An Approach of Highway Toll Revenue Calculation Based on the Prediction of OD Matrix

Journal of Advanced Transportation

WILEY

RESEARCH ARTICLE OPEN ACCESS

An Approach of Highway Toll Revenue Calculation Based on the Prediction of OD Matrix

Jian Wan[1] | Yinghao Chen[2] | Mengyu Jiang[3] | Hua Tong[4] | Runsheng Wang[3]

[1]Jinling Institute of Technology, Nanjing 211169, China | [2]Chalmers University of Technology, Gothenburg 412 96, Sweden | [3]School of Transportation, Southeast University, Jiulonghu Campus, Nanjing 211100, China | [4]School of Foreign Language, Nanjing University, Xianlin Campus, Nanjing 210023, China

Correspondence: Yinghao Chen (213210474@seu.edu.cn)

Received: 4 March 2025 | **Revised:** 16 October 2025 | **Accepted:** 27 December 2025

Academic Editor: Rui Jiang

Keywords: origin–destination matrix estimation | recurrent neural networks | revenue forecasting | temporal convolutional networks | traffic prediction

ABSTRACT

Accurate toll revenue calculation is essential for highway operation management and construction planning. Reliable forecasting requires a thorough understanding of evolving traffic flow patterns and regional origin–destination (OD) distributions. However, challenges such as incomplete data collection and difficulties in capturing transit traffic patterns often hinder accurate OD estimation. To address this, we propose a bidirectional long short-term memory (Bi-LSTM)–based method for OD matrix estimation in highway revenue prediction. Using real-world data from the Hanghui Highway in China, the proposed model is evaluated against eight benchmark approaches, including traditional machine learning and deep learning models. Results show that Bi-LSTM achieves the best performance, with a root mean squared error (RMSE) of 2.8922 and a mean absolute error (MAE) of 1.1890, outperforming all comparison methods. These findings demonstrate not only the precision and robustness of the Bi-LSTM approach but also its novelty in bridging OD estimation with revenue forecasting, providing new insights for the intelligent highway management and operation as well as the highway revenue prediction and analysis.

1 | Introduction

The rapid development of highway networks worldwide has introduced new challenges and opportunities in the field of highway toll revenue management. Highway revenue is primarily generated by collecting tolls from highway users. From the perspective of investors, highway construction requires significant capital investment, involves lengthy construction periods, and has long investment recovery cycles. As such, highway revenue plays a crucial role in recovering the costs associated with road construction and maintenance, serving as a key metric for evaluating the economic viability of highway projects. From the perspective of operations management and project development, the continuous annual expansion of China's operational highway network makes highway revenue a key factor in shaping the operation, maintenance, and management of existing highways, which in turn influences the strategic planning and development of the overall highway network. As highways continue to expand and traffic volumes increase, ensuring accurate forecasts is essential for sustainable financial planning and effective decision-making within traffic and highway management systems. Therefore, it is necessary to accurately estimate highway revenue to provide decision support for investors, government decision-making bodies, and project construction management departments.

A critical component in toll revenue forecasting is the estimation of the origin–destination (OD) matrix, which captures the flow of vehicles traveling between different regions and helps model and predict traffic patterns on specific highway segments. This OD matrix reflects how traffic demand varies across regions and times, providing insight into both local traffic and transit traffic flows. However, accurate OD matrix estimation remains a significant challenge due to limitations in data availability, gaps in data collection, and the inherent complexity of tracking transit traffic patterns. Traditional methods for OD estimation, such as statistical modeling and simple machine learning approaches, often fall short in capturing the complexity of traffic flows on highways, particularly when faced with limited data.

In recent years, deep learning techniques have emerged as a promising approach to enhance OD matrix estimation and, by extension, toll revenue forecasting. Recurrent neural networks (RNNs), and particularly long short-term memory (LSTM) networks, have shown considerable success in capturing temporal dependencies in sequential data. Building on these advancements, this study proposes a Bidirectional LSTM (Bi-LSTM) model to address the challenges of OD matrix estimation for highway toll revenue forecasting. The Bi-LSTM model is designed to leverage both past and future traffic flow information, providing a more comprehensive understanding of traffic dynamics and OD distributions. This dual-directional approach allows for more accurate modeling of the intricate temporal patterns inherent in highway traffic data, ultimately enhancing revenue prediction accuracy.

Although RNNs and other deep learning networks have been widely applied in traffic prediction tasks, most existing studies focus primarily on traffic flow or speed forecasting. These approaches often treat prediction as a generic sequence learning problem and overlook the structural characteristics of toll revenue forecasting, which depends on the perspective of accurate OD matrix estimation. Moreover, prior works rarely address the real-world operational challenges in highway toll systems, such as incomplete transaction data, class-level vehicle aggregation, and the translation of OD flows into revenue outcomes.

This paper presents a comprehensive evaluation of the proposed Bi-LSTM–based OD matrix estimation method. We compare its performance against eight benchmark models, including both traditional machine learning and other deep learning models, to validate its effectiveness and reliability for highway revenue prediction.

The contributions of this study are twofold. First, we propose an effective short-term revenue forecasting method, based on the Bi-LSTM approach to OD matrix estimation, tailored specifically for the context of highway toll revenue forecasting. Our contributions lie not in proposing a new neural network architecture, but in tailoring deep learning methods to the specific requirements of toll revenue analysis. Second, we demonstrate the model's efficacy and robustness by evaluating it against multiple performance metrics and benchmark models. The proposed method offers a new perspective on highway toll revenue prediction, providing theoretical and data-driven support for decision-making by investors, government authorities, and project construction management departments.

The remainder of the paper is organized as follows: In the next section, we present a brief review of the related work. In Section 3, we claimed our problem and related feature engineering and exposed included machine learning approaches. Section 4 is dedicated to the data acquisition and construction process. Section 5 presents the experimental results from the case study of real-world highway traffic flow datasets in Hangzhou, Zhejiang, China, and followed by the comparison with traditional methods as well as other deep learning models. Finally, a conclusion is presented in Section 6.

2 | Related Work

2.1 | Prediction Based on Traffic Revenue Data

The primary factors influencing highway revenue include traffic volume, vehicle toll classification, toll mileage, toll rates, and policies [1]. Generally, revenue forecasting focuses on two main aspects: toll rate calculation methods and traffic volume. Traditional highway revenue forecasting predominantly relies on the four-step method [2], where a base year is selected to predict future highway traffic volumes, and long-term revenue forecasts are made using a revenue calculation formula [3]. However, Wang et al. [4] summarized the shortcomings of the revenue forecasting methods and models based on the four-step method for traffic forecasting, highlighting that this approach is mainly used for regional planning and may not be suitable for toll highway scenarios. They demonstrated the rationale and necessity of using statistical regression methods to establish highway revenue forecasting models through case studies. With the development of highway informatization, the accumulation of highway data, and advancements in forecasting techniques, machine learning-based methods have become the primary approach for short-term revenue forecasting. These include nonparametric regression [5, 6], neural networks [7], time series models [8], and ensemble forecasting models [9].

In addition to indirectly forecasting highway revenue through traffic volume predictions, direct forecasting is also feasible by utilizing revenue data to establish relationships between current and historical data. Sun [10] utilized quarterly revenue data from previous years to predict highway toll revenue using two methods: the moving average trend elimination method and the GM (1, 1) seasonal adjustment model. Based on the Gene Expression Programming (GEP) algorithm, Qian et al. [11] established a system of ordinary differential equations to model the dynamic characteristics of revenue data and other factors, using numerical solutions for revenue forecasting. They also accounted for the impact of major holiday toll exemption policies on revenue levels, applying local adjustments to the model. Building on this approach, Liu et al. [12] developed a hybrid highway revenue forecasting model by using historical growth rate data to establish a functional relationship for revenue prediction. Accurate modeling of traffic flow and speed characteristics is also essential for revenue-related analysis. Shi et al. [13] proposed a driving-style-aware parameter calibration method for microscopic traffic simulation, demonstrating that incorporating heterogeneous driving behaviors can significantly improve traffic speed estimation accuracy, thereby providing more reliable inputs for traffic volume and revenue forecasting.

In general, the limited availability of highway revenue data has resulted in a scarcity of studies that directly predict revenue using this data. Moreover, since revenue is influenced by multiple factors, building a forecasting model directly from revenue data often reduces the model's interpretability, making it challenging to apply in real-world business scenarios. In revenue forecasting research based on traffic volume prediction models, most studies focus on sectional traffic flow. However, revenue forecasting typically targets either a highway network within a specific region or certain segments of a highway. To ensure the accuracy of the model's predictions, different traffic volume prediction models need to be trained according to the characteristics of specific highway segments.

