



The HWAM-EMD-GRU Forecasting Model for Short-Term Passenger Flow in an Airport Light Rail Transit Line

Qian Qin¹ · Ziji'an Wang² · Bing Li² · Ailing Huang²

Received: 21 September 2023 / Revised: 4 February 2024 / Accepted: 16 February 2024 / Published online: 3 May 2024
© The Author(s) 2024

Abstract Accurate forecasting of airport light rail transit line (ALRTL) outbound passenger flow is critical to the optimal operations of both light rail and airport systems. Considering the nonlinearity, non-stationarity, uncertainty, and periodicity of outbound passenger flow in the ALRTL, we propose a combined forecasting model that integrates the Holt and Winters additive model (HWAM), empirical mode decomposition (EMD) and gated recurrent unit (GRU). Firstly, the edge effect of EMD will greatly affect the performance of the forecasting model. To overcome this, we extend the passenger flow by HWAM. After that, the decomposition method, EMD, can be applied to passenger flow, and several intrinsic mode function (IMF) components can be extracted. After extracting all the IMFs, the remaining part is referred to as the residual (Res) component. Then, a correlation test is performed on all the components, followed by their aggregation. Finally, the GRU is used to predict each of the aggregated components, and the prediction of aggregated components requires reconstruction. To verify the performance of the HWAM-EMD-GRU, we conducted a comparative study on the hourly passenger flow data for Beijing Daxing International Airport Express and set the autoregressive integrated moving average model, HWAM, Prophet,

and GRU as the baseline. Predictions of the HWAM-EMD-GRU combined model demonstrated higher accuracy than baseline models, with a root mean square error of 83.52 (Prophet is 110.21) and mean absolute percentage error of 8.32% (Prophet is 12.48 %). The experimental result shows that the HWAM-EMD-GRU forecasting model offers more accurate predictions.

Keywords Passenger flow forecasting · Airport light rail transit line (ALRTL) · Empirical mode decomposition (EMD) · Gated recurrent unit (GRU) · Holt–Winters additive model (HWAM)

1 Introduction

The number of air passenger trips has continued to increase in recent years. In 2023, the number of Chinese air passengers exceeded 620 million. The airport light rail transit line (ALRTL), as one of the airport's land-based transportation methods, typically handles 10–30% of passenger transfer operations [1], for example reaching 32.1% at Beijing Daxing International Airport [2]. The growing demand for air travel has posed a challenge for short-term forecasting in terms of passenger flow, which is a crucial aspect of ALRTL operations [3]. In urban rail transit systems, the short-term forecasting of passenger flow has garnered significant attention, but research on the ALRTL is still insufficient. Passenger flow is characterized by nonlinearity, non-stationarity, uncertainty, and periodicity [4]. However, due to the different composition, passenger flow differs between the ALRTL and normal urban rail transit, with strong weekend/weekday differences.

The passenger flow on the ALRTL can be further categorized into inbound flow and outbound flow. The inbound

✉ Ziji'an Wang
wangzijian_bjtu@163.com

✉ Ailing Huang
alhuang@bjtu.edu.cn

¹ International Technology Department of Capital Airports Holding Company, Beijing 100621, China

² Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China

Communicated by Baoming Han.

direction is linked to the arrival flight wave, which is usually used as the main feature for prediction, while the correlation between outbound passenger flow and departure flight waves is weaker, is more challenging, and requires further research.

Forecasting the outbound passenger flow of ALRTL is meaningful for both passengers and airport managers. The forecast results reflect the congestion level of the ALRTL, as the airport is one of the main transfer nodes, which may help passengers in planning travel in advance [5]. Meanwhile, the forecast results help airport managers deploy security measures in advance and avoid the risk presented by large passenger flow.

From the perspective of short-term passenger flow forecasting, the primary research avenues include time series techniques, artificial intelligence, and hybrid models combining multiple methods. The time series technique applies smoothing technology to reveal the potential trend, periodicity, and seasonality of data, and includes the autoregressive integrated moving average model (ARIMA) [6] and exponential smoothing methods [7]. In recent years, artificial intelligence methodologies such as statistical learning techniques [8] and artificial neural networks [9, 10] have gained widespread use. Liu et al. [11] introduced a forecasting model in which particle swarm optimization (PSO) was used for the long short-term memory (LSTM) model, to make predictions based on the automated fare collection (AFC) data. Within this model, PSO is employed for optimizing the hyperparameters of the LSTM, contributing to enhanced predictive performance. A study by Du et al. [12] still primarily revolves around the LSTM model. The combination of LSTM with deep irregular convolution residual (DST-ICRL) considers the irregularity of multiple channels. Artificial intelligence methods are more widely used and can adaptively learn parameters from data. Nonetheless, they often demand substantial amounts of data for effective learning.

To address the limitations of single models and enhance predictive performance, researchers in recent years have shifted their attention toward the integration of various models [13–15]. Combined models are those in which two or more models are combined to achieve higher accuracy in passenger flow prediction by learning from each other. This approach aims to leverage the strengths of different models by combining them, thereby achieving more robust and accurate predictions.

