*sustainability*

MDPI

*Article*

# A Model-Data Dual-Driven Approach for Predicting Shared Bike Flow near Metro Stations

**Zhuorui Wang [1,*], Dexin Yu [2,*], Xiaoyu Zheng [3], Fanyun Meng [1] and Xincheng Wu [2,4]**

[1] School of Transportation, Jilin University, Changchun 130022, China; mfy20@mails.jlu.edu.cn
[2] Navigation College, Jimei University, Xiamen 361021, China; wuxincheng@xmoc.edu.cn
[3] BIT—Barcelona Innovative Transportation Research, Civil Engineering School, UPC–Barcelona Tech, 08034 Barcelona, Spain; xiaoyu.zheng@upc.edu
[4] Navigation College, Xiamen Ocean Vocational College, Xiamen 361012, China
* Correspondence: wangzr17@mails.jlu.edu.cn (Z.W.); yudx@jmu.edu.cn (D.Y.)

**Abstract:** Bike-sharing has emerged as an innovative green transportation mode, showing promising potential in addressing the 'last-mile' transportation challenge in an eco-friendly manner. However, shared bikes around metro stations often face supply–demand imbalance problems during peak hours, causing bike shortages or congestion that compromise user experience and bike utilization. Accurate prediction enables operators to develop rational dispatch strategies, improve bike turnover rate, and promote synergistic metro–bike integration. However, state-of-the-art research predominantly focuses on improving complex deep-learning models while overlooking their inherent drawbacks, such as overfitting and poor interpretability. This study proposes a model–data dual-driven approach that integrates the classical statistical regression model as a model-driven component and the advanced deep-learning model as a data-driven component. The model-driven component uses the Seasonal Autoregressive Integrated Moving Average (SARIMA) model to extract periodic patterns and seasonal variations of historical data, while the data-driven component employs an Extended Long Short-Term Memory (xLSTM) neural network to process nonlinear relationships and unexpected variations. The fusion model achieved R-squared values of 0.9928 and 0.9770 for morning access and evening egress flows, respectively, and reached 0.9535 and 0.9560 for morning egress and evening access flows. The xLSTM model demonstrates an 8% improvement in $R^2$ compared to the conventional LSTM model in the morning egress flow scenario. For the morning egress and evening access flows, which exhibit relatively high variability, classical statistical models show limited effectiveness (SARIMA's $R^2$ values are 0.8847 and 0.9333, respectively). Even in scenarios like morning access and evening egress, where classical statistical models perform well, our proposed fusion model still demonstrates enhanced performance. Therefore, the proposed data–model dual-driven architecture provides a reliable data foundation for shared bike rebalancing and shows potential for addressing the challenges of limited robustness in statistical regression models and the susceptibility of deep-learning models to overfitting, ultimately enhancing transportation ecosystem sustainability.

**Keywords:** shared bike; demand prediction; sustainable transportation; model–data dual-driven

check for updates

Received: 9 December 2024
Revised: 22 January 2025
Accepted: 22 January 2025
Published: 27 January 2025

**Citation:** Wang, Z.; Yu, D.; Zheng, X.; Meng, F.; Wu, X. A Model-Data Dual-Driven Approach for Predicting Shared Bike Flow near Metro Stations. *Sustainability* **2025**, *17*, 1032. https://doi.org/10.3390/su17031032

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

The rapid proliferation of shared bike systems has fundamentally transformed urban mobility, offering a sustainable and flexible alternative to conventional transportation modes [1,2]. As a promising solution to the 'last mile' connectivity challenge, bike-sharing

services have emerged as a pivotal mechanism for enhancing urban transport efficiency and accessibility [3,4]. In China, the shared bike market demonstrated significant growth dynamics, with a notable compound annual growth rate (CAGR) of 10% between 2017 and 2022. During this period, the market size expanded substantially from CNY 13.03 billion (USD 1.78 billion) to CNY 30.4 billion (USD 4.15 billion), reflecting robust market development. Concurrently, the user base experienced remarkable growth, increasing from 310 million to 460 million users, representing a CAGR of 8.21% [5]. It is noteworthy that the intermodal integration has reached approximately 54% of shared bike users leveraging these services to facilitate multimodal transportation connections, with an overwhelming 91% of such trips serving as links to public transit infrastructure [6]. Despite technological advancements in shared bicycle systems, complex challenges persist. In terms of road safety, cyclists are considered vulnerable road users, prone to being involved in accidents. Therefore, it is essential to design dedicated road facilities for these users to ensure their safety [7]. Regarding supply–demand imbalance issues, spatial-temporal imbalances in shared bike distribution are common, especially during peak hours, particularly in areas surrounding metro stations. During these rush hours, shared bike systems frequently experience demand–supply mismatches, manifesting as either bike shortages or excessive bicycle accumulations. These spatiotemporal mismatches in bike distribution disrupt the user experience and lead to suboptimal resource utilization. Consequently, accurate demand prediction for shared bikes near metro stations is pivotal for optimizing deployment and scheduling approaches. Moreover, precise and robust demand forecasting results empower operators to develop more sophisticated rebalancing strategies, enhancing bike redistribution and system operational efficiency [8]. Furthermore, predictive scheduling can effectively mitigate common operational challenges such as insufficient bike resources and parking congestion, thereby improving service levels and promoting seamless integration between shared bike and rail transit systems [9–12].

Although classical model-driven prediction methods offer good interpretability, they demonstrate poor robustness when handling complex data [13]. For instance, ARIMA-based models are sensitive to anomalous fluctuations, while probability-based prediction methods may be constrained by specific datasets [14,15]. In contrast, data-driven prediction methods, such as deep learning approaches, exhibit greater adaptability and can effectively handle nonlinear characteristics of data. However, due to their black-box nature, these methods suffer from limited interpretability. Moreover, they are prone to overfitting when dealing with small datasets or multiple input features [16]. In response to the distinctive characteristics of bike flow during peak hours, this study proposes a novel data–model dual-driven approach that integrates SARIMA and extended Long Short-Term Memory (xLSTM) neural network aiming to achieve more precise prediction of access and egress shared-bike trips around metro stations during weekday peak hours. The proposed fusion model comprises two primary components. The model-based component (SARIMA); a classical approach that captures the periodic patterns and seasonal variations of bike flow during peak hours through a SARIMA modeling technique. The Data-driven component (xLSTM); an advanced neural network architecture designed to extract complex, non-linear features from historical peak-hour bike flow and relevant external features, including weather and calendar factors. The fusion mechanism dynamically adjusts the weights of two components based on the distinctive characteristics of different peak flow scenarios (morning or evening, access or egress). By combining the strengths of both modeling approaches, the proposed method seeks to enhance prediction precision and model robustness.

The contribution of this paper is twofold. First, from a theoretical perspective, this study advances bike-sharing demand prediction research by challenging the existing

paradigm of complex deep learning models. The research breaks through existing methodological limitations by proposing an innovative 'data–model dual-driven' fusion model framework. By designing a mechanism that adaptively adjusts the weights of classical and neural network components across different peak flow scenarios (morning/evening, access/egress), this study offers a robust solution that simultaneously captures periodic patterns and complex non-linear features. Second, in terms of methodological and practical contributions, the research develops a sophisticated fusion model that dynamically integrates SARIMA and xLSTM neural networks. By critically addressing the overfitting risks of advanced neural networks in data-limited scenarios, the study emphasizes the continuing importance of classical statistical methods and demonstrates their potential to provide nuanced insights. By bridging the theoretical gap between classical statistical modeling and data-driven methodologies, our study introduces a novel theoretical approach that balances model complexity with generalization capabilities.

The remainder of this paper is organized as follows: Section 2 provides a detailed description of the proposed fusion model architecture. Section 3 introduces the study area, describes the dataset, presents empirical results, and reports ablation experiments. Finally, Section 4 summarizes the key findings and discusses potential directions for future research.

## 2. Literature Review

Due to the reasons outlined above, shared bike demand prediction has been extensively investigated, representing a critical research domain for operators and policymakers to optimize service efficiency and enable timely rebalancing strategies. We review the existing literature from two perspectives: model-driven methods and data-driven techniques.

### *2.1. Model-Driven Methods*

Regarding the model-driven approaches, the main approaches primarily encompass ARIMA-based models, logistic regression, and probability-based models (e.g., Markov-based and Bayesian-based models). For instance, Kaltenbrunner et al. [17] applied the ARIMA model to analyze and predict bike usage patterns in the Barcelona Bicing community bike-sharing program. Yoon et al. [18] further refined the ARIMA model, focusing on station-level shared bike demand forecasting by effectively addressing the non-stationarity of time series data through differentiation and seasonal adjustments, and validating the model using Dublin's shared bike system data. While logistic regression has also been employed to explore bike-sharing usage patterns, Caulfield et al. [19], for instance, applied logistic regression to study bike-sharing behavior in Cork, Ireland, incorporating key factors such as travel distance, weather conditions, and temporal variables. Furthermore, probability-based models have gained traction as well, Zhou et al. [20] proposed a Markov chain-based demand prediction model, claiming enhanced forecasting accuracy and generalization capabilities, which was validated using shared bike data from Zhongshan City, China. Nistor and Dias [21] developed a Markovian model to forecast bike distributions considering weather and event impacts, applied to New York City data. Bayesian approaches have also proven effective in this domain, Cagliero et al. [22] utilized Bayesian and associative classifiers to determine bike-sharing stations as critical or non-critical in the near future, aiding in bike redistribution and improving user experience. Ma et al. [23] proposed a demand forecasting model based on Bayesian theory and validated it using bike-sharing data from Xiamen, China. Their approach integrates geographic network encoding, historical demand data, and ARIMA for time series prediction to estimate bike demand for specific periods. Expanding the scope of Bayesian methods, Duan et al. [24] employed a Bayesian spatiotemporal model to analyze shared bike demand, effectively quantifying the influence of meteorological conditions, population density, and

per capita GDP, thus enhancing the accuracy of spatiotemporal bike demand predictions. Further advancing this line of research, Ramkumar and Nannapaneni [15] introduced a Quantum Bayesian Network (QBN) for bike-sharing demand forecasting. By leveraging quantum computing principles, their model improves upon traditional Bayesian Networks by capturing complex dependencies and uncertainties, offering more precise predictions.

