

Research Article

Prediction and Impact Analysis of Passenger Flow in Urban Rail Transit in the Postpandemic Era

Guifang Shi ¹ and Limei Luo²

¹Jiangsu AI Transportation Innovations and Applications Engineering Research Center, Jinling Institute of Technology, Nanjing, China

²Nanjing Institute of City & Transport Planning Company Limited, Nanjing, China

Correspondence should be addressed to Guifang Shi; sgf@jit.edu.cn

Received 14 January 2023; Revised 2 June 2023; Accepted 13 July 2023; Published 2 August 2023

Academic Editor: Rui Jiang

Copyright © 2023 Guifang Shi and Limei Luo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the postpandemic era, exploring the relationship between the daily new COVID-19 cases and passenger flow in urban rail transit can help effectively predict the impact of future pandemic situations on rail transit. In this study, based on a gated recurrent unit (GRU) neural network model, the daily passenger flow in urban rail transit in the postpandemic era was predicted, and the results were compared with those obtained using the long short-term memory (LSTM) neural network and other conventional time series analysis models such as SARIMA (seasonal autoregressive integrated moving average). Based on the trained GRU model, a partial dependence plot (PDP) was adopted to explore the quantitative relationship between the daily passenger flow and the daily new cases or weather attribute. The results showed that (1) the prediction accuracy of the GRU neural network model was 95.25%, which was the highest among the prediction models studied, indicating that the GRU could achieve the best performance. (2) The GRU model did not fluctuate significantly in the initial training stage, and its convergence rate was higher than that of the LSTM. (3) The number of daily new cases was negatively correlated with the daily passenger flow. For every new case on the previous day, the daily passenger flow fell by an average of 54,600 person-times. (4) Compared with no rain condition, the daily passenger flow decreased by 207,600 person-times on an average on rainy days. In summary, the neural network could achieve accurate prediction, while the PDP could compensate for the “black box” disadvantage of nonparametric models, owing to which the quantitative relationship between the number of new cases and daily passenger flow could be successfully explored. Our study can serve as a basis for demand prediction, operational organization, and policy implementation related to urban rail transit.

1. Introduction

The COVID-19 pandemic (hereinafter referred to as the pandemic) has had a huge impact on urban rail transit. At the national level, data from the China Urban Rail Transit Operation Development Report (2020-2021) showed that the average passenger intensity of urban rail transit in China would be 4,500 person-time/(km-day) in 2020, a decrease of 2,700 person-time/(km-day) or 36.9% compared with the same period last year. In Shanghai, according to the Annual Report of Shanghai Comprehensive Transportation Development in 2021, the average daily passenger flow of rail transit in Shanghai was 7.75 million passenger trips per day in 2020, down 27.2% year-on-year.

Since the outbreak of the pandemic in January 2020, passenger flow in the rail transit of Shanghai can be categorized into three stages: sharp decline, gradual recovery, and normal stabilization [1]. On January 24, the Shanghai Municipal People's Government activated the Level 1 response mechanism for major public health emergencies, which, combined with the Spring Festival effect, led to a sudden drop in passenger flow. Since February 9, work and production resumed, and the total number of passengers on working days started to recover gradually. On May 9, the emergency response level for major public health emergencies in Shanghai was raised to Level 3, and the total passenger flow stabilized to over 9 million person-times, entering the normal stabilization phase. This is followed by

the postpandemic era. This is the period in which COVID-19 cases are expected to be under control but will continue to have a lasting and significant impact on the public's daily choice of trip modes [2, 3]. In the postpandemic era, the number of new COVID-19 cases will tend to stabilize but minor eruptions are likely.

Passenger flow prediction considering the impact of the pandemic can help accurately assess urban rail transit demand and provide an important reference for predicting the operation state of urban rail transit and in the formulation of organizational management strategies in the postpandemic era. To predict the passenger flow in rail transit, existing research methods can be divided into parametric and nonparametric models. Parametric models are mostly based on a self-regression time-series model, where historical passenger flow estimation model parameters are used to forecast the future passenger flow. For instance, Wang et al. [4] used the seasonal autoregressive integrated moving average (SARIMA) model to predict the daily inbound passenger flow in Beijing. The SARIMA model can predict periodic time series more accurately than the conventional ARIMA model. Similarly, Kumar and Vanajakshi [5] constructed a SARIMA model for a time series analysis of short-term traffic flow based on limited input data. Milenković et al. [6] adopted a SARIMA model to predict the monthly passenger flows of Serbian railways. Li et al. [7] used the SARIMA model to predict the hourly passenger flow of the Guangzhou–Zhuhai intercity railroad.

However, the self-regression model only considers the variation in the passenger flow over time, ignoring the influence of external factors such as holidays and weather. To overcome the above shortcomings, Cheng et al. [8] incorporated the holiday effect in SARIMA with the exogenous factors (ARIMAX) method to build a daily passenger flow prediction model for Hongqiao hub. Xu et al. [9] used the SARIMAX model to explain the effect of different weather factors on subway passenger flow. Bai et al. [10] proposed a combined ARIMA and multiple linear regression model for a nonconventional short-term passenger flow prediction of urban rail transit.

Compared with parametric models, nonparametric models are more flexible and can effectively deal with the nonlinear relationship between passenger flow and multidimensional influencing factors, thus producing a better prediction performance. The main methods for predicting the daily passenger flow in rail transit using nonparametric models include the hybrid deep neural network model [11], long short-term memory (LSTM) neural network [12], random forest model [13], support vector machine [14], and bilayer parallel wavelet neural networks [15].

LSTM neural networks can learn long-term information by introducing gated units, such as forgetting gates, memory gates, and output gates, based on recurrent neural networks for an effective prediction of nonlinear time-series data [16]. Li et al. [12] divided the factors influencing passenger flow into external and internal and used the LSTM neural network for a 15-min real-time prediction of the passenger flow in rail transit. The prediction accuracy of the LSTM was higher than that of the multiple linear regression and back

propagation (BP) neural networks. Teng and Li [17] combined the LSTM neural network with the particle swarm optimization (PSO) algorithm to predict the daily passenger flow in the Shanghai–Nanjing one-way railroad, considering date attributes and weather factors. Liu et al. [18] established an LSTM neural network to forecast hourly metro passenger flows, and the effects of weather variables on the model's performance were analyzed. Yang et al. [19] built a spatio-temporal LSTM to analyze the time series outbound passenger volume at urban rail stations using historical passenger volume data, a station origin-destination matrix, and rail transit operation data.

A variant of the LSTM model has been applied to the prediction of passenger flow in rail transit. Hou et al. [20] used a gated recurrent unit (GRU) neural network to predict the short-term passenger flow in urban rail transit. The results showed that the GRU has faster convergence, lower prediction error, and better stability than the LSTM. Huang et al. [21] used the gray relation analysis (GRA) to filter the weather factor with a high correlation with the passenger flow and the bidirectional LSTM (BiLSTM) to predict the hourly passenger flow in rail transit on weekdays and nonweekdays, respectively. The BiLSTM outperformed the conventional LSTM in terms of the prediction performance.

In summary, in terms of the influencing factors, the weather and holiday attributes are external factors significantly affecting the passenger flow; there has been no research on the influence of the pandemic on passenger flow. This study considered incorporating the daily number of local new COVID-19 cases, weather, temperature, and holidays in the daily passenger flow prediction process for an accurate prediction of the passenger flow. In terms of the model performance, the parametric model is superior at handling time series with significant trends and seasonal variations, while the nonparametric model is more effective at handling multidimensional nonlinear inputs. Passenger flow is significantly influenced by nonperiodic factors such as holidays and weather. In the pandemic era, the sudden factor of daily new cases should also be incorporated in the passenger flow prediction so that the nonparametric model can achieve a higher prediction accuracy. However, the “black box” characteristic of the nonparametric model prevents it from assessing the quantitative relationship between the input and output variables. The above studies that used nonparametric models focused on the passenger flow prediction performance without explaining the influence degree of each factor on the passenger flow.