In multidimensional data scenarios, the various types of data must be processed by different models and then reproduced organically to formulate combined models. Ma et al. [16] used the daily and historical passenger flow data as input, applied the Kalman filter algorithm followed by the K-nearest neighbor fusion algorithm to formulate predictions, and weighted and regrouped the prediction results to construct a forecasting model. He et al. [17] introduced a

combined model of multi-image convolution and a recurrent neural network (RNN) considering the temporal–spatial correlation between the stations of the urban metro to make predictions about passenger flow.

In one-dimensional data scenarios, the combined model uses different methods at different stages. Several scholars have proposed prediction models that combine coarse-grained and fine-grained data to obtain more accurate results locally. Jing et al. [18] proposed a hybrid model for short-term forecasting, including a lightweight gradient decision-making enhanced tree, LSTM, and dynamic regression approaches, for passenger flow at Chengdu East Railway Station.

The application of decomposition techniques to passenger flow time series is also a research hotspot, in which a combination of decomposition algorithms and predictive models are typically used, and where the decomposition algorithm is used to decompose simple periodic data from complex passenger flow time series, thereby reducing the cost of training. Fu et al. [19] presented a combination of ensemble empirical mode decomposition (EEMD) and backpropagation (BP) neural network approaches to forecast incoming passenger flow in a subway station. Li et al. [20] studied monthly granular passenger flow at railway stations and built the basic model with extreme learning machine (ELM) and BP. Then, pairing them in combination with empirical mode decomposition (EMD) or variational mode decomposition (VMD), the VMD-BP method performed the best in the comparison experiment. Chen et al. [21] used an EMD-LSTM hybrid model to forecast passenger flow, in which Kendall correlation analysis was used to select components with higher correlations.

To summarize, the combined model exhibits superior performance relative to the individual model. In this work, we concentrate on the short-term forecasting of passenger flow with nonlinearity, uncertainty, and periodicity in the ALRTL, and we propose a forecasting model that combines the Holt–Winters additive model (HWAM), EMD, and gated recurrent unit (GRU) model, which we call HWAM-EMD-GRU. Firstly, HWAM is used to extend passenger flow on the right side (the historical dataset is to extend the left side) to eliminate the edge effect of EMD. Secondly, the EMD method is applied to decompose passenger flow and obtain decomposed components including the intrinsic mode functions (IMFs) and residual (Res). Then, we use correlation analysis to incorporate low-correlation IMFs into the Res to form aggregated components. The aggregated components have simpler periodicity, which is easier to predict by general predictive models. Finally, the GRU network is employed for forecasting each component, and the predictions are aggregated and reconstructed to finalize predictions. To verify the predictive performance of the HWAM-EMD-GRU model, we performed empirical studies. Hourly data for outbound passengers from

the Daxing Airport Station spanning 12 days were selected for this empirical study.

The remainder of this paper is structured as follows. In the next section, we propose a HWAM-EMD method to eliminate the edge effects and decompose the passenger flow. In Section III, we introduce the HWAM-EMD-GRU forecasting model and provide a general forecasting process. In this model, the GRU, as the predictive module, is used to forecast each component of the original passenger flow, which will be reconstructed to obtain the forecasting value. In Section IV, we conducted a case study of Beijing Daxing International Airport Express Daxing airport station to evaluate the predictive performance of the combined HWAM-EMD-GRU model. Section V presents the conclusion.

2 Hybrid HWAM-EMD Decomposition Method

In this section, we employ the EMD method to decompose the passenger flow to extract the underlying periodicity and trend, which can facilitate passenger flow forecasting. However, the edge effect of EMD will lead to data drift at both ends of the IMF components. To address this issue, we utilize the HWAM technique to extend the passenger flow data. Subsequently, we discard the extended part after the decomposition process to mitigate the impact of the edge effect.

2.1 Triple Exponential Smoothing Addition Model

Triple exponential smoothing is a time series prediction method proposed by Holt and Winters [22], which is widely used in time series forecasting when the series has trend characteristics and seasonal characteristics. The seasonal feature takes the operating time of the ALRTL in one day as a cycle. This feature remains roughly constant over time, so we chose the additive model, called the Holt–Winters additive model (HWAM), whose core formula is as follows:

$$\hat{y}_{i+h|i} = l_i + hb_i + s_{i-m+1+[(h-1) \bmod m]} \quad (1)$$

The level component l_i is expressed as

$$l_i = \theta_l(y_i - s_{i-m}) + (1 - \theta_l)(l_{i-1} + b_{i-1}) \quad (2)$$

The trend component t_i is expressed as

$$t_i = \theta_t(l_i - l_{i-1}) + (1 - \theta_t)b_{i-1} \quad (3)$$

The seasonal component s_i is expressed as

$$s_i = \theta_s(y_i - l_{i-1} - b_{i-1}) + (1 - \theta_s)s_{i-m} \quad (4)$$

The initial value condition is expressed as

$$s_0 = y_0 \quad (5)$$

where i, h is the index of time series, y_i is the historical passenger flow at time i , $\hat{y}_{i+h|i}$ is the prediction at time $i + h$, which is predicted using the data at time i , and m is the seasonal frequency; as there are 52 weeks in a year, $m = 52$. $\theta_l, \theta_t, \theta_s$ are smoothing parameters, which are level, trend, and seasonal, respectively.