Furthermore, recent data-driven traffic flow models based on Gaussian processes have been proposed to extend classical fundamental diagrams by incorporating multidimensional traffic states and supply-side factors, offering improved capacity estimation and traffic flow modeling [14], which indirectly supports more refined revenue forecasting.

2.2 | Prediction Based on OD Data

In China, the development of intelligent toll systems and the fully enclosed management model of highways provide detailed vehicle detection data at toll stations. This allows for the direct identification of vehicle categories and travel paths based on entry and exit data. When the toll rates and calculation methods are fixed, highway revenue within a specified study area can be calculated using OD data, which aggregates trip information. Under these circumstances, the challenge of revenue forecasting is effectively transformed into an OD prediction problem, which is fundamentally linked to travel demand forecasting and closely related to traffic volume prediction methods. Recent studies have explored various OD forecasting methods. Ren and Xie [15] employed time series models (auto-regressive integrated moving average [ARIMA], Holt-Winters), support vector regression (SVR), and artificial neural networks (ANN) based on CANDECOMP/PARAFAC tensor decomposition to construct an efficient and accurate multidimensional OD prediction model. In recent years, researchers have extensively studied the use of deep learning methods for traffic forecasting and OD matrix estimation, significantly enhancing the options available for selecting traffic forecasting methods. Yin et al. [16, 17] comprehensively examined traffic forecasting methods based on deep learning from multiple perspectives. Convolutional neural networks (CNNs) [18, 19] comprehensively examined traffic forecasting methods based on deep learning from multiple perspectives. CNNs [20, 21], LSTM networks [22], and gated recurrent units (GRUs) [23] can capture temporal dependencies. Wang et al. [24] used deep learning models trained on traffic flow and rainfall data to predict traffic flow and speed on target road sections, achieving promising results and demonstrating that deep learning methods can capture more hidden traffic features. To further address limitations in traffic state representation and multisource knowledge fusion, Liu et al. [25] proposed the TRIP framework, which integrates physical traffic dynamics and semantic reasoning through a dual state space and hierarchical decision-making, providing a new foundation for complex traffic and OD prediction problems.

In air traffic management, Zhang et al. [26] proposed an attention-based graph convolutional LSTM (AGC-LSTM) model that captures complex spatiotemporal dependencies among air route segments, achieving a 14.4% reduction in MAE compared to traditional GCN-LSTM models. Similarly, Wang et al. [27] developed the IHPO-VMD-LSTM-Informer model for highway traffic flow prediction, which combines dimensionality reduction, optimized time-series decomposition, and a hybrid LSTM-Informer architecture to improve prediction accuracy. In the context of intelligent vehicles, Yang et al. [28] introduced CCTP-Net, a multimodal trajectory prediction model incorporating causal interventions and road rule constraints, demonstrating superior performance in complex traffic scenarios. Beyond prediction accuracy, network resilience has also been considered: Hong et al. [29] proposed a method for complex traffic networks that integrates an SIRD-R fault propagation model with LSTM-based resilience trend forecasting, enhancing both network security and recovery capacity. These studies highlight the increasing complexity of prediction models that not only capture temporal and spatial dependencies but also contain causal relationships, external features, and system resilience.

3 | Methods

3.1 | Problem Definition

Highway toll revenue forecasting can be formulated as a supervised learning problem. Let X_t denote the feature vector at time t, which includes OD traffic flows and auxiliary variables (e.g., time of day and vehicle classes). The objective is to predict the toll revenue R_t at time t, or future horizon $t+h$.

Formally, the task can be written as follows:

$$\widehat{R}_{t+h} = f\left(X_t, X_{t-1}, \ldots, X_{t-L}\right), \tag{1}$$

where L represents the historical time window and $f()$ is the predictive model.

In practice, toll revenue depends on total inflows and outflows but also on the specific OD pairs traveled, since toll charges vary with both distance and vehicle class. Therefore, revenue forecasting can be transformed into an OD matrix estimation problem. Let $V^k_{i,j,t}$ denote the number of vehicles of class k traveling from entry station i to exit station j during time interval t, and let $\tau^k_{i,j,t}$ be the corresponding toll rate. The total toll revenue at time t is then:

$$R_t = \sum_i \sum_j \sum_k V^k_{i,j,t} \cdot \tau^k_{i,j,t}. \tag{2}$$

Thus, by estimating the OD flows $V^k_{i,j,t}$, the toll revenue R_t can be directly computed. This formulation establishes a clear link between OD estimation and revenue forecasting, positioning OD prediction as the essential intermediate step for accurate toll revenue analysis.

3.2 | Feature Engineering

The OD matrix describes the number of vehicles traveling between each entry-exit toll station pair within a specific time interval. Since toll revenue is determined not only by overall traffic volume but also by the distance traveled and the class of vehicles, accurate OD estimation is a prerequisite for reliable revenue forecasting. To achieve this, several preprocessing and feature engineering steps were implemented:

- OD Division: Each highway trip is mapped to an OD pair based on entry and exit records obtained from the toll collection system. By aligning entry and exit timestamps, we reconstructed complete trip trajectories, thereby forming an

TABLE 1 | Summary of benchmark model characteristics.

Model	Model type	Characteristics
ARIMA	Statistical	A linear time series model that captures autoregressive, differencing, and moving average components, suitable for stationary or differenced stationary sequences
MLR	Statistical	A linear regression model that establishes the relationship between multiple input features and the output, interpretable but limited in capturing nonlinear or complex temporal dependencies
HA	Statistical	A simple time series prediction model that uses the average traffic flow from a historical time window as the predicted flow for the next time step
SVR	Machine learning	A regression model based on support vector machines, which fits the relationship between input and output by finding an optimal hyperplane in high-dimensional space
MLP	Deep learning	A feedforward neural network that learns complex mappings between input features and output through multiple layers of nonlinear transformations
GRU	Deep learning	A variant of the LSTM model, retaining LSTM's ability to model temporal dependencies, but only considers time features while ignoring spatial features
LSTM	Deep learning	A special neural network that captures long- and short-term dependencies in time series through memory cells and gate mechanisms (input gate, forget gate, output gate)

OD flow table. This procedure captures the spatial distribution of trips and enables the calculation of distance-dependent tolls.

- Vehicle Class Aggregation: Although the Chinese toll system officially distinguishes more than 10 vehicle categories, certain classes have relatively low sample sizes, leading to data sparsity. To ensure statistical stability and model robustness, we aggregated vehicles into four broader categories: small, medium, large, and extra-large. This aggregation retains the most critical heterogeneity of toll pricing while reducing the dimensionality of the OD matrix.
- Time Slot Division: To capture the short-term temporal dynamics of traffic demand, the continuous transaction records were segmented into fixed-length time slots (e.g., 15 min). Each slot aggregates OD flows across all vehicle categories, enabling the model to learn both peak and off-peak traffic fluctuations as well as recurrent daily patterns.
- Feature Construction: For each OD pair (i, j) and time slot t, we constructed a feature vector including lagged OD volumes to capture temporal dependencies and vehicle class proportions within the OD pair, ensuring class-level heterogeneity is preserved. Contextual indicators, such as workday/holiday flags and time-of-day encodings, to represent external periodic or policy-driven effects; distance-based attributes, i.e., toll rates, explicitly link OD flows to revenue.

3.3 | Prediction Models

In this study, we selected the baseline models ARIMA, multiple linear regression (MLR), and historical average (HA), as well as the deep learning models multilayer perceptron (MLP), GRU, and LSTM as the prediction models. The characteristics of each model are summarized in Table 1.

Three deep learning models are used in this study: MLP, GRU, and LSTM. The following sections will provide a more detailed introduction to the model construction [30].

3.3.1 | MLP

MLP is the basic form of a deep neural network, which consists of multiple layers of neurons. In an MLP, each neuron in a layer is fully connected to the neurons in the next layer, receiving inputs from the lower layer and influencing the neurons in the upper layer.