In this study, a partial dependence function was used to compensate for the poor interpretation of the nonparametric model. An effective GRU neural network model was constructed for passenger flow prediction based on the daily passenger flow data and the daily number of local new cases in Shanghai urban rail transit. A partial dependence plot (PDP) was then employed to explore the external factors affecting passenger flow and to investigate the quantitative relationship between the pandemic and daily passenger flow. This study provides a basis for urban rail transit demand prediction, operation organization, and policy implementation.

2. Data Preparation

Since the outbreak of the pandemic, passenger flow in the rail transit of Shanghai has experienced three stages, as shown in Figure 1 (where the unit in the y -axis is 10,000 person-time). The focus of this study is the daily passenger flow in the rail transit of Shanghai in the postpandemic era, i.e., the third period in Figure 1. The postpandemic era is the period in which COVID-19 cases will be under control but will have a lasting and significant impact on the public's daily choice of trip modes until the coronavirus becomes less harmful. The reasons for selecting the postpandemic era are as follows: (1) among the three periods shown in Figure 1, the postpandemic era is the most long-lasting one and has the most long-lasting impact on the public's daily life and (2) this lasting and sustained impact will make the quantitative results calculated by partial dependence methods more practical.

The analysis period was from June 1, 2020, to December 31, 2021 (a total of 579 days). The obtained information included the daily passenger flow in rail transit, the daily number of local new COVID-19 cases in Shanghai, and the corresponding weather and holiday attributes of Shanghai on that day. The data were collected from Weibo.

Figure 2 illustrates the time-series curves of the daily passenger flow and the daily number of new cases. As observed, there is a correlation between the daily number of local new cases and the daily passenger flow. The passenger flow corresponding to the number of days where new cases appeared tends to be the local minimum in the period before and after. Notably, due to the impact of Severe Typhoon In-Fa, on July 26th, 2021, Shanghai saw a massive suspension of classes, home offices, and some rail transit lines; therefore, the passenger flow on that day reached the minimum: 1.814 million person-time.

The relevant data will be used as input and output of the subsequent prediction model. Table 1 shows the variable definitions and descriptive statistics. All the variables are divided into two categories: external and internal factors. As observed, local new cases occurred on 24 of the 579 days analyzed, accounting for 4%. In terms of the weather attribute, there were 249 days with rain, accounting for 43%; the mean minimum temperature and mean maximum temperature were 17°C and 23°C, respectively. In terms of the holiday attribute, the number of holidays is 399 days, including weekends, Spring Festival, Qingming Festival, Labor Day, Dragon Boat Festival, Mid-Autumn Festival, and National Day vacation, accounting for 31%. In 579 days, the mean daily passenger flow is 9.587 person-time, and SD is 2.287 person-time.

3. Method and Model

3.1. Technology Route. Figure 3 shows the data structure on the i^{th} day. Herein, the input variable dimension is $t \times n$, where t represents the time step and n represents the feature dimension. Considering a cycle of 7 days a week, to predict the daily passenger flow of rail transit on the i^{th} day, the features of that day and the previous 7 days (8 days in total)

are used as input, i.e., the time step is set to 8, $t = 1, \dots, 8$. The features for each day include the number of new cases on that day, weather attribute, minimum temperature, maximum temperature, holiday attribute, and passenger flow yesterday, with six feature dimensions, i.e., $n = 1, \dots, 6$. Since the first seven days of data are used as input, the complete data structure is available from June 8, 2020, with a total of 572 days of valid data, i.e., $i = 1, \dots, 572$. In summary, the input variable on the i^{th} day has a dimension of $8 \times 6 = 48$. The output variable on the i^{th} day is the number of passengers on that day, which corresponds to the variable $X_i^{0 \times 6}$ with a time step of 0 and a feature dimension of 6.

Figure 4 shows the technical route employed in this study. First, the acquired raw data were divided into training and testing sets in the ratio of 3:1: training set for the first 429 days (June 8, 2020, to August 10, 2021) and testing set for the second 143 days (August 11, 2021, to December 31, 2021). The neural network model is then trained using the training set, and the prediction performance of the trained model is evaluated on the test set. During the training process, the model parameters are continuously optimized based on the loss function until the maximum number of iterations is reached. During the impact evaluation, the training and test sets are combined to train the final model. Based on the final model to determine the partial dependence function, the PDPs of the input and output variables are plotted, and the quantitative impact of the daily number of new cases on the daily passenger flow is evaluated.

3.2. GRU Neural Network Model. Owing to the advantage of the nonparametric model in handling multidimensional nonlinear input, the GRU neural network is used to predict the daily passenger flow of rail transit. The GRU neural network is a variant of the LSTM neural network, which is a special type of recursive neural network (RNN). The RNN uses temporal-dimensional information to process data with temporal characteristics. However, it cannot solve the long-term dependency problem and has the disadvantages of gradient disappearance and gradient explosion, which led to the development of the LSTM neural network. The LSTM neural network introduces various gated units (e.g., forgetting gates, memory gates, and output gates), retains information that requires long-term memory, and forgets information with decaying value for an accurate prediction of time-series data [22].

Compared with the LSTM, the GRU neural network is simpler and more efficient; Figure 5 shows its structure. In this approach, data input x_t and hidden state h_{t-1} of the previous time step at time step t are received, and the hidden state h_t of the current time step is outputted. Unlike the LSTM, the GRU neural network has only two gated units: the reset gate r_t and the update gate z_t . The reset gate r_t is used to discard irrelevant historical information and control how much of the previous time step's hidden state h_{t-1} needs to be retained by the candidate hidden state \tilde{h}_t . The candidate hidden state \tilde{h}_t is used to assist in the computation of the hidden state h_t . The update gate z_t is used to control how

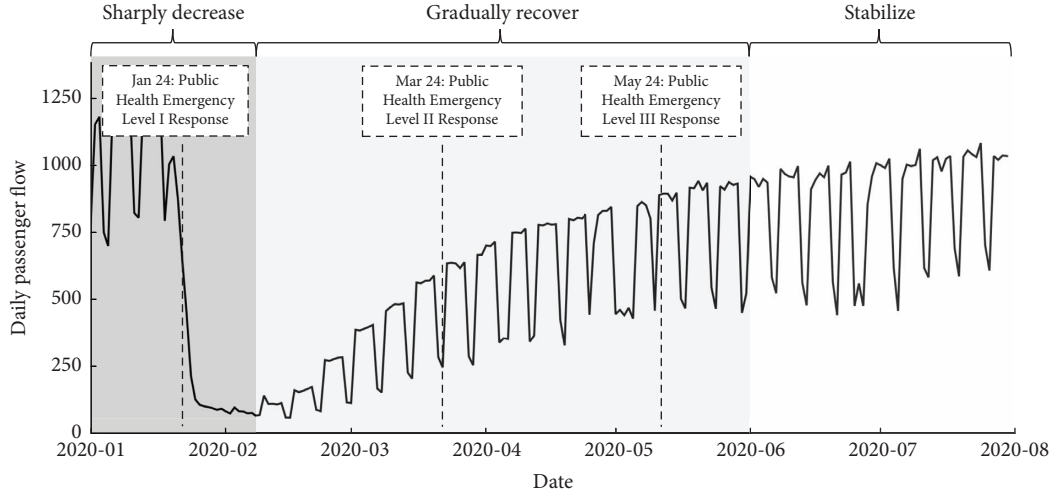


FIGURE 1: Stages of passenger flow in the rail transit of Shanghai.

