

# A Novel Short-Time Passenger Flow Prediction Method for Urban Rail Transit: CEEMDAN-CSSA-LSTM Model Based on Station Classification

Jinbao Zhao, Jiawei Jiang, Wenjing Liu, Mingxing Li, Yuejuan Xu, Keke Hou, Shengli Zhao

**Abstract**—For urban rail transit (URT) operating companies, short-time passenger flow forecasting is a complex and critical task that determines the formulation and arrangement of operation plans and timetables. The automatic fare collection (AFC) system of URT provides detailed passenger flow data, which supports short-time passenger flow forecasting. This study proposes a combined model of Chaotic Sparrow Search Algorithm and Long Short-term memory artificial neural network (CSSA-LSTM) based on station classification. The model classifies stations based on various indicators such as point of interest (POI) data in the radiation area of Hangzhou, utilizes complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to denoise and smooth passenger flow data. It also conducts short-time passenger flow prediction and cross-validation for different stations by using the proposed model. To demonstrate the accuracy of the model, evaluation metrics such as R-squared and RMSE are introduced, and the results of the CSSA-LSTM model are compared with those of the LSTM, PSO-LSTM, and SSA-LSTM models. The experimental results show that the CSSA-LSTM model can effectively improve prediction accuracy, with R-squared increasing by 14.80%, 8.60%, and 6.82% compared to the other three algorithms, respectively. In addition, the cross-validation results of different stations prove the wide applicability of the CSSA-LSTM model, and this study has practical significance for URT planning and management.

**Index Terms**—Urban Rail Transit, Passenger Flow Forecasting, Station Classification, CSSA, LSTM, Cross-validation

Manuscript received May 5th, 2023; revised September 27th, 2023.

This work was supported in part by the National Natural Science Foundation of China under Grant 51608313, Natural Science Foundation of Shandong Province under Grant ZR2021MF109, Science and Technology Project of Shandong Expressway Group under Grant 2020-SDHS-GSJT-024, Shandong Province Graduate Education Quality Improvement Program SDYJG19110.

Jinbao Zhao is a lecturer engaged in the research of Urban Rail Transit, Shandong University of Technology, Zibo 255000, China. (e-mail: jinbao@sdut.edu.cn)

Jiawei Jiang is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (e-mail: jiawei.jiang0609@gmail.com)

Wenjing Liu is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (e-mail: wenjingliu719@163.com)

Mingxing Li is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (Corresponding author, e-mail: 1217193158@qq.com)

Yuejuan Xu is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (e-mail: 1874018039@qq.com)

Keke Hou is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (e-mail: 2404879184@qq.com)

Shengli Zhao is a postgraduate student at Shandong University of Technology, Zibo 255000, China. (e-mail: 841213463@qq.com)

## I. INTRODUCTION

WITH the continuous improvement of economic development and urbanization construction in China, people's travel demands have been steadily increasing, leading to a series of urban traffic problems, such as traffic congestion and excessive traffic pollution emissions. If these issues are not effectively resolved, they could obstruct the sustainable growth of society and the economy. Since the turn of the century, huge population concentrations have caused an issue known as "urban disease" that is visible in Europe, the United States, Japan, and other nations. Urban public transit, which draws on their development expertise, is one of the efficient solutions to several issues, including traffic congestion [1, 2]. Urban public transit is clearly preferable to private automobile travel in terms of increasing the effectiveness of land use, lowering energy usage, and other factors. Therefore, giving public transportation construction priority has emerged as a key strategy for solving urban traffic issues [3].

In this context, China's URT industry is booming, bringing new impetus and vitality to urban development. It has successfully enhanced the volume and speed of urban mobility, increased urban space, decreased oil consumption and exhaust emissions, and lessened plenty of urban stresses. But as URT has grown, a variety of flaws have also come to light, drawing the attention of academics.