2.2 Empirical Mode Decomposition Method

The EMD [23] technique can adaptively decompose a signal into its constituent IMFs and a Res. The IMFs capture the different scales or modes of variation present in the data. The Res represents the remaining trend or low-frequency components that cannot be further decomposed into IMFs.

Application of the IMFs satisfies the following two characteristics [23]:

2.2.1 Extreme and Zero-Crossing Points

Extreme points are local maxima and minimal points in the original signal, and zero-crossing points are the points where the signal changes sign. In EMD, the necessary conditions for a signal to be considered an IMF is that the difference between the number of extreme and zero-crossing points should be at most 1.

2.2.2 Upper and Lower Envelopes

EMD also involves the determination of upper and lower envelopes. The cubic spline interpolation method is applied to identify the upper and lower envelopes of the original signal, in which the upper (lower) envelope is determined by identifying the local maxima (minima) of the data, and the average value should be zero.

The hourly passenger flow data of the ALRTL can be treated as a time series, which can be decomposed into IMFs and Res. IMFs are independent of each other, and the quantity is determined by the local characteristic time scale of the passenger flow; Res retains the trend information of the original signal and the period information beyond the current data scale. Applying EMD technology, the passenger flow data $y(\tau)$ can be decomposed into n IMFs and one residual Res, expressed as follows:

$$y(\tau) = \sum_{j=1}^n c_j(\tau) + r(\tau) \quad (6)$$

where $y(\tau)$ is a time series, τ is the index of time, $c_j(\tau)$ is the IMF, $r(\tau)$ is the Res and j is the index of the IMFs.

2.3 Empirical Mode Decomposition Edge Effect

The EMD edge effect [24] is manifested as data drift at both ends of the IMF component, since the two ends of the series may not be exactly extreme points when using cubic spline interpolation. This results in envelope distortion and affects the accuracy of the IMF components at both ends, which in turn will cause a large deviation in the subsequent prediction results of each component. To solve the edge effect of EMD, researchers have proposed a diverse range of continuation methods, including the linear regression method [25], the Coughlin method [26], and the Rato method [27]. The principle of these methods is to extend the original signal, then decompose it by EMD and delete the part of the IMFs corresponding to the range of extension, thereby reducing the influence of EMD edge effects on the signal decomposition results.

To address the specific characteristics of hourly passenger flow data on the ALRTL, this paper introduces a bidirectional extension technique to eliminate the edge effect utilizing HWAM technology and historical data. The left edge extension can be filled with one period (usually 24 hours) of historical data, and the right edge extension can be filled with predictions. Utilizing the HWAM method, we employ all available historical data in the research period (including the data of the left edge) to forecast one period of data as the extension of the right edge. The hourly passenger flow of the ALRTL has obvious periodicity, so the HWAM method can obtain higher prediction accuracy.

3 HWAM-EMD-GRU Combined Model

In this section, we propose the framework of the HWAM-EMD-GRU combined forecasting model. Using the hourly flow as the input, the framework applies the HWAM-EMD method for decomposition, generating multiple IMF components and a Res component. Next, a correlation analysis is conducted between the original passenger flow, its IMFs, and its Res. Components with high correlation are retained, while components with low correlation are merged. Finally, the individual predictions of each reconstructed component are made using the GRU. The predictions of each component are added together to obtain the prediction of the original passenger flow.

3.1 GRU Neural Network

The GRU network as proposed by Cho et al. [28] is a variant of RNN, but more lightweight, and is widely applied in predictive applications such as time series. As a variant of LSTM, GRU also overcomes the long-term dependence and the gradient in direction propagation, and has better

performance on time series prediction problems [29]. However, the GRU can achieve predictive performance comparable to LSTM with less computational overhead.

The GRU unit structure is shown in Fig. 1. Its main structure is a gate structure, which is a combination of an activation function and a matrix bitwise multiplication operator, including an update gate and a reset gate.

The function of the update gate z_t is analogous to the combination of the input gate and forget gate in LSTM, which is used to filter the information of the previously hidden layer and pass it to the next layer. The reset gate r_t is applied to determine the volume of forgotten information. The functions of these gates are expressed as follows:

$$r_t = \sigma(w_r[h_{t-1}, x_t] + b_r) \quad (7)$$

$$z_t = \sigma(w_z[h_{t-1}, x_t] + b_z) \quad (8)$$

where h_{t-1} is the hidden state of time $t - 1$, x_t is the time series, σ is the sigmoid function, w_r, w_z is the weight of the reset and update gate, and b is the bias.

The hourly passenger flow time series data of ALRTL has obvious periodicity, so RNN has great adaptability in passenger flow forecasting. Considering the limited availability of historical data in this study, we chose to use the GRU network, which is more lightweight.