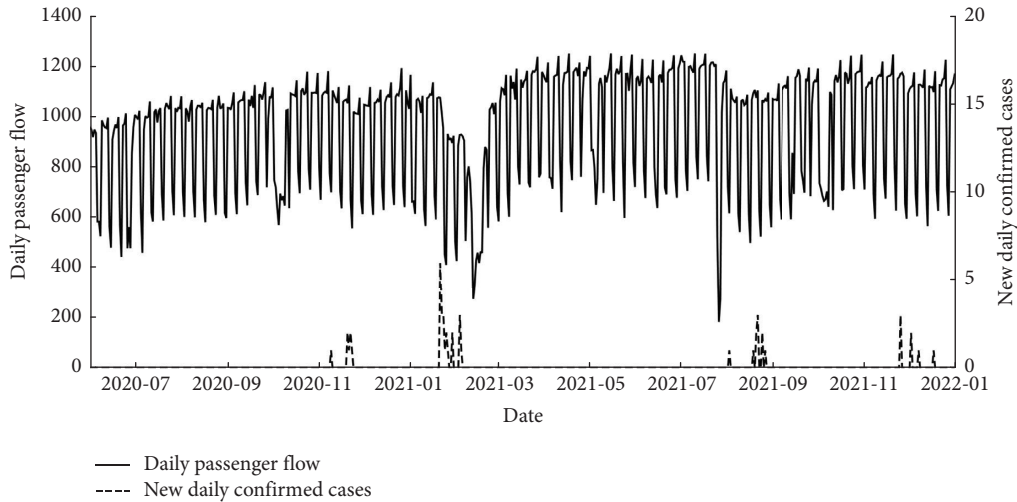


FIGURE 2: Daily passenger flow time series of Shanghai urban rail transit.

the hidden state h_t is updated by the candidate hidden state \tilde{h}_t [23].

The expressions for the GRU neural network are as follows:

$$\begin{aligned}
 r_t &= \sigma(W_{ir}x_t + b_{ir} + W_{hr}h_{t-1} + b_{hr}), \\
 z_t &= \sigma(W_{iz}x_t + b_{iz} + W_{hz}h_{t-1} + b_{hz}), \\
 \tilde{h}_t &= \tanh(W_{in}x_t + b_{in} + r_t * (W_{hn}h_{t-1} + b_{hn})), \\
 h_t &= (1 - z_t) * n_t + z_t * h_{t-1},
 \end{aligned} \tag{1}$$

where x_t is the input of the current time step t , h_t is the hidden state of time step t , and h_{t-1} is the hidden state of the previous time step. r_t and z_t are the reset and update gates, respectively. \tilde{h}_t is the candidate hidden state. W_{ir} , W_{iz} , and W_{in} are the weights of the input x_t and reset gate, update gate, and candidate hidden state, respectively. W_{hz} and W_{hn} are the weights of the hidden state and update gate, candidate hidden state, respectively. b_{ir} , b_{iz} , b_{in} , b_{hz} , and b_{hn} are

the bias vectors. σ is the Sigmoid activation function, \tanh is the tanh activation function, and $*$ is the Hadamard product.

The GRU neural network is trained using PyTorch with the parameter settings shown in Table 2. The optimizer is a method to update the parameters in neural networks, where the goal is to make the parameters approximate or reach the optimal, thus minimizing the network loss. The Adam optimizer used in this study combines the momentum algorithm with the root mean square propagation (RMSProp) algorithm, using the momentum cumulative gradient, for faster convergence and smaller fluctuations [24]. The loss function used is the mean square error (MSE) loss, which is calculated as follows:

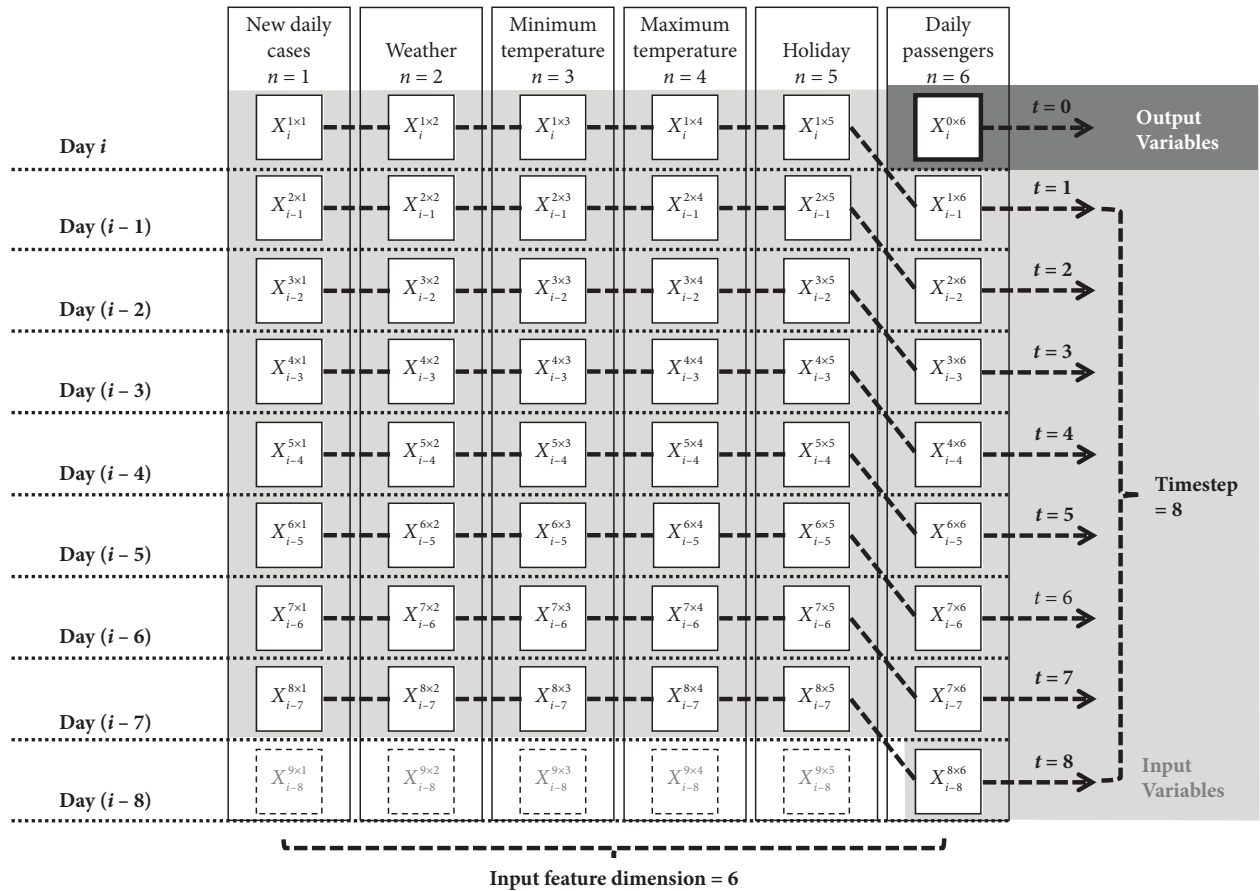
$$E_{MS} = \frac{1}{n} \sum_{i=1}^n (X_i^{0 \times 6} - \tilde{X}_i^{0 \times 6})^2, \tag{2}$$

where n is the number of predicted samples and $X_i^{0 \times 6}$ and $\tilde{X}_i^{0 \times 6}$ are the predicted and theoretical daily passenger flows on the i^{th} day, respectively.