### *2.2. Data-Driven Techniques*

In the context of data-driven techniques, the rapid development of machine learning technologies has made research in shared bike demand prediction increasingly popular. Researchers have extensively applied state-of-the-art machine learning algorithms, especially deep neural networks, to explore the potential features and patterns within datasets. Multiple studies in existing literature have investigated various machine learning methods' predictive performance. For instance, Gu and Lin [25] compared multiple linear regression and random forest models in London's shared bike demand prediction, revealing the random forest model's enhanced performance over conventional linear regression approaches. Sathishkumar et al. [26] further explored the Gradient Boosting Machine (GBM) in the prediction of hourly demand for shared bicycles in Seoul, identifying key variables influencing bike demand and verifying that ensemble learning methods can provide more accurate predictions.

More recently, deep learning has garnered researchers' attention due to its powerful capabilities in handling complex non-linear relationships and capturing intricate spatial-temporal patterns. Chen et al. [27] developed a real-time bike rental and return demand prediction system using recurrent neural networks (RNN), comparing five RNN architectures on the New York Citi Bike dataset and demonstrating satisfactory results at both global and station levels. Wang et al. [28] proposed a three-dimensional residual neural networks (3D ResNet) model for short-term traffic flow prediction by extracting spatiotemporal features from shared bike networks, validating the approach using datasets from New York and Suzhou. Zhai et al. [29] introduced a self-supervision spatiotemporal part-whole convolutional neural network (STPWNet) for traffic prediction, demonstrating its effectiveness in handling non-stationary time series data using Beijing taxi and New York City shared bike datasets. Pan et al. [30] developed a deep LSTM model for real-time prediction of bike rental and return demands across different urban areas, utilizing historical data and meteorological information from the Citi Bike System Data to effectively address bike-sharing demand prediction and provide insights for resource allocation. Sardinha et al. [31] proposed a stacked RNN model integrating spatial, meteorological, and contextual information to predict shared bike station demands, highlighting the importance of diverse contextual factors using Lisbon's GIRA bike-sharing system data. Collini et al. [32] developed a bi-directional LSTM model to predict available bikes at shared bike stations, validating the model with data from two Italian cities, Siena and Pisa. Jin et al. [33] proposed a temporal convolutional network (TCN) model for predicting dynamic travel demand in dockless bike-sharing systems across Traffic Analysis Zones, using shared bicycle data from Shanghai to confirm TCN's advantages in demand prediction and provide strategic support for system rebalancing. Lin et al. [34] introduced a graph convolutional neural network with a data-driven graph filter (GCNN-DDGF) model for predicting hourly station-level demands in large-scale shared bike networks, utilizing the New York City Citi Bike dataset to demonstrate the model's more precise prediction performance and its potential in capturing hidden heterogeneous correlations between stations. Qin et al. [35] proposed the Residual Graph Convolutional Network (RESGCN) method to address the challenge of accurately predicting available dock numbers in bike-sharing systems, effectively capturing complex spatiotemporal correlations and external factors like weather by utilizing historical ride

data from Boston's Hubway system and demonstrating more accurate prediction results compared to state-of-the-art approaches.

Building upon existing deep learning approaches, some researchers further explored attention mechanisms to enhance predictive performance. He and Shin [36] introduced a graph attention convolutional neural network (GACNN) for fine-grained flow forecasting in bike-sharing systems, employing attention mechanisms to differentiate station-to-station correlations and capture nuanced spatial relationships in the GBikes dataset. Lee and Ku [37] proposed a dual attention-based recurrent neural network (RNN) model for short-term bike sharing usage demand prediction using YouBike data from Taipei, demonstrating improved predictive performance compared to existing methods and offering insights for optimizing bicycle allocation to address bike shortage issues. Zi et al. [38] developed a novel deep graph convolutional network with temporal attention (TAGCN) for station-level demand prediction, consistently outperforming state-of-the-art methods across four Divvy Bike System datasets in Chicago by incorporating temporal attention mechanisms that focus on critical time periods and effectively capture the spatial and temporal dependencies for accurate bike-sharing demand forecasting.

Furthermore, several studies have delved into the application of hybrid models to enhance prediction performance. Mehdizadeh Dastjerdi and Morency [39] proposed a hybrid CNN-LSTM model to address bike-sharing demand prediction at the community level during the COVID-19 pandemic. By incorporating additional input features, the model effectively captured complex and dynamic demand patterns, demonstrating the adaptability of deep learning models in uncertain environments. Zhou et al. [40] developed a short-term prediction model combining TCN and GRU to tackle the bike-sharing demand forecasting task. Through multidimensional analysis of travel characteristics, such as weather, temperature, and humidity, using bike-sharing data from London, the model was validated for its efficiency and robustness.

To summarize, while the aforementioned model-based methods have shown promising predictive performance and offer better interpretability compared to data-driven machine learning approaches, they face challenges in handling complex, non-linear relationships and exhibit limited robustness to anomalous fluctuations. Likewise, probability-based models stand out for their simplicity and accuracy but are restricted by their limited applicability to specific datasets. In terms of data-driven methods, most research endeavors have focused on integrating diverse external features, including weather conditions, calendar information, spatial characteristics, and multimodal transportation data (such as public transit), to predict bike-sharing demand more precisely, with a primary emphasis on model accuracy improvement. However, a critical aspect often overlooked is the inherent overfitting risk of complex deep learning models when confronted with limited data. Notably, shared bike systems are most prone to supply–demand imbalances during peak hours, where bike flow remains relatively stable with minimal external influence. Consequently, classical model-driven approaches remain meaningful, effectively capturing peak-hour bike flow characteristics and intrinsic regularity. While conventional model-driven methods exhibit limitations in handling anomalous fluctuations, machine learning techniques, particularly deep learning, excel at processing nonlinear relationships and addressing unexpected variations. Therefore, the integration of machine learning techniques with classical model-based methods could provide more robust and adaptable demand prediction results. Addressing these potential improvements could lead to more efficient bike-sharing systems, ultimately enhancing urban mobility and sustainability.

## 3. Methodology

This paper proposed a data–model dual-driven fusion model framework for predicting bike-sharing flow around metro stations during weekday peak hours. The framework integrates a SARIMA model and an extended LSTM (xLSTM) network, with dynamically adjusted weights based on different flow patterns (morning/evening peaks, access/egress trips). The motivation behind this hybrid approach stems from the complementary strengths of both kinds of models: for the model-driven part, SARIMA demonstrates strong effectiveness in capturing regular temporal dependencies and seasonal patterns in commuting bike flows, while for the data-driven part, the xLSTM component exhibits enhanced capability in modeling complex non-linear relationships and incorporating external factors. Through the dynamic weight adjustment scheme, the fusion model optimizes prediction accuracy by leveraging the statistical robustness of the SARIMA model and the adaptive learning ability of the xLSTM model under different peak-flow scenarios, thereby providing more robust and reliable support for bike-sharing system management during peak hours. The following part will detail the model architecture, and the method for determining dynamic weights, as well as proposing evaluation metrics.

### *3.1. SARIMA-xLSTM Fusion Model Architecture*

The proposed prediction framework consists of three major components: (1) a SARIMA model that captures linear temporal dependencies and seasonal patterns in historical peak-time commuting bike flow data, (2) an xLSTM model that learns complex non-linear relationships from historical flow patterns, and incorporates external features such as weather and calendar features, and (3) a weighted combination mechanism that dynamically adjusts the weights for different peak-flow types, integrating predictions from both components to minimize error and achieve optimal forecasting performance. The overall architecture of the proposed fusion model is illustrated in Figure 1.