Each component of the original passenger flow time series is also a time series with a relatively certain periodicity, so GRU can also be used for prediction. Since the periodic regularity after decomposition is simpler than that before decomposition, when the decomposed hourly passenger flow data is predicted by GRU, the model can be trained more fully, and better prediction accuracy can be obtained with the same amount of data.

The GRU network requires the determination of four key parameters, including lookback and the size of all layers. In

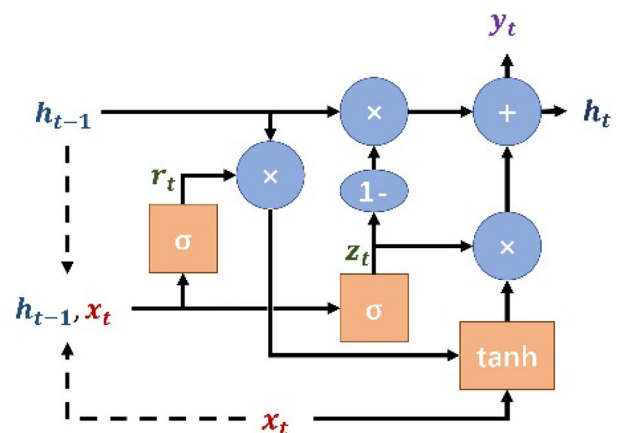


Fig. 1 The structure of the GRU cell

the input layer, the size is set to 1, which corresponds to the dimension of the features employed for prediction; in the output layer, the size is set to 1, which corresponds to the dimension of labels for prediction; in the hidden layer, the size is set according to the experiments; the lookback is the amount of historical data used for prediction, which also must be determined experimentally.

Determining the optimal values for the lookback and the size of the hidden layer is crucial for achieving high prediction accuracy. However, it is important to note that the relationship between these parameters and prediction accuracy is not linear. Additionally, increasing the number of neurons can lead to higher computational overhead. Hence, it is essential to perform repeated experiments to determine the optimal values.

To discover the optimal combination of parameters, the root mean square error (RMSE) is commonly used as the selected index, shown as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

where y_i is the actual hourly flow of the ALRTL in time i , \hat{y}_i is the predicted flow of the ALRTL in time i , and n is the time scale of the study.

3.2 HWAM-EMD-GRU Forecasting Model

The HWAM-EMD-GRU is proposed to realize the short-term forecasting of outboard passenger flow in the ALRTL. The model is composed of three basic methods: HWAM, EMD, and GRU. The EMD method decomposes the original hourly passenger flow data into multiple components with simple periodicity and a longer-term trend, which are easier to predict than the original data. The HWAM model is used to extend the EMD model to eliminate the component data distortion caused by the edge effect. The algorithm flow of the HWAM-EMD-GRU combined model is as follows:

Step 1. Time series extension of hourly flow:

Extend the flow in both directions. The left extension uses a period of historical data, and the right extension uses the prediction result of the HWAM method.

Step 2. Time series decomposition of hourly passenger flow by EMD:

Decompose the extended data into several IMFs and a Res.

Step 3. Extension part cutting of decomposed data:

Remove the extension part of the decomposed data.

Step 4. Important component identification:

Apply correlation analysis to identify important components. The Pearson correlation coefficient between the

original data and each component is calculated separately, and the components with weak correlations are merged into Res.

Step 5. Component prediction and reconstruction:

GRU is used to predict the components separately, and then the predicted values of all components are reconstructed to form the predictions.

4 Experiment

In this section, to evaluate the predictive performance of the HWAM-EMD-GRU combined model, we conducted a numerical experiment on the hourly outbound passenger flow desensitized dataset for Daxing Airport Station in Beijing Daxing International Airport Express. Firstly, we performed data cleaning on the desensitized dataset. Subsequently, following the prediction framework of the HWAM-EMD-GRU combined model, we conducted predictive experiments on the dataset. During the training process, we focused on optimizing certain hyperparameters. To determine the best combination of hyperparameters, we utilized a grid search method.

To fully validate the performance of the prediction framework, we also performed comparative experiments. We compared HWAM-EMD-GRU with the classic models GRU [28], ARIMA [6], HWAM [22], and Prophet [30] by applying these models to the same dataset for forecasting, serving as baseline models.

4.1 Passenger Flow Desensitized Dataset

The data in this article use the hourly flow of outbound passengers at the Daxing Airport Station of the Beijing Daxing International Airport Express from September 7 to 18, 2020, for a total of 12 days.

Considering the security of passenger travel information, the dataset is aggregated and desensitized. The details of the dataset are shown in Table 1. STATIONNAME is the name of the station in the Beijing Daxing International Airport Express; STARTTIME is the start time of the time slot (e.g., “2020-08-01 02” means that these data reflect the passenger flow between 2 o’clock and 3 o’clock on August 1, 2020); INSTATIONFLOW and OUTSTATIONFLOW are the

Table 1 Details of desensitized data

Field name	Data type	Example
STATIONNAME	Character	DXJC
STARTTIME	YYYYMMDD HH	2020-08-01 02
INSTATIONFLOW	Integer	159
OUTSTATIONFLOW	Integer	258

numbers of inbound and outbound passenger flow, respectively, aggregated from AFC data. Due to privacy concerns, the original AFC data have not been provided.