TABLE 1: Descriptive statistics of variables.

Factor types	Variable description	Variable types	Statistics	
			Value: frequency	
			0: 555	
			1: 11	
			2: 7	
			3: 5	
			4: 0	
			5: 0	
			6: 1	
			Value: frequency	
			0 (clear): 57%	
			1 (rainy): 43%	
External factors	Weather	Categorical		
	Minimum air temperature (°C)	Continuous	Min: -6	Mean: 17
			Max: 29	STD*: 8
	Maximum air temperature (°C)	Continuous	Min: -3	Mean: 23
			Max: 38	STD: 9
			Value: frequency	
			0 (workday): 69%	
			1 (holiday): 31%	
Internal factors	Daily passenger flow (ten thousand passengers)	Continuous	Min: 181.4	Mean: 958.7
			Max: 1255.4	STD: 228.7

Note. STD* denotes the standard deviation.

FIGURE 3: Data structure on the i^{th} day.

3.3. *Partial Dependence Plot (PDP)*. The PDP is essentially a machine learning visualization method with “black box” features, which is widely used as a technique to increase the

interpretation of input variables in machine learning or deep learning models [25, 26]. The plotting of the PDP relies on the fitted model to describe the average marginal effect of

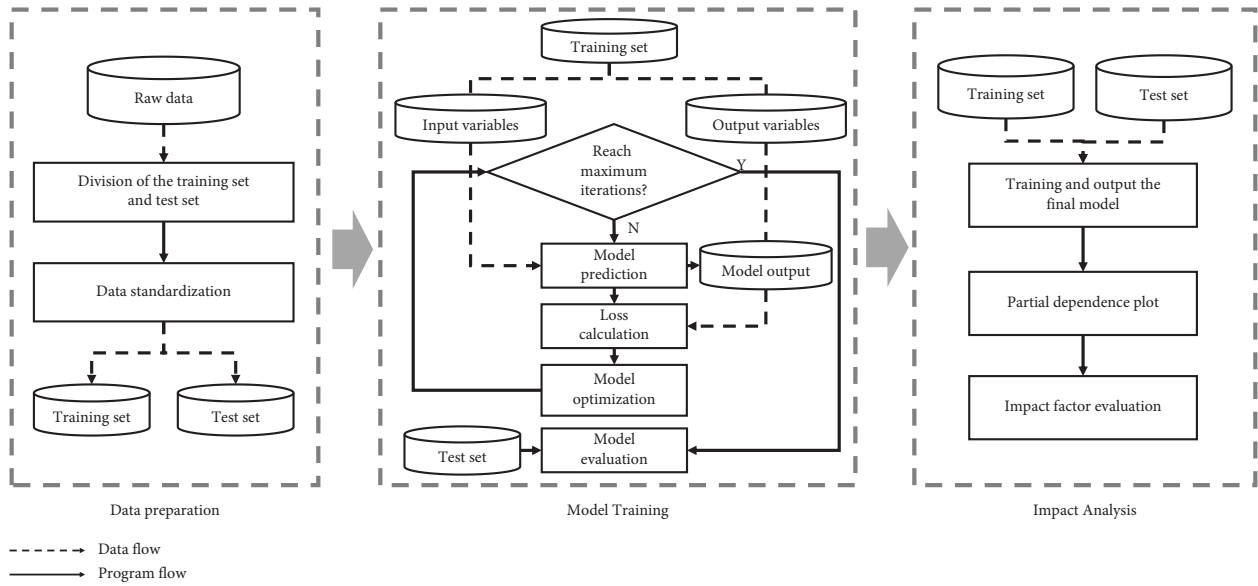


FIGURE 4: Technical route.

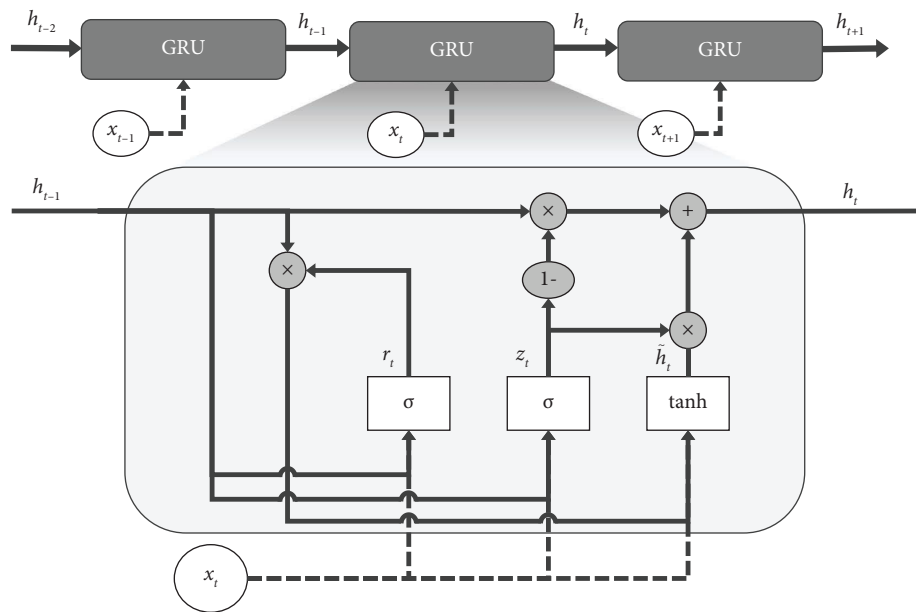


FIGURE 5: GRU structure.

TABLE 2: Parameter settings of the GRU neural network.

Parameters	Description	Values
epoch	Number of iterations	100
time_step	Number of timesteps	8
input_size	Dimensions of the input feature vector	6
hidden_size	Dimensions of the hidden feature vector	64
num_layers	Number of hidden layers	1
lr	Learning rate	0.02
optimizer	Type of optimizer	Adam
weight_decay	Coefficient of weight decay	0.001
loss_func	Type of loss function	MSELoss

a feature variable on the predicted outcome through a variable intervention approach [27]. To plot the PDP for feature variable x and output variable y , the steps are as follows:

- (a) Determine the possible values $[x_1, x_2, \dots, x_k]$ of the features x and set $i = 1$
- (b) Force the value of x of the dataset to be x_i
- (c) Based on the fitted model, calculate the mean y_i of the output of the dataset
- (d) Plot the points (x_i, y_i) in the PDP
- (e) Make $i = i + 1$, repeat steps (b), (c), and (d), until $i = k$, and the points (x_k, y_k) are plotted
- (f) Obtain the final PDP

Substituting into the current daily passenger flow prediction problem, if it is necessary to plot the PDP of the number of new cases $X_i^{1 \times 1}$ on the i^{th} day and the daily passenger flow $X_i^{0 \times 6}$ on that day, when $X_i^{1 \times 1}$ is equal to z , the PDP function can be expressed as follows:

$$f(X_i^{1 \times 1} = z) = \frac{1}{n} \sum_{i=1}^n N_{\text{GRU}} \left(\begin{array}{cccc} X_i^{1 \times 1} = z & \cdots & X_i^{1 \times 5} & X_i^{1 \times 6} \\ \cdots & \cdots & \cdots & \cdots \\ X_{i-7}^{8 \times 1} & \cdots & X_{i-7}^{8 \times 5} & X_{i-8}^{8 \times 6} \end{array} \right), \quad (3)$$

where z is the possible value of the number of new cases $X_i^{8 \times 1}$, based on the statistical results listed in Table 1, $z = 0, 1, \dots, 6$. n is the number of samples, and according to the technical route shown in Figure 4, the full dataset is used to draw the PDP; therefore, $n = 572$. N_{GRU} is the GRU neural network trained with the full sample.