For instance, one of the difficulties is the rail transit system's capacity restriction, where the demand for passenger flow occasionally exceeds the system's capacity. As seen in Hangzhou rail transit during working days where the passenger saturation rate of some sections exceeds 100%, a large influx of passengers within a short period of time during peak hours can increase operational pressure on the transit line. There are times when the demand for passenger flow exceeds the carrying capacity of rail transit, resulting in empty seats on trains during off-peak hours, resource waste, and harm to sustainable development. Therefore, the solution to this issue lies in understanding short-term changes in passenger flow and achieving passenger flow prediction.

There are three types of passenger flow prediction: medium- and long-term predictions, short-term predictions, and short-time predictions. Short-term passenger flow prediction is typically used for station size renovation design and other content, while medium- and long-term passenger

flow prediction is used for macro content, such as network planning and design, in the early stages of rail transit. Both macro predictions, though, might not have an impact on how rail transit functions on a daily basis. When discussing the short-time passenger flow prediction, which is used to support work planning, train scheduling, personnel deployment, and other decision-making tasks, it is typically used to refer to the prediction of passenger flow within the next 15 minutes. Strong randomness, non-stationarity, and non-linearity are characteristics of short-time passenger flow. At present, the train interval tracking in many first-tier cities in China's URT is mostly less than 5 minutes. Hence, the prediction of passenger flow with a time granularity of 5 minutes is usually performed, which has profound significance for the operation and scheduling of URT systems [4, 5].

Deep learning and other optimization algorithms are currently being used to predict short-term passenger flow at various URT stations based on passenger flow sequence data [6]. To fully understand the patterns of passenger flow between URT stations, it is not enough to predict the amount of traffic at a single station. Various factors, including geographic location, land use characteristics within the station's radiation area, and road network density, have an impact on the passenger flow at each URT station. As a result, a model developed for one station may not be applicable to accurately predict passenger flow at other stations. The study suggests a station classification-based short-term passenger flow prediction model to solve this issue. Firstly, to categorize all URT stations, this study uses clustering algorithms to analyze their attributes. Following that, representative stations from each distinct category are chosen for in-depth analysis, enabling a more thorough and detailed study. The CEEMDAN algorithm is used to eliminate noise and smooth the passenger flow data of the chosen stations to increase the model's predictive accuracy. In the meantime, the CSSA algorithm is used for parameter optimization, resolving the problem of traditional prediction algorithms' slow convergence speed for URT passenger flow. High precision, good stability, quick convergence, and robust search capabilities are all benefits of this algorithm. The accuracy of the model can be significantly increased when used in conjunction with deep learning algorithms. Finally, individual evaluations of short-term passenger flow prediction for stations of various categories are performed using various indicators, and cross-validation between models of stations of dissimilar categories is evaluated to demonstrate the proposed model's usability and effectiveness.

The remaining structure of this study is as follows. Section 2 expands on commonly used methods for short-term passenger flow prediction and introduces the study's research ideas. Section 3 describes the study area and provides relevant details about the methods used. Section 4 presents the experimental findings and results. Finally, in Section 5, this study summarizes the findings and discusses current issues and future research directions.

## II. LITERATURE REVIEW

Numerous statistical models, including the autoregressive integrated moving average (ARIMA) and seasonal

autoregressive integrated moving average (SARIMA) models, are frequently used for forecasting. Williams et al. [7] used a seasonal time series model and an exponential smoothing model to predict traffic flow in a single interval of urban highways in 1998. Lee and Fambro [8] investigated the use of the subset ARIAM model in short-term traffic volume forecasting and discovered that it was more stable and accurate than the full ARIMA model. Williams and Hoel [9] proposed a theoretical basis for modeling single-variable traffic data as a SARIMA process using a SARIMA model of single-variable traffic data. Milenkovic et al. [10] forecasted the time series of Serbian railway passenger volume from January 2004 to June 2014 using the SARIMA model, and the experimental results demonstrated good predictive performance.