Using the matrix $\mathbf{X} \in R^{n\times d}$ to represent a mini-batch of n samples, where each instance has d inputs (features), we can define a single hidden layer MLP with h hidden units. Let $\mathbf{H} \in R^{n\times h}$ represent the output of the hidden layer, i.e., the hidden representation. Since both the hidden layer and the output layer are fully connected, we have hidden layer weights $\mathbf{W}^{(1)} \in R^{h\times q}$, hidden layer biases $\mathbf{b}^{(1)} \in R^{1\times h}$, output layer weights $\mathbf{W}^{(2)} \in R^{h\times q}$, and output layer biases $\mathbf{b}^{(1)} \in R^{1\times h}$. This allows us to calculate the output $\mathbf{O} \in R^{n\times q}$ of the single-layer MLP as follows:

$$\begin{aligned} H &= XW^{(1)} + b^{(1)}, \\ O &= XW^{(2)} + b^{(2)}. \end{aligned} \tag{3}$$

By folding the hidden layer, we obtain an equivalent single-layer model with parameters $W = W^{(1)}W^{(2)}$ and $b = b^{(1)}W^{(2)} + b^{(2)}$:

$$O = \left(XW^{(1)} + b^{(1)}\right)W^{(2)} + b^{(2)} = \mathbf{X}W^{(1)}W^{(1)} + b^{(1)}W^{(2)} + b^{(2)} = XW + b. \tag{4}$$

To fully harness the potential of a multilayer architecture, a nonlinear activation function σ is applied to each hidden unit after the affine transformation. With the activation function in place, the MLP can no longer be simplified into a linear model:

$$H = \sigma\left(XW^{(1)} + b^{(1)}\right),$$
$$O = XW^{(2)} + b^{(2)}. \tag{5}$$

To build more generalized MLPs, we can continue stacking such hidden layers, for example, $H^{(1)} = \sigma_1(XW^{(1)} + b^{(1)})$ and $H^{(2)} = \sigma_1(XW^{(2)} + b^{(2)})$, layer by layer, resulting in a more expressive model.

3.3.2 | GRU

Traditional RNNs often suffer from vanishing or exploding gradients, which limits their ability to capture long-term dependencies in traffic time series. In practice, traffic data may also contain missing values or interruptions due to facility maintenance, further challenging sequence modeling. To address these issues, advanced recurrent architectures such as LSTM [22] and the GRU [23] introduce gating mechanisms that enable more effective information storage and update. GRU, a simplified variant of LSTM, uses reset and update gates to balance past and new information efficiently, often achieving comparable performance with reduced complexity.

The computation formulas for the reset gate $\mathbf{R}_t \in R^{n\times h}$ and the update gate $\mathbf{Z}_t \in R^{n\times h}$ at time step t are as follows:

$$R_t = \sigma\left(X_t W_{xr} + H_{t\text{-}1} W_{hr} + b_r\right),$$
$$Z_t = \sigma\left(X_t W_{xz} + H_{t\text{-}1} W_{hz} + b_z\right). \tag{6}$$

In the equations, h represents the number of hidden units, $X_t \in R^{n\times d}$ is the mini-batch input at the given time step t (where n is the number of samples and d is the number of input features), and $H_{t\text{-}1} \in R^{n\times h}$ is the hidden state from the previous time step. $W_{xr}, W_{xz} \in R^{d\times h}$ and $W_{hr}, W_{hz} \in R^{h\times h}$ are weight parameters, $b_r, b_z \in R^{1\times h}$ are bias parameters, σ is the sigmoid function that maps input values to the interval (0, 1).

Next, we will integrate the calculation of the reset gate $\mathbf{R}_t$ with the conventional hidden state update mechanism to obtain the candidate hidden state $\widetilde{H}_t \in R^{n\times h}$ at time step t.

$$\widetilde{H}_t = tanh\left(X_t W_{xh} + \left(R_t \odot H_{t\text{-}1}\right) W_{hh} + b_h\right). \tag{7}$$

In the equation, $W_{xh} \in R^{d\times h}$ and $W_{hh} \in R^{h\times h}$ are weight parameters, while $b_h \in R^{1\times h}$ is a bias parameter, the symbo ⊙ denotes the element-wise multiplication operator. The tanh function is used as the activation function to ensure that the values in the candidate hidden state remain within the interval (–1, 1).

Finally, by combining the effects of the update gate Z_t, the information from the old state $H_{t\text{-}1}$ the new candidate hidden state $\widetilde{H}_t$ is merged to determine the new hidden state $H_t \in R^{n\times h}$.

$$H_t = Z_t \odot H_{t\text{-}1} + \left(1 - Z_t\right) \odot \widetilde{H}_t. \tag{8}$$

3.3.3 | LSTM Network

LSTM introduces a memory cell with the same shape as the hidden state, along with three gates. These gates include the input gate, which determines when to read data into the cell; the forget gate, which controls when to reset the cell's contents; and the output gate, which determines what information to output from the cell. We refer to this as the output gate.

The input gate $\mathbf{I}_t \in R^{n\times h}$, forget gate $\mathbf{F}_t \in R^{n\times h}$, and output gate $\mathbf{O}_t \in R^{n\times h}$ at time step t are computed as follows:

$$I_t = \sigma\left(X_t W_{xi} + H_{t\text{-}1} W_{hi} + b_i\right), \tag{9}$$

$$F_t = \sigma\left(X_t W_{xf} + H_{t\text{-}1} W_{hf} + b_f\right), \tag{10}$$

$$O_t = \sigma\left(X_t W_{xo} + H_{t\text{-}1} W_{ho} + b_o\right), \tag{11}$$

where h is the number of hidden units, $X_t \in R^{n\times d}$ represents the mini-batch input at time step (with n as the number of samples and d as the number of input features), and $H_{t\text{-}1} \in R^{n\times h}$ is the hidden state from the previous time step. The weight parameters are $W_{xi}, W_{xf}, W_{xo} \in R^{d\times h}$ and $W_{hr}, W_{hz}, W_{ho} \in R^{h\times h}$, while the bias parameters are $\mathbf{b}_r, \mathbf{b}_z \in R^{1\times h}$, The function σ is the sigmoid function, which maps input values to the range (0, 1).

Next, the candidate memory cell $\widetilde{\mathbf{C}}_t \in R^{n\times h}$ is calculated.

$$\widetilde{C}_t = tanh\left(X_t W_{xc} + H_{t\text{-}1} W_{hc} + b_c\right). \tag{12}$$

In this equation, $W_{xc} \in R^{d\times h}$ and $W_{hc} \in R^{h\times h}$ are weight parameters, and $b_c \in R^{1\times h}$ is the bias parameter. The tanh function is used as the activation function to ensure that the values in the candidate memory cell remain within the range (–1, 1).

Similar to the two gates in GRU, LSTM uses the input gate $\mathbf{I}_t$ and the forget gate $\mathbf{F}_t$ to selectively incorporate information from the old memory cell $\mathbf{C}_{t\text{-}1}$ and the candidate memory cell $\widetilde{\mathbf{H}}_t$.

$$C_t = F_t \odot C_{t\text{-}1} + I_t \odot \widetilde{C}_t. \tag{13}$$

In this equation, ⊙ represents element-wise multiplication.

Finally, the output gate O_t controls the flow of information from the memory cell $\mathbf{C}_t$ to the hidden state $H_t \in R^{n\times h}$:

$$H_t = O_t \odot tanh\left(C_t\right). \tag{14}$$

It can be observed that the larger the value of the input gate in LSTM, the more effectively it absorbs information at the current time step, enabling the model to capture short-term dependencies in the sequence. Conversely, the larger the value of the forget gate, the more firmly it retains historical information, allowing the model to capture long-term dependencies in the sequence.

4 | Data Acquisition and Construction

4.1 | Dataset Description

This study selects the G56 Hanghui Expressway as the research object, focusing on the section from the Liuxia Hub (east side) to the Yuqian Hub (west side), which includes a total of eight toll stations along the route. Traffic flow data was collected from these eight toll stations and 14 pairs of gantries, with the data collection taking place on February 7, 2024. The specific research scope is indicated by the blue segment in Figure 1.

FIGURE 1 | Research scope of the expressway. (*Note: the road segment in this figure is extracted from Google Maps imagery).