The PDP reflects how the predicted mean daily passenger flow changes when a certain input variable changes. The PDP analyzes the quantitative change in the output variable with the input variable and intuitively analyzes the causal relationship between them, increasing the explanatory power of the machine learning model [28]. However, the independence assumption is the main problem with the PDP. According to its principle, the premise of PDP is that there is no correlation between the model input and output, while in the real world, there are almost no two variables that are completely independent of each other. In this study, variables with a poor correlation with the daily passenger flow are considered to be used in the PDP calculation. Considering the strong correlation between the minimum and maximum temperatures, as well as the time series of the temperature itself, the analysis of the minimum and maximum temperatures is not included in the subsequent analysis.

It should be noted that besides PDP, there are also other methods that can be utilized for model interpretation, such as attention mechanism and saliency map. The attention mechanism enables the model to selectively focus on a certain part of the input by adding an attention layer to the network. The attention layer assigns a global alignment weight to the hidden layer of the encoder network,

indicating which input component should be allocated with more attention. However, though the attention mechanism is able to indicate which input component is more important, it is not able to interpret the quantitative relationship between input and output with realistic meanings as PDP does. The saliency map is a virtualization technique that can highlight the most important regions or features of an input that contribute most to the output prediction, which is widely used in image captioning and object detection. The method used by saliency maps for identifying the important features is based on gradient calculation, the mechanism of which is similar to PDP, that is, to add a disturbance to the input and then examine the changes brought to model prediction results. Thus, PDP is finally chosen as the model interpretation method due to its realistic significance.

4. Results and Discussion

4.1. Model Training and Prediction. Based on the training set (2020), the GRU neural network was used for iterative training, and the training results of the LSTM network were used for comparison. Figure 6 shows the change in the MSE with the iteration number during the iteration. Notably, since the standardized variables are used in the network training process, the value range of the MSE is $[0, 1]$. As shown in Figure 6, both the GRU and LSTM achieve good convergence after 100 iterations. However, the LSTM fluctuates more at the beginning, and the MSE decreases to the same level as that in the case of the GRU after 80 iterations. In contrast, the GRU corresponds to a smoother curve that converges quickly at the beginning of training and stabilizes after 40 iterations. Therefore, the GRU neural network benefits from its simplicity and effectiveness and outperforms the LSTM in terms of the convergence speed and stability in the proposed daily passenger flow prediction problem.

Using the trained GRU neural network model, the prediction of the testing set was conducted. Based on the same testing set, the prediction result of the GRU was compared with the results outputted by LSTM, SARIMA, and SARIMAX (SARIMA with the exogenous factors). SARIMA is a conventional autoregressive model for time-series prediction, while exogenous variables are added to the SARIMAX model on the basis of SARIMA [5]. The SARIMA model can be expressed as $(p, d, q) \times (P, D, Q)_s$, where p, d, q, P, D, Q , and s are the orders of the model. Akaike's Information Criteria (AIC) are used to determine the optimal order set for SARIMA: $(0, 1, 6) \times (0, 1, 1)$.

Figure 7 shows the actual passenger flows and the results predicted using the aforementioned four models. As observed, both the GRU and LSTM can represent the periodic variation in the daily passenger flow on a weekly basis and effectively reflect the sudden changes in the passenger flow generated by holidays or new cases. In comparison, the SARIMA and SARIMAX show a relatively poor performance in the case of such sudden changes. This suggests that nonparametric models, such as the GRU and LSTM, show better adaptation to external emergency events such as

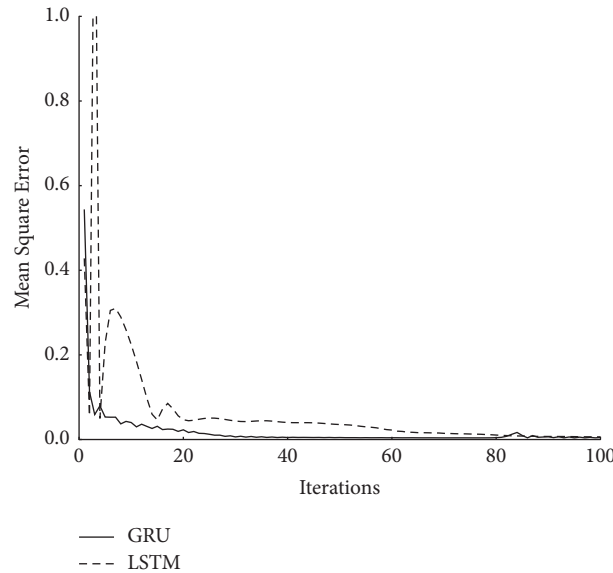


FIGURE 6: Training set iteration process.

COVID-19 new cases. In addition, the passenger flows predicted by the GRU are closer to the true value, with a mean square error (MSE) of 40.76 million person-time² and a prediction accuracy of 95.25% (calculated by $(1/n)\sum_{i=0}^n (|Y_i - \hat{Y}_i|/Y_i)$, where n is the sample size, Y_i is the actual value of Sample i , and \hat{Y}_i is the predicted value of Sample i). In contrast, the LSTM has a slightly greater deviation in the prediction of local peaks, with an MSE of 49.21 million person-time² and a prediction accuracy of 94.40%, while the MSE values of SARIMA and SARIMAX are much greater, respectively, reaching 194.64 million person-time² and 187.24 million person-time². The results suggest that the GRU has the best prediction performance.

To explore the effects of new cases on the prediction performance, a new GRU model was trained and tested on the same dataset after removing “new cases” from the independent variables. The MSE of the new model (without new cases as the independent variable) was 41.39 million person-time², which was higher than that obtained by the original model (40.76 million person-time²) (with new cases as the independent variable). This indicates that the prediction performance of the original model is improved when taking new cases as the independent variable.

4.2. Partial Dependence Plot (PDP). The partial dependence diagrams of the external factors are obtained, and the training and test sets are combined to train the GRU model and plot the partial dependence of the daily passenger traffic on each input variable through PDP. The premise of the PDP calculation is that the two variables are independent of each other, while two completely independent variables are practically unavailable. In this study, it is considered that the variables that are not strongly correlated with the other input variable can be used in the calculation of the PDP. A lack of strong correlation is defined when the absolute value of Pearson’s correlation coefficient is less than 0.5 [29]. After

the analysis, both the number of new cases and the weather attribute of each time step satisfy this precondition. The minimum temperature, maximum temperature, and holiday attribute have strong correlations in their own time series, and the minimum and maximum temperatures of the same time step are highly correlated with each other; therefore, these three variables were excluded from the subsequent analysis.