With the rapid development of computer technology and various algorithms, machine learning has become a very effective tool for predicting passenger flow. Support vector machines (SVM) and artificial neural network (ANN) models are examples of common methods. Wei and Chen [11] predicted passenger flow using empirical mode decomposition (EMD) and neural network models, with the previous six-time steps as input. Jeong et al. [12] proposed a new online learning weighted support vector regression prediction model. Jiao et al. [13] improved the traditional Kalman filter method and proposed three models, KF-ECC, KF-HD, and KF-BCNR, which they validated using passenger flow data from Beijing Metro Line 13. The KF-BCNR model performed the best, according to the results. Li et al. [14] proposed a new multiscale radial basis function (MSRBF) network. When combined with the analysis of Beijing bus data, the algorithm demonstrated an exceptional ability to predict non-traditional demand within 30 minutes. Ouyang et al. [15] proposed a new bus passenger flow prediction model based on XGBoost feature extraction and decoding of multilayer neural networks, employing LSTM.

The models mentioned above only consider passenger flow within the URT, ignoring potential URT passenger flow, which influences short-term passenger flow prediction. Mobile phones, on the other hand, provide a reliable data source that can supplement existing AFC system data and improve the accuracy of passenger flow prediction [16]. Researchers can obtain travel trajectories of passengers as they pass by base stations using mobile phone data, providing additional insights into passenger movements. Dai et al. [17] proposed a data-driven short-term subway passenger flow prediction framework that uses spatial and temporal correlation information to combine potential passenger flow with AFC card-swipe passenger flow, improving the performance of short-term traffic prediction.

Single models are inherently flawed, and it is difficult to say that one model is superior to others under all conditions. As a result, more researchers are focusing on combining models, which can combine the strengths of multiple models and significantly improve the accuracy of passenger flow prediction [18]. Jia et al. [19] explored a deep learning-based model for predicting short-term passenger flow at each station that combined Long Short-Term Memory neural networks (LSTM-NN) and stacked autoencoders (SAE). Shahriari et al. [20] combined the Bootstrap and ARIMA

models to create an E-ARIMA ensemble, which improved prediction accuracy for traffic data from major roads in Sydney, Australia. Zeng et al. [21] proposed a combined model for short-term passenger flow prediction in URT based on Adaptive Noise Empirical Mode Decomposition (CEEMDAN) and LSTM neural networks, and calculated LSTM hyperparameters using Improved Particle Swarm Optimization (IPSO). To eliminate noise and improve short-term predictions, Xiu et al. [22] proposed a new framework that integrates Empirical Mode Decomposition (EEMD) and Bidirectional Gated Recurrent Unit (Bi-GRU) models.

Based on the above research, it can be concluded that commonly used methods for short-term passenger flow prediction involve combination forecasting models that combine global optimization algorithms with deep learning algorithms. Traditional optimization algorithms, on the other hand, have some drawbacks, such as reduced population diversity and susceptibility to local optima when approaching the minimum value. Furthermore, the nonlinearity and randomness of URT passenger flow make it ideal for chaotic algorithms. Therefore, this study proposes a CSSA-LSTM model for short-term passenger flow prediction based on the Tent chaotic algorithm. Furthermore, many previous studies have only chosen a few URT stations in a city, which may result in incomplete findings and make accurately identifying passenger flow patterns difficult. In contrast, this study classified 79 Hangzhou stations and chose representative stations from each category for research, yielding more comprehensive and reliable results.

### III. MATERIALS AND METHODS

#### A. Study area and Data processing

##### Study area and data sources

Hangzhou, China's capital and administrative center, is located between 118°1' E and 120°31' E longitude and 29° 11' N and 30°33' N latitude. It is one of eastern China's most important cities. Hangzhou is a vital component of the Yangtze River Delta and is adjacent to Shanghai, making it a core city of the Yangtze River Delta region. Furthermore, Hangzhou is an important node for the North-South Economic Corridor and the Maritime Silk Road, acting as a major transportation hub connecting mainland China with Southeast Asia, South Asia, the Middle East, and Europe. Hangzhou's urban area covered 728.46 square kilometers and had a population of 9.115 million as of early 2019. As shown in Fig.1, Hangzhou's transportation system is primarily supported by its three operational metro lines: Line 1, Line 2, and Line 4, which play a critical role in facilitating transportation throughout the city.