4.2 | Feature Engineering

4.2.1 | OD Division

The toll station OD and road OD can be specifically divided based on different interchanges (or junctions). Take the Wangjiabu interchange as an example, whose input and output toll stations and roads are shown in Figure 2. Vehicles can enter from the Jiufeng Toll Station (marked as Jiufeng Toll-O and Jiufeng Toll-D in the figure) or from the west ring expressway of Hangzhou (O2-initial and D2-initial in the figure, including both upper and lower positions). The former can be directly recorded by the toll station, while the latter can only be detected by the nearest gantries, such as Gantry 5 and Gantry 27. Therefore, for vehicles entering from the west ring expressway of Hangzhou, the origin point (O) can be located at Gantry 5 and Gantry 27, while the destination point (D) can be located at Gantry 6 and Gantry 26.

Therefore, based on the location information of the gantries and toll stations throughout the road network, the O points within the study area mainly include all the toll station entrances shown in Figure 3, O1 to O3. The D points mainly include all the toll station exits shown in Figure 3, D1 to D3. Thus, all OD pairs can be categorized into four types: toll station to toll station, toll station to other road sections, other road sections to toll station, and other road sections to other road sections, as summarized in Table 2.

4.2.2 | OD Distance Calculation

The calculation of OD distance can mainly be divided into the following steps: (1) Identify the edge corresponding to each O/D based on the toll station ID. (2) Use the Dijkstra algorithm to compute the shortest path from the origin to the destination. (3) Traverse the lengths of all edges between OD pairs and calculate the total OD distance. The calculation

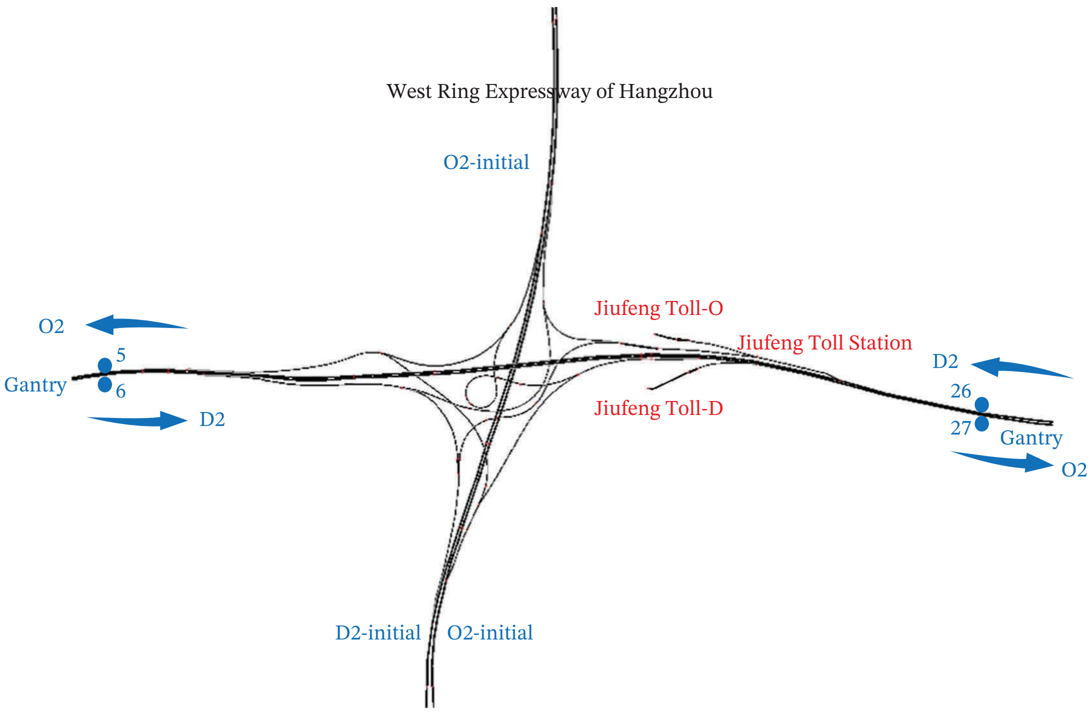

FIGURE 2 | Wangjiabu interchange roads and toll stations. (*Note: the road network in this figure is an original model constructed by the authors using SUMO).

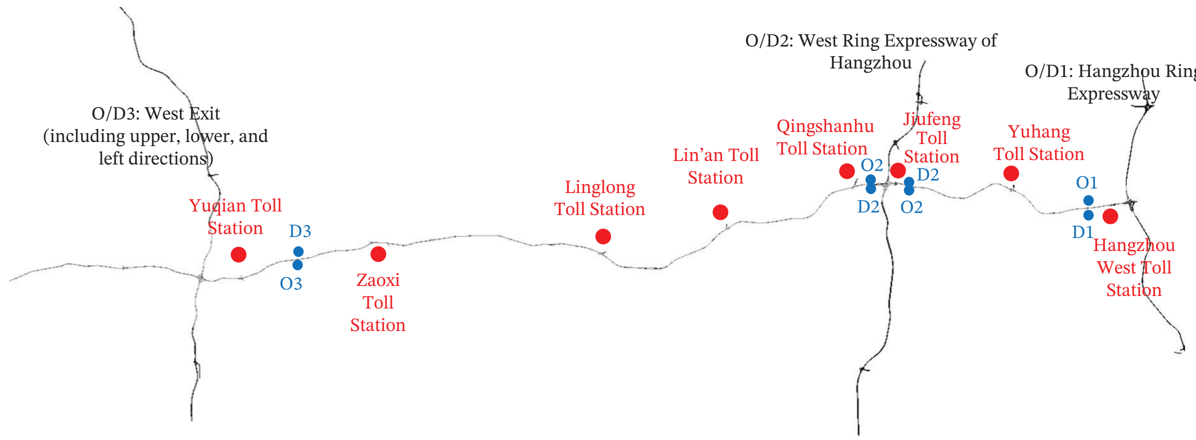

FIGURE 3 | Gantries, toll station locations, and OD (where the toll station can serve as both O and D). (*Note: the road network in this figure is an original model constructed by the authors using SUMO).

process can be represented by the pseudocode shown in Table 3, and the results of distance calculations for some OD pairs are shown in Table 4.

4.2.3 | Time Slot Division and Feature Extraction

First, the raw data from the toll stations must be cleaned, primarily addressing duplicate and missing data. Then, OD-based data aggregation is conducted, including selecting all vehicles departing from each O point within a 15 min interval and, based on their corresponding D points, calculating the OD flow. Different vehicle types are classified into four categories (1.0, 1.5, 2.5, and 4.0) based on conversion standards, and the OD flow for each vehicle category is recorded in a dictionary. Next, the dataset is divided into time slots with a 15 min interval as the time step. Finally, time series features and external features are extracted. The former includes traffic characteristics for the current time slot, as well as the previous six 15 min time slots, the same time slot from the previous day, and the same time slot from the previous week, as shown in Figure 4. The latter involves external information such as whether it is a working day and the day of the week.

4.3 | Evaluation Metrics

This study uses two evaluation metrics to measure the performance of the model:

1. RMSE

$$\mathrm{RMSE} = \sqrt{\frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\left(\mathbf{y}_i^j - \widehat{\mathbf{y}}_i^j\right)^2}. \tag{15}$$

2. MAE

$$\mathrm{MAE} = \frac{1}{MN}\sum_{j=1}^{M}\sum_{i=1}^{N}\left|\mathbf{y}_i^j - \widehat{\mathbf{y}}_i^j\right|, \tag{16}$$

where y_i^j represents the actual value of the OD pair i at time slice j, $\widehat{y}_i^j$ represents the corresponding predicted value, M represents the number of test time slices, and N represents the number of OD pairs. RMSE and MAE are used to measure prediction errors, with smaller values indicating better predictive performance.

5 | Experimental Results and Analysis

5.1 | Model Results

In this study, we selected time series lag features (vol_t-6 to vol_t-1) as well as the external time-slice feature (ts_15) and plotted the correlation coefficient heatmap, shown in Figure 5. The results show that the correlation coefficients among different lag features are generally above 0.88, with the highest correlations observed between adjacent lags, indicating significant short-term continuity and inertia in traffic flow over time. Meanwhile, the external feature ts_15 exhibits relatively low correlations with the lag features (around 0.18–0.24), suggesting that it provides complementary information to the time series features. These findings provide a basis for the rational combination of lagged variables and external variables in subsequent modeling.

Using the grid search method for small-scale hyperparameter tuning on the above deep learning methods, the candidate parameters are shown in Table 5. Each deep learning model uses

TABLE 2 | OD pair classification.