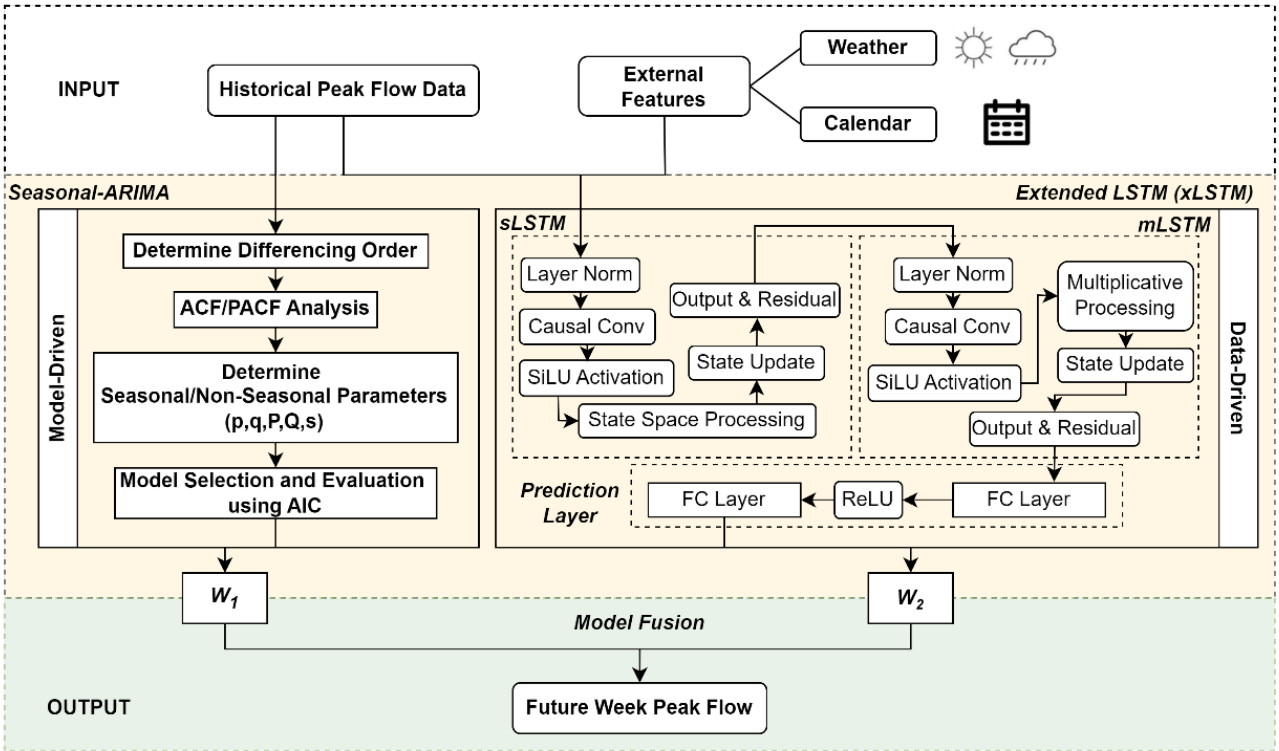


**Figure 1.** The architecture of the SARIMA-xLSTM fusion model.

This study aims to predict bike-sharing trips around metro stations during peak hours. Since different shared bike trips are driven by various travel activity purposes, we categorize peak-hour bike flows into four types for separate prediction: morning peak access and egress, and evening peak access and egress, to capture the characteristics of peak flow precisely. The historical bike-sharing flow dataset is partitioned chronologically into three subsets: the training set (first 9 weeks), the validation set (week 10) for weight optimization, and the testing set (week 11) for final fusion model evaluation. This structured approach ensures robust training and thorough evaluation of the fusion model's predictive effectiveness.

The fusion model consists of both model-driven and data-driven components. The SARIMA model serves as the model-driven component, providing regression-based predictions of the abovementioned four types of shared-bike peak flow. Peak-hour passenger flows are predominantly commuting-purpose trips, exhibiting strong periodicity and regularity, which aligns well with SARIMA's pattern-capturing strengths. The data-driven xLSTM component extracts features from the historical ridership data and leverages external factors for enhanced forecasting accuracy. Compared to classical ARIMA-based regression models, the xLSTM model demonstrates enhanced capability in capturing nonlinear relationships within the data.

The integration of these two components aims to combine their respective advantages, allowing the fusion model to dynamically adjust the weights of both components based on the data characteristics of different time periods and flow directions (access or egress), thereby achieving more accurate predictions and enhanced robustness. The following sections will detail the specific structures of both SARIMA and xLSTM components.

### *3.2. SARIMA Component*

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is employed to capture both the temporal dependencies and seasonal patterns in peak-time shared-bike flows around metro stations. The SARIMA (p, d, q) (P, D, Q) s model can be expressed as:

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D Y_t = c + \theta(B)\Theta(B^S)\varepsilon_t \quad (1)$$

where $Y_t$ is the time series, here is the historical peak flow around each metro station, B is the backshift operator, S represents the seasonal period (s = 5 for weekday patterns). (p, d, q) are the non-seasonal orders for autoregressive (AR), differencing, and moving average (MA) components. (P, D, Q) are the seasonal orders for AR, differencing, and MA components.

$$\Phi(B^S) = 1 - \beta_1 B^S - \beta_2 (B^S)^2 - \cdots - \beta_P (B^S)^P \quad (2)$$

$$\Theta(B^S) = 1 - \alpha_1 B^S - \alpha_2 (B^S)^2 - \cdots - \alpha_Q (B^S)^Q \quad (3)$$

where $\Phi(B^S)$ is the seasonal AR operator, $\Theta(B^S)$ is the seasonal MA operator, $\varepsilon_t$ represents white noise.

Initially, the time series data of morning and evening peak bicycle flows are tested for stationarity using the Augmented Dickey-Fuller (ADF) test. If non-stationarity is detected, appropriate differencing (d, D) is applied to achieve stationarity. Subsequently, the optimal orders (p, d, q) (P, D, Q) s are determined by minimizing the Akaike Information Criterion (AIC):

$$AIC = -2\ln(L) + 2k \quad (4)$$

where L is the likelihood of the model and k is the total number of parameters. A grid search is performed over possible combinations of orders ($p \leq 3$, $d \leq 2$, $q \leq 3$, $P \leq 2$, $D \leq 1$, $Q \leq 2$) to identify the parameter set that yields the minimum AIC value.

Finally, numerical forecasting is made by the eligible SARIMA (p, d, q) (P, D, Q) s.

### *3.3. Extended LSTM (xLSTM) Component*

For the data-driven component, we leverage the Extended LSTM (xLSTM) architecture for our task. Compared to conventional LSTM, xLSTM introduces exponential gating and proposes two novel memory units: Scalar LSTM (sLSTM) and Matrix LSTM (mLSTM). Scalar LSTM improves the flexibility of the model by enhancing gating mechanisms and memory blending, while Matrix LSTM increases the memory capacity and parallel processing capability of the model by extending the memory units to matrices [41]. Unlike the original xLSTM architecture that employs multiple alternating layers of sLSTM and mLSTM. In this study, we adopt a simplified sequential structure with one sLSTM layer followed by one mLSTM layer to balance model complexity and effectiveness, considering our dataset size and overfitting prevention. In our implementation, input features are first processed through the sLSTM layer, and its outputs are then fed into the mLSTM layer, followed by a fully connected layer for final predictions.

The sLSTM block incorporates layer normalization and causal convolution for preprocessing, followed by state-space processing that includes state transformation and gating mechanisms. Historical information is processed through recurrent connections. The mLSTM block shares a similar preprocessing pipeline but utilizes a multiplicative interaction mechanism that enhances feature extraction through multiplicative gating, processing the intermediate features from sLSTM.

The implemented xLSTM architecture provides several advantages for our application. The causal convolution enables effective temporal modeling capabilities. The multiplicative mechanisms in mLSTM facilitate feature interactions, while the state-space characteristics of sLSTM support long-term dependency modeling. Moreover, the residual connections and normalization techniques contribute to training stability.

#### 3.3.1. Causal Convolution Layer

Both the sLSTM block and mLSTM block begin with a causal convolution layer, which captures local temporal dependencies and ensures no information leakage from future time steps:

$$\tilde{h}_t = \text{Conv1D}(x_t) \in \mathbb{R}^d \tag{5}$$

where Conv1D represents a one-dimensional causal convolution with a kernel size of d, and appropriate padding is used to maintain causality. $\tilde{h}_t$ refers to the feature output obtained through one-dimensional convolution. $x_t$ represents the feature vector of the input sequence at time step t. $\mathbb{R}^d$ denotes a d-dimensional vector.

#### 3.3.2. Scalar LSTM (sLSTM) Block

The sLSTM block incorporates explicit state-space modeling through recurrent connections. The block updates are defined as follows:

Intermediate state update: computes the new information added to the cell state.

$$z = \tanh(W_z x_t + R_z h_{t-1}) \tag{6}$$

Input gate: controls how much of the new input $x_t$ is added to the cell state.

$$i = \sigma(W_i x_t + R_i h_{t-1}) \tag{7}$$

Forget gate: controls how much of the previous cell state $c_{t-1}$ is retained.

$$f = \sigma(W_f x_t + R_f h_{t-1}) \tag{8}$$

Output gate: controls how much information from the current cell state $c_t$ is passed to the hidden state $h_t$.

$$o = \sigma(W_o x_t + R_o h_{t-1}) \tag{9}$$

where W$*$ and R$*$ are learnable parameters, σ denotes the sigmoid activation function, and $h_{t-1}$ is the hidden state from the previous time step. The cell state and hidden state are updated as:

Cell state update: the forget gate decides how much past information is retained, and the input gate decides how much new information is added.

$$c_t = f \odot c_{t-1} + i \odot z \tag{10}$$

Hidden state update: combines the output gate with the cell state to determine the hidden state.

$$h_t = o \odot \tanh(c_t) \tag{11}$$

where $\odot$ denotes element-wise multiplication.

3.3.3. Matrix LSTM (mLSTM) Block

The mLSTM block enhances feature interactions through multiplicative mechanisms. The key components are computed as follows:

Query vector: maps the input to a query vector q.

$$q = W_q x_t \tag{12}$$

Key vector: maps the input to a key vector k.

$$k = \frac{W_k x_t}{\sqrt{d}} \tag{13}$$

Value vector: maps the input to a value vector v.

$$v = W_v x_t \tag{14}$$

where d is the hidden dimension to ensure numerical stability. The state updates incorporate multiplicative interactions:

$$\text{Input gate}: \ i = \sigma(W_i x_t) \tag{15}$$

$$\text{Forget gate}: \ f = \sigma(W_f x_t) \tag{16}$$

$$\text{Output gate}: \ o = \sigma(W_o x_t) \tag{17}$$

Cell state update: incorporates key-value interactions $v \odot k$ into cell state update.

$$c_t = f \odot c_{t-1} + i \odot (v \odot k) \tag{18}$$

Hidden state update: the output state considers both the cell state $c_t$ and the query vector q.

$$h_t = o \odot \tanh(c_t) \odot q \tag{19}$$

3.3.4. Hierarchical Output Layer

The final prediction is generated through a two-layer feedforward neural network:

$$\hat{y}_t = W_2(\mathrm{ReLU}(W_1 h_t + b_1)) + b_2 \tag{20}$$

where $W_1 \in \mathbb{R}^{d \times \frac{d}{2}}$ and $W_2 \in \mathbb{R}^{\frac{d}{2} \times 1}$ are the weight matrices responsible for linear transformations in the first and second layers, respectively. $b_1$ and $b_2$ are learnable biases. $W_1$ projects the input vector $h_t$ (of dimension $d$) into a lower-dimensional hidden representation (of dimension $\frac{d}{2}$), while $W_2$ maps the hidden representation to the final output $\hat{y}_t$.

