow rate can correspond to two distinct traffic states, uncongested and congested, due to the non-monotonic relationship between flow and density. This inherent ambiguity highlights the limitations of traditional data-driven approaches that rely solely on flow measurements, posing a significant challenge for accurate traffic prediction. To overcome this, we propose a novel predictive framework that integrates DL with TFDs. Our framework operates in two streams: In the supplementary stream, we utilize a limited dataset containing flow, density, and speed data to calibrate TFDs, which then provide congestion labels. These labels are used to train a traffic congestion prediction model that operates on flow data. Meanwhile, in the main stream, which reflects real-world scenarios where only traffic flow data is typically available, we use a DL model to forecast future flow values. Concurrently, the congestion predictor from the supplementary stream predicts the congestion state associated with these future flow values. Finally, using the calibrated TFDs, we derive unique density and speed values from the predicted flow and its corresponding congestion label. Compared with previous classical traffic flow predictions, the proposed framework has the following merits:

- This approach goes beyond traditional traffic flow prediction methods by predicting additional traffic attributes even when only traffic flow time series is accessible.
- It introduces an innovative method for identifying the optimal representative road TFD, which captures the intricate relationship between traffic attributes. This method is applicable in prediction, particularly in scenarios with constrained data resources.
- This approach addresses a significant gap by bridging communication between ML and macroscopic traffic model-based approaches. By integrating these approaches into a unified modeling framework, it harnesses the strengths of each, leading to more comprehensive results.
- This method's adaptability allows seamless integration with various data-driven TFPMs, providing a versatile and scalable solution for traffic management.

The rest of the article is organized as follows: The next section provides a summary of relevant background literature on the traffic prediction models. Section 3 describes the methodology for a road TFD calibration and the proposed framework to predict traffic congestion states and flow values. Section 4 presents the case study and some statistics related to

TFD calibration step. The results from our proposed framework are analyzed in Section 5. The final section provides a conclusion and recommends future research directions.

2. Background

This section starts by examining the fundamental goals of traffic prediction, followed by a summary of traditional data-driven TFPMs. Additionally, it provides an overview of other relevant studies that employ model-based approaches within data-driven frameworks.

2.1. Focus areas in traffic prediction models

Data-driven traffic prediction involves utilizing a learnable function that leverages historical data from multiple previous time stamps to forecast traffic attributes for subsequent time stamps. The accuracy of prediction is significantly influenced by traffic data collection, preprocessing, and training techniques (Wu et al., 2018). In the literature, forecasting within one or a few hours is commonly referred to as short-term prediction (Gao et al., 2025; Vlahogianni et al., 2014). Short-term road traffic prediction addresses three interlinked challenges: predicting travel time, traffic congestion, and traffic flow (Yang, 2013; Banani Ardecani et al., 2025). Each of these facets contributes uniquely to the comprehensive understanding of traffic dynamics, requiring tailored approaches and models to manage their complexities effectively.

The distinctions among these prediction challenges lie in their direct relevance and applications. While travel time prediction directly benefits end-users such as drivers and commuters by providing timely and user-specific benefits, predictions of traffic congestion and flow predominantly serve traffic managers and urban planners (Shafiei et al., 2023).

The prediction of traffic congestion involves identifying areas where areas where congestion has propagated or is likely to propagate. One common qualitative measure used in these analyses is binary congestion prediction, which categorizes areas as congested or uncongested. Alternatively, the Level of Service (LOS) scale is often employed, ranging from A (indicating the best conditions) to F (representing the worst conditions), describing the operational conditions within a traffic stream as perceived by motorists and passengers. Additionally, predicting the length of vehicle queues is another aspect of congestion prediction models (Akhtar & Moridpour, 2021; Yuan, Zhang, et al., 2021).

Traffic flow prediction quantifies the number of vehicles passing a specific point on a road network within a set time frame. This is crucial for the design of traffic control and management strategies, necessitating accurate continuous monitoring and analysis of traffic flows. Recent literature reveals a focus on reducing prediction errors in traffic flow studies. However, many overlook the nuances that distinguish low traffic flow due to congestion from low flow in uncongested conditions, which could provide deeper insights into traffic conditions. Recognizing these differences is crucial and can have a more significant impact on the overall traffic analysis than merely minimizing prediction errors (Pan et al., 2024).

2.2. Data-driven traffic flow prediction models

Early data-driven TFPMs were primarily statistical. For example, autoregressive Integrated Moving Average (ARIMA) and its variants are widely used as parametric approaches for time-series forecasting (Vlahogianni et al., 2014). This model considers a linear combination of past observations, reflecting the autoregressive and moving average nature of the historical data. The ARIMA-family models can be built on limited data. This capability was particularly valuable when traffic surveillance and computation resources were limited. Although these models are easy to interpret, their prediction power is relatively weak, and they are inefficient for big traffic data analysis. There are comprehensive reviews of different statistical models found in the survey papers (Panayiotou et al., 2023; Vlahogianni et al., 2014).

The escalating size of traffic data has led to a growing adoption of DL models by researchers and practitioners, showcasing successful applications in time-series traffic flow forecasting tasks (Kumar & Raubal, 2021). Deep learning (DL) and combinations of neural networks have led to several innovations for applications on traffic flow prediction (Aouedi et al., 2025; Panayiotou et al., 2023). The models can utilize latent information hidden in traffic data captured from thousands of active and passive sensors in the transport network. DL methods are highly capable of expressing patterns in data and alleviating the overfitting problem. In the last decade, the convolutional neural network (CNN) model has been widely applied to traffic flow prediction because this approach can mine the spatial features progressively. In CNN methods, road traffic flow is considered as an image and the input traffic flow data is preprocessed by partitioning the network into a set of cells for each time interval. Hence, the entire traffic data is fed to the model as several images referring to different time stamps (Tedjopurnomo et al., 2020).

Recurrent Neural Networks (RNNs) are commonly utilized to extract the temporal features of traffic data (Kumar & Raubal, 2021). In the feedforward network, an input is fed in one direction *via* successive hidden layers to estimate the output. On the contrary, RNNs include a feedback loop at every time step in the sequence, developing cyclic connections between neurons so that outputs can be fed back as inputs. Therefore, the structures of RNN models incorporate time dependency using sequences of inputs and correlation between time steps. The limitation of conventional RNN models is that they lose the long-term dependencies (Wu et al., 2018). Researchers have proposed various new architectures to address this limitation (Liu et al., 2021; Zhang et al., 2020). The traffic prediction landscape has witnessed the effectiveness of models like Long Short-Term Memory (LSTM) in managing long-term memories through integrated gates in conventional RNN architectures (Zheng et al., 2019). While LSTMs have demonstrated robust performance in traffic flow forecasting, recent studies have introduced modifications to enhance their core structure. Notably, Bidirectional LSTMs (BiLSTM) have leveraged both forward and backward dependencies for improving predictive accuracy (Abduljabbar et al., 2021a, 2021b). Another widely adopted model, Gated Recurrent Units (GRU), rivals LSTM's performance with fewer trainable parameters (Cho et al., 2014). In addition to recurrent and CNNs, attention mechanisms have gained prominence in enhancing the performance of TFPMs. Attention mechanisms enable models to focus on specific parts of the input sequence, allowing for more effective handling of long-term dependencies and capturing intricate patterns within the data (Huang et al., 2023; Lan et al., 2023; Wang, Qin, et al., 2023; Wang, Tian, et al., 2023; Zhang, Mao, et al., 2024; Zhang, Wen, et al., 2024).

In recent years, the realm of DL applications has provided innovative tools for spatiotemporal prediction tasks, fostering deeper comprehension of the complex interplay between spatial and temporal factors (Kong et al., 2024). The current traffic flow prediction frameworks exhibit a tripartite categorization: (1) Spatial–temporal Sequential Model (SSM) (Jana et al., 2023), (2) Spatial–temporal Grid Model (SGM) (Lv et al., 2023; Zhang, Mao, et al., 2024; Zhang, Wen, et al., 2024), and (3) the Spatial–temporal Graph Sequential Model (SGSM) (Wang, Qin, et al., 2023; Wang, Tian, et al., 2023). The SSM class focuses on sequential patterns, capturing the temporal evolution of traffic phenomena. Meanwhile, the SGM paradigm adopts a grid-based approach to spatial-temporal modeling, offering a distinct perspective on urban traffic dynamics. Lastly, the SGSM framework integrates graph-based structures to model sequential interactions within a spatial-temporal context, showcasing the adaptability and diversity within contemporary DL-based traffic prediction methodologies (Zheng et al., 2023).