4.2 Desensitized Data Preprocessing

Although the data have been processed in terms of aggregation and desensitization, numerous issues still persist for technical reasons. Hence, preprocessing of the data is crucial for enhancing data quality, which in turn benefits subsequent predictions. The preprocessing is performed as follows:

4.2.1 Data Filtering

For selecting the outbound passenger flow of Beijing Daxing International Airport Express at Daxing Airport Station from the desensitized dataset, we selectively preserve the data within the operating hours (6:00–23:00) and discard the data for other periods, which consistently remain at 0.

4.2.2 Outlier Detection

The interquartile range (IQR) is used to detect the outlier, and the outlier data will be set as missing values.

4.2.3 Missing Data Imputation

The Lagrange interpolating polynomial is utilized for missing data imputation in the raw data and the missing value marked in the outlier detection step.

4.2.4 Dataset Partitioning

The desensitized passenger flow data from September 7 to 16, 2020, are used as the training set for a total of 10 days, where the data for September 7 are set as the left extension; the data from September 17 to 18, 2020, are used as the test set for a total of 2 days.

4.3 Evaluation Metrics

To assess the predictive performance of the HWAM-EMD-GRU combined model, besides RMSE, we also considered the following three performance metrics:

Mean squared error (MSE):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Mean absolute error (MAE):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

Mean absolute percentage error (MAPE) [31]:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (12)$$

4.4 Ablation Study

To investigate the impact of the HWAM method used in the forecasting model on the predictive performance, an ablation study was conducted by canceling the HWAM relative component shown in Section 3.3. Specifically, we cancel Step 1 and Step 2 in the HWAM-EMD-GRU model to obtain the EMD-GRU model. The evaluation metrics of the ablation study are shown in Table 2, and the effect of HWAM on improving the predictive performance is verified.

4.5 Experimental Process

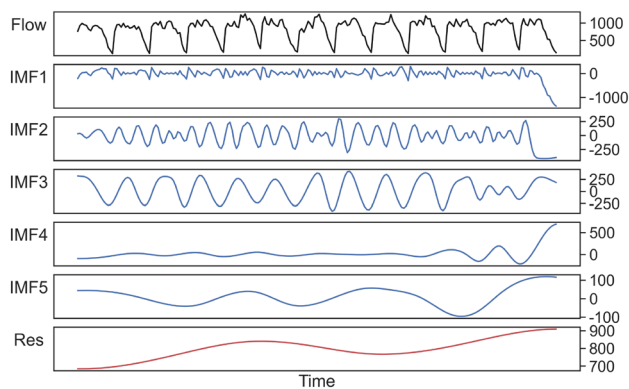
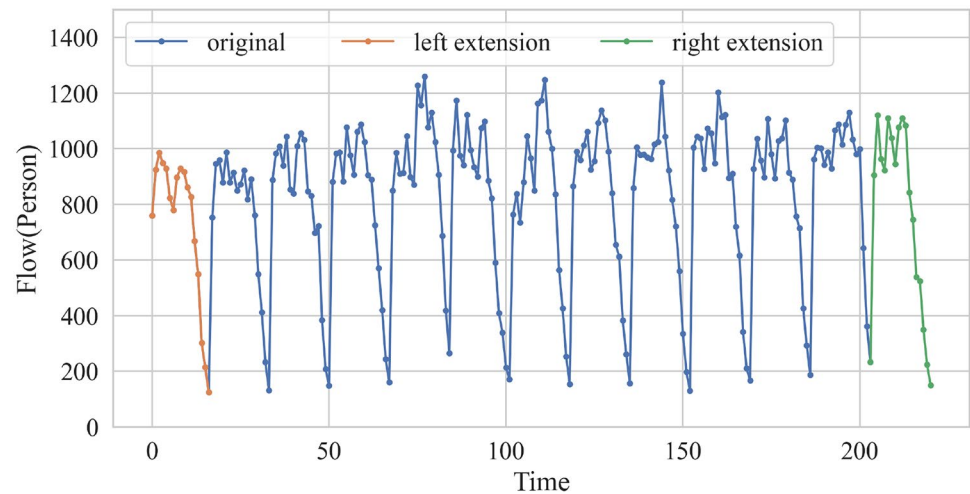
Figure 2 shows the bidirectional extension of the passenger flow by applying HWAM technology.

The extended hourly passenger flow is decomposed by EMD, as shown in Fig. 3. The hourly flow time series are decomposed into five IMFs and a Res. The IMF components show different periodic characteristics of the original series, while the Res component retains the trend. The decomposition results also further show that the edge effect of EMD, especially IMF 1 and IMF 2, are more obvious, and the left and right ends are severely distorted. Applying this distorted data will reduce the prediction accuracy when making predictions by GRU.