In addition to its economic and transportation importance, Hangzhou is an important tourist destination, with 170 million domestic and foreign visitors in 2019. The tourism industry's rapid growth has increased demand for URT operations, necessitating short-term passenger flow prediction. Accurate passenger flow prediction can help URT operators with scheduling, optimizing station layouts, and increasing transportation efficiency, resulting in more convenient travel services for both citizens and tourists.

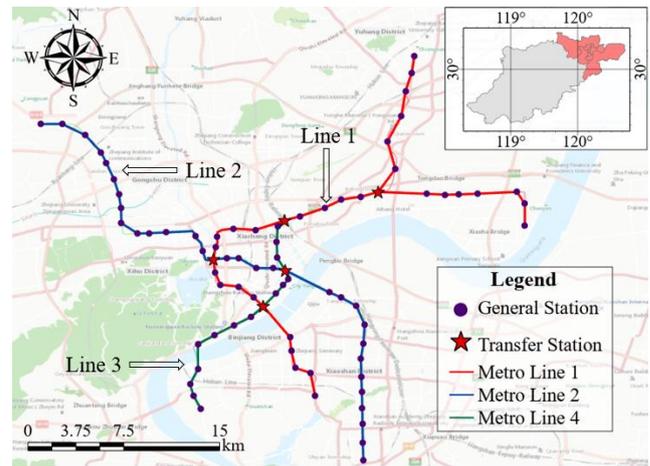


Fig. 1. Routes of Hangzhou's URT Lines (2019)

#### Data processing

The AFC system of Hangzhou URT system is the primary source of data for passenger flow prediction. This system is an advanced public transportation management tool that can collect and process data on ticket sales, ticket inspection, and passenger travel. It offers critical management and operational decision support for forecasting passenger flow in the URT. This study's data was obtained from the Internet and includes swipe card records from Hangzhou Metro Lines 1, 2, and 4 from January 1st to January 25th, 2019, with an average of approximately 2.6 million swipes per day. The information is saved in a.csv file and includes the swipe time, user ID, line ID, station ID, device ID, entry/exit status, and payment method. As shown in Fig.2, we eliminated illogical data by combining the same user ID's entry and exit card swiping time intervals (long intervals may indicate station staff, while short intervals may indicate non-use of URT). We also removed data that had missing values. Finally, to ensure data consistency, we only kept passenger flow data for each station from 6:00 a.m. to 11:00 p.m. daily. The final average daily data volume for each station was approximately 4800 entries after studying with a 5-minute time granularity.

Card Swipe ID	Card Swipe Time	Card Swipe Location	Entry/Exit Status	Card Swipe ID	Card Swipe Time	Card Swipe Location	Entry/Exit Status
1	1	1	1	1	1	1	1
1	1	0	2	2	1	1	0
2	1	2	2	2	1	2	0
2	0	2	2	2	0	2	1
2	1	2	2	2	1	2	0
2	0	3	3	3	0	3	1
3	1	3	3	3	1	3	0
3	0	3	3	3	0	3	1
3	1	3	3	3	1	3	0
3	0	3	3	3	0	3	1

Fig. 2. Combination of Passenger Flow Data from URT system

POI data is available from a variety of open-source websites on the Internet. Open-source websites offer open, non-commercial data that users can freely upload and download. Amap is a Chinese digital map content and navigation provider that provides global earth services such as POI data. The POI data spans many cities and regions worldwide and covers a wide range of categories such as catering, attractions, healthcare, shopping, traffic, and so on. The POI data in this study is based on the Amap open platform, and a total of 534,596 POI points in Hangzhou were crawled. We retained data only within a 1000-meter

radius of the stations and removed some outliers to focus on land use within the radiation range of rail transit stations, resulting in 100,232 data points across 11 categories, such as catering and shopping, as shown in Table I.

TABLE I  
POI STATISTICS FOR HANGZHOU

Categorization	Counting
Catering	34197
Company	18288
Shopping	52136
Traffic	17616
Hotel	5751
Education	10335
Attraction	1112
Residence	7360
Leisure	2913
Healthcare	7192
Sport	2570

We can use the POI points and other indicators (such as peak passenger flow ratio, number of peaks, and so on) from table 1 to perform K-means clustering analysis and classify the URT stations [23]. This classification method enables more accurate prediction of passenger flow for various types of URT stations.