In addition to methodological improvements, several studies have considered external factors, such as weather conditions, calendar information, and incident details to predict non-recurrent traffic flow. Weather conditions have been shown as a critical external feature that can impact traffic flow. Previous studies have focused more on the direct rainfall intensity and visibility of the road traffic flow (Nigam & Srivastava, 2023; Zhang et al., 2023). The studies considered the immediate effects of weather conditions at the predicted time window. For instance, drivers reduce their speed during intense rainfall due to lower visibility and higher sliding risk. However, in addition to the immediate effects, the weather condition can significantly impact vehicular demand patterns and travelers' route choices (Maze et al., 2006). Weather conditions also affect mode choice and the departure time. Travelers tend to use more private vehicles in adverse weather conditions instead of active modes or public transport with long access and egress walking legs. Time and calendar information is also an intuitive factor. Time of the day, day of the week, and public holidays significantly influence the traffic flow data. Finally, another crucial external parameter is incident information. Training TFPMs based on limited historical road incidents-impacted traffic data is challenging since traffic incidents occur in different locations and time intervals. Moreover, incidents take place in various levels of severity. They range from simple vehicle breakdowns to severe chain accidents, which affect the road condition differently. In addition, there are multiple reasons for traffic flow variations and patterns in a multimodal system, making it challenging to train models. Therefore, more advanced data-driven and model-based approaches seem essential in traffic flow prediction under incident conditions (Kumar & Raubal, 2021).

2.3. Hybrid data- and model-based traffic flow prediction models

Traditional model-based traffic flow analyses offer valuable insights into traffic principles but often require extensive efforts in parameter calibration and rely on strong modeling assumptions. In contrast, purely data-driven approaches encounter challenges in result interpretation and are highly dependent on data quality, particularly when predicting traffic conditions under non-recurrent circumstances (Yuan et al., 2021).

Some recent studies have addressed challenges associated with sole data-driven or model-based approaches by integrating ML methods into traffic models. For example, an integrated data- and model-driven framework has been proposed, predicting travel demand inputs for traffic simulation using ML methods (Shafiei et al., 2019, 2022). The simulation captures the complex interaction between travelers and the road network, allowing for the evaluation of unknown scenarios, including considerations for factors like lane closures and speed reductions due to incidents.

Certain studies have utilized traffic models to generate non-recurrent data and train data-driven models with synthetic traffic information to overcome limitations in historical data availability (Abduljabbar et al., 2021a, 2021b). Importantly, there is a growing acknowledgment of the necessity to incorporate traffic flow principles into data-driven model structures (Shi et al., 2021; Yuan, Wang, et al., 2021; Yuan et al., 2025), drawing inspiration from Physics-Informed Neural Network (PINN) (Raissi et al., 2019). The paradigm, combining physics-based models and deep neural networks, has emerged as a potent solution. PINN uniquely excels in achieving robust predictive power and sample-efficient training by synergizing the strengths of physics-based and data-driven approaches. This innovative approach provides a solution for developing reliable data-driven models with limited historical data, ensuring the consistency of data-driven models with the underlying physics of the system. The key challenge of PINNs lies in effectively integrating domain knowledge or physical laws into the neural network architecture and training process. This involves ensuring that the neural network not only learns from data but also adheres to the fundamental principles governing the physical system being modeled. Achieving this integration requires careful design of the neural network architecture, choice of loss functions that enforce physical constraints, and selection of appropriate training data and strategies to effectively capture the underlying physics. Balancing the flexibility of neural networks with the constraints imposed by physics can be a complex and non-trivial task, making PINNs a challenging area of research. Some advanced training methods are employed in recent studies to address this challenge (Di et al., 2023; B. Wang, Qin, et al., 2023; Zhang, Mao, et al., 2024; Zhang, Wen, et al., 2024).

Recently, increasing emphasis has been placed on enhancing the interpretability of DL-based traffic prediction models. Although many contemporary approaches offer strong predictive capabilities, their limited transparency has raised concerns regarding their practical applicability in transportation systems. In response, explainable frameworks have been introduced to retain high accuracy while improving model interpretability. Techniques such as the integration of large language models and stacking-based learning architectures have been employed to support more transparent decision-making processes in traffic forecasting. These advancements highlight the growing trend toward explainable artificial intelligence in the field of transportation modeling (Chen et al., 2024; Guo et al., 2024).

3. Methodology

This section introduces the proposed framework that aims to predict multiple road traffic attributes based solely on available flow data. The framework encompasses three major components TFD calibration, traffic flow prediction, and congestion prediction. The proposed framework employs a two-pronged data stream approach to handle both primary and supplementary traffic datasets. Figure 1 illustrates the overall data flow architecture which consists of training and deployment (i.e., testing) stages.

1. The training stage includes two streams: “supplementary stream” for calibrating TFDs, which estimate traffic attributes like speed and density based on flow and congestion data, and for training a congestion prediction model, which takes as input a flow series and outputs the congestion status for multiple times ahead in a future horizon based on a certain collection of traffic data with both flow and density information; and the “main stream” for training a flow predictor to forecast future flow values, based on historical flow data (which can be different from those used in the “supplementary stream”).
2. In the deployment stage, the flow predictor predicts flow values for multiple future time steps using a window of historical flow data. Meanwhile, the congestion predictor will predict the congestion status (i.e., 0/1 for non-congestion

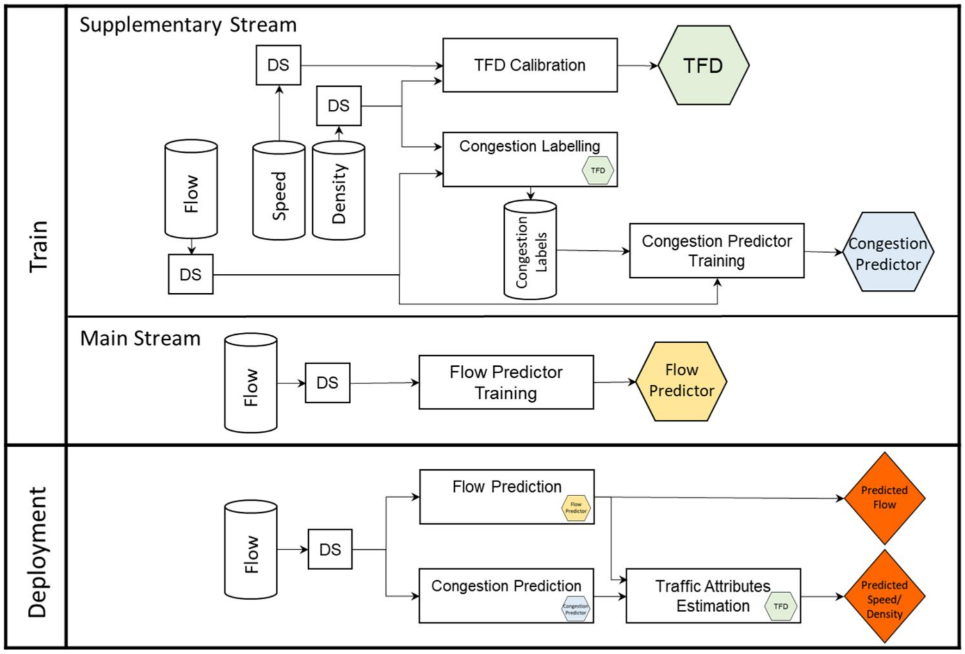

Figure 1. Overview of the traffic prediction framework data flow architecture.

and congestion) for those predicted flow values based on the same window of historical flow values. After that, the calibrated TFDs will take the predicted flow values and their associated predicted congestion statuses as inputs to estimate corresponding density and speed values.

The ensuing sections provide a thorough examination of the proposed methodology, elucidating the intricacies of data preprocessing, traffic flow prediction, and congestion estimation. The goal is to furnish a comprehensive guide to the implementation of the methodology, emphasizing clarity and transparency at every stage.