OD type	OD pair (example)
Toll Station-Toll Station	Hangzhou West Toll Station-Lin'an Toll Station
Toll Station-Other Road Segment	Yuhang Toll Station-West Exit (i.e., Yuhang Toll Station-D3)
Other Road Segment-Toll Station	Hangzhou Ring Expressway-Linglong Toll Station (i.e., O2-Linglong Toll Station)
Other Road Segment-Other Road Segment	Hangzhou Ring Expressway West Line-West Exit (i.e., O2-D3)

TABLE 3 | Pseudocode for OD distance calculation.

```
Algorithm: OD distance calculation
Input: Toll Station Id, directed road graph G = (V, E, W), edge length w
Output: OD Distance
Begin
  S ← (Liu et al.)
  dist [s, s] ← 0
  for each OD in all ODs do
    Find the start node s and end node e corresponding to the OD pair
    for v in V–S do
      dist [s, v_i] ← w (s, v_i)
    while V-S ≠ ∅ do
      Find the minimum distance [s, v_i] in V-S
      S ← S ∪ {v_j}
      for v_i ∈ V-S do
        if dist [s, v_j] + w_ji < dist [s, v_i] then
          dist [s, v_i] ← dist [s, v_j] + w_ji
    OD Distance ← dist [s, e]
End
```

the ReLU activation function and the Adam optimizer, with each parameter set undergoing 100 training epochs.

The optimal hyperparameters selected for the Bi-LSTM model were a hidden size of (128, 128), a learning rate of 0.001, and a batch size of 32. All models used the ReLU activation function and the Adam optimizer with 100 training epochs, providing a balance between predictive performance and training stability.

In this study, the last 10 time slices (86–95) of time slice 2.7 were selected as the test set, and the previous time slices were used as the training set. Various models were trained on the dataset, and the best parameters were selected based on the size of the loss function. The model with the best parameters was then used for predictions. Taking category 1.0 vehicles as an example, the predictive accuracy of each model is shown in Table 6.

According to Table 6, traditional non-deep learning models such as ARIMA, HA, and SVR perform poorly in predicting OD volume, with both RMSE and MAE relatively large. These models are generally limited in capturing the complex nonlinear and temporal dependencies inherent in traffic flow data. Interestingly, MLR demonstrates relatively competitive performance, which indicates that, despite its linear nature, MLR is able to capture a substantial portion of the variability in OD volume by leveraging multiple lagged and external features simultaneously. Its ability to combine several predictors linearly allows it to partially account for temporal and cross-feature relationships, which may explain its surprisingly good performance compared with other non-deep learning baselines. In contrast, deep learning models like MLP, GRU, and LSTM demonstrate higher precision in OD volume prediction, better handling the complex features of the task. Among them, the bidirectional GRU and LSTM models further improve predictive accuracy, indicating that bidirectional models capture more comprehensive feature information from time series. Bidirectional models process the sequence in both forward and backward directions, enabling the model to incorporate information from both earlier and later time steps. This is particularly beneficial for highway OD demand and revenue prediction, where certain traffic patterns (e.g., peak flow periods and congestion propagation) can be influenced not only by past volumes but also by anticipated downstream conditions. As a result, bidirectional models can better capture temporal correlations and abrupt fluctuations, leading to improved predictive accuracy in this context.

The prediction results for each OD pair are visualized in Figure 6. As shown, traditional models like HA and SVR underperform, with smoother prediction curves that fail to capture short-term fluctuations in OD volume. This result is interpretable, as HA and SVR do not differentiate the importance of features; they simply attempt to discover linear relationships and approximate mappings from the input domain to the target domain. ARIMA shows almost no predictive capability, while MLR performs

TABLE 4 | Distance calculations for selected OD pairs.

OD	O	D	Distance
Hangzhou West_Jiufeng	1124884051	562685056#1.325	16.68599
Hangzhou West_Lin'an	1124884051	240470455#1	29.68211
Hangzhou West_Linglong	1124884051	712593536#3	39.58748
Hangzhou West_Qingshanhu	1124884051	240470458#1	19.97802
Hangzhou West_Yuhang	1124884051	307809645.1	8.22997
Hangzhou West_Yuqian	1124884051	163269898	65.47337

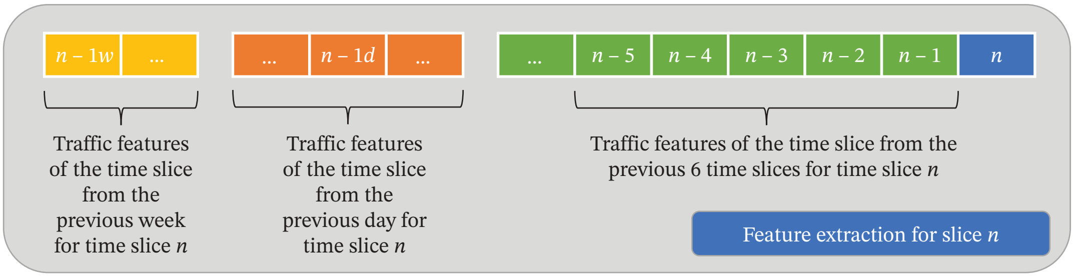

FIGURE 4 | Time series feature extraction.

reasonably well, capturing major trends despite its linear nature. In comparison, deep learning models such as MLP, GRU, LSTM, and their bidirectional variants show significantly better performance. In particular, bidirectional GRU and LSTM capture flow trends more accurately.

5.2 | Model Analysis

Based on the above modeling results, Shapley Additive exPlanations (SHAP), a powerful interpretable tool for black-box models, were used to obtain the importance of each feature in the Bi-LSTM model based on the marginal contribution. It effectively assesses the influence of each feature on the final prediction values. The features used for prediction are mentioned in Section 4.2.3. Since the external features remain the same throughout the day's data, the SHAP values of the flow over the first six time periods were calculated. The importance and influence of the six features in the model output are shown in Figure 7. In the figure, the *X*-axis represents the SHAP value, indicating the impact of each feature on the model output, with positive and negative values representing positive and negative influences, respectively. The *Y*-axis represents the six analyzed features, while the color indicates the feature value. It can be observed that the previous time step (i.e., vol_t-1) has the greatest impact on the prediction result. When the feature value is larger, its positive influence on the model is more apparent, indicating that this feature plays a critical role in the prediction. The SHAP values for the first two and three time steps are relatively smaller but still have a significant impact on the model. The influence of the remaining features is comparatively smaller. The average SHAP values for all samples are shown in Table 7, reflecting the average impact of each feature on the results.

5.3 | Revenue Calculation

According to Article 16 of "Regulations on the Administration of Toll Roads" issued by the State Council of the People's Republic of China, "the toll rate for vehicles shall be determined based on factors such as the technical grade of the road, the total investment, the local price index, the repayment period of loans or funds raised through compensation, the recovery period of the investment, and traffic volume." This shows that the toll rates of different highways may vary. According to the "Vehicle Classification for Toll Roads" standard published by the Ministry of Transport of the People's Republic of China, vehicles are categorized into four classes of passenger vehicles, six classes of trucks, and six classes of

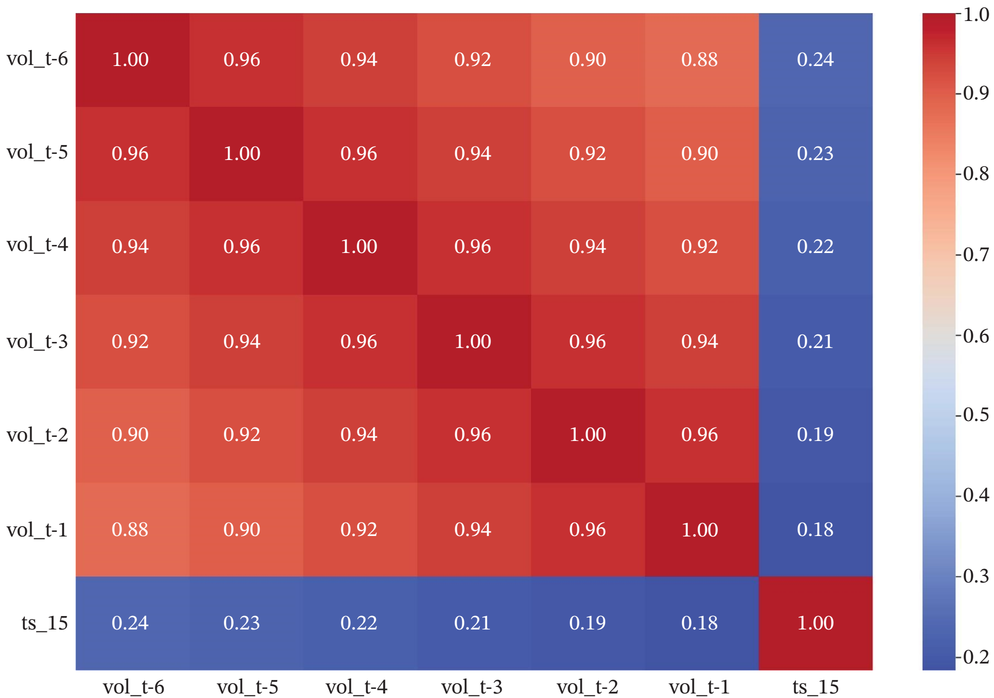

FIGURE 5 | Feature correlation heatmap.