Tables 3 and 4 show the correlations of the daily number of new cases $X_i^{1 \times 1}$ and the weather attribute $X_i^{1 \times 2}$ with the other input variables on the i^{th} day, respectively. As shown in Table 3, the table headers represent the time step and feature dimension corresponding to the variable for which the correlation coefficient is calculated with $X_i^{1 \times 1}$. The table contents represent the absolute value of Pearson’s correlation coefficient between $X_i^{1 \times 1}$ and that variable. If $t = 2$ and $n = 2$, the corresponding value is 0.04, which means that the correlation coefficient between $X_i^{1 \times 1}$ and variable $X_{i-1}^{2 \times 2}$ (weather attribute of time step 2, i.e., weather attribute on the $(i - 1)^{\text{th}}$ day) is 0.04. From Tables 3 and 4, the correlation coefficients of the variables $X_i^{1 \times 1}$ and $X_i^{1 \times 2}$ with the other input variables are less than 0.5, except for the correlation coefficient of 1 with itself, which is low.

The PDP in Figure 8 shows the relationship between the daily passenger flow and the number of new cases at each time step, where t represents the time step. Based on the data structure shown in Figure 3, for the variable “daily number of new cases,” $t = 1$ corresponds to the number of new cases on that day, $t = 2$ corresponds to the previous day, and $t = 8$ corresponds to a week ago.

As shown in Figure 8, there is a negative correlation between the daily passenger flow and the number of new cases, regardless of the time step. The number of new cases on the current day ($t = 1$) and the previous day ($t = 2$) has the greatest effect on daily passenger flow. As the number of new cases increases, the daily passenger flow decreases significantly, suggesting that rail transit trips will decrease if there

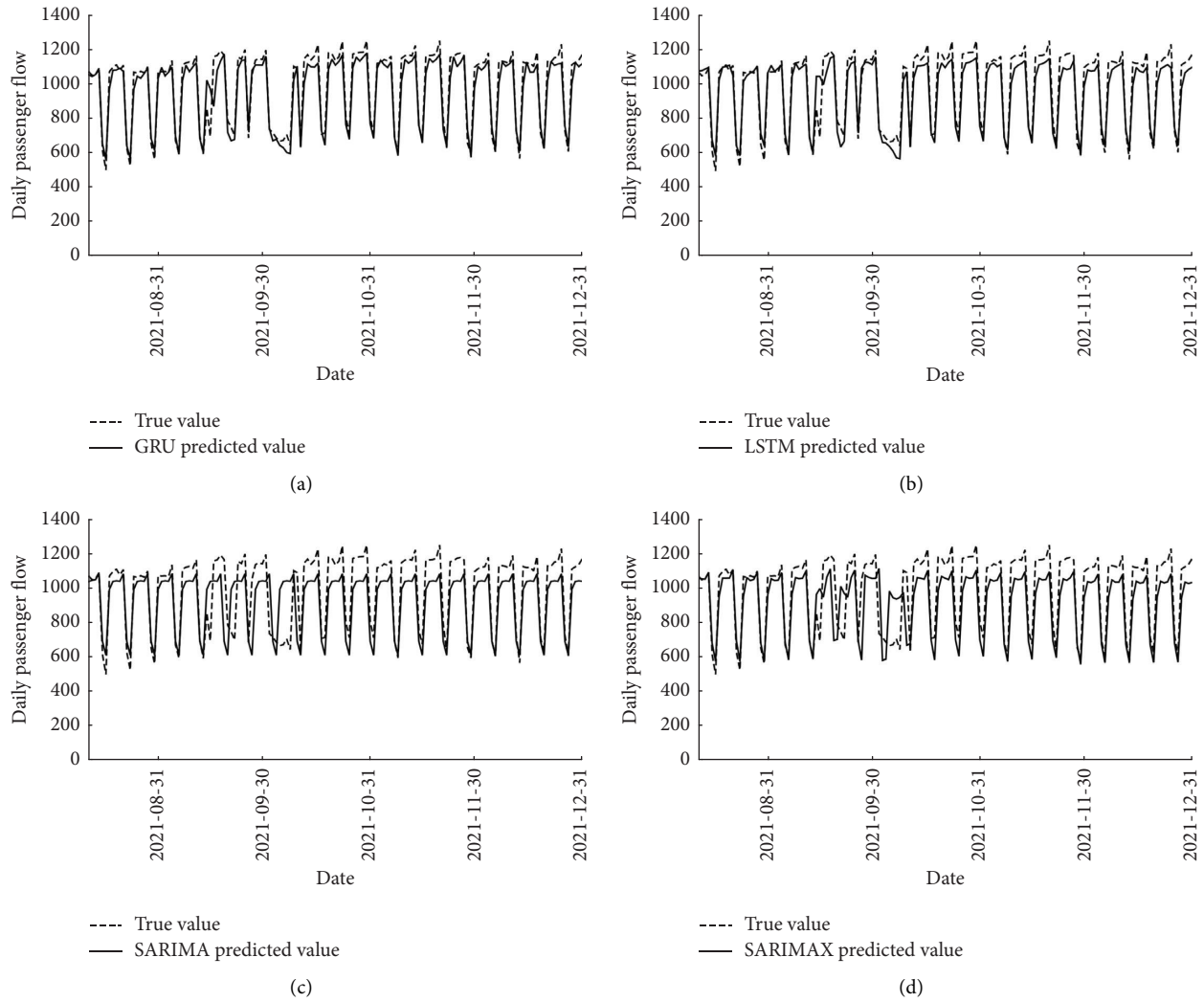


FIGURE 7: Prediction result of the test set. (a) GRU. (b) LSTM. (c) SARIMA. (d) SARIMAX.

TABLE 3: Correlation coefficients of new cases and other input variables.

Timestep	New daily cases $n = 1$	Weather $n = 2$	Minimum temperature $n = 3$	Maximum temperature $n = 4$	Holiday $n = 5$	Daily passengers $n = 6$
$t = 1$	1.00	0.07	0.11	0.11	0.04	0.00
$t = 2$	0.39	0.04	0.10	0.11	0.05	0.01
$t = 3$	0.29	0.01	0.11	0.11	0.05	0.06
$t = 4$	0.21	0.01	0.10	0.11	0.05	0.09
$t = 5$	0.17	0.02	0.11	0.12	0.02	0.04
$t = 6$	0.13	0.02	0.13	0.13	0.07	0.01
$t = 7$	0.07	0.05	0.12	0.14	0.02	0.03
$t = 8$	0.07	0.07	0.11	0.13	0.04	0.02

TABLE 4: Correlation coefficients of weather attributes and other input variables.

Timestep	New daily cases $n = 1$	Weather $n = 2$	Minimum temperature $n = 3$	Maximum temperature $n = 4$	Holiday $n = 5$	Daily passengers $n = 6$
$t = 1$	0.07	1.00	0.20	0.29	0.05	0.01
$t = 2$	0.08	0.46	0.25	0.33	0.01	0.01
$t = 3$	0.07	0.26	0.25	0.30	0.04	0.03
$t = 4$	0.04	0.20	0.26	0.28	0.02	0.02
$t = 5$	0.03	0.13	0.27	0.28	0.01	0.01
$t = 6$	0.01	0.11	0.26	0.27	0.00	0.01
$t = 7$	0.04	0.12	0.26	0.25	0.04	0.03
$t = 8$	0.04	0.07	0.26	0.25	0.04	0.01