We compute the Pearson correlation coefficient for each component with respect to the original hourly passenger flow data and perform a correlation test as depicted in Table 3. The results show that the correlation between the various IMF components is low, which is consistent with the relatively independent characteristics of IMFs after EMD decomposition. IMF 1 to IMF 4 have strong correlations with the original data, while IMF 5 and residual Res have low correlations. Therefore, IMF 1 to IMF 4 mark input components C_1 to C_4 , respectively, and IMF 5 and residual Res will be merged and mark input component C_5 .

Table 2 Evaluation metrics of the ablation study

	MSE	RMSE	MAE	MAPE (%)
EMD-GRU	32627.20	180.63	98.75	22.03
HWAM-EMD-GRU	6976.70	83.53	64.70	11.43

Fig. 2 Bidirectional extension of original passenger flow**Fig. 3** IMF and Res of decomposed passenger flow time series

GRU is employed for predicting each component as well as directly forecasting the original series, serving as a baseline.

We set one layer of GRU hidden layers, with a dropout layer to avoid overfitting, and the random inactivation probability is set to 0.1. The excitation function is set to the linear rectification (ReLU) function. The initialization learning rate is 0.001, and Adamax is considered as the optimizer. We adopt the learning rate scaling mechanism and set the factor to 0.9 and step size to 100, which means that the learning

rate will be adjusted every 100 training steps by multiplying it by 0.9.

The parameters are used with the same settings for both component prediction and original series prediction. Other hyperparameters of each component of GRU and original series, including epoch, batch size, lookback, and the size of hidden, are determined through experiments.

We conduct a grid search experiment to determine the optimal hyperparameter combination on the training set. Specifically, we set the RMSE as evaluation metrics. The search space is set as follows: [100, 200] for epoch, [6, 12, 18, 24] for batch size, [1, 2, ..., 20] for lookback, and [8,

Table 4 Optimal hyperparameter combination for each component

Component	Epoch	Batch size	Lookback/ hour	Hidden layer size
C_1 (IMF 1)	200	12	17	8
C_2 (IMF 2)	100	12	10	16
C_3 (IMF 3)	100	6	2	8
C_4 (IMF 4)	100	6	3	8
C_5 (IMF 5+res)	100	6	5	16
Original	200	6	14	32

Table 3 Pearson correlation coefficient

	IMF2	IMF3	IMF4	IMF5	Res	Original data
IMF1	−0.02	−0.11	0.05	0.03	−0.08	0.3
IMF2	—	0.07	0.07	−0.01	−0.01	0.52
IMF3	—	—	0.00	−0.09	−0.06	0.75
IMF4	—	—	—	−0.09	−0.05	0.23
IMF5	—	—	—	—	−0.08	0.05
Res	—	—	—	—	—	0.04

16, 32, 64] for hidden layer size. The grid search results are shown in Table 4.

4.6 Comparison and Analysis

Figure 4 shows the comparison results between the HWAM-EMD-GRU and other baselines; forecasting passenger flow curves are plotted in the same figure. Except for the ARIMA model, these models exhibit favorable predictive performance, indicating their compliance with the patterns of outbound passenger flow in the ALRTL.

The evaluation metrics are shown in Table 5. In the comparison of GRU with other baselines in short-term ALRTL passenger flow, the predictive performance of GRU is worse than Prophet and HWAM in almost all indexes. This phenomenon may be caused by the small size of the training set (even though we deliberately increased the proportion of the training set during the data partitioning) and insufficient training of the artificial neural network, leading to underfitting.

Comparing the predictive performance of HWAM-EMD-GRU and other baselines, the HWAM-EMD-GRU model is superior in all indexes. The evaluation metrics of HWAM-EMD-GRU are much lower than Prophet, which is the best model among the baselines of this paper. Specifically, the MSE is 42.56% lower, the RMSE is 24.21% lower, the MAE is 16.01% lower, and the MAPE is 8.41% lower. Thus, the predictive performance of the HWAM-EMD-GRU combined model is demonstrated by the above evaluation metrics. The results also indicate that HWAM-EMD can decompose the original signal into multiple simpler signals, reducing the complexity of fitting the components and thus reducing the data volume requirements when applying GRU for prediction.

Table 5 Comparison of evaluation metrics between HWAM-EMD-GRU and baseline

	MSE	RMSE	MAE	MAPE (%)
HWAM	13958.24	118.14	83.77	14.08
ARIMA	82124.33	286.57	196.95	115.12
Prophet	12146.77	110.21	77.03	12.48
GRU	16961.05	130.23	94.85	13.52
HWAM-EMD-GRU	6976.70	83.53	64.70	11.43

5 Conclusion

EMD technology was employed to decompose the original time series of the ALRTL passenger flow into multiple components. To address the endpoint effect inherent in EMD, HWAM extension was utilized. Subsequently, GRU was employed to independently predict each component, and by reconstructing the predictions, the forecasted passenger flow of the ALRTL was obtained.

Experiments were conducted on the hourly outbound passenger flow dataset for Beijing International Daxing Express Daxing Airport Station, and verified that the accuracy of the HWAM-EMD-GRU is much higher than the baselines in the short-term passenger flow forecast. The bidirectional extension employed in this study effectively addresses the edge effect and improves the decomposition quality of each component. Another notable advantage of the HWAM-EMD-GRU model is its ability to achieve high accuracy in passenger flow prediction using minimal historical data, without the need for other features.