3.1. TFD

Traffic flow dynamics are described through three primary traffic attributes: speed (v), density (k), and flow (q). If the cumulative flow $N(x,t)$ is defined as the total number of vehicles that passed the point x at time interval t, the three primary traffic attributes are defined as:

$$\begin{aligned} q &= \frac{\partial N(x,t)}{\partial t} \\ k &= \frac{\partial N(x,t)}{\partial x} \\ v &= \frac{q}{k} \end{aligned} \tag{1}$$

Equation (1) links the q, k, and v and is named the traffic road fundamental identity. The bivariate relationships between flow and density, speed and density, and speed and flow are of great value in traffic engineering and are presented in TFDs. This section proposes a data-driven approach to identify road TFDs. TFDs have different parameters that need to be estimated accurately to better understand the road traffic dynamics. Since traffic attributes change over time and space, we propose a methodology to determine the parameters of TFDs for each site in the studied area. The proposed approach captures the variations of road TFDs over multiple days and overcomes the identification of the "typical day" problem.

An essential application of TFDs studies is in traffic simulation and analytical models. A vast body of literature has been devoted to calibrating such models against historical traffic data (Gu et al., 2017; Qu et al., 2017). The TFD parameters play a pivotal role in modeling traffic in the road network and are considered an essential part of traffic model calibration. Since traffic models are often calibrated for a typical day, most research focuses on normal situations. However, some have considered a set of TFDs for different days with poor weather conditions (Mahmassani et al., 2012). Fundamental diagrams are divided into single-regime and multi-regime. Single-regime models describe the relationship through a single mathematical equation, while multi-regime models consider different functions for a subdivided range of traffic conditions (Qu et al., 2017). Furthermore, some empirical studies have revealed the "capacity drop" phenomenon (Kontorinaki et al., 2017; Saberi & Mahmassani, 2013). The phenomenon in traffic flow

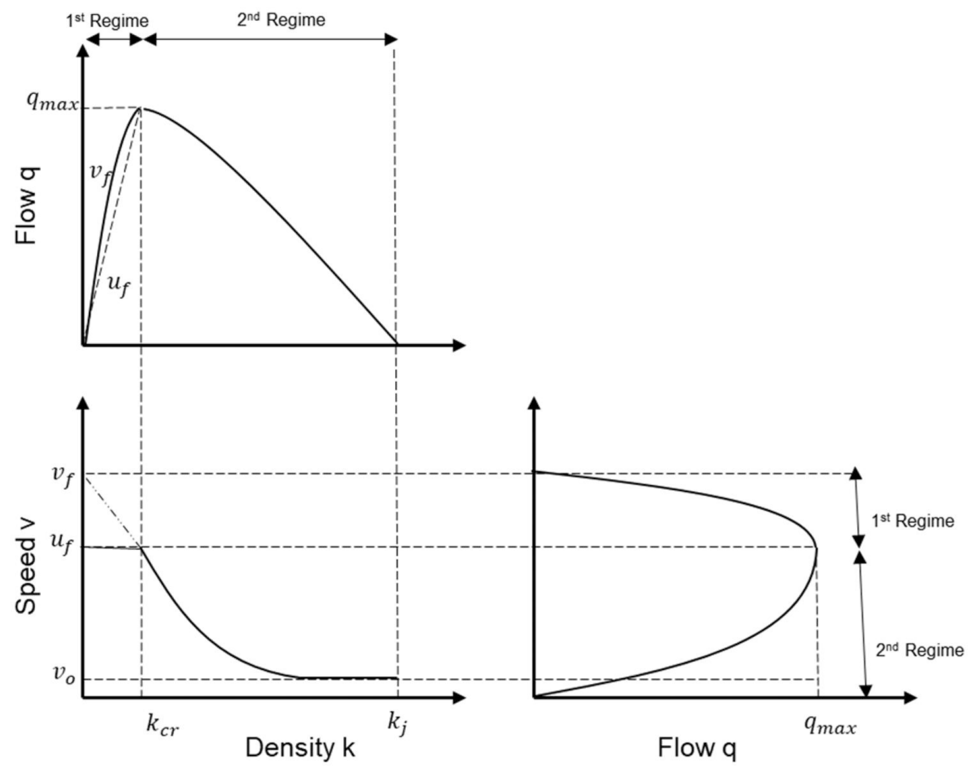

Figure 2. Dual regime modified Greenshields TFD.

theory refers to the unexpected decrease in road capacity when traffic density exceeds a critical threshold. This decrease occurs due to factors such as reduced vehicle speeds, increased lane changing, and heightened driver reactions, leading to congestion and decreased traffic flow efficiency. The capacity drop phenomenon results in disconnection in the TFDs around critical density, and the maximum flow reduces after moving to the congested regime. This phenomenon can be modeled in multi-regime models.

This study employs the dual-regime modified Greenshields TFD model (Mahmassani et al., 2012). The road traffic conditions were divided into two regimes: congested and uncongested (or under and oversaturated). The mathematical relationship between traffic speed and density is expressed in Equation (2) and depicted in Figure 2.

$$v = \begin{cases} u_f & 0 < k < k_{cr} \\ v_0 + (v_f - v_0) \cdot \left(1 - \frac{k}{k_j}\right)^{\alpha} & k_{cr} < k < k_j \end{cases} \tag{2}$$

where, u_f, v_o, and v_f are the free-flow speed, the jam speed, and the intercept speed; k_{cr} and k_j are the critical (breakpoint) density and the jam density, and α is a shape parameter. If k_j and v_o are assumed approximately 143 vpkmpl and zero, respectively, with no road capacity drop at k_{cr}, Equation (2) can be further written in Equation (3). The model includes three parameters (k_{cr}, v_f, α) which jointly determine the shape of the TFDs.

$$v = \begin{cases} v_f \cdot \left(1 - \frac{k_{crt}}{k_j}\right)^{\alpha} & 0 < k < k_{cr} \\ v_f \cdot \left(1 - \frac{k}{k_j}\right)^{\alpha} & k_{cr} < k < k_j \end{cases} \tag{3}$$

To estimate TFD parameters from field data, Equation (3) curve fitting was applied daily for each detection site using Damped Least-Squares (DLS) to solve nonlinear least squares problems. This curve-fitting algorithm is a combination of the gradient descent and the Gauss-Newton methods. The DLS method acts more like a gradient-descent method when the parameters are far from the optimal value and acts more like the Gauss-Newton method when the parameters are close to the optimal value. Therefore, the method is quite robust for nonlinear curve-fitting and converges after limited iterations. Let's assume a set of n empirical pairs of speed and density at a specific site. The TFDs parameters of Equation (3) (f) are calibrated through the Euclidean distance minimization:

$$argmin_{k_{crt}, v_f, \alpha} = \sum_{i-1}^{n} ||v_i - f(k_i; k_{crt}, v_f, \alpha)|| \tag{4}$$

The calibration for each road section includes a $3 \times l$ matrix that shows the TFD parameters (v_f, k_j, and v_0) and l is the number of days.

Traffic flows are influenced by external factors such as weather, work zone, or road surface condition. Therefore, the next step is to select a day that represents a typical day. The Fréchet distance method measures the similarity between daily estimated TFDs for each site to extract the TFD that reflects the typical daily traffic dynamics. The Fréchet distance (δ_F) measures the similarity between curves approximated by a polygonal curve. Thus, $\delta_F(G, H)$ is a measure of similarity between a polygonal curve G of length N and a polygonal curve H of length M. A large $\delta_F(G, H)$ indicates that the two curves are far away from each other and present a large dissimilarity, and vice versa. Interested readers can find further details on the Fréchet distance calculation and its variants in the original references (Alt & Godau, 1992; Eiter & Mannila, 1994).

3.2. Road traffic congestion and flow prediction

We employ two distinct data-driven models: the Traffic Flow Prediction Model (TFPM) and the traffic congestion prediction model (TCPM). TFPM forecasts future traffic flows, while TCPM predicts congestion states based on historical traffic flow data (q_i^τ) during $\tau \in (t-n, t-1)$ obtained from site i (where $i \in \{1, \ldots, p\}$). Both models utilize machine learning (ML) techniques, specifically tailored to process temporal sequences of traffic flow data for accurate predictions.