TABLE 5 | Candidate hyperparameters for small-scale tuning.

Parameters	Range
Hidden layer size	(32, 32), (64, 64), (128, 128)
Learning rate	$1e-4$, $1e-3$
Batch size	32, 64

TABLE 6 | The predictive accuracy of each model.

Model	MAE	RMSE
ARIMA	2.3417	6.5374
MLR	1.2367	2.8887
HA	1.5862	3.7554
SVR	2.1288	3.6455
MLP	1.3074	3.1314
GRU	1.3047	3.1267
LSTM	1.2774	3.0001
Bi-GRU	1.2218	2.9840
Bi-LSTM	**1.1890**	**2.8922**

Note: The bold values indicate the results obtained by the model proposed in this study.

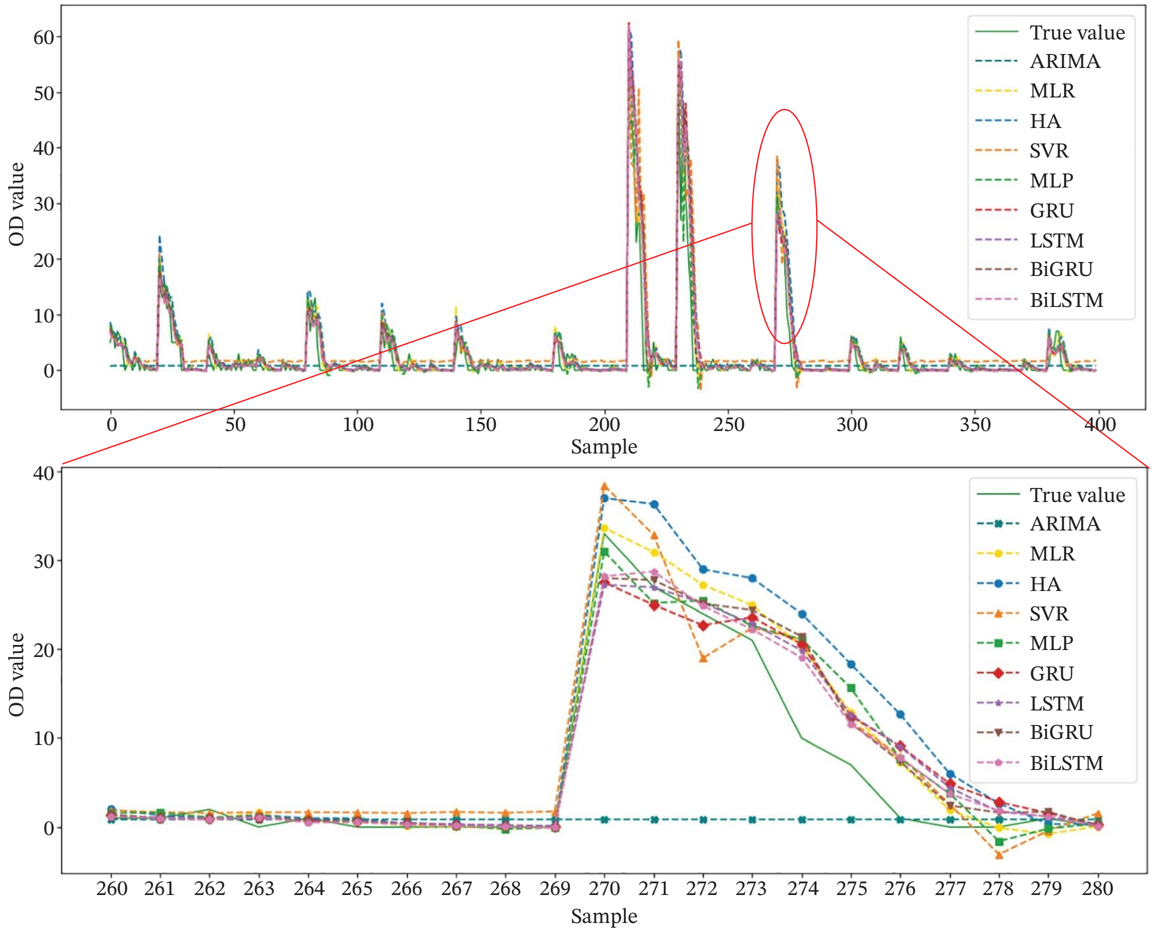

FIGURE 6 | Visualization of prediction results.

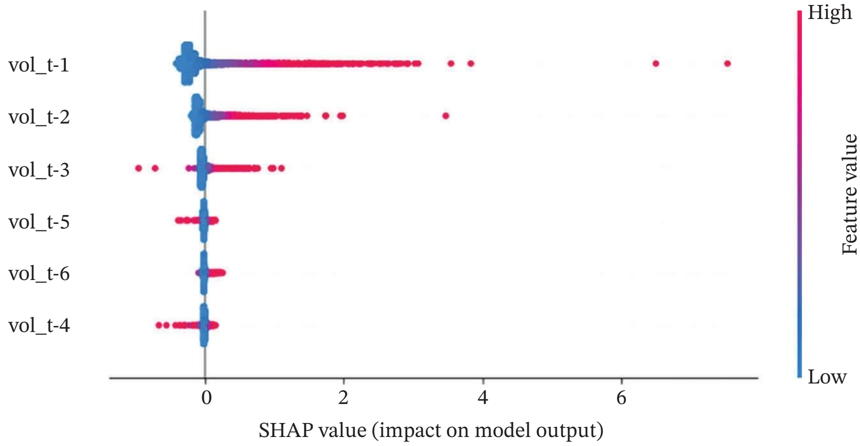

FIGURE 7 | Feature impact plot.

specialized vehicles, each with different toll rates. Furthermore, the "Implementation Plan for the Comprehensive Promotion of Differential Tolling on Highways," jointly released by the Ministry of Transport and the National Development and Reform Commission, encourages local governments to develop differentiated tolling policies based on regional characteristics while ensuring the financial viability of highways. The goal is to promote traffic flow optimization, cost reduction, and efficiency improvements, aiming for a win-win scenario for all parties involved.

As the highway under study is located in Zhejiang Province, the toll calculation follows the guidelines outlined in the "Notice on Continuing to Implement the Vehicle Tolling Policies for Toll Roads across the Province" issued by the General Office of the People's Government of Zhejiang Province. The toll rates for vehicles are calculated as follows:

$$\begin{aligned} T_{\text{pas}} &= P_f + P_k \times D + P_t, \\ T_{\text{tru/spe}} &= P_k \times D + P_t. \end{aligned} \tag{17}$$

In this context, T_{pas} refers to the toll for passenger vehicles, $T_{\text{tru/spe}}$ refers to the toll for trucks/specialized vehicles, P_f refers to the base rate per trip, P_k refers to the per-kilometer rate, D refers to the actual distance traveled by the vehicle, and P_t refers to the additional toll for tunnels or bridges.

The per-trip fee is set at 5 yuan, and the per-kilometer rate depends on the vehicle type. Based on the vehicle classifications, the toll rates for the four categories of vehicles (1.0, 1.5, 2.5, and 4.0) can be approximated as 0.45, 1.0, 1.5, and 2.0 yuan per kilometer, respectively. Due to the lack of actual revenue data, only toll station OD data is used to predict OD flows and estimate projected revenue. Using all remaining models for prediction, excluding ARIMA and SVR due to their poor performance, the final revenue forecast results for time slices 86 to 95 (i.e., from 21:30 to 24:00) are presented in Table 8.

TABLE 7 | Mean SHAP values of features.