#### B. CEEMDAN for Passenger Flow Decomposition

Non-stationary passenger flow data directly input into a prediction model may reduce prediction accuracy due to data noise. As a result, before entering passenger flow data, the original data must be denoised. Empirical Mode Decomposition (EMD) and wavelet denoising are currently popular denoising methods. EMD is a powerful nonlinear time-domain and time-frequency analysis method that can adaptively decompose a signal into a superposition of multiple Intrinsic Mode Functions (IMFs), each representing a different frequency component of the signal. EMD denoises signals by removing noise-containing IMF components or smoothing the IMF components, making it one of the most effective methods for signal denoising [11, 22].

Traditional EMD algorithms, on the other hand, have some limitations, such as the issue of mode mixing in signal decomposition. Scholars proposed Ensemble Empirical Mode Decomposition (EEMD) and Complementary Ensemble Empirical Mode Decomposition (CEEMD) to solve this problem by adding paired positive and negative Gaussian white noise to the signal. The intrinsic mode functions (IMFs) obtained from these two algorithms, however, may still contain residual white noise, which can affect the subsequent signal analysis and processing.

Torres et al. proposed CEEMDAN, an improved algorithm, to overcome these limitations. It is an improved method based on EMD that incorporates Gaussian noise and averaging multiple iterations from EEMD [24, 25]. The CEEMDAN method's calculation procedure is as follows:

Calculate the IMF1 component. Add a set of signals  $\{x(t)+E_1(n_i)\}$  consisting of paired positive and negative Gaussian white noise  $E_1(n_i)$  ( $i=1, 2, \dots, I$ ) to the original waveform signal  $x(t)$ . Using EMD to decompose each signal

separately, obtain the first IMF component of the decomposition and combine them to form a set. Take the average of the set to obtain the  $IMF_1$  component, denoted as  $h_1$ , and simultaneously calculate the residual of  $IMF_1$ , denoted as  $r_1$ , as shown in equations (1)-(2).

$$h_1 = \frac{1}{N} \sum_{i=1}^N h_1^{(i)} \quad (1)$$

$$r_1 = x(t) - h_1 \quad (2)$$

Similarly, calculate the  $IMF_2$  component to obtain the  $IMF_2$  component  $h_2$ , and calculate the residue  $r_2$  as shown in equations (3)-(4).

$$h_2 = \frac{1}{N} \sum_{i=1}^N h_2^{(i)} \quad (3)$$

$$r_2 = r_1 - h_2 \quad (4)$$

Following the same procedure, the  $IMF_k$  component  $h_k$  and the residue  $r_k$  of the original signal can be computed, as shown in equations (5)-(6).

$$h_k = \frac{1}{N} \sum_{i=1}^N h_k^{(i)} \quad (5)$$

$$r_k = r_{k-1} - h_k \quad (6)$$

Repeat the above steps until the residual signal becomes a monotonic function, completing the decomposition. The decomposition result of the original signal  $x(t)$  at this point is shown in equation (7).

$$x(t) = \sum_{j=1}^K h_j + r_k \quad (7)$$

#### C. SSA and CSSA Algorithms

##### SSA Algorithm

The central concept of the swarm intelligence optimization algorithm is to simulate the movement and behavior patterns observed in various natural phenomena and organisms, with the goal of finding optimal solutions within a given range of solution space. Observing the collective behaviors of birds, ants, whales, and other organisms has led to the development of numerous swarm intelligence optimization algorithms. Beni et al. [26] proposed the concept of "swarm intelligence" for the first time in 1989. Colorni et al. [27] proposed the ant colony optimization (ACO) algorithm in 1991 by simulating how ants avoid obstacles and choose the shortest path from their nest to food sources. Kennedy et al. [28] proposed the particle swarm optimization (PSO) algorithm in 1995, which was inspired by bird hunting behavior. Other researchers then proposed the bat algorithm, whale optimization algorithm (WOA), seagull optimization algorithm (SOA), and other algorithms. Among these algorithms, the sparrow search algorithm (SSA), proposed by Xue and Shen [29] in 2020, has gotten a lot of attention because of its advantages such as simple implementation, ease of expansion, and self-organization.