Traffic flow data, q_i^τ, serve as the primary input for our models. The temporal input features are concatenated to a feature matrix as follows:

$$Q_{t-n,t-1} = \begin{bmatrix} q_1^{t-n} & q_1^{t-n+1} & \cdots & q_1^{t-1} \\ q_2^{t-n} & q_2^{t-n+1} & \cdots & q_2^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ q_p^{t-n} & q_p^{t-n+1} & \cdots & q_p^{t-1} \end{bmatrix} \tag{5}$$

This matrix represents the input features for both TFPM and TCPM, where each row corresponds to a different site and each column to a specific time lag relative to t. The outputs from our models are designed for use over a prediction horizon from t to h. The predicted traffic congestion output ($S_{t,h}$) and the traffic flow output ($Q_{t,h}$) are derived from the TCPM and the TFPM, respectively, using the feature matrix as input ($Q_{t-n,t-1}$). In uncongested states, traffic flow increases as traffic density rises. However, once density reaches a critical level, the increase in density begins to negatively impact flow (see Figure 2). The road density at which the maximum traffic flow is observed is called the critical density (k_{crt}). The traffic dynamics change significantly before and after the critical density. Therefore, for the supplementary data where density information is available, we labeled the road binary states based on the critical density obtained from the determined TFD for the road. The traffic congestion state of the road i at time stamp t is defined as:

$$s_i^t = \begin{cases} 1, if \ k_i^t \geq k_{crt\,i} \\ 0, if \ k_i^t < k_{crt\,i} \end{cases} \tag{6}$$

The road traffic congestion state output ($S_{t,h}$) for the prediction horizon will be determined through the binary decision congestion prediction model as:

$$S_{t,h} = TCPM(Q_{t-n,t-1})$$
$$S_{t,h} = \begin{bmatrix} s_1^t & s_1^{t+1} & \ldots & s_1^h \\ s_2^t & s_2^{t+1} & \ldots & s_2^h \\ \vdots & \vdots & & \vdots \\ s_p^t & s_p^{t+1} & \ldots & s_p^h \end{bmatrix} \tag{7}$$

Similarly, traffic flow output ($Q_{t,h}$) for the prediction horizon (t, h) will be obtained through the TFPM as follows:

$$Q_{t,h} = TFPM(Q_{t-n,t-1})$$
$$Q_{t,h} = \begin{bmatrix} q_1^t & q_1^{t+1} & \ldots & q_1^h \\ q_2^t & q_2^{t+1} & \ldots & q_2^h \\ \vdots & \vdots & & \vdots \\ q_p^t & q_p^{t+1} & \ldots & q_p^h \end{bmatrix} \tag{8}$$

Later, in Section 5, we compare the performance of several widely used ML and DL methods in predicting road traffic congestion and flow. Our analysis includes a diverse set of models to determine which performs best for each problem. We then select the highest-performing model to be integrated into our proposed framework. This selection process ensures that our framework leverages the most effective prediction method, enhancing its reliability and accuracy in overall proposed framework.

3.3. Metrics

Overall, our approach consists in firstly, predicting the traffic congestion based on incoming data streams, and secondly, to predict traffic flow in a time-series approach the traffic flow along our network. Therefore, two sets of criteria for these tasks are used to evaluate the models. Because the binary congestion prediction is an imbalanced problem, we use the F_1 score. If the number of true positives is N_{tp}, the number of false-positive N_{fp} and the number of false-negative N_{fn}; furthermore, the precision (P), recall (R), and F_1 scores are calculated by using the following formulas:

$$P = \frac{N_{tp}}{N_{tp} + N_{fp}}$$
$$R = \frac{N_{tp}}{N_{tp} + N_{fn}} \tag{9}$$
$$F_1 = 2 * \frac{P * R}{P + R}$$

To evaluate the traffic flow prediction, the mean relative error (MRE), the mean absolute error (MAE), and the root mean square error (RMSE) are the most common measurements to compare the accuracy of different prediction methods. MRE provides a percentage-based comparison of errors, while MAE and RMSE offer absolute and squared error comparisons, respectively. Using all three metrics offers a comprehensive assessment of prediction accuracy, accounting for different aspects of error evaluation such as magnitude and relative performance across varied conditions. Suppose $\hat{y}_i$ and y_i denotes the prediction and the true value, respectively, and I is the total number of prediction points then the measurements are defined as follows:

$$MAE = \frac{\sum_i^I |\hat{y}_i - y_i|}{I}$$
$$MRE = \frac{\sum_i^I |\hat{y}_i - y_i|/y_i}{I} \tag{10}$$
$$RMSE = \sqrt{\frac{\sum_i^I |\hat{y}_i - y_i|^2}{I}}$$

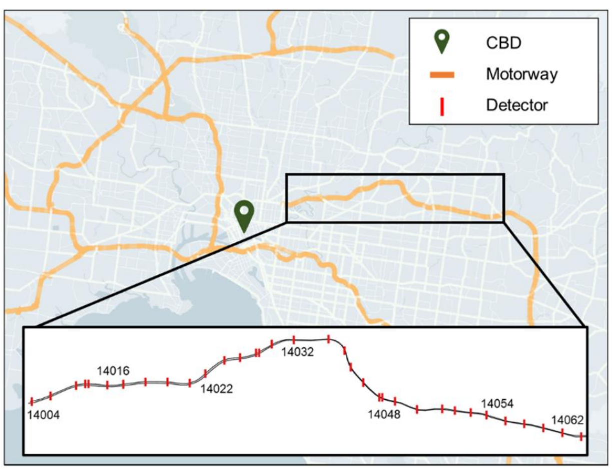

Figure 3. Study corridor – M3 freeway, Melbourne, Australia.

4. Case study

In this study, we used three months' worth of archived freeway loop detector data at 10-min intervals from a section of the freeway in Melbourne, Australia. The layout of the study area is presented in Figure 3. The studied corridor stretches from Melbourne CBD to the eastern residential suburbs as well as in the opposite direction.

The traffic data used in this study comes from freeway mid-blocks and includes traffic flow, speed, and occupancy. The occupancy (o) data shows the percentage of time that vehicles occupy a detector. This information can be changed to traffic density (k, the number of vehicles per length unit, here, vehicles per kilometer) as follows:

$$k = \frac{8.58}{L_v + L_d} \times o \tag{11}$$

where L_v and L_d are the average lengths for vehicles and detectors. L_v and L_d are assumed to be 4.5 and 1.5 m, respectively (Gu et al., 2017).

The data was collected from 33 sites with both inbound (toward CBD) and outbound (toward the eastern residential suburbs). For sites with multiple detectors, traffic data collected by disparate detectors is averaged to obtain the mean traffic data for the freeway. Figure 4 shows the box plots of average traffic flow through the corridors for both weekday and weekend approaches, indicating clearly that weekday peaks differ from those on the weekend. On weekdays, morning peak hours for the inbound direction are between 6:00 am and 10:00 am; in the opposite direction, weekday peak hours occur in the afternoon between 3:00 pm and 5:00 pm. On weekends, daily traffic profiles are similar for both directions, with a peak from 4:00 pm to 6:00 pm.

Figure 5 shows a sample of raw traffic data over a month for a specific observation station (14,061 WB). The speed remains constant during the uncongested regime, whereas it decreases once the density exceeds the critical value and moves to the congested regime. As part of the data preprocessing, we should filter out the low-speed or density observations. These

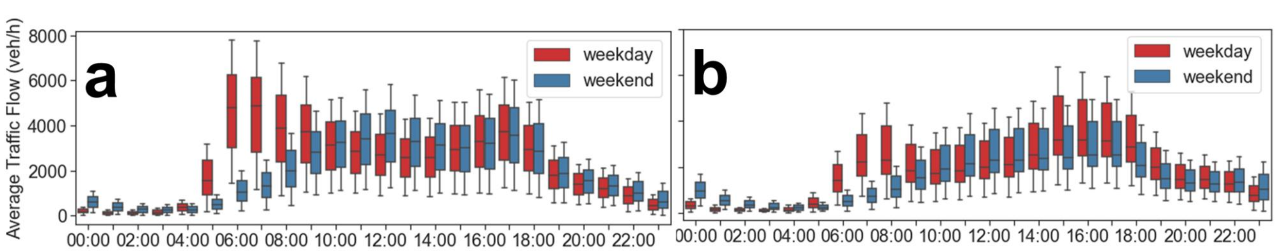

Figure 4. Average traffic flow for weekdays and weekends in March 2019: (a) inbound (toward CBD) and (b) outbound.