Feature	Mean SHAP value
Vol_t-1	$1.95648193e-05$
Vol_t-2	$-1.29988781e-05$
Vol_t-3	$3.58583142e-05$
Vol_t-4	$1.88776363e-04$
Vol_t-5	$2.11226637e-04$
Vol_t-6	$-2.42327396e-04$

To further interpret the reported error metrics in practical terms, we translated the model's prediction errors into their impact on toll revenue. Based on the official tolling formulas adopted in Zhejiang Province (equations (10) and (11)), an RMSE of 2.8922 and an MAE of 1.1890 correspond to deviations equivalent to only a few vehicles per OD pair per 15 min interval. When converted into toll charges using the vehicle-class-based rates, these deviations amount to very minor fluctuations in projected revenue. From an operational perspective, this level of accuracy ensures that revenue forecasts can reliably support short-term toll collection planning, financial reporting, and traffic management with minimal risk of bias.

Revenue prediction results for each model in Table 8 indicate significant differences in predictions across different time periods. Overall, all models capture the trend of decreasing revenue over time. Additionally, the predictions from various models show smaller discrepancies in the earlier periods and larger discrepancies in the later periods. Specifically, the MLR and SVR models significantly underestimate the revenue results, even displaying a pattern of first increasing and then decreasing during the 22:15 to 23:00 time periods. The HA model shows slight overestimations, while the MLP model shows slight underestimations. The unidirectional GRU and LSTM exhibit larger fluctuations in predicted revenue values, with significantly higher predictions in the later periods, while the estimates from the bidirectional GRU and LSTM are more reasonable.

Using the Bi-LSTM model, which showed the best prediction performance, the estimated revenue for time slices 86 to 95 is shown in Figure 8. It can be observed that overall revenue shows a declining trend from 21:30 to 24:00. Notably, during the 23:00 to 24:00 time periods, revenue decreases significantly compared to earlier periods. The time-based variation in predicted revenue reflects the typical decrease in highway traffic and revenue at night, consistent

TABLE 8 | The revenue predictions of each model.

Time periods	HA	MLR	SVR	MLP	GRU	LSTM	Bi-GRU	Bi-LSTM
21:30–21:45	4573.702	3861.826	4740.239	4100.244	3864.835	3861.826	3885.678	3756.506
21:45–22:00	4180.382	3443.851	3867.424	3503.518	3434.773	3443.851	3419.023	3277.943
22:00–22:15	3374.997	2928.243	3373.968	2877.730	2971.391	2928.243	2901.723	2733.206
22:15–22:30	2922.445	2717.548	2982.490	2755.657	2699.486	2717.548	2639.751	2525.649
22:30–22:45	2524.240	2273.933	3353.714	2365.067	2300.393	2273.933	2262.181	2071.869
22:45–23:00	2139.840	1806.262	2760.271	1775.952	1914.807	1806.262	1786.962	1620.003
23:00–23:15	1543.296	1255.304	2562.040	1213.294	1357.857	1255.304	1200.533	1047.642
23:15–23:30	894.643	773.093	1905.918	481.815	778.537	773.093	490.352	517.869
23:30–23:45	431.271	525.582	1552.602	99.565	538.487	525.582	296.946	328.335
23:45–24:00	204.060	371.870	1427.487	61.747	387.861	371.870	269.575	211.189

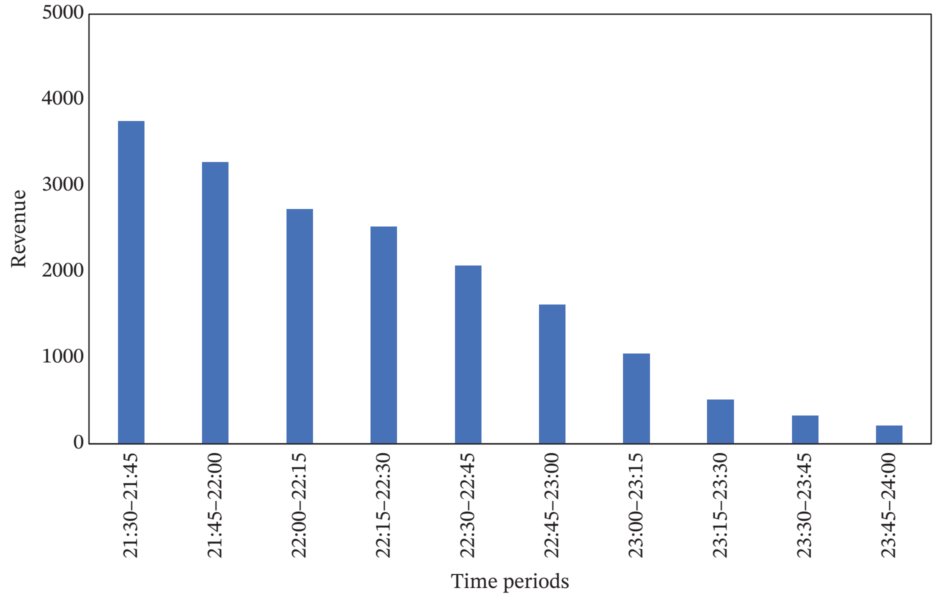

FIGURE 8 | Predicted revenue for the next ten time periods.

with changes in travel demand for passenger cars, trucks, and other vehicle types. This indicates that the revenue predictions are reasonably accurate, reflecting changes in traffic conditions, and can provide important insights for highway management authorities.

6 | Conclusion

Accurate revenue forecasting is crucial for the management and development of toll roads, as it provides a scientific basis for the approval of transportation projects and helps optimize toll rate adjustments. This, in turn, helps to mitigate financial risks arising from external factors and ensures the long-term sustainability of highways. This paper constructs prediction models, including ARIMA, MLR, HA, SVR, MLP, GRU, and LSTM, combining various data processing and learning mechanisms to better capture the temporal features and complex patterns of traffic flow. By comparing the predictive accuracy of different models, this study aims to improve the accuracy of OD flow predictions, providing a more precise foundation for traffic flow analysis and highway revenue forecasting.

Beyond prediction accuracy, the novelty of this research lies in tailoring deep learning techniques to the operational context of toll systems: integrating OD flows with vehicle class aggregation, translating OD predictions into revenue outcomes, and analyzing feature contributions via SHAP. It should be noted that the SHAP analysis in this study was limited to traffic volume features. This restriction arises because the available dataset covered only a single day of toll transactions, during which contextual features such as workday/holiday flags did not vary, and external information such as road incidents was unavailable. Consequently, incorporating these features into the SHAP framework would not have provided meaningful interpretability. While this represents a limitation of the current analysis, we emphasize that the proposed approach is fully compatible with multiday and multisource datasets. These domain-oriented adaptations demonstrate that deep learning models, though not new in themselves, can be made innovative when aligned with practical challenges in toll collection and highway management. Nevertheless, some limitations remain. The absence of external feature data (e.g., accidents, weather, or control strategies) constrains the models' ability to reflect sudden fluctuations, and spatial correlations between different road segments are not explicitly considered. Future research could integrate multisource traffic data and spatial dependencies to further enhance forecasting performance.

Looking forward, future work will explore the integration of spatial interaction features and multisource traffic data to further enhance OD estimation and revenue forecasting. This will deepen the operational relevance of data-driven methods and expand their applicability to broader intelligent transportation systems. For future directions of this prediction task, the following areas can be explored:

1. Deep Feature Extraction: The final accuracy of the prediction model is not only related to the model structure but also closely tied to the input data. Future research can attempt to further extract information from toll station data or combine multisource data, such as gantry data collected on highways and accident data, along with traffic engineering theory, to provide more prior knowledge for the model.
2. Incorporating Spatial Modules: Solely considering time-series data often neglects spatial information, while traffic flow data have clear spatial characteristics. Future models can integrate spatial modules to better capture the interactions between different OD points and enhance the model's understanding of traffic flow features.
3. Introducing More Diverse Data Sources: The integration of multisource data can enhance the adaptability of prediction models and facilitate the establishment of a more comprehensive revenue prediction evaluation system, which enables more effective management of complex traffic environments and fluctuating economic conditions.

Author Contributions

Jian Wan: conceptualization, methodology development, investigation design and execution, validation, project administration, and original draft writing. Yinghao Chen: supervision (corresponding author), formal analysis, validation, visualization, and manuscript review and editing. Mengyu Jiang: investigation execution, formal analysis, and manuscript editing and revision assistance. Hua Tong: formal analysis, visualization, manuscript review and editing, and language refinement. Runsheng Wang: formal analysis, manuscript review and editing, and language editing and polishing.