3.3.5. Model Training

The xLSTM model is trained by minimizing the mean squared error (MSE) loss between the predicted and actual values:

$$L = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{21}$$

where N is the number of training samples. We apply layer normalization and residual connections in both sLSTM and mLSTM blocks to facilitate stable training and improve gradient flow.

### *3.4. Weight Optimization of the Fusion Model*

In the training set (first nine weeks), we combined the predictions of the SARIMA model and the xLSTM model for each of the four peak-flow types. The sum of the weights for the two components was constrained to be 1. We tested different weight combinations of each peak flow with a step size of 0.1, selecting the combination that minimized prediction error on the validation set (10th week). The optimal weights were then applied to the last week (11th week) test set for final evaluation.

### *3.5. Evaluation Metrics*

To comprehensively assess the predictive performance of our proposed model, we evaluated its accuracy using four widely recognized statistical metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R-squared). The detailed mathematical formulations for these metrics are presented in Equations (22)–(25). MAE provides a direct and intuitive quantification of the average absolute deviation between predicted and actual values, offering a straightforward interpretation of prediction error magnitude. RMSE, calculated as the square root of the Mean Squared Error (MSE), normalizes the error metric to the original data scale while exhibiting heightened sensitivity to larger prediction discrepancies. MAPE evaluates the average percentage deviation of predictions from actual values, delivering a normalized accuracy measure that enables robust performance comparisons across diverse datasets. R-squared quantifies the proportion of variance in the dependent variable effectively explained by the proposed model, thereby providing a comprehensive assessment of model fit and predictive capability. These metrics collectively offer a comprehensive perspective on the model's predictive accuracy and robustness.

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{22}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{23}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{24}$$

$$R^2 = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \overline{y})^2} \tag{25}$$

where $\hat{y}_i$ represents the predicted value of the model, $y_i$ represents the actual value, and n represents the number of samples, $\overline{y}$ is the mean of the actual values.

## 4. Experiment

In this section, we present the empirical study of this paper, focusing on the prediction of peak-period bike-sharing access and egress flows around metro stations in Shenzhen. First, we provide an overview of the dataset and study area. Then, we validate the effectiveness of the proposed fusion model using real-world data from Shenzhen, including a series of ablation experiments to analyze the contributions of different model components. Additionally, we conduct a detailed case study and discussion on several representative metro station areas to further explore the model's performance and insights.

### *4.1. Data and Study Area*

Shenzhen is a major metropolis located in southern China, recognized as a key economic hub and a first-tier city. As one of the core cities in the Greater Bay Area, Shenzhen plays a pivotal role in the region's transportation network, with its extensive metro system handling significant passenger volumes daily. In parallel, the city has seen substantial usage of dockless bike-sharing services, particularly as a means for first-mile and last-mile connectivity to metro stations. For our study, we utilized bike-sharing trip data from the Shenzhen Government Open Data Platform [42], selecting records from June to August 2021 as the shared bike riding order data during this period was relatively stable. The raw dataset included the start and end times of trips, as well as the GPS coordinates of the trip origins and destinations. To ensure data quality and focus on actual commuting behavior, we performed several preprocessing steps. Specifically, we filtered out trips that were less than 100 m or more than 10 km in length, trips with durations shorter than 2 min or longer than 1 h, and trips with average speeds below 5 km/h or exceeding 30 km/h, as these were unlikely to represent typical commuter behavior.

The analysis included all 237 operational metro stations in Shenzhen during the study period. We utilized ArcGIS (version 10.3) to define buffer zones with a 300-m radius around each metro station, as shown in Figure 2. Trips ending within these buffer zones were classified as access trips (indicating an inbound connection to the rail transit system), while trips starting within the buffer zones were classified as egress trips (indicating an outbound connection from the metro). Based on the total number of trips across the city, we applied K-means clustering to identify time periods with distinct travel patterns, defining the morning peak as 7:00–9:59 and the evening peak as 17:00–19:59. Figure 3 illustrates the distribution of the average hourly bike-sharing rides per day across Shenzhen. It is evident that the number of rides during the defined peak hours is noticeably higher than during off-peak hours.

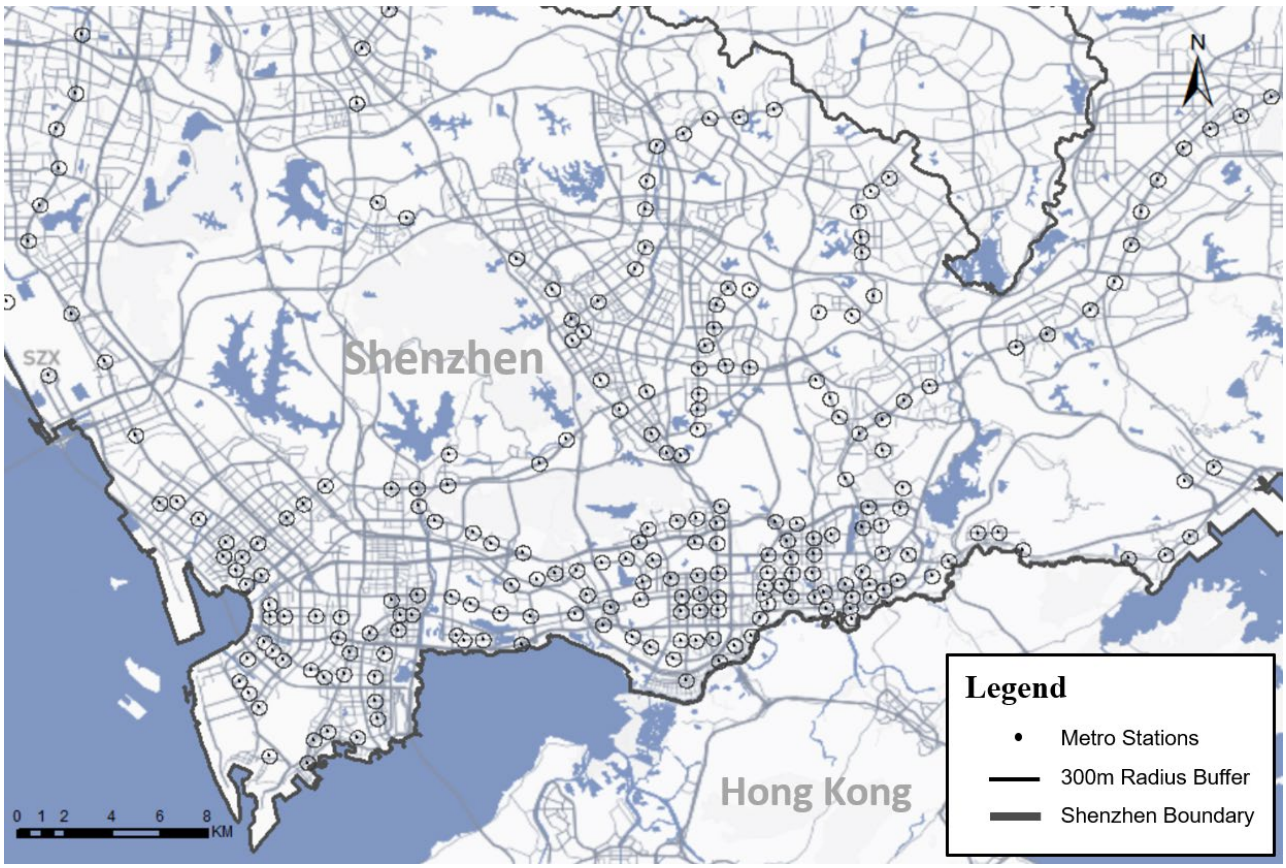


**Figure 2.** Buffer zones of 237 metro stations in Shenzhen.

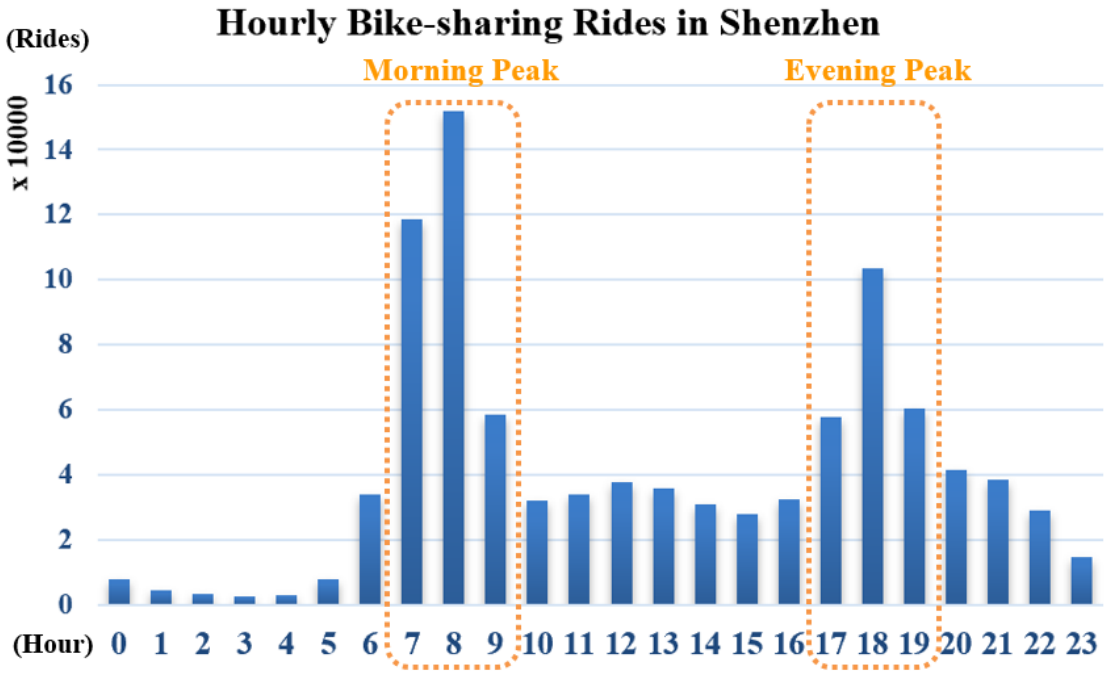


**Figure 3.** Ride times of bike-sharing service per hour in Shenzhen.

We compiled historical morning and evening peak access and egress flows for all 237 metro stations from 7 June to 27 August, focusing on weekdays (Monday to Friday) over 11 weeks. These data served as the input for training, validation, and testing of our fusion model. Additionally, for the xLSTM model, we incorporated external features,

including calendar and weather characteristics. Calendar features represented the day of the week, while weather features were simplified to indicate sunny or rainy conditions, as temperature variations during these three months were relatively stable in Shenzhen.