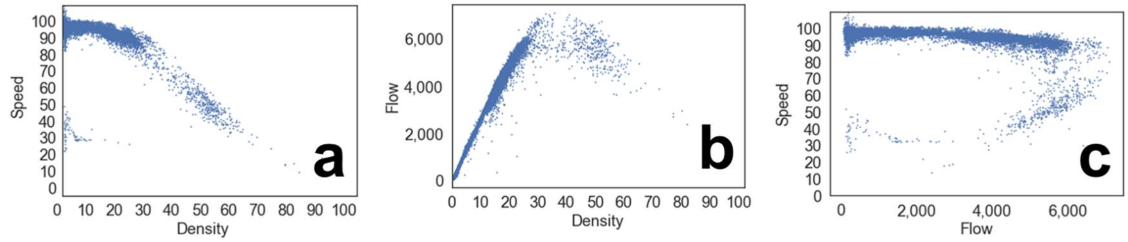

Figure 5. Speed, flow, and density plots for a sample observation site (14,061 WB).

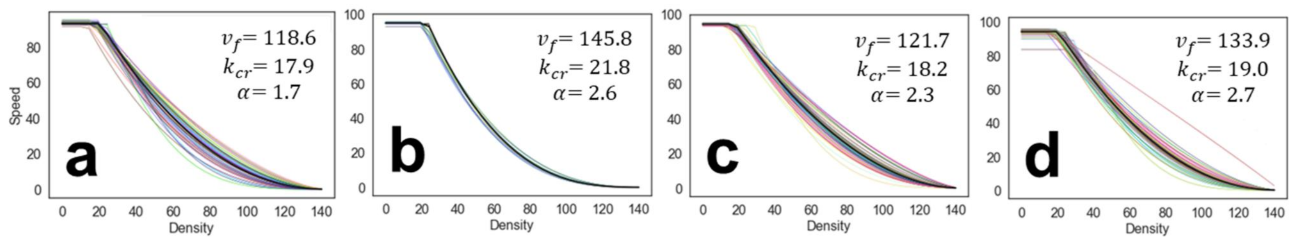

Figure 6. Four sets of daily TFDs between speed *vs.* density and the final selected TFD (in black) with the most similarity to others.

Table 1. Statistics of estimated TFD parameters for observed sites in the studied area.

TFD parameters	Min	Max	Mean	SD
k_{crt}	12.41	22.74	17.04	3.42
v_f	117.79	148.18	129.63	10.68
α	1.92	2.67	2.42	0.25

observations cannot present normal freeway traffic flow and often correspond to the temporary midnight lane closures caused by road maintenance.

In this study, we divided the dataset into two parts: 25% supplementary and 75% main stream datasets. The supplementary data was used for TFD calibration and subsequently for traffic congestion prediction. We applied DLS to the daily data and used the Fréchet distance method to identify the best representative TFD, which plots speed against density for each site. Figure 6 presents the estimated TFDs for four random locations. For some, like Figure 6(b), smaller daily variations can be observed, while for others, like Figure 6(d), there are certain outliers. These outlier TFDs are due to insufficient observation of the congested regime.

The main TFD parameters, including v_f, k_{crt}, and α, are determined by applying the proposed methodology to each site (road section). Table 1 summarizes the statistics of the estimated parameters for all sites in the corridor. By inserting Equation (1) in the estimated TFDs, traffic flow is also determined. Therefore, the derived equations will serve to calculate all three key traffic attributes for future predictions.

5. Evaluation

Considerable attention was given to data preprocessing to ensure the identification and removal of corrupted data for both supplementary and main datasets. Such missing or noisy data affects the accuracy of the traffic congestion and traffic flow prediction, particularly if the percentage of faulty data reaches a certain ratio. Results of raw data preprocessing indicate a lack of data for all sites on five random days during the three-month interval. In addition, due to road maintenance, the traffic patterns significantly change for several sites, particularly in the outbound direction. Therefore, for prediction applications, we focus only on the inbound approach. Furthermore, if traffic data was found to be continuously missing for a period of 12 daytime intervals (2 h), it was removed for that date.

We apply a generic noisy point smoothing approach to the raw traffic data. Observations with a significant deviation from the historical data are considered noisy points if the following conditions are met. Since the daily traffic pattern for weekdays and weekends is substantial, two sets of the mean are calculated. The corresponding historical averages for weekdays and weekends are presented as m_t and m'_t, respectively. Accordingly, thresholds A and A' are determined to control the proportion of noise as follows:

$$\begin{gathered}\frac{|m_t - x_t|}{m_t} > A \\ \frac{|m'_t - x_t|}{m'_t} > A'\end{gathered} \tag{12}$$

After conducting some data engineering testing, A and A' are considered 40% for both weekdays and weekends. It is worth noting that small A and A' values result in the remaining data noise, and a large amount of the thresholds remove the day-to-day data variations. The noisy or random missing points can be replaced by the average rolling method presented in Equation (13).

$$x_t = 2/3x_{t-1} + 1/3x_{t-2} \tag{13}$$

Because some ML methods are sensitive to the range of input data, the next step of data preprocessing normalizes traffic flow to the 0 and 1 range for purposes of training all models. To do this, the MAX–MIN scaling method is applied to normalize the data (Pedregosa et al., 2011). Finally, since the evaluation becomes increasingly biased with the same data incorporated into the model configuration, the dataset is divided into training/calibration and testing/verification. In the model training, the objective is to calculate the model parameters that minimize the error between the estimated and historical data. For the testing data set, the data sample was used to provide an unbiased evaluation of a model fit on the training data.

In this article, the traffic data is sorted according to experiments. The last 20% of data is utilized as the testing set and the other as the training set. We test the precision of our prediction models for a future prediction horizon of 10-, 30-, and 60-min intervals. For the traffic flow and congestion prediction models, we test several widely used models as:

- Decision Trees (DT): DT contains recursive binary splits based on certain cutoff values in the features to minimize a specific loss function. The target variable is predicted by splitting data subsets and placing them in the tree's leaf nodes. By moving down the tree, the value of the leaf node is determined.
- Support Vector Machines (SVMs): The primary idea of SVMs is to find a decision function that, for all instances, deviates less than the required value from the target variable. Therefore, slack variables should be defined to account for deviations outside the accepted range.
- Vanilla Recurrent Neural Networks (RNN): RNNs are explicitly designed for sequential data. RNNs include a feedback loop at each time step in the sequence, forming cyclic connections between nodes and allowing outputs to be ingested back as inputs. The output of an RNN is influenced by the currently fed input and prior inputs because each hidden state has memory on the previous state, and thus every other preceding state.
- LSTM: it has been used to model traffic flow owing to its superior performance in capturing long short-term dependencies. In a standard RNN, the input and hidden states from the previous time stamps is passed through the activation layer to estimate a new state. However, LSTM incorporates a gating mechanism to keep useful information from the previous time stamps. The LSTM gating mechanism includes three gates, namely Forget Gate, Input Gate, and Output Gate.
- Bidirectional LSTM (BiLSTM): BiLSTM has a similar architecture to LSTM, but with the added capability of simultaneously capturing information from forward and backward instances.
- GRU: GRU utilizes a gating mechanism, such as LSTM, to keep old information without the information vanishing through time. The key difference between GRU and LSTM is that GRU has two gates, namely reset and update.
- Attention-based (ATT): The ATT model introduces a dynamic mechanism that enhances the adaptability of traffic flow prediction. Unlike the fixed-weight approaches of traditional models, attention assigns varying degrees of importance to different time intervals during the prediction process. This enables the model to automatically focus on the critical temporal patterns, thereby improving its predictive accuracy.

The DT and SVM models are implemented using the Python packages Scikit-Learn (Pedregosa et al., 2011). For the DL models, we utilize Keras API (Chollet, 2007). Our primary architecture for the recurrent models consists of three layers of either RNN, LSTM, bidirectional LSTM, or GRU followed by one fully connected layer. In the ATT model, the recurrent layers are replaced with one attention layer. Multiple sets of hyperparameters were tested to find the combination of values that resulted in the best performance with parameters selected according to our engineering experience. The input has the dimensions of 64 neurons, and the size of the hidden layers is 64 and 32 neurons throughout our experiments. In addition, the activation function for the DL layers is the tanh function. We use the Adam algorithm for optimization and the learning rate is $r = 0.001$. The batch size remains 32, and the hyperparameters for DT and SVM are optimized *via* cross-validation module in the Python packages Scikit-Learn.

Table 2. Quantitative comparisons on traffic road congestion prediction of the testing data.