Acknowledgments

This study was supported by the Natural Science Foundation of Jiangsu Province (No. BK20232019) and the Jiangsu Provincial Scientific Research Center of Applied Mathematics (No. BK20233002).

The authors also acknowledge Yunhai Liu (removed from the author list) for his assistance in data collation, auxiliary investigation, and manuscript revision. His contributions did not meet the criteria for authorship, and he has agreed to be acknowledged accordingly.

Declaration of Generative AI. During the writing process of this work, the authors did not employ generative AI technologies.

Funding

This study was supported by the Natural Science Foundation of Jiangsu Province, No. BK20232019 and Jiangsu Provincial Scientific Research Center of Applied Mathematics, No. BK20233002.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The traffic data used in this study were obtained from the Zhejiang Expressway system through authorized access for research purposes. Permission was required and obtained prior to data usage. Due to data confidentiality and management regulations, the data are not publicly available. However, some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. X. Cai, Y. Li, and P. Li, *Research on Forecasting Methods of Highway Toll Revenues* (Road and Motor Transportation, 2018).

2. Y. Pan, "Research and Application of Higway Toll Prediction," (Xi'an, China: Changan University, 2015), https://d.wanfangdata.com.cn/thesis/D748673.

3. Q. Zhou, "Research on Risk Evaluation of Highway PPP Project Investors," *Highway and Motor Transportation* 02 (2016): 223–226, https://link.cnki.net/urlid/43.1362.u.20160401.1646.122.

4. S. Wang, J. Yang, F. Song, Y. Song, G. Zhao, and S. Li, "Provincial Highway Toll Forecasting Methods and Models," *Chinese and Foreign Highways* 37, no. 03 (2017): 290–294, https://doi.org/10.14048/j.issn.1671-2579.2017.03.066.

5. T. Andrysiak, Ł. Saganowski, M. Choraś, and R. Kozik, *Network Traffic Prediction and Anomaly Detection Based on ARFIMA Model* (Springer, 2014).

6. W.-C. Hong, "Traffic Flow Forecasting by Seasonal SVR With Chaotic Simulated Annealing Algorithm," *Neurocomputing* 74, no. 12 (2011): 2096–2107, https://doi.org/10.1016/j.neucom.2010.12.032.

7. R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction," in *Paper Presented at the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)* (Wuhan, China: IEEE, 2016).

8. S. V. Kumar and L. Vanajakshi, "Short-Term Traffic Flow Prediction Using Seasonal ARIMA Model with Limited Input Data," *European Transport Research Review* 7, no. 3 (2015): 1–9, https://doi.org/10.1007/s12544-015-0170-8.

9. J. Wang, W. Deng, and Y. Guo, "New Bayesian Combination Method for Short-Term Traffic Flow Forecasting," *Transportation Research Part C: Emerging Technologies* 43 (2014): 79–94, https://doi.org/10.1016/j.trc.2014.02.005.

10. S. Sun, C. Li, A. H. Paterson, et al., "Exploration of Highway Toll Revenue Forecasting Method Based on Moving Average and GM (1, 1) Gray Model," *Statistics and Management* 9, no. 12 (2018): 16–21, https://doi.org/10.16722/j.issn.1674-537x.2018.12.005.

11. W. Qian, B. Ge, L. Liu, and Z. Huang, "Application of Highway's Toll Prediction Based on GEP Differential Equation Model," *Journal of Wuhan University of Technology* 40, no. 08 (2018): 55–62+69, https://doi.org/10.3963/j.issn.1671-4431.2018.08.010.

12. N. Liu, Z. Huang, and Q. Tan, "Research on Freeway Toll Prediction Method Based on GEP," *Application Research of Computers* 36, no. 07 (2019): 1998–2002, https://doi.org/10.19734/j.issn.1001-3695.2018.01.0023.

13. Y. Shi, T. Wu, T. Guo, et al., "Traffic Simulation Optimization Considering Driving Styles," *Communications in Transportation Research* 5 (2025): 100181, https://doi.org/10.1016/j.commtr.2025.100181.

14. Z. Liu, C. Lyu, Z. Wang, S. Wang, P. Liu, and Q. Meng, "A Gaussian-Process-Based Data-Driven Traffic Flow Model and its Application in Road Capacity Analysis," *IEEE Transactions on Intelligent Transportation Systems* 24 (2023): 1544–1563, https://doi.org/10.1109/TITS.2022.3223982.

15. J. Ren and Q. Xie, "Efficient OD Trip Matrix Prediction Based on Tensor Decomposition," in *Paper Presented at the 2017 18th IEEE International Conference on Mobile Data Management (MDM)* (Daejeon, South Korea: IEEE, 2017).

16. Z. Ma and P. Zhang, "Individual Mobility Prediction Review: Data, Problem, Method and Application," *Multimodal Transportation* 1, no. 1 (2022): 100002, https://doi.org/10.1016/j.multra.2022.100002.

17. X. Yin, G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin, "Deep Learning on Traffic Prediction: Methods, Analysis, and Future Directions," *IEEE Transactions on Intelligent Transportation Systems* 23, no. 6 (2022): 4927–4943, https://doi.org/10.1109/TITS.2021.3054840.

18. N. Kalchbrenner and P. Blunsom, "Recurrent Continuous Translation Models," in *Paper Presented at the Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (Seattle, WA: Association for Computational Linguistics, 2013).

19. A. Nigam and S. Srivastava, "Hybrid Deep Learning Models for Traffic Stream Variables Prediction During Rainfall," *Multimodal Transportation* 2, no. 1 (2023): 100052, https://doi.org/10.1016/j.multra.2022.100052.

20. J. L. Elman, "Distributed Representations, Simple Recurrent Networks, and Grammatical Structure," *Machine Learning* 7, no. 2-3 (1991): 195–225, https://doi.org/10.1023/a:1022699029236.

21. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Representations by Back-Propagating Errors," *Nature* 323, no. 6088 (1986): 533–536, https://doi.org/10.1038/323533a0.

22. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation MIT-Press* 9, no. 8 (1997): 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.

23. K. Cho, B. Van Merrienboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches," *Computation and Language* 1 (2014): 103–111, https://doi.org/10.3115/v1/w14-4012.

24. T. Wang, A. Hussain, Q. Sun, S. E. Li, and C. Jiahua, "The Prediction of Urban Road Traffic Congestion by Using a Deep Stacked Long Short-Term Memory Network," *IEEE Intelligent Transportation Systems Magazine* 14, no. 4 (2021): 102–120, https://doi.org/10.1109/mits.2021.3049383.

25. Z. Liu, Z. Zhou, Z. Gu, et al., "TRIP: Transport Reasoning With Intelligence Progression—A Foundation Framework," *Transportation Research Part C: Emerging Technologies* 179 (2025): 105260, https://doi.org/10.1016/j.trc.2025.105260.

26. Y. Zhang, S. Xu, L. Zhang, W. Jiang, S. Alam, and D. Xue, "Short-Term Multi-Step-Ahead Sector-Based Traffic Flow Prediction Based on the Attention-Enhanced Graph Convolutional LSTM Network (AGC-LSTM)," *Neural Computing and Applications* 37, no. 20 (2025): 14869–14888, https://doi.org/10.1007/s00521-024-09827-3.

27. R. Wang, Y. Cao, X. Ji, and D. Qiao, "A Prediction Method for Highway Traffic Flow Based on the IHPO-VMD-LSTM-Informer Model," *Information Technology and Control* 54, no. 2 (2025): 380–395, https://doi.org/10.5755/j01.itc.54.2.39228.

28. Z. Yang, J. Yang, Y. Zhou, Q. Xu, and M. Ou, "Multimodal Trajectory Prediction for Intelligent Connected Vehicles in Complex Road Scenarios Based on Causal Reasoning and Driving Cognition Characteristics," *Scientific Reports* 15, no. 1 (2025): 7259, https://doi.org/10.1038/s41598-025-91818-y.

29. S. Hong, T. Yue, Y. You, et al., "A Resilience Recovery Method for Complex Traffic Network Security Based on Trend Forecasting," *International Journal of Intelligent Systems* 2025, no. 1 (2025): 3715086, https://doi.org/10.1155/int/3715086.

30. A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive Into Deep Learning* (Cambridge University Press, 2021), arXiv preprint arXiv:2106.11342.