### *4.2. Empirical Results and Discussions of Fusion Model*

In this section, we introduced our proposed fusion model and validated its performance using peak-time access/egress shared-bike flow data collected from areas surrounding Shenzhen metro stations. We obtained optimal weight combinations and final prediction results, performed ablation studies to validate the model's effectiveness, and conducted detailed multi-dimensional analyses of the experimental outcomes.

#### 4.2.1. Optimal Weight Combinations for Peak Flows

The optimal weight combinations of our fusion model are presented in Table 1. For the SARIMA component, the weights are 0.6 and 0.7 for morning access and evening egress flows, respectively, indicating strong regularity and minimal fluctuations in these patterns. Conversely, the xLSTM component demonstrates higher weights of 0.8 and 0.6 for morning egress and evening access flows, respectively. Notably, the SARIMA component weight for morning egress flow is only 0.2, suggesting significant volatility and susceptibility to external factors. Through weight optimization of both components, we achieved notable prediction performance, as shown in Table 2. The $R^2$ values reached 0.9928 and 0.9770 for morning access and evening egress flows while achieving 0.9535 and 0.9560 for morning egress and evening access flows, respectively.

**Table 1.** The weight combinations of the fusion model for four different peak flow types.

| Peak Flow | SARIMA Weight | xLSTM Weight |
|---|---|---|
| Morning Access | **0.6** | 0.4 |
| Morning Egress | 0.2 | **0.8** |
| Evening Access | 0.4 | **0.6** |
| Evening Egress | **0.7** | 0.3 |

**Table 2.** Performance metrics (mean error and $R^2$) of the fusion model for four peak flow types.

| Peak Flow | MAE | RMSE | MAPE | $R^2$ |
|---|---|---|---|---|
| Morning Access | 33.12 | 37.09 | 4.79 | 0.9928 |
| Morning Egress | 52.17 | 59.77 | 5.45 | 0.9535 |
| Evening Access | 27.11 | 32.75 | 3.93 | 0.9560 |
| Evening Egress | 28.43 | 32.89 | 5.27 | 0.9770 |

#### 4.2.2. Ablation Experiments

We conducted ablation experiments on peak-hour flows by implementing several comparative models. First, we replaced the xLSTM component in the fusion model with conventional LSTM using identical weight determination methods. We then compared this with standalone xLSTM and conventional LSTM models, followed by comparisons with individual SARIMA and ARIMA models. The ablation experiment results for four distinct peak flow types are presented in Tables 3–6. The proposed fusion model consistently outperformed other models across all scenarios. While the performance difference between our proposed fusion model and the LSTM-based fusion model was minimal for morning peak access and evening peak egress flows, our xLSTM-based fusion model showed notably improved performance in scenarios where the xLSTM component had higher weights. This

validates the enhanced capability of the xLSTM model in predicting complex patterns, with xLSTM consistently outperforming LSTM across all peak flows, thereby confirming its structural advantages. Tables 3–6 have shown the comparative performance analysis from ablation studies.

**Table 3.** The comparative performance analysis from ablation studies for **morning access** flow.

| Model | $R^2$ | MAE | RMSE | MAPE (%) |
|---|---|---|---|---|
| **SARIMA-xLSTM** | **0.9928** | **33.12** | **37.09** | **4.79** |
| xLSTM | 0.9862 | 51.51 | 54.69 | 8.80 |
| SARIMA-LSTM | 0.9912 | 33.08 | 39.34 | 4.72 |
| LSTM | 0.9790 | 56.77 | 63.95 | 10.79 |
| SARIMA | 0.9892 | 33.12 | 41.31 | 4.41 |
| ARIMA | 0.9815 | 44.86 | 53.13 | 5.86 |

**Table 4.** The comparative performance analysis from ablation studies for **evening access** flow.

| Model | $R^2$ | MAE | RMSE | MAPE (%) |
|---|---|---|---|---|
| **SARIMA-xLSTM** | **0.9560** | **27.11** | **32.75** | **3.93** |
| xLSTM | 0.9140 | 37.60 | 44.07 | 5.41 |
| SARIMA-LSTM | 0.9467 | 29.27 | 34.98 | 4.06 |
| LSTM | 0.8593 | 46.45 | 54.57 | 6.25 |
| SARIMA | 0.9333 | 33.07 | 40.17 | 4.64 |
| ARIMA | 0.8849 | 40.77 | 52.63 | 5.82 |

**Table 5.** The comparative performance analysis from ablation studies for **morning egress** flow.

| Model | $R^2$ | MAE | RMSE | MAPE (%) |
|---|---|---|---|---|
| **SARIMA-xLSTM** | **0.9535** | **52.17** | **59.77** | **5.45** |
| xLSTM | 0.9497 | 56.47 | 65.34 | 6.07 |
| SARIMA-LSTM | 0.9169 | 65.33 | 78.45 | 6.69 |
| LSTM | 0.8691 | 85.22 | 101.69 | 8.67 |
| SARIMA | 0.8847 | 73.02 | 87.99 | 7.89 |
| ARIMA | 0.8979 | 73.96 | 89.74 | 7.74 |

**Table 6.** The comparative performance analysis from ablation studies for **evening egress** flow.

| Model | $R^2$ | MAE | RMSE | MAPE (%) |
|---|---|---|---|---|
| **SARIMA-xLSTM** | **0.9770** | **28.43** | **32.89** | **5.27** |
| xLSTM | 0.9591 | 35.96 | 41.82 | 7.10 |
| SARIMA-LSTM | 0.9744 | 30.64 | 36.08 | 5.89 |
| LSTM | 0.9497 | 46.40 | 51.05 | 9.70 |
| SARIMA | 0.9730 | 29.05 | 35.86 | 5.55 |
| ARIMA | 0.9324 | 46.98 | 54.33 | 8.50 |

Figures 4 and 5 illustrate the prediction results of all models for peak access and egress flows at two stations (station 51 and station 206), with actual values represented by

black solid lines. The proposed fusion model's predictions consistently demonstrated the closest alignment with actual values. Non-fusion models frequently exhibited deviations or anomalous fluctuations, such as ARIMA's underestimation of evening egress flow at Station 206 and LSTM's overestimation of morning access flow at Station 51. Additionally, SARIMA displayed anomalous fluctuations in morning egress flow predictions at Station 51. Both fusion models maintained closer predictions to actual values across all scenarios, with our xLSTM-based fusion model showing superior performance.