Model		10 min			30 min			60 min		
		p	R	F_1	p	R	F_1	p	R	F_1
DT	Uncongested	0.96	0.97	0.97	0.96	0.97	0.97	0.96	0.97	0.96
	Congested	0.74	0.67	0.71	0.74	0.67	0.70	0.72	0.65	0.68
	Overall			0.94			0.94			0.94
	Uncongested	0.99	0.97	0.98	0.99	0.97	0.98	0.99	0.96	0.98
SVM	Congested	0.74	0.87	0.80	0.72	0.87	0.78	0.66	0.86	0.75
	Overall			0.96			0.96			0.96
	Uncongested	0.97	0.97	0.97	0.98	0.97	0.98	0.98	0.95	0.97
RNN	Congested	0.76	0.74	0.75	0.72	0.81	0.76	0.54	0.72	0.62
	Overall			0.95			0.95			0.94
LSTM	Uncongested	0.98	0.97	0.98	0.98	0.97	0.98	0.99	0.96	0.97
	Congested	0.76	0.83	0.8	0.73	0.82	0.78	0.66	0.82	0.73
	Overall			0.96			0.96			0.95
BiLSTM	Uncongested	0.98	0.98	0.98	0.98	0.97	0.98	0.98	0.97	0.98
	Congested	0.78	0.82	0.8	0.75	0.83	0.79	0.73	0.82	0.77
	Overall			0.96			0.96			0.96
GRU	Uncongested	0.98	0.98	0.98	0.99	0.97	0.98	0.98	0.97	0.97
	Congested	0.77	0.83	0.80	0.72	0.84	0.78	0.71	0.79	0.74
	Overall			0.96			0.96			0.95

The experimental results for traffic road congestion prediction model on the testing data are provided in Table 2. The results show that all models could predict uncongested states well in different prediction horizons. Because most road state observations are associated to uncongested state, the data is imbalanced and thus, the overall accuracy of all models are high (ranging between 0.94 and 0.96). Therefore, this model performance measurement is less intuitive for imbalanced databases. A reliable model for traffic congestion prediction module should be able to predict the congested state. The results show that the GRU and BiLSTM generally have slight superior performances compared to the other models. In general, the overall models' performances gradually decline for longer horizons (from 10 min to 60 min), however, this trend is less significant for GRU and BiLSTM models. As the prediction window extends, the other models tend to predict an uncongested state for road traffic.

In Table 3, the performance metrics for various models across different prediction horizons are presented. As expected, shorter prediction horizons yield lower errors compared to longer ones, as indicated by the MAE, MRE, and RMSE values. Notably, the ATT model exhibits a significant increase in errors as the prediction horizon extends to 60 min, suggesting a decrease in prediction accuracy over longer time intervals. The trend of increasing errors with longer prediction horizons is consistent across all models, with RMSE values generally higher than MAE because RMSE penalizes larger errors more severely due to the squaring of residuals. To account for differences in flow values, MRE is also considered, revealing that

Table 3. Quantitative comparisons of traffic flow prediction on the testing data.

Model	10 min			30 min			60 min		
	MAE	MRE	RMSE	MAE	MRE	RMSE	MAE	MRE	RMSE
DT	56.54	0.21	86.56	63.69	0.33	102.60	70.02	0.38	115.52
SVM	67.86	0.24	85.12	75.08	0.30	98.37	89.20	0.36	114.03
RNN	53.41	0.14	74.51	55.79	0.16	82.04	59.80	0.18	86.80
LSTM	48.45	0.13	68.52	55.04	0.15	79.48	57.82	0.16	85.19
BiLSTM	50.40	0.16	70.02	53.65	0.15	77.31	56.81	0.16	84.28
GRU	49.88	0.14	69.65	52.98	0.15	76.37	54.24	0.15	79.93
ATT	48.14	0.12	67.43	54.12	0.15	78.24	82.92	0.79	110.59

neural network structures slightly improve overall prediction accuracy. For instance, the GRU model demonstrates RMSE values ranging from 69.65 for 10-min forecasting horizons to 79.93 for 60-min forecasting horizons.

It is noted that our study data is collected from multiple sites, and a particular road-based model is developed for each. It is well-acknowledged that neural network models are sensitive to the dataset, and the number of hidden neurons and layers, the number of epochs, and the loss function – all of which affect performance – are frequently changed to induce improved performance. Therefore, these parameters require adjustment for each site. Considering the plethora of sites in this study, proposing different structures to obtain the best outcome for each requires substantial engineering effort. Therefore, we apply a similar structure for all sites in the study.

To demonstrate the performance of the whole proposed framework in predicting the traffic flow and the road congestion state, we use GRU model (which has a better performance in both prediction modules) and plot the spatial-temporal heatmaps for road flow and density in Figure 7. To avoid redundancy, only two traffic attributes are presented, as the third can be readily derived using the fundamental traffic identity (Equation (1)). In contrast to the traffic flow, the traffic density can be used as an appropriate measurement to present congestion throughout the corridor. Figure 7 shows the spatial-temporal diagram of the averaged measured and predicted traffic flow and the density on weekdays and weekends. The results show that the overall predicted flow diagrams are like measured traffic flow plots. The maximum traffic flow occurs in the morning peak hours between 6:00 am and 10:00 am on weekdays (Figure 7(a,b)).

Figure 7(c) illustrates the measured traffic density along the corridor. The figure presents the corridor as having two major recurrent weekday bottlenecks located within a few kilometers of each other. The more significant bottleneck occurs close to the CBD, starting from 7:00 am and continuing for almost 3 h

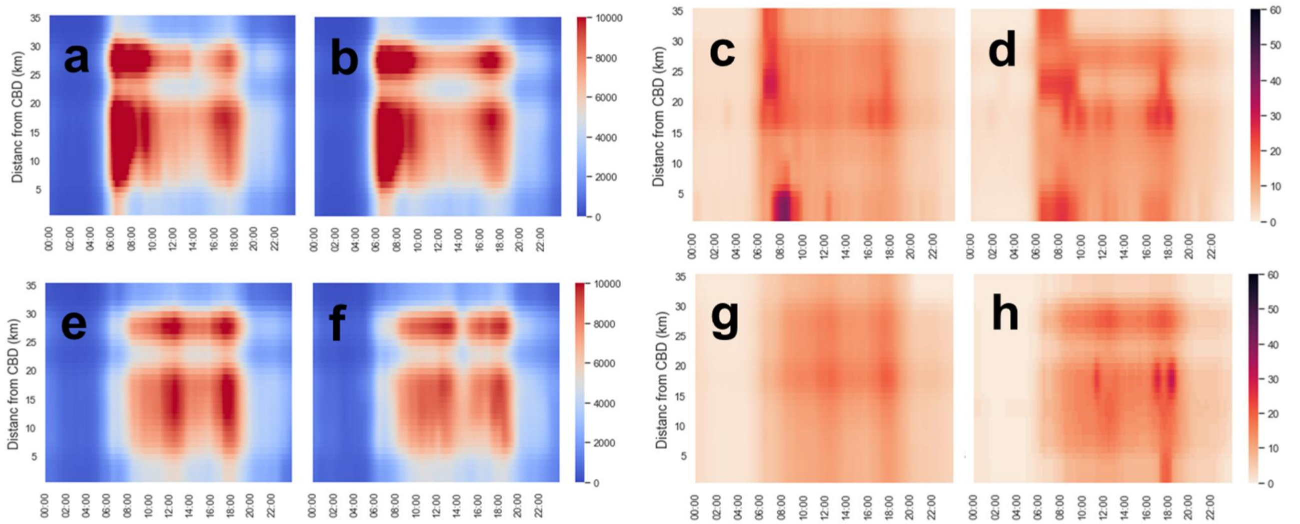

Figure 7. Inbound spatial-temporal traffic flow and density snapshots on the average of weekdays and weekends: (a) predicted and (b) observed traffic flow, (c) predicted and (d) observed traffic density associated with the averaged weekdays, (e) predicted and (f) observed traffic flow, (g) predicted, and (h) observed traffic density associated with the averaged weekend.

to 10:00 am. The congestion reaches its peak at 8:30 am, 7 km from the CBD. The other congestion point appears at a 20–25 km distance from the CBD and occurs from 7:00 to 8:00 am. Such information about congestion is hard to observe in the measured traffic flow figure.