In future research, we will further evaluate ALRTL passenger flow forecasting at finer granularity. At the same time, it is also possible to consider collecting multidimensional

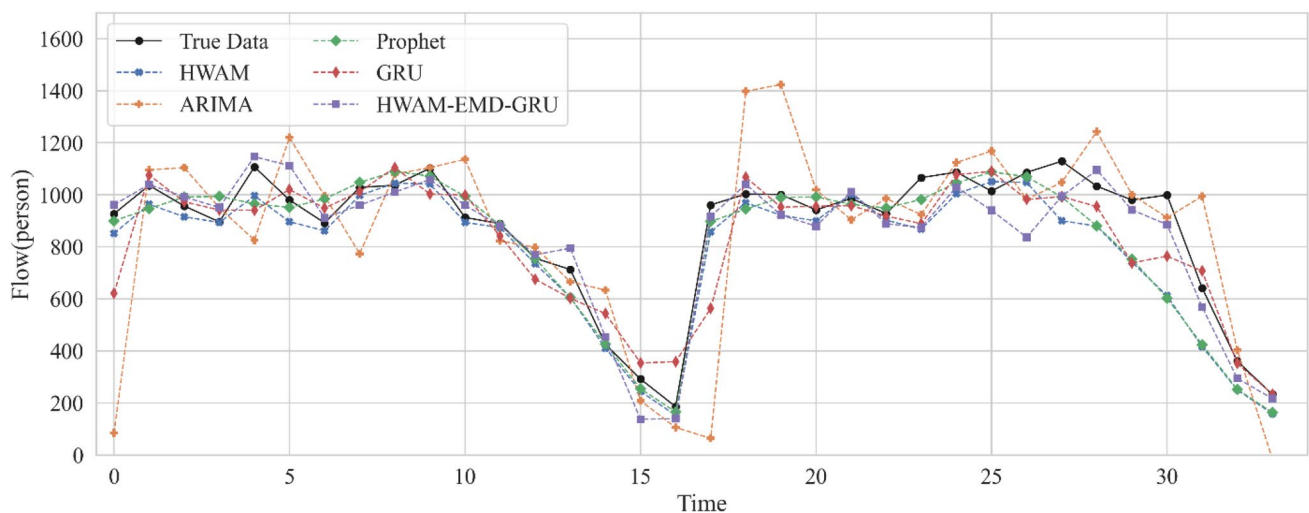


Fig. 4 Comparison of prediction results between HWAM-EMD-GRU combined mode and baseline

data and using multisource data fusion technology to obtain higher prediction accuracy.

Funding This work is supported by the National Key Research and Development Program of China (2018YFB1601200) and the National Natural Science Foundation of China (72271018).