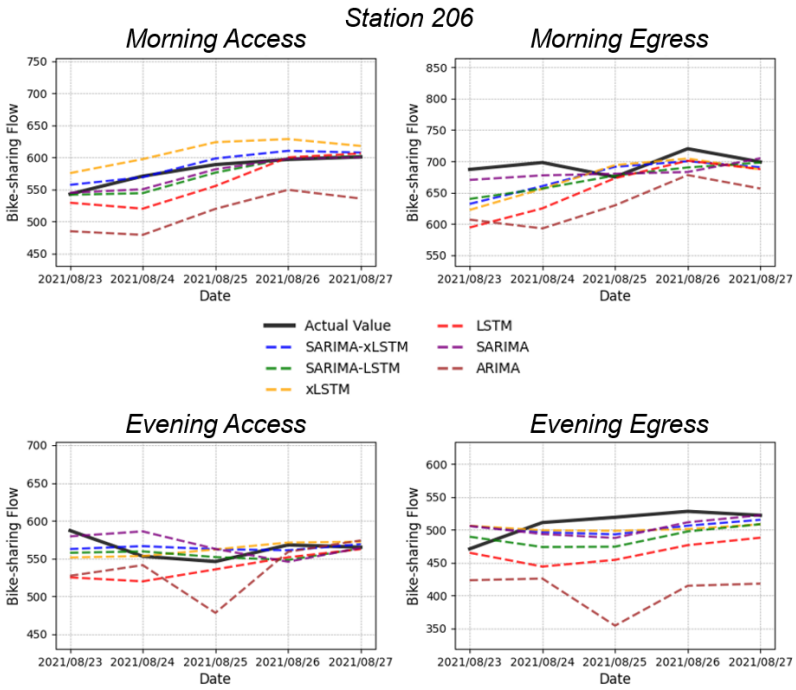


**Figure 4.** The prediction results of different models for different peak flows around Station 206.

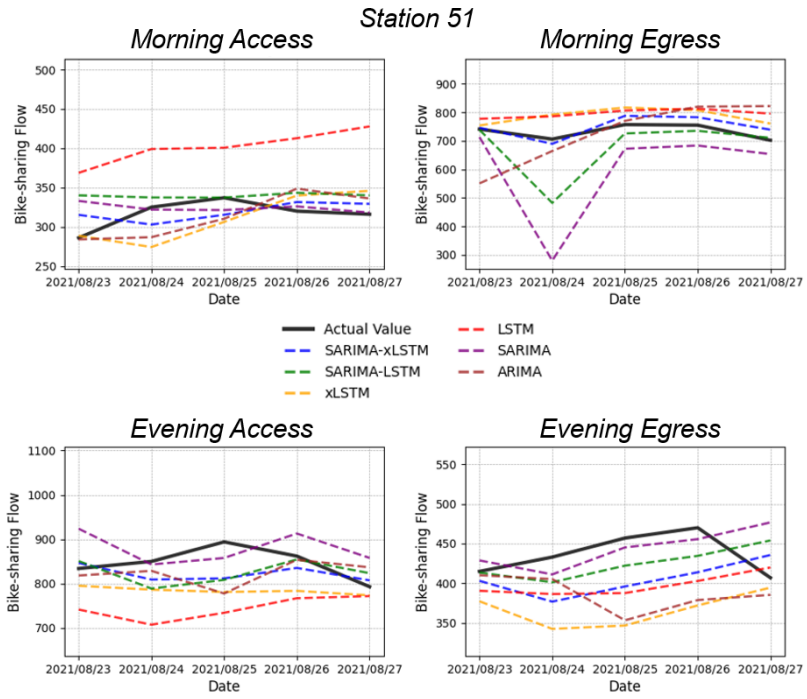


**Figure 5.** The prediction results of different models for different peak flows around Station 51.

#### 4.2.3. Discussions

To facilitate more intuitive visualization and in-depth analysis of prediction results of different models, we visualized MAE across three dimensions: (1) Figure 6 illustrates the daily error comparisons across four peak flow types for different models, with fusion models (blue), deep learning models (orange), and statistical models (green). The fusion models demonstrated enhanced overall performance. Morning egress flows exhibited the highest composite error among all peak flow types, particularly on August 24th, where xLSTM showed significantly lower prediction errors compared to LSTM for both morning egress and evening access flows, validating xLSTM's effectiveness in complex scenarios. (2) Figure 7 clearly illustrates that our proposed fusion model maintained minimal errors across all flow types over the five days, while SARIMA, LSTM, and the SARIMA-LSTM fusion model showed notable errors in morning egress flows. (3) Figure 8 shows error heatmaps for different stations (10 selected metro stations) and time periods. The selection of stations was based on the following criteria: ensuring relatively balanced traffic volumes across stations, achieving a spatially uniform distribution to the greatest extent possible, and including an equal proportion of regular stations and transfer stations. These error heatmaps provide a more intuitive comparison of prediction performance across different stations and flow types. Notably, in Figure 8, an analysis of the daily average MAE across 10 different stations reveals that certain stations exhibit lower average errors for specific flow types under alternative models. This observation warrants further explanation: while the proposed model achieves the lowest average error across all 10 stations for each passenger flow type, there are individual stations where it does not yield the minimum error. However, such instances are relatively infrequent, and the differences in error margins are minimal. Taking Station 51 as an example, the SARIMA model demonstrates superior performance in predicting morning access and evening egress flows. Nevertheless, it shows significantly higher prediction errors for the other two flow types (morning egress and evening access). In contrast, the proposed fusion model performs exceptionally well in these scenarios. Therefore, when considering the comprehensive performance across all flow types and stations, the proposed model demonstrates superior overall effectiveness.