In contrast, Figure 7(a) shows that observed traffic flow during more congested periods is relatively low, misleadingly making it seem as if there is not much traffic during these periods. This is the main reason that the study proposes a complimentary analysis allowing for more intuitive traffic flow predictions. Figure 7(d) illustrates the predicted congestion pattern as the secondary outcome of the prediction model. Though the predicted density presents the congestion location, it still differs from the observed plot in Figure 7(c). The prediction error occurs because the density is indirectly predicted according to the expected traffic flow, road state, and pre-estimated TFDs. These inputs may include errors that lead to miscalculated density values. Similarly, Figure 7(e–h) presents the traffic flow and the density on weekends. There is no critical bottleneck traveling inbound on weekends; overall, the observed and predicted traffic flow and density are similar.

We implemented the entire framework on Collab dashboard and observed that the training model time ranges from approximately 2 min for the DT model to almost 5 min for the bidirectional LSTM model. Notably, this computational time remains consistent for both traffic congestion and flow prediction tasks. While it is true that our framework entails a slightly longer training process compared to traditional prediction tasks due to the inclusion of both traffic congestion and traffic flow prediction components, we emphasize that this does not pose any significant challenge for real-time applications. In fact, the training can be conducted prior to real-time application, ensuring seamless integration and efficient operation in practical scenarios.

6. Conclusion

This study introduced a novel framework that significantly enhances traditional traffic flow prediction methods by also predicting additional traffic attributes using solely traffic flow data. Previous research predominantly focused on enhancing the accuracy of traffic flow predictions without addressing the crucial non-monotonic relationship between traffic flow and congestion. Recognizing and distinguishing between low traffic flow caused by congestion and low flow in uncongested conditions is vital for a deeper understanding of traffic dynamics. This differentiation is essential for comprehensive traffic analysis and management.

Our proposed framework comprises three key components:

TFD Calibration: This involves defining the most representative TFD for each road segment using a limited set of historical data. This calibration considers variations in traffic behavior over multiple days, allowing for more precise predictions of traffic conditions.

Traffic Congestion Prediction: Utilizes historical traffic flow data to determine road traffic congestion states – whether congested or uncongested. This module enhances the framework's ability to predict and manage traffic flow effectively.

Traffic Flow Prediction: At the core of the framework, this component predicts future traffic flows using a DL model trained on mainstream datasets, closely reflecting real-world scenarios.

Experimental results, utilizing a real-world freeway traffic dataset from Melbourne, Australia, highlight the framework's effectiveness. Access to historical traffic flow data, complemented by limited supplementary road data, provides a richer understanding of other critical traffic attributes, such as density and speed. Furthermore, our comparative analysis reveals that GRU models, in particular, demonstrate superior performance in both traffic congestion and flow prediction.

This framework not only surpasses traditional methods by predicting multiple traffic attributes from minimal data but also bridges a significant gap in data-driven traffic flow prediction research. By integrating ML with TFD within a single predictive framework, it leverages the strengths of both approaches, resulting in more accurate and comprehensive traffic predictions.

This study can be extended in several potential future research directions. First, the presence of scatter in real traffic data introduces stochastic characteristics and randomness, which classical deterministic traffic fundamental diagrams (TFDs) fail to capture. Therefore, a promising direction for future research is to explore replacing these proposed stochastic models (Bai et al., 2021) with deterministic counterparts.

Second, the incorporation of a graph structure in the modeling approach holds potential for enhancing results and achieving superior performance. Exploring the full potential of spatial-temporal correlations in traffic flow for short-term prediction applications remains an open challenge. Investigating more efficient deep architectures that seamlessly integrate traffic flow theory and applying them to traffic data could yield valuable insights.

Lastly, the scalability of the proposed method needs further exploration across diverse weather conditions, road types, and other environmental factors. A comprehensive investigation into the method's performance under varied conditions will contribute to a more robust understanding of its applicability and limitations.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work is fully supported by the Australian Research Council (ARC) under grant LP180100114.

ORCID

Sajjad Shafiei http://orcid.org/0000-0002-0155-6866
Simona Adriana Mihaita http://orcid.org/0000-0001-7670-5777
Hanna Grzybowska http://orcid.org/0000-0003-2614-5964
Hussein Dia http://orcid.org/0000-0001-8778-7296

Data availability statement

The code used for the analysis in this article is available on GitHub at https://github.com/sshafiei-tech/Traffic-flow-prediction-enhanced-by-FD.git. The code was developed using Python. Users can access the code, documentation, and issue tracker at the GitHub repository. For questions and support related to the code, please contact corresponding author.

References

Abduljabbar, R. L., Dia, H., & Tsai, P. W. (2021a). Unidirectional and bidirectional LSTM models for short-term traffic prediction. *Journal of Advanced Transportation*, *2021*, 1–16. https://doi.org/10.1155/2021/5589075

Abduljabbar, R., Dia, H., & Tsai, P. W. (2021b). Development and evaluation of bidirectional LSTM freeway traffic forecasting models using simulation data. *Scientific Reports*, *11*(1), 23899. https://doi.org/10.1038/s41598-021-03282-z

Akhtar, M., & Moridpour, S. (2021). A review of traffic congestion prediction using artificial intelligence. *Journal of Advanced Transportation*, *2021*(1), 1–18. https://doi.org/10.1155/2021/8878011

Alt, H., & Godau, M. (1992). Measuring the resemblance of polygonal curves. In *Proceedings of the Eighth Annual Symposium on Computational Geometry* (pp. 102–109). ACM Press.

Aouedi, O., Le, V. A., Piamrat, K., & Ji, Y. (2025). Deep learning on network traffic prediction: Recent advances, analysis, and future directions. *ACM Computing Surveys*, *57*(6), 1–37. https://doi.org/10.1145/3703447

Bai, L., Wong, S. C., Xu, P., Chow, A. H. F., & Lam, W. H. K. (2021). Calibration of stochastic link-based fundamental diagram with explicit consideration of speed heterogeneity. *Transportation Research Part B: Methodological*, *150*, 524–539. https://doi.org/10.1016/j.trb.2021.06.021

Banani Ardecani, F., Mahmoudzadeh, A., & Mesbah, M. (2025). Fuzing multiple erroneous sensors to estimate travel time. *Journal of Intelligent Transportation Systems*, *29*(5), 491–504. https://doi.org/10.1080/15472450.2024.2315514

Chen, C., Liu, J., Li, Y., & Zhang, Y. (2024). Explainable stacking-based learning model for traffic forecasting.

Journal of Transportation Engineering, Part A: Systems, *150*(4), 04024006. https://doi.org/10.1061/JTEPBS.TEENG-8208

Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv: 1409.1259*.

Chollet, F. (2007). others. Keras: Deep learning for humans, computer software, 2025. GitHub. https://github.com/keras-team/keras

Di, X., Shi, R., Mo, Z., & Fu, Y. (2023). Physics-informed deep learning for traffic state estimation: A survey and the outlook. *Algorithms*, *16*(6), 305. https://doi.org/10.3390/a16060305

Eiter, T., & Mannila, H. (1994). *Computing discrete Fréchet distance*. Technische Universitat Wien Technical report.

Gao, J., Ozbay, K., & Hu, Y. (2025). Real-time anomaly detection of short-term traffic disruptions in urban areas through adaptive isolation forest. *Journal of Intelligent Transportation Systems*, *29*(3), 269–286. https://doi.org/10.1080/15472450.2024.2312809

Gu, Z., Saberi, M., Sarvi, M., & Liu, Z. (2017). A big data approach for clustering and calibration of link fundamental diagrams for large-scale network simulation applications. *Transportation Research Procedia*, *23*, 901–921. https://doi.org/10.1016/j.trpro.2017.05.050

Guo, X., Zhang, Q., Jiang, J., Peng, M., Zhu, M., & Yang, H. F. (2024). Towards explainable traffic flow prediction with large language models. *Communications in Transportation Research*, *4*, 100150. https://doi.org/10.1016/j.commtr.2024.100150

Huang, X., Zhang, B., Feng, S., Ye, Y., & Li, X. (2023). Interpretable local flow attention for multi-step traffic flow prediction. *Neural Networks: The Official Journal of the International Neural Network Society*, *161*, 25–38. https://doi.org/10.1016/j.neunet.2023.01.023

Jana, D., Malama, S., Narasimhan, S., & Taciroglu, E. (2023). Edge-based graph neural network for ranking critical road segments in a network. *PLoS One*, *18*(12), e0296045. https://doi.org/10.1371/journal.pone.0296045

Kong, J., Fan, X., Zuo, M., Deveci, M., Jin, X., & Zhong, K. (2024). ADCT-Net: Adaptive traffic forecasting neural network via dual-graphic cross-fused transformer. *Information Fusion*, *103*, 102122. https://doi.org/10.1016/j.inffus.2023.102122