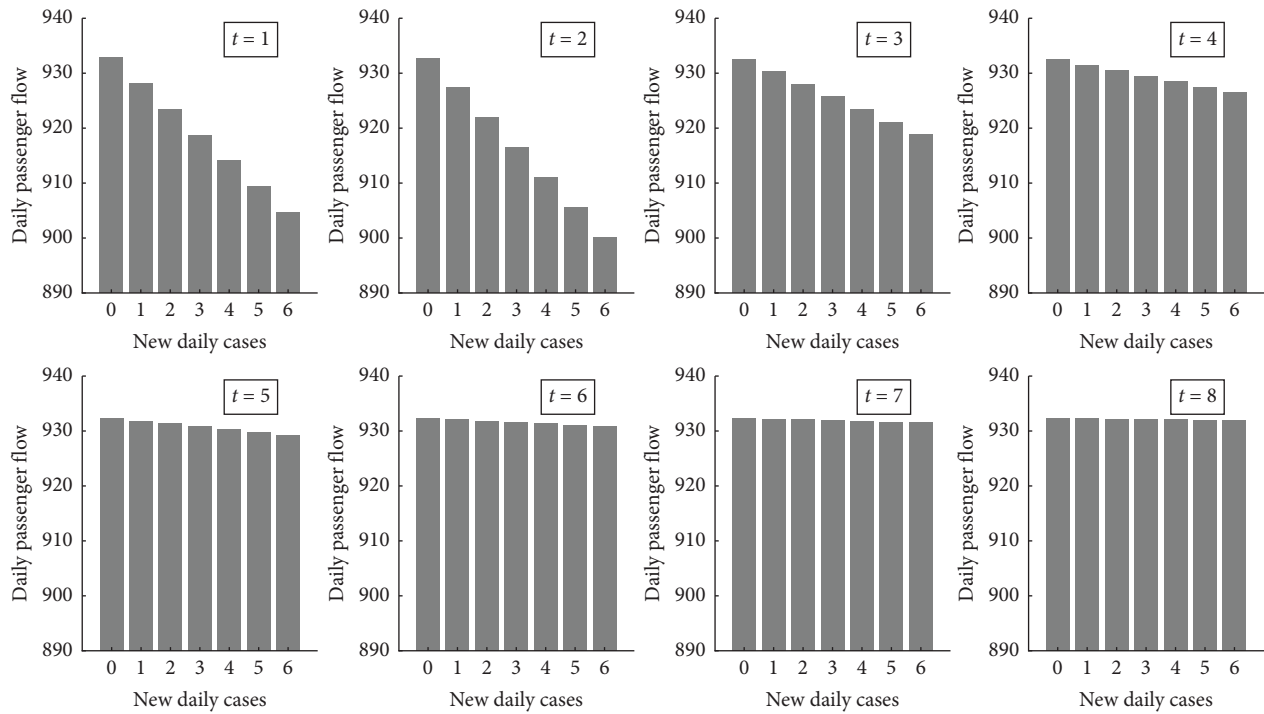


FIGURE 8: Partial dependence plots between daily passenger flow and new cases.

are local new cases. Quantitatively, for each additional number of new cases on the previous day, the daily passenger flow decreases by 54,600 person-times on average, while for each additional daily number of new cases, the daily passenger flow decreases by 46,800 person-times on average. Hence, the number of new cases yesterday has a greater impact on passenger flow today, while the passenger flow today is slightly less sensitive to the pandemic on that day, suggesting that rail transit on that day will be adjusted to a greater extent based on the pandemic situation yesterday.

As time passes, the number of new cases from the previous day ($t = 3, 4, \dots, 8$) will have an increasingly small impact on today's passenger flow.

Likewise, Figure 9 shows the PDP of the daily passenger flow with respect to the weather attribute at each time step. As shown, the weather attribute of the current day ($t = 1$) has the greatest effect on the daily passenger flow. With increasing time steps, the effect of the weather attribute on the daily passenger flow decreases. Compared with the case of no rain (weather attribute = 0), the average daily passenger

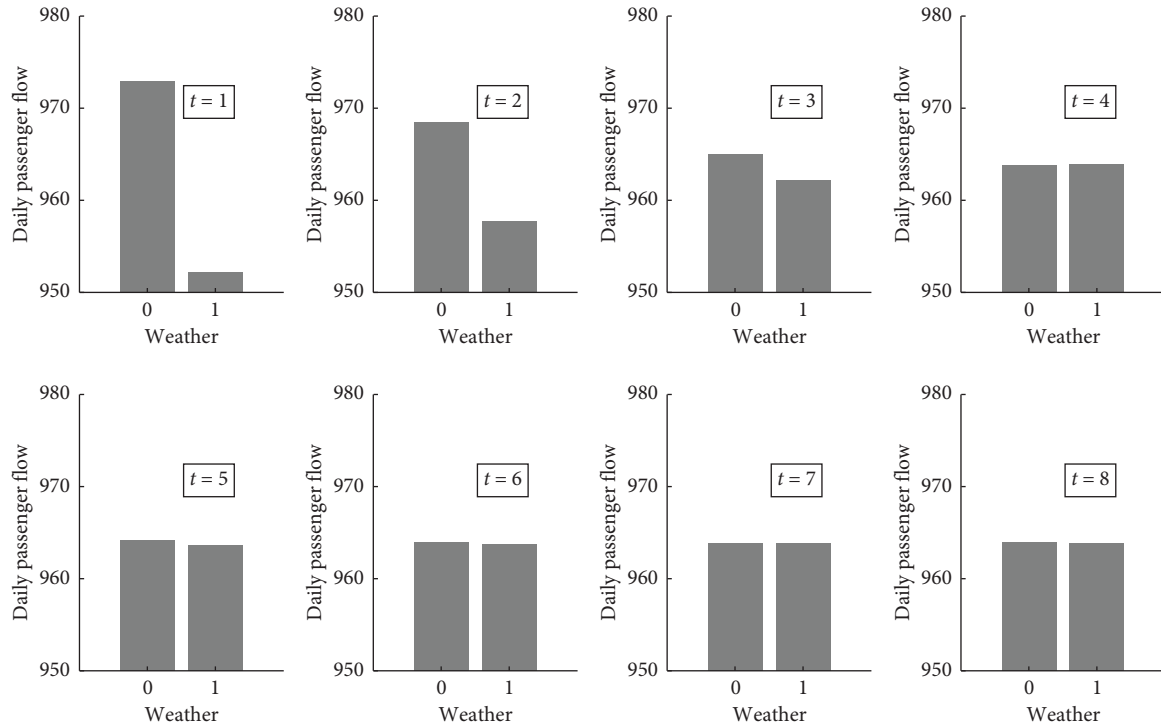


FIGURE 9: PDPs of the daily passenger flow and weather attribute.

flow decreases by 207,600 person-times when there is rain (weather attribute = 1), suggesting that rainy days significantly reduce the number of rail transit trips.

5. Conclusions

Based on a GRU neural network model, the daily number of new cases, weather attribute, temperature, holiday attribute, and historical passenger flow were used as input parameters to predict the daily passenger flow in urban railways in the postpandemic era. The results showed that the GRU neural network can produce an accurate prediction of the daily passenger flow and exhibit faster convergence and lower MSE than the LSTM neural network, consistent with previous studies (Hou et al. [20]), further demonstrating that the GRU, as a variant of the LSTM, retains the ability of the LSTM in dealing with long-term dependence problems, while converging and stabilizing faster owing to its simplified structure.

Based on the trained GRU neural network model, a partial dependency graph of the daily passenger flow, daily number of new cases, and weather attribute was drawn. The results showed that (1) daily passenger flow was negatively correlated with the number of new cases. In all the eight time steps, the number of new cases yesterday had the greatest impact on the daily passenger flow. For each additional case, the daily passenger flow decreased by 54600 person-time on average. (2) The weather attribute of the day also significantly influenced the daily passenger flow. The daily passenger flow on rainy days decreased by 207600 person-time on average compared with that on nonrainy days.