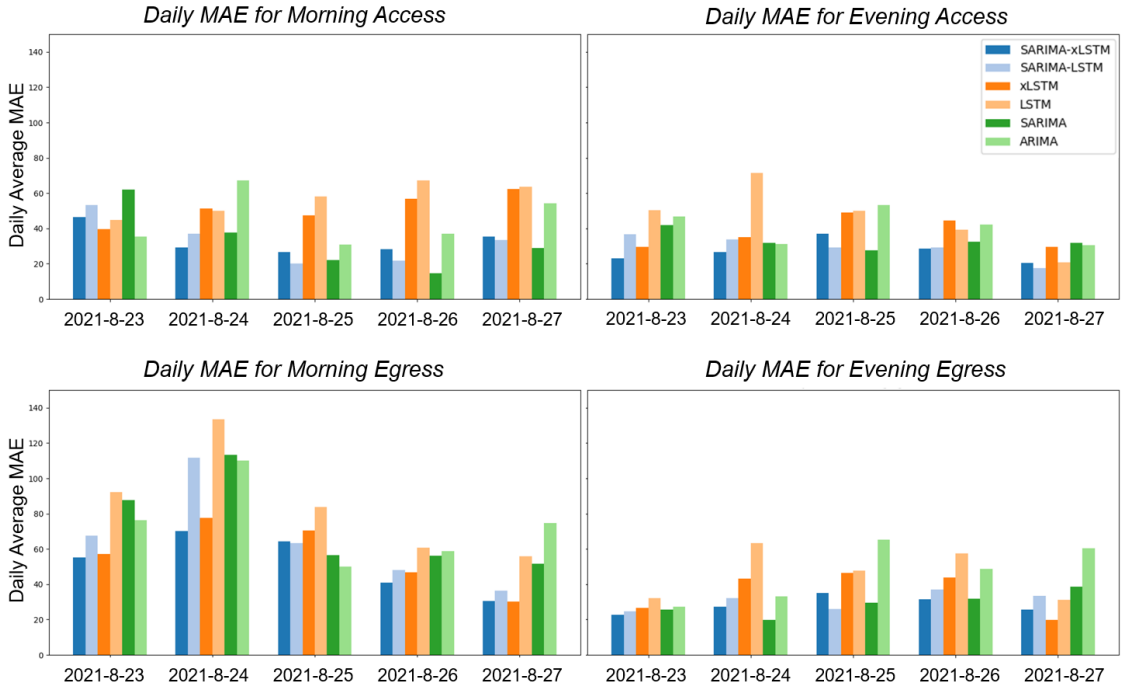


**Figure 6.** Daily MAE trend of the four peak flow types for different models.

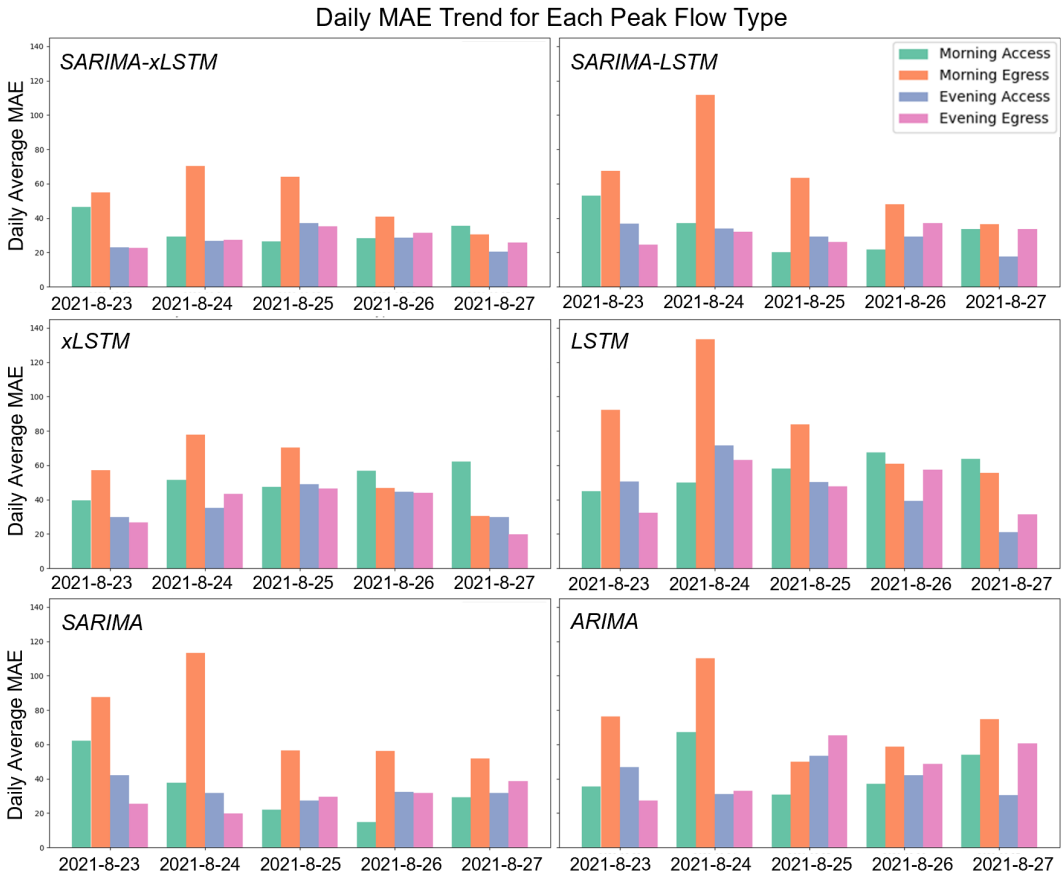


**Figure 7.** Daily MAE trend of different models across four peak flow types.

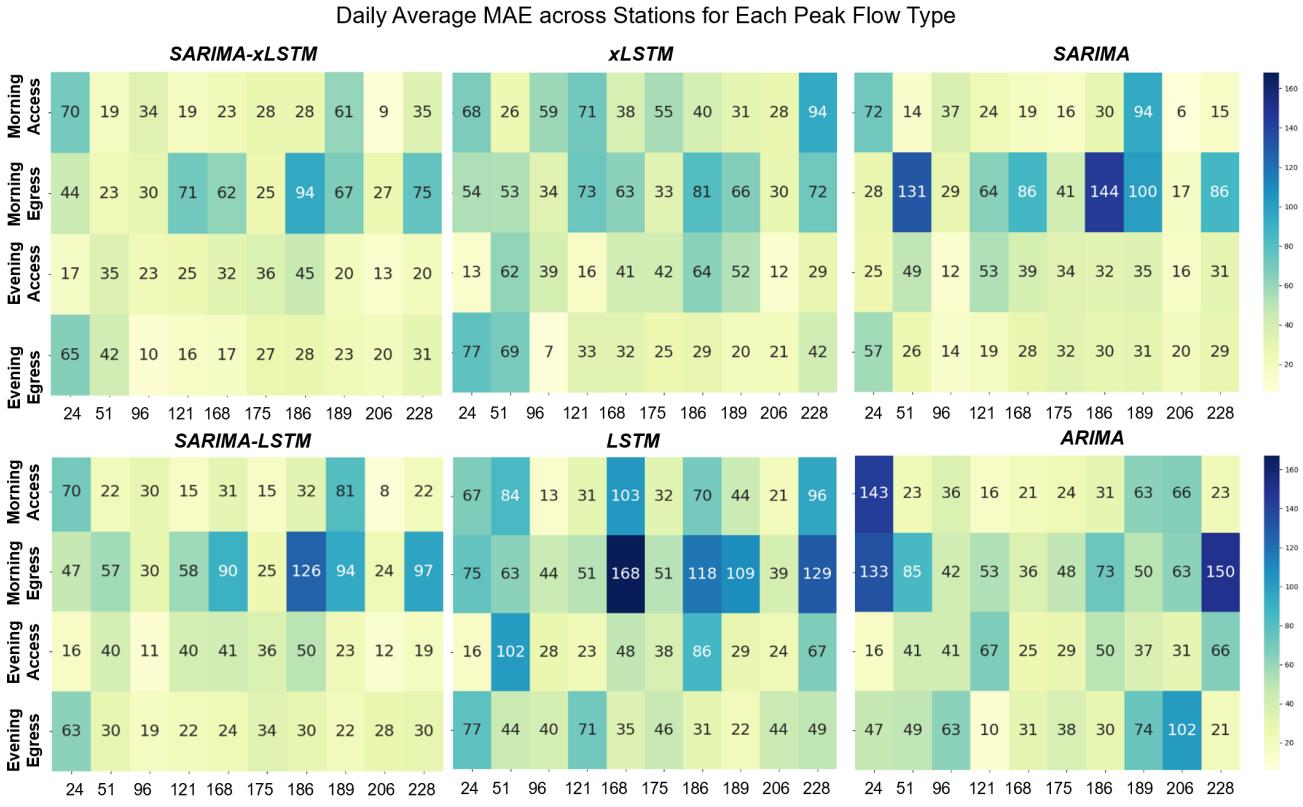


**Figure 8.** Daily average MAE for different models across four types of peak flow at different stations.

In summary, the SARIMA-xLSTM fusion model demonstrated superior error performance across all peak flows, while SARIMA performed well in morning access and evening egress flows. Notably, stations with higher morning egress flow errors showed distinct station characteristics, such as Stations 186 and 228 exhibiting larger errors compared to minimal errors at Stations 96 and 206. We hypothesize that these variations may be attributed to station characteristics and the surrounding built environment, suggesting the need for deeper feature extraction considering additional factors to achieve enhanced prediction accuracy.

## 5. Conclusions

This paper proposes a data–model dual-driven fusion approach for predicting bike-sharing flow around metro stations during peak hours. This approach separately predicts access and egress trips during morning and evening peak hours to better capture the patterns of different bike flow types. Through ablation studies, our experimental results demonstrate that the proposed fusion model achieves the highest R-squared values and lowest MAE, MSE, and RMSE across all four types of peak-hour bike flow prediction tasks. Specifically, the fusion model outperforms both individual SARIMA and xLSTM models, as well as the classical ARIMA and LSTM models. It achieved R-squared values of 0.9928 and 0.9770 for morning access and evening egress flows, respectively, and 0.9535 and 0.9560 for morning egress and evening access flows. Moreover, the xLSTM model shows an 8% increase in $R^2$ compared to the conventional LSTM model in the morning egress flow scenario, highlighting its greater robustness. For the morning egress and evening access flows, which exhibit relatively high variability, classical statistical models demonstrate limited effectiveness, as evidenced by SARIMA's $R^2$ values of 0.8847 and 0.9333, respectively. Even in scenarios such as morning access and evening egress, where classical models perform relatively well, our proposed fusion model still delivers enhanced performance.

Our study provided the following key merits and implications. First, our model achieves satisfactory prediction performance and computational efficiency while utilizing a fusion model structure integrating SARIMA and a simplified xLSTM model with only historical bike flow data and limited external features. SARIMA is adopted as the model-driven component to capture the periodic patterns of historical bike flow, while xLSTM serves as the data-driven component to handle the linear features and anomalous fluctuations in the data. Second, by separately predicting access and egress flows during peak hours, the fusion model better captures the distinct patterns of different bike flow types. Third, by dynamically adjusting the weights for different flow types, the fusion model achieves higher prediction accuracy and improved robustness. Even for highly regular and stable flow types like morning peak access flow, where SARIMA alone performs well, the fusion model further improves the R-squared score and reduces prediction errors. For more challenging flow types, such as morning peak egress and evening peak access trips where SARIMA struggles, the fusion model shows significant gains in prediction accuracy.

The proposed fusion model accurately predicts shared bike trips surrounding metro stations during peak hours, providing a reliable data foundation for subsequent shared bike rebalancing, thereby enhancing the overall efficiency and sustainability of commuting transportation. Moreover, the proposed dual-driven model framework is not only applicable to shared bike demand prediction but can also be extended to other related domains, particularly in engineering and management applications with high model interpretability requirements or where intrinsic data characteristics need to be explored. Such applications include public transit ridership prediction, travel time prediction, and other similar scenarios. By leveraging the respective advantages of the two components in the fusion model, we improve the overall prediction accuracy and robustness, thereby offering more reliable

data support for multimodal transportation. This approach facilitates the optimization and integration of different transportation modes, ultimately promoting the sustainability of intelligent transportation systems.

Our model has several limitations that point to important avenues for future research. First, from the error analysis of the fusion model, we observed that the prediction errors of the morning egress bike flow of certain metro stations are relatively high, while the errors at other stations are comparatively small. Therefore, in future work, incorporating spatial information about metro stations and factors related to the built environment around metro stations (such as POI, etc.) could further improve the performance of the prediction model [43,44]. Second, while our study only focused on predicting bike-sharing flow during peak hours, future research could extend the temporal scope to cover the entire day, adjusting the weight of the fusion model according to different time period characteristics to achieve accurate prediction. Third, regarding weather conditions, this study only considered precipitation because the temperature and humidity in Shenzhen remained relatively stable during the period covered by our dataset. Future studies should incorporate additional factors, such as temperature and humidity, for a more in-depth analysis. Furthermore, with the recent emergence of Large Language Models (LLMs) and their powerful capabilities in nonlinear modeling and data integration, future studies could explore using LLMs to integrate multimodal inputs (weather, POIs, social media data, etc.) for more intuitive and interpretable predictions [45].

In conclusion, bike-sharing demand prediction remains a worthy topic for in-depth research, as accurate demand forecasting provides a reliable data foundation for subsequent shared-bike rebalancing operations. Precise rebalancing strategies can prevent bicycle accumulation or shortage around metro stations during peak hours. This not only enhances the service level of shared bikes but also improves their utilization rate, thereby reducing resource waste. Furthermore, effectively addressing bike scheduling around metro stations is crucial for promoting the integration of two vital public transport modes, rail transit and bike-sharing systems. This integration can enhance the overall efficiency of the public transportation system and further contribute to reducing related carbon emissions.

**Author Contributions:** Z.W. contributed to the conceptualization, methodology, and writing of the original draft. D.Y. oversaw project administration and supervision. X.Z. was responsible for data curation and software development. F.M. and X.W. contributed to visualization and writing, including reviewing and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Natural Science Foundation of Fujian Province, grant number 2024J01703 (D.Y.) and Natural Science Foundation of Xiamen Municipality, grant number 3502Z20227211 (D.Y.) and Natural Science Foundation of Xiamen Municipality, grant number 3502Z202474018 (X.W.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available by the authors on request.