Kontorinaki, M., Spiliopoulou, A., Roncoli, C., & Papageorgiou, M. (2017). First-order traffic flow models incorporating capacity drop: Overview and real-data validation. *Transportation Research Part B: Methodological*, *106*, 52–75. https://doi.org/10.1016/j.trb.2017.10.014

Kumar, N., & Raubal, M. (2021). Applications of deep learning in congestion detection, prediction and alleviation: A survey. *Transportation Research Part C: Emerging Technologies*, *133*, 103432. https://doi.org/10.1016/j.trc.2021.103432

Lan, T., Zhang, X., Qu, D., Yang, Y., & Chen, Y. (2023). Short-term traffic flow prediction based on the optimization study of initial weights of the attention mechanism. *Sustainability*, *15*(2), 1374. https://doi.org/10.3390/su15021374

Liu, L., Zhen, J., Li, G., Zhan, G., He, Z., Du, B., & Lin, L. (2021). Dynamic spatial-temporal representation learning for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *22*(11), 7169–7183. https://doi.org/10.1109/TITS.2020.3002718

Lv, Z., Li, J., Dong, C., & Xu, Z. (2023). DeepSTF: A deep spatial–temporal forecast model of taxi flow. *The Computer Journal*, *66*(3), 565–580. https://doi.org/10.1093/comjnl/bxab178

Mahmassani, H. S., Kim, J., Hou, T., Zockaie, A., Saberi, M., Jiang, L., Verbas, Ö., Cheng S., Chen, Y., & Haas, R. (2012). *Implementation and evaluation of weather responsive traffic estimation and prediction system (No. FHWA-JPO-12-055)*. Joint Program Office for Intelligent Transportation Systems.

Maze, T. H., Agarwal, M., & Burchett, G. (2006). Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation Research Record: Journal of the Transportation Research Board*, *1948*(1), 170–176. https://doi.org/10.1177/0361198106194800119

Nigam, A., & Srivastava, S. (2023). Hybrid deep learning models for traffic stream variables prediction during rainfall. *Multimodal Transportation*, *2*(1), 100052. https://doi.org/10.1016/j.multra.2022.100052

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

Panayiotou, T., Michalopoulou, M., & Ellinas, G. (2023). Survey on machine learning for traffic-driven service provisioning in optical networks. *IEEE Communications Surveys & Tutorials*, *25*(2), 1412–1443. https://doi.org/10.1109/COMST.2023.3247842

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

Qu, X., Zhang, J., & Wang, S. (2017). On the stochastic fundamental diagram for freeway traffic: Model development, analytical properties, validation, and extensive applications. *Transportation Research Part B: Methodological*, *104*, 256–271. https://doi.org/10.1016/j.trb.2017.07.003

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Saberi, M., & Mahmassani, H. S. (2013). Hysteresis and capacity drop phenomena in freeway networks: Empirical characterization and interpretation. *Transportation Research Record: Journal of the Transportation Research Board*, *2391*(1), 44–55. https://doi.org/10.3141/2391-05

Shafiei, S., Mihăiță, A.-S., Nguyen, H., & Cai, C. (2022). Integrating data-driven and simulation models to predict traffic state affected by road incidents. *Transportation Letters*, *14*(6), 629–639. https://doi.org/10.1080/19427867.2021.1916284

Shafiei, S., Mihăiță, A., & Cai, C. (2019). Demand estimation and prediction for short-term traffic forecasting in existence of non-recurrent incidents. In *ITS World Congress 2019 (ITSWC2019), Singapore.*

Shafiei, S., Wang, E., Grzybowska, H., & Cai, C. (2023). Arterial corridor travel time prediction under non-recurring conditions. *Journal of Intelligent Transportation Systems*, *27*(3), 335–346. https://doi.org/10.1080/15472450.2021.2023017

Shi, R., Mo, Z., Huang, K., Di, X., & Du, Q. (2021). A physics-informed deep learning paradigm for traffic state and fundamental diagram estimation. *IEEE transactions on intelligent transportation systems* (pp. 1–11). IEEE.

Tedjopurnomo, D. A., Bao, Z., Zheng, B., Choudhury, F. M., & Qin, A. K. (2020). A survey on modern deep neural network for traffic prediction: Trends, methods and challenges. *IEEE Transactions on Knowledge and Data Engineering*, *34*(4), 1544–1561. https://doi.org/10.1109/TKDE.2020.3001195

Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies*, *43*, 3–19. https://doi.org/10.1016/j.trc.2014.01.005

Wang, B., Qin, A. K., Shafiei, S., Dia, H., Mihaita, A.-S., & Grzybowska, H. (2023, June). Training physics-informed neural networks via multi-task optimization for traffic density prediction. In *2023 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–7). IEEE.

Wang, C., Tian, R., Hu, J., & Ma, Z. (2023). A trend graph attention network for traffic prediction. *Information Sciences*, *623*, 275–292. https://doi.org/10.1016/j.ins.2022.12.048

Wu, Y., Tan, H., Qin, L., Ran, B., & Jiang, Z. (2018). A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, *90*, 166–180. https://doi.org/10.1016/j.trc.2018.03.001

Yang, S. (2013). On feature selection for traffic congestion prediction. *Transportation Research Part C: Emerging Technologies*, *26*, 160–169. https://doi.org/10.1016/j.trc.2012.08.005

Yuan, Y., Wang, Q., & Yang, X. T. (2021). Traffic flow modeling with gradual physics regularized learning. *IEEE Transactions on Intelligent Transportation Systems*, *23*(9), 14649–14660.

Yuan, Y., Zhang, Z., Yang, X. T., & Zhe, S. (2021). Macroscopic traffic flow modeling with physics regularized Gaussian process: A new insight into machine learning applications in transportation. *Transportation Research Part B: Methodological*, *146*, 88–110. https://doi.org/10.1016/j.trb.2021.02.007

Yuan, Y., Zhang, W., Yang, X., Liu, Y., Liu, Z., & Wang, W. (2025). Traffic state classification and prediction based on trajectory data. *Journal of Intelligent Transportation Systems*, *29*(4), 365–379. https://doi.org/10.1080/15472450.2021.1955210

Zhang, J., Mao, S., Yang, L., Ma, W., Li, S., & Gao, Z. (2024). Physics-informed deep learning for traffic state estimation based on the traffic flow model and computational graph method. *Information Fusion*, *101*, 101971. https://doi.org/10.1016/j.inffus.2023.101971

Zhang, K., Liu, Z., & Zheng, L. (2020). Short-term prediction of passenger demand in multi-zone level: Temporal convolutional neural network with multi-task learning. *IEEE Transactions on Intelligent Transportation Systems*, *21*(4), 1480–1490. https://doi.org/10.1109/TITS.2019.2909571

Zhang, W., Yao, R., Du, X., Liu, Y., Wang, R., & Wang, L. (2023). Traffic flow prediction under multiple adverse weather based on self-attention mechanism and deep learning models. *Physica A: Statistical Mechanics and Its Applications*, *625*, 128988. https://doi.org/10.1016/j.physa.2023.128988

Zhang, X., Wen, S., Yan, L., Feng, J., & Xia, Y. (2024). A hybrid-convolution spatial–temporal recurrent network for traffic flow prediction. *The Computer Journal*, *67*(1), 236–252. https://doi.org/10.1093/comjnl/bxac171

Zheng, W., Yang, H. F., Cai, J., Wang, P., Jiang, X., Du, S. S., Wang, Y., & Wang, Z. (2023). Integrating the traffic science with representation learning for city-wide network congestion prediction. *Information Fusion*, *99*, 101837. https://doi.org/10.1016/j.inffus.2023.101837

Zheng, Z., Yang, Y., Liu, J., Dai, H.-N., & Zhang, Y. (2019). Deep and Embedded Learning Approach for Traffic Flow Prediction in Urban Informatics. *IEEE Transactions on Intelligent Transportation Systems*, *20*(10), 3927–3939. https://doi.org/10.1109/TITS.2019.2909904

Zhuang, L., Wu, X., Chow, A. H. F., Ma, W., Lam, W. H. K., & Wong, S. C. (2025). Reliability-based journey time prediction via two-stream deep learning with multi-source data. *Journal of Intelligent Transportation Systems*, *29*(2), 134–152. https://doi.org/10.1080/15472450.2023.2301707