In the postpandemic era, the previously established daily passenger flow prediction model is no longer applicable; the number of new cases should be incorporated as an influencing factor to effectively predict and prevent the impact of future pandemic situations on urban rail transit. To the best of our knowledge, this is the first study to use the daily number of local new COVID-19 cases for daily passenger flow prediction. The quantitative relationship between the number of new cases and daily passenger flow was investigated using PDP while using a nonparametric model for an accurate prediction. Both the research method and the results provide important references for subsequent studies.

The shortcomings of the paper and suggestions for future work are explained as follows: (1) as time passes, the public will become more tolerant to health-related emergencies, and the coping mechanisms of society will improve. Whether the number of new cases will have a lesser impact on daily passenger flow and whether the proposed model will be applicable to passenger flow prediction in the long term remain to be verified. Future studies can consider adopting the methodology of this paper to model and analyze the daily passenger flow in different periods and mining the variation laws of the pandemic's impact on the daily passenger flow. (2) The PDP is limited by the independence assumption, and the quantitative relationship between all the input variables and daily passenger flow cannot be explored accurately. The impact of the remaining input variables on passenger flow prediction performance can be evaluated in combination with the variable importance calculation methods like attention mechanism and gradient calculation, which can also provide a further insight into the model's prediction behavior. (3) The proposed

methodology can be further extended to other transportation modes (e.g., intercity rail transit and bus travel) to analyze the impact of the pandemic on urban transportation travel structures.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was jointly supported by research grants from the National Social Science Foundation of China (18CFX062) and Research Start-Up Foundation for High-Level Talents of Jinling Institute of Technology (jit-rcyj-202304).

References

- [1] Y. Zhang, "Passenger flow impact analysis of COVID-19 on Shanghai rail transit," *Modern Urban Transit*, vol. 12, pp. 66–71, 2021.
- [2] X. Zhu, X. Ao, Z. Qin et al., "Intelligent financial fraud detection practices in post-pandemic era," *Innovation*, vol. 2, no. 4, Article ID 100176, 2021.
- [3] J. R. Saura, D. Ribeiro-Soriano, and P. Zegarra Saldaña, "Exploring the challenges of remote work on Twitter users' sentiments: from digital technology development to a post-pandemic era," *Journal of Business Research*, vol. 142, pp. 242–254, 2022.
- [4] Y. Wang, B. Han, and Q. Zhang, "Forecasting of entering passenger flow volume in Beijing subway based on SARIMA model," *Journal of Transportation Systems Engineering and Information Technology*, vol. 15, no. 6, pp. 205–211, 2015.
- [5] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *European Transport Research Review*, vol. 7, no. 3, pp. 21–29, 2015.
- [6] M. Milenković, L. Švadlenka, V. Melichar, N. Bojović, and Z. Avramović, "SARIMA modelling approach for railway passenger flow forecasting," *Transport*, vol. 33, no. 5, pp. 1–8, 2016.
- [7] J. Li, Y. L. Bi, X. N. Shen et al., "Cerebrospinal fluid α -synuclein predicts neurodegeneration and clinical progression in non-demented elders," *Translational Neurodegeneration*, vol. 9, no. 1, pp. 41–51, 2020.
- [8] C. Cheng, Y. C. Du, and X. Liu, "A passenger volume prediction model of transportation hub considering holiday effects," *Journal of Transportation Systems Engineering and Information Technology*, vol. 15, no. 5, pp. 202–207, 2015.
- [9] M. L. Xu, X. Fu, J. Y. Tang, and Z. Liu, "Effects of weather factors on the spatial and temporal distributions of metro passenger flows: an empirical study based on smart card data," *Progress in Geography*, vol. 39, no. 1, pp. 45–55, 2020.
- [10] L. Bai, L. Yang, D. P. Duan, X. Y. Xu, Z. Li, and Y. M. Liu, "Urban rail transit normal and abnormal short-term passenger flow forecasting method," *Journal of Transportation Systems Engineering and Information Technology*, vol. 35, no. 2, pp. 127–130, 2017.
- [11] L. Liu and R. C. Chen, "A novel passenger flow prediction model using deep learning methods," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 74–91, 2017.
- [12] M. Li, J. Li, and Z. J. Wei, "Short-time passenger flow forecasting at subway station based on deep learning LSTM structure," *Urban Mass Transit*, vol. 21, no. 11, pp. 42–46+77, 2018.
- [13] C. Ding, S. Q. Ni, and H. X. Lu, "Forecast and analysis of urban rail transit passenger flow based on gradient boosting," *Urban Mass Transit*, vol. 21, no. 9, pp. 60–63, 2018.
- [14] K. Zhu, P. Xun, W. Li, Z. Li, and R. Zhou, "Prediction of passenger flow in urban rail transit based on big data analysis and deep learning," *IEEE Access*, vol. 7, pp. 142272–142279, 2019.
- [15] T. J. Wei, X. Q. Yang, and G. M. Xu, "Medium-term forecast method for daily passenger flow of high-speed railway based on DLP-WNN," *China Railway Science*, vol. 42, no. 6, pp. 194–204, 2021.
- [16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [17] J. Teng and J. Y. Li, "Short-term forecast method for intercity railway passenger flow considering date attributes and weather factors," *China Railway Science*, vol. 41, no. 5, pp. 136–144, 2020.
- [18] L. Liu, R. C. Chen, and S. Zhu, "Impacts of weather on short-term metro passenger flow forecasting using a deep LSTM neural network," *Applied Sciences*, vol. 10, no. 8, p. 2962, 2020.
- [19] X. Yang, Q. Xue, M. Ding, J. Wu, and Z. Gao, "Short-term prediction of passenger volume for urban rail systems: a deep learning approach based on smart-card data," *International Journal of Production Economics*, vol. 231, Article ID 107920, 2021.
- [20] X. Y. Hou, L. P. Shao, and J. Li, "Urban rail transit short-time passenger flow OD forecasting based on deep learning modeling," *Urban Mass Transit*, vol. 23, no. 1, pp. 55–58, 2020.
- [21] H. C. Huang, J. Y. Chen, and S. Wang, "Multi-factor rail transit passenger flow prediction model," *Journal of East China Jiaotong University*, vol. 38, no. 3, pp. 61–66, 2021.
- [22] B. Shao, M. Li, Y. Zhao, and G. Bian, "Nickel price forecast based on the LSTM neural network optimized by the improved PSO algorithm," *Mathematical Problems in Engineering*, vol. 2019, Article ID 1934796, 15 pages, 2019.
- [23] J. Chung, C. Gulcehre, and K. H. Cho, "Empirical evaluation of gated recurrent neural networks on sequence modeling," 2014, <https://arxiv.org/abs/1412.3555>.
- [24] D. P. Kingma and B. A. J. Adam, "A method for stochastic optimization," 2014, <https://arxiv.org/abs/1412.6980>.
- [25] P. M. Johnson, W. Barbour, J. V. Camp, and H. Baroud, "Using machine learning to examine freight network spatial vulnerabilities to disasters: a new take on partial dependence plots," *Transportation Research Interdisciplinary Perspectives*, vol. 14, Article ID 100617, 2022.
- [26] M. Hassani Niaki, M. Ghorbanzadeh Ahangari, and M. Pashaian, "A material-independent deep learning model to predict the tensile strength of polymer concrete," *Composites Communications*, vol. 36, Article ID 101400, 2022.
- [27] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [28] Q. Zhao and T. Hastie, "Causal interpretations of black-box models," *Journal of Business & Economic Statistics*, vol. 39, no. 1, pp. 272–281, 2021.
- [29] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Routledge, London, UK, 2013.