The basic modeling steps of SSA can be summarized as follows:

(1) Constructing the population. The population consists of  $N$  sparrows represented in the following form:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \quad (8)$$

In the above equation,  $n$  represents the number of sparrow populations, and  $d$  represents the dimensionality of the problem to be optimized.

(2) Calculate the fitness value. Like other heuristic algorithms, the SSA algorithm requires computing the fitness value of every individual in the swarm.

$$F_x = \begin{bmatrix} f(x_{1,1}, x_{1,2}, \dots, x_{1,d}) \\ f(x_{2,1}, x_{2,2}, \dots, x_{2,d}) \\ \cdots \\ f(x_{n,1}, x_{n,2}, \dots, x_{n,d}) \end{bmatrix} \quad (9)$$

Here,  $f$  represents the fitness value of each sparrow, and the fitness function can be chosen as the average mean squared error of the training and testing sets.

(3) Update the position of the discovery agent. In the SSA algorithm, the discovery agent provides the swarm with foraging direction, and its position update is described as follows:

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot t_m}\right) & \text{if } R_2 < S_T \\ x_{i,j}^t + Q \cdot L & \text{if } R_2 > S_T \end{cases} \quad (10)$$

Here,  $t$  represents the iteration number,  $x_{ij}^t$  represents the  $i$ -th sparrow's  $j$ -th dimensional value in the  $t$ -th iteration and can be seen as different parameters of the problem to be optimized.  $t_m$  is the maximum preset number of iterations,  $Q$  is a random number generated from a standard normal distribution, and  $L$  is a  $1 \times d$  matrix of all ones.  $R_2$  and  $S_T$  are the safety threshold and alarm threshold, respectively. When  $R_2 < S_T$ , the discovery agent enters the search mode and updates its position. When  $R_2 > S_T$ , it indicates that the sentinel agents in the swarm have detected danger, and the discovery agent will stop foraging and fly to a safe location.

(4) Updating follower positions. In the SSA algorithm, the roles of discoverers and followers switch, but the proportion of the two roles remains fixed within the population. Typically, the  $m$  top-performing sparrows are assigned the role of discoverers, while the remaining  $n-m$  sparrows act as followers. As the energy level of the followers diminishes, their foraging position deteriorate, and they may eventually exit the population. The update formula for follower positions is expressed as follows:

$$y_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{y_w^t - y_{i,j}^t}{i^2}\right) & i > \frac{n}{2} \\ y_p^{t+1} + \frac{1}{D} \sum_{d=1}^D (\text{rand}\{-1,1\} \cdot |y_{i,j}^t - y_p^t|) & i \leq \frac{n}{2} \end{cases} \quad (11)$$

Where  $y_p$  represents the best position within the population at the current time, and  $y_w$  represents the worst position. When  $i > n/2$ , it indicates that the  $i$ -th newcomer fails to acquire food and may depart from the population to explore alternative food sources.

(5) Regarding anti-predator behavior. It is presumed that

during the initial stages, 10% to 20% of the sparrows possess awareness of potential risks and promptly maneuver closer to their counterparts to mitigate the likelihood of predation. The formula is as follows:

$$z_{i,j}^{t+1} = \begin{cases} z_b^t + \beta \cdot |z_{i,j}^t - z_b^t| & f_i \neq f_g \\ z_{i,j}^t + K \cdot \left( \frac{|z_{i,j}^t - z_w^t|}{(f_i - f_w) + e} \right) & f_i = f_g \end{cases} \quad (12)$$

Where  $z_b^t$  denotes the best solution at time  $t$ ,  $z$  is either  $x$  or  $y$ ,  $f_i$  is the fitness value. When  $f_i > f_g$ , it indicates that the sparrows located at the edges of the population become aware of the risk, they perceive the position  $z_b$ , which represents the central position of the population, as a safe area. When  $f_i < f_g$ , the sparrows located in the middle of the population tend to move closer to their companions. Here,  $K$  is a constant that governs the step size, and  $e$  is a small non-zero value introduced to prevent division by zero in the denominator. Compared to other swarm intelligence optimization algorithms, SSA exhibits several advantages, such as high search precision, rapid convergence, good stability, and strong robustness. However, like other swarm intelligence optimization algorithms, SSA may encounter problems such as reduced population diversity and the risk of getting trapped in local optima during the search process [30].