**Acknowledgments:** The authors are grateful to Shenzhen Government Open Data Platform for providing bike-sharing order data to conduct this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References


1. Liang, Y.; Zhao, Z.; Webster, C. Generating sparse origin–destination flows on shared mobility networks using probabilistic graph neural networks. *Sustain. Cities Soc.* **2024**, *114*, 105777. [CrossRef]
2. Caggiani, L.; Camporeale, R.; Ottomanelli, M.; Szeto, W.Y. A modeling framework for the dynamic management of free-floating bike-sharing systems. *Transp. Res. Part C Emerg. Technol.* **2018**, *87*, 159–182. [CrossRef]
3. Meng, F.; Zheng, L.; Ding, T.; Wang, Z.; Zhang, Y.; Li, W. Understanding dockless bike-sharing spatiotemporal travel patterns: Evidence from ten cities in China. *Comput. Environ. Urban Syst.* **2023**, *104*, 102006. [CrossRef]
4. Liu, Z.; Fang, C.; Li, H.; Wu, J.; Zhou, L.; Werner, M. Efficiency and equality of the multimodal travel between public transit and bike-sharing accounting for multiscale. *Sustain. Cities Soc.* **2024**, *101*, 105096. [CrossRef]
5. Huajing Information Network. Analysis of the Current Situation and Development Trends of China's Bike-Sharing Industry in 2023. 2024. Available online: https://www.huaon.com/channel/trend/962946.html (accessed on 30 November 2024).
6. World Resources Institute. How Dockless Bike-Sharing Changes Lives: An Analysis of Chinese Cities. 2024. Available online: https://wri.org.cn/research/how-dockless-bike-sharing-changes-lives (accessed on 30 November 2024).
7. Macioszek, E.; Jurdana, I. Bicycle Traffic in the Cities. *Zesz. Nauk. Politech. Śląskiej* **2022**, *117*, 115–127. [CrossRef]
8. Guo, R.; Jiang, Z.; Huang, J.; Tao, J.; Wang, C.; Li, J.; Chen, L. BikeNet: Accurate Bike Demand Prediction Using Graph Neural Networks for Station Rebalancing. In Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 19–23 August 2019.
9. Bei, Y.; Ge, Y.; Zhang, D. A machine learning based shared bikes scheduling method. In Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing, Virtual, 26–28 August 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 32–36.
10. Cao, M.; Liang, Y.; Zhu, Y.; Lü, G.; Ma, Z. Prediction for Origin-Destination Distribution of Dockless Shared Bicycles: A Case Study in Nanjing City. *Front. Public Health* **2022**, *10*, 849766. [CrossRef]
11. Zheng, L.; Meng, F.; Ding, T.; Yang, Q.; Xie, Z.; Jiang, Z. The effect of traffic status on dockless bicycle-sharing: Evidence from Shanghai, China. *J. Clean. Prod.* **2022**, *381*, 135207. [CrossRef]
12. Luo, X.; Gu, W.; Fan, W. Joint design of shared-bike and transit services in corridors. *Transp. Res. Part C Emerg. Technol.* **2021**, *132*, 103366. [CrossRef]
13. Garcia, E.; Calvet, L.; Carracedo, P.; Serrat, C.; Miró, P.; Peyman, M.T. Predictive Analyses of Traffic Level in the City of Barcelona: From ARIMA to eXtreme Gradient Boosting. *Appl. Sci.* **2024**, *14*, 4432. [CrossRef]
14. Wei, Z.; Wang, K.; Gao, J. *Research on Traffic Safety Accident Prediction Based on ARIMA*; Atlantis Press: Amsterdam, The Netherlands, 2024.
15. Harikrishnakumar, R.; Nannapaneni, S. Forecasting Bike Sharing Demand Using Quantum Bayesian Network. *Expert Syst. Appl.* **2023**, *221*, 119749. [CrossRef]
16. Afandizadeh, S.; Abdolahi, S.; Mirzahossein, H. Deep Learning Algorithms for Traffic Forecasting: A Comprehensive Review and Comparison with Classical Ones. *J. Adv. Transp.* **2024**, *2024*, 9981657. [CrossRef]
17. Kaltenbrunner, A.; Meza, R.; Grivolla, J.; Codina, J.; Banchs, R. Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. *Pervasive Mob. Comput.* **2010**, *6*, 455–466. [CrossRef]
18. Yoon, J.W.; Pinelli, F.; Calabrese, F. Cityride: A Predictive Bike Sharing Journey Advisor. In Proceedings of the 2012 IEEE 13th International Conference on Mobile Data Management, Bengaluru, India, 23–26 July 2012.
19. Caulfield, B.; O'Mahony, M.; Brazil, W.; Weldon, P. Examining usage patterns of a bike-sharing scheme in a medium sized city. *Transp. Res. Part A Policy Pract.* **2017**, *100*, 152–161. [CrossRef]
20. Zhou, Y.J.; Wang, L.L.; Zhong, R.; Tan, Y.L. A Markov Chain Based Demand Prediction Model for Stations in Bike Sharing Systems. *Math. Probl. Eng.* **2018**, *2018*, 8028714. [CrossRef]
21. Nistor, M.; Dias, A. Bike distribution model for urban data applications. *Int. J. Transp. Dev. Integr.* **2019**, *3*, 67–78. [CrossRef]
22. Cagliero, L.; Cerquitelli, T.; Chiusano, S.; Garza, P.; Xiao, X. Predicting critical conditions in bicycle sharing systems. *Computing* **2017**, *99*, 39–57. [CrossRef]
23. Ma, H.; Peng, T.; Sun, Y. Prediction Model of Demand for Shared Bikes based on Bayesian Theory. In Proceedings of the 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shanghai, China, 15–17 October 2021.
24. Duan, Y.M.; Zhang, S.; Yu, Z.R. Applying Bayesian spatio-temporal models to demand analysis of shared bicycle. *Phys. A Stat. Mech. Its Appl.* **2021**, *583*, 126296. [CrossRef]
25. Gu, K.; Lin, Y. Prediction for bike-sharing demand in London using multiple linear regression and random forest. In Proceedings of the 2023 5th International Conference on Artificial Intelligence and Computer Science (AICS 2023), Wuhan, China, 26–28 July 2023; SPIE: Bellingham, WA, USA, 2023.
26. Sathishkumar, V.E.; Park, J.; Cho, Y. Using data mining techniques for bike sharing demand prediction in metropolitan city. *Comput. Commun.* **2020**, *153*, 353–366.

27. Chen, P.-C.; Hsieh, H.-Y.; Su, K.-W.; Sigalingging, X.K.; Chen, Y.-R.; Leu, J.-S. Predicting station level demand in a bike-sharing system using recurrent neural networks. *IET Intell. Transp. Syst.* **2020**, *14*, 554–561. [CrossRef]
28. Wang, B.; Vu, H.L.; Kim, I.; Cai, C. Short-term traffic flow prediction in bike-sharing networks. *J. Intell. Transp. Syst.* **2022**, *26*, 461–475. [CrossRef]
29. Zhai, L.; Yang, Y.; Song, S.; Ma, S.; Zhu, X.; Yang, F. Self-supervision Spatiotemporal Part-Whole Convolutional Neural Network for Traffic Prediction. *Phys. A Stat. Mech. Its Appl.* **2021**, *579*, 126141. [CrossRef]
30. Pan, Y.; Zheng, R.C.; Zhang, J.; Yao, X. Predicting bike sharing demand using recurrent neural networks. *Procedia Comput. Sci.* **2019**, *147*, 562–566. [CrossRef]
31. Sardinha, C.; Finamore, A.C.; Henriques, R. Context-aware demand prediction in bike sharing systems: Incorporating spatial, meteorological and calendrical context. *arXiv* **2021**, arXiv:2105.01125.
32. Collini, E.; Nesi, P.; Pantaleo, G. Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations. *IEEE Access* **2021**, *9*, 124337–124347. [CrossRef]
33. Jin, K.; Wang, W.; Li, S.; Liu, P.; Sun, H. Dockless Shared-Bike Demand Prediction with Temporal Convolutional Networks. In Proceedings of the 20th COTA International Conference of Transportation Professionals (CICTP 2020), Xi'an, China, 14–16 August 2020; American Society of Civil Engineers: Reston, VA, USA, 2020; pp. 2851–2863.
34. Lin, L.; He, Z.; Peeta, S. Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 258–276. [CrossRef]
35. Qin, T.; Liu, T.; Wu, H.; Tong, W.; Zhao, S. RESGCN: RESidual Graph Convolutional Network based Free Dock Prediction in Bike Sharing System. In Proceedings of the 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 30 June–3 July 2020.
36. He, S.N.; Shin, K.G.; Assoc Comp, M. Towards Fine-grained Flow Forecasting: A Graph Attention Approach for Bike Sharing Systems. In Proceedings of the 29th World Wide Web Conference (WWW), Taipei, Taiwan, 20–24 April 2020.
37. Lee, S.H.; Ku, H.C. A Dual Attention-Based Recurrent Neural Network for Short-Term Bike Sharing Usage Demand Prediction. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 4621–4630. [CrossRef]
38. Zi, W.; Xiong, W.; Chen, H.; Chen, L. TAGCN: Station-level demand prediction for bike-sharing system via a temporal attention graph convolution network. *Inf. Sci.* **2021**, *561*, 274–285. [CrossRef]
39. Mehdizadeh Dastjerdi, A.; Morency, C. Bike-Sharing Demand Prediction at Community Level under COVID-19 Using Deep Learning. *Sensors* **2022**, *22*, 1060. [CrossRef]
40. Zhou, S.; Song, C.; Wang, T.; Pan, X.; Chang, W.; Yang, L. A Short-Term Hybrid TCN-GRU Prediction Model of Bike-Sharing Demand Based on Travel Characteristics Mining. *Entropy* **2022**, *24*, 1193. [CrossRef]
41. Beck, M.; Pöppel, K.; Spanring, M.; Auer, A.; Prudnikova, O.; Kopp, M.; Klambauer, G.; Brandstetter, J.; Hochreiter, S. xLSTM: Extended Long Short-Term Memory. *arXiv* **2024**, arXiv:2405.04517.
42. Shenzhen Government Open Data Platform. Daily Order from Shared Bike Enterprises. 2021. Available online: https://opendata.sz.gov.cn/data/api/toApiDetails/29200_00403627 (accessed on 30 November 2024).
43. Zhang, Y.; Hu, X.; Wang, H.; An, S. How does the built environment affect the usage efficiency of dockless-shared bicycle? An exploration of time-varying nonlinear relationships. *J. Transp. Geogr.* **2024**, *118*, 103908.
44. Li, Z.; Tang, J.; Yu, T.; Liu, B.; Cao, J. Understanding the interplay among urban public transport modes: Spatial variation and built environment effects. *J. Clean. Prod.* **2024**, *479*, 144038. [CrossRef]
45. Qu, X.; Lin, H.; Liu, Y. Envisioning the future of transportation: Inspiration of ChatGPT and large models. *Commun. Transp. Res.* **2023**, *3*, 100103. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.