Declarations

Conflict of interest The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Lu F, Zhou X, Qiao YL, Li WY, Wang B (2020) China's Airport Rail Lines in 2020: statistics and analysis of cooperative operation. *Urban Transp China* 33(6):1–6 (in Chinese)
- Bian Y, Jia DZ, Feng XF, Zhao JJ, Li J (2022) Characteristics of landside passenger flow of Beijing Daxing International Airport. *Urban Transp China* 20(4):63–69 (in Chinese)
- Gkiotsalitis K, Cats O, Liu T, Bult JM (2023) An exact optimization method for coordinating the arrival times of urban rail lines at a common corridor. *Transp Res Part E* 178:103265. <https://doi.org/10.1016/j.tre.2023.103265>
- Liu Y, Liu ZY, Jia R (2019) DeepPF: a deep learning based architecture for metro passenger flow prediction. *Transp Res Part C Emerg Technol* 101:18–34. <https://doi.org/10.1016/j.trc.2019.01.027>
- Deng YJ, Chen ML (2021) Impacts of real-time transit information on bus passengers' travel choices based on travel behavior survey. *Promet Traffic Transp* 33(3):539–550 (in Chinese)
- Zhang G, Jin H (2022) Research on the prediction of short-term passenger flow of urban rail transit based on improved ARIMA model. *Comput Appl Softw* 39(1):339–344 (in Chinese)
- Wang X (2019) The short-term passenger flow forecasting of urban rail transit based on Holt-Winters' seasonal method. In: 2019 4th international conference on electromechanical control technology and transportation (ICECTT). IEEE, pp 265–268. <https://doi.org/10.1109/ICECTT.2019.00067>
- Zhu K, Xun P, Li W, Li Z, Zhou RC (2019) Prediction of passenger flow in urban rail transit based on big data analysis and deep learning. *IEEE Access* 7:142272–142279. <https://doi.org/10.1109/ACCESS.2019.2944744>
- Zhang JL, Chen F, Cui ZY, Guo YN, Zhu YD (2020) Deep learning architecture for short-term passenger flow forecasting in urban rail transit. *IEEE Trans Intell Transp Syst* 22(11):7004–7014. <https://doi.org/10.1109/TITS.2020.3000761>
- Zhang JL, Chen F, Guo YN, Li XH (2020) Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit. *IET Intell Transp Syst* 14(10):1210–1217. <https://doi.org/10.1049/iet-its.2019.0873>
- Liu JX, Jiang R, Zhu D, Zhao JD (2022) Short-term subway inbound passenger flow prediction based on AFC Data and PSO-LSTM optimized model. *Urban Rail Transit* 8(1):56–66. <https://doi.org/10.1007/s40864-022-00166-x>
- Du BW, Peng H, Wang SZ, Bhuiyan MZA, Wang LH, Gong QR, Liu L, Li J (2020) Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction. *IEEE Trans Intell Transp Syst* 21(3):972–985. <https://doi.org/10.1109/TITS.2019.2900481>
- Shuai CY, Xie YW, Shan J, Oouyang X (2022) Prediction of short-term inbound passenger flow of urban rail transit based on the singular spectrum analysis and support vector regression model. *Urban Rapid Rail Transit* 35(5):76–83 (in Chinese)
- Guo ZQ, Zhao X, Chen YX, Wu W, Yang J (2019) Short-term passenger flow forecast of urban rail transit based on GPR and KRR. *IET Intell Transp Syst* 13(9):1374–1382. <https://doi.org/10.1049/iet-its.2018.5530>
- Wu J, He Y, Zhang N, Wu H (2022) Forecast of short-term passenger flow of urban rail transit based on VMD-GRU. *Urban Rapid Rail Transit* 35(1):79–86 (in Chinese)
- Ma XD, Wu JW, Liang SD, Zhao TY (2019) Short-term passenger flow prediction method based on dynamic weight of fusion model. *Transp Res* 5(4):127–132 (in Chinese)
- He YX, Li LS, Zhu XT, Tsui KL (2022) Multi-graph convolutional-recurrent neural network (MGC-RNN) for short-term forecasting of transit passenger flow. *IEEE Trans Intell Transp Syst* 23(10):18155–18174. <https://doi.org/10.1109/TITS.2022.3150600>
- Jing Y, Hu HT, Guo SY, Wang X, Chen FQ (2020) Short-term prediction of urban rail transit passenger flow in external passenger transport hub based on LSTM-LGB-DRS. *IEEE Trans Intell Transp Syst* 22(7):4611–4621. <https://doi.org/10.1109/TITS.2020.3017109>
- Fu CL, Huang M, Sha ZR (2020) Short-term forecast of passenger flow into an urban rail transit station based on EEMD-BP. *Railw Transp Econ* 42(3):7 (in Chinese)
- Li YB, Zhao Z (2019) Forecast of monthly passenger flow transport volume at railway station based on VMD-BP neural network model. *Traffic Stand* 5(2):52–59 (in Chinese)
- Chen JX, Cao LB, Yu KH (2021) EMD-LSTM based deep learning inbound and outbound passenger flow prediction. In: 2021 2nd international conference on artificial intelligence and information systems, pp 1–6. <https://doi.org/10.1145/3469213.3470381>
- Holt CC (2004) Forecasting seasonals and trends by exponentially weighted moving averages. *Int J Forecast* 20(1):5–10. <https://doi.org/10.1016/j.ijforecast.2003.09.015>
- Huang NE, Shen Z, Long SR et al (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond Ser A* 454:903–995. <https://doi.org/10.1098/rspa.1998.0193>
- Tao X, Bao YK, Hu ZY (2014) Does restraining end effect matter in EMD-based modeling framework for time series prediction? Some experimental evidences. *Neurocomputing* 123:174–184. <https://doi.org/10.1016/j.neucom.2013.07.004>
- Wang ZJA, Chen C, Li XL, Li J (2019) Short-term urban rail transit passenger flow forecasting based on empirical mode decomposition and LSTM. In: Proceedings of the 2019 international conference on modeling, analysis, simulation technologies and applications, pp 119–126. <https://www.atlantis-pess.com/proceedings/masta-19/125913203>
- Coughlin KT, Tung KK (2004) 11-Year solar cycle in the stratosphere extracted by the empirical mode decomposition method.

- Adv Space Res 34(2):323–329. <https://doi.org/10.1016/j.asr.2003.02.045>
27. Rato RT, Ortigueira MD, Batista AG (2008) On the HHT, its problems, and some solutions. *Mech Syst Signal Process* 22(6):1374–1394. <https://doi.org/10.1016/j.ymssp.2007.11.028>
 28. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint <https://arxiv.org/abs/1406.1078>
 29. Sha S, Li J, Zhang K, Yang Z, Wei Z, Li X, Zhu X (2020) RNN Based subway passenger flow rolling prediction. *IEEE Access* 8:15232–15240. <https://doi.org/10.1109/ACCESS.2020.2964680>
 30. Taylor SJ, Letham B (2018) Forecasting at scale. *Am Stat* 72(1):37–45. <https://doi.org/10.1080/00031305.2017.1380080>
 31. Zhao YY, Xia L, Jiang XG (2020) Short-term metro passenger flow prediction based on EMD-LSTM. *J Traffic Trans Eng* 20(04):194–204 (in Chinese)