#### CSSA Algorithm

Researchers have proposed several improvement strategies to overcome the challenges of reduced population diversity and local optima in swarm intelligence optimization algorithms [31, 32]. For example, Zhang et al. [33] redesigned the SOA algorithm's representation and update strategy for seagull positions, making it discrete and introducing a random mutation factor to improve individuals' ability to escape local optima. This improved SOA algorithm solved optimal path problems with good stability. Yang et al. [34] discovered that chaos theory could endow the PSO algorithm's inertia weight with chaos search ability, and they proposed a new simplified chaotic PSO algorithm based on the Logistic map to reduce the likelihood of being trapped in local optima. Han et al. [35] added a mutation operator with a Gaussian function and chaotic properties to the PSO algorithm to help the population jump out of local optima and improve global search ability.

These optimization strategies offer useful insights and methodological guidance for improving the Sparrow Search Algorithm. Inspired by the Gaussian distribution's good local search ability and the Tent chaotic sequence's uniform traversal and rapid convergence, a new algorithm called Chaos-based Sparrow Search Algorithm (CSSA) is proposed. The Tent chaotic mapping is used to initialize the population for an even distribution of initial positions. Then, Gaussian mutation and chaotic disturbance are used to adjust individuals when the population exhibits "aggregation" or "divergence" tendencies. These mechanisms aid in escaping local optima and improving model accuracy. The following are the main steps in the process:

**Step 1** Initialization, including the population size  $N$ , the number of explorers  $p_{num}$ , the number of sentinels  $s_{num}$ , the dimension of the objective function  $D$ , the lower and upper bounds  $lb$  and  $ub$  for initial values, and the maximum

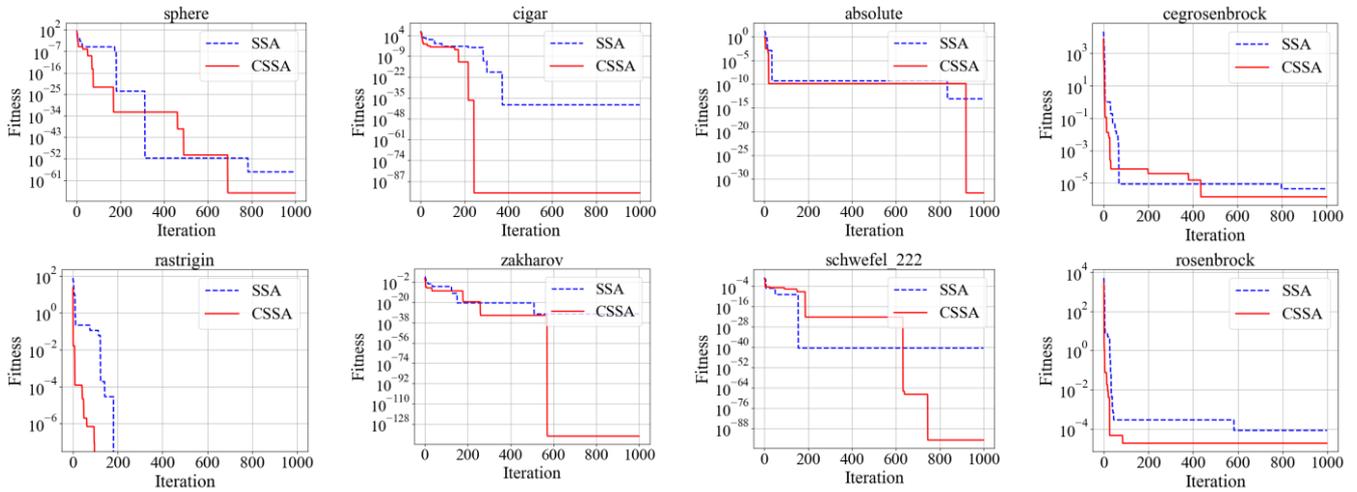


Fig. 3. Comparison of Different Benchmark Function Tests

number of iterations  $t_m$ .

**Step 2** Initialize the population using the Tent chaotic sequence. Generate  $N$   $D$ -dimensional vectors  $W_i$  and map their components onto the variable range of the original problem space.

**Step 3** Calculate the fitness value  $f_i$  of each sparrow and select the current best fitness value  $f_g$  and its corresponding position  $z_p$ , as well as the current worst fitness value  $f_w$  and its corresponding position  $z_w$ .

**Step 4** Select the top  $p_{num}$  sparrows with better fitness values as explorers and assign the remaining sparrows as followers and update the positions of the explorers and followers according to equations (10)-(11).

**Step 5** Randomly select  $s_{num}$  sparrows from the sparrow population for reconnaissance and warning. Update their positions according to formula (12).

**Step 6** After each iteration, recalculate the fitness value  $f_i$  of each sparrow and the average fitness value  $f_{avg}$  of the sparrow population.

1) If  $f_i < f_{avg}$ , it indicates the occurrence of "aggregation" phenomenon and requires Gaussian mutation processing. If the mutated individual is superior to the original individual, replace the original individual with the mutated individual; otherwise, keep the original individual unchanged.

2) If  $f_i \geq f_{avg}$ , it indicates the occurrence of "divergence" trend. Perform Tent chaotic disturbance on individual  $i$  in the population. If the individual after disturbance has better performance, replace the original individual with the disturbed individual; otherwise, keep the original individual unchanged.

**Step 7** Update the overall best position  $z_p$  and its fitness  $f_g$ , as well as the worst position  $z_w$  and its fitness  $f_w$ , based on the current state of the sparrow population.

**Step 8** Check if the maximum number of iterations has been reached. If so, terminate the loop; otherwise, go back to step 4.

To validate the performance of the improved CSSA algorithm, this study used benchmark functions to run a performance test on both the CSSA and the SSA algorithms. Benchmark functions are well-known mathematical functions with well-known optimal solutions that are used to assess and compare the performance of optimization algorithms. These functions cover a wide range of problems, such as continuous optimization, discrete optimization,

multi-objective optimization, and others. Eight benchmark functions were used in this study to compare the performance of the two algorithms. The experiments used a population of 30, a maximum iteration of 1000, and a dimension of 30. As shown in Fig.3, the CSSA algorithm outperformed the SSA algorithm in terms of optimization performance. The CSSA algorithm's superiority was manifested in its faster convergence speed and easier attainment of the optimal solution. As a result, this study concludes that the CSSA algorithm performs better in terms of optimization when dealing with practical problems [36].

#### D. LSTM

The LSTM, which was proposed in 1997, is a type of artificial neural network that is commonly used for sequential data processing. It outperforms traditional recurrent neural networks, particularly in addressing challenges related to processing long sequences and mitigating the vanishing gradient problem. Each neuron in the LSTM model has three gates: the input gate, the output gate, and the forget gate, which regulate the flow of information for input, output, and forgetting, respectively.

These gated units allow LSTM to selectively store and retrieve essential information over long sequences while preventing irrelevant information from propagating throughout the network [37, 38].

The input gate, forget gate, output gate, and memory cell are the four main components of an LSTM. The input gate determines whether the current input data should be stored in the memory cell. The forget gate decides whether or not to discard previous memory information. The output gate determines which memory cell information should be outputted at the current time step. The memory cell functions as a long-term data storage unit. LSTM effectively processes long sequences and captures long-term dependencies within the sequence by utilizing these gated units. Because of its effectiveness in dealing with long-term data, LSTM has been widely used in fields such as natural language processing, time series prediction, and image analysis [39-41].

(1) The forget gate employs an activation function to determine how much of the previous cell state  $C_{t-1}$  is retained in the current cell state  $C_t$ , as shown in equations (13)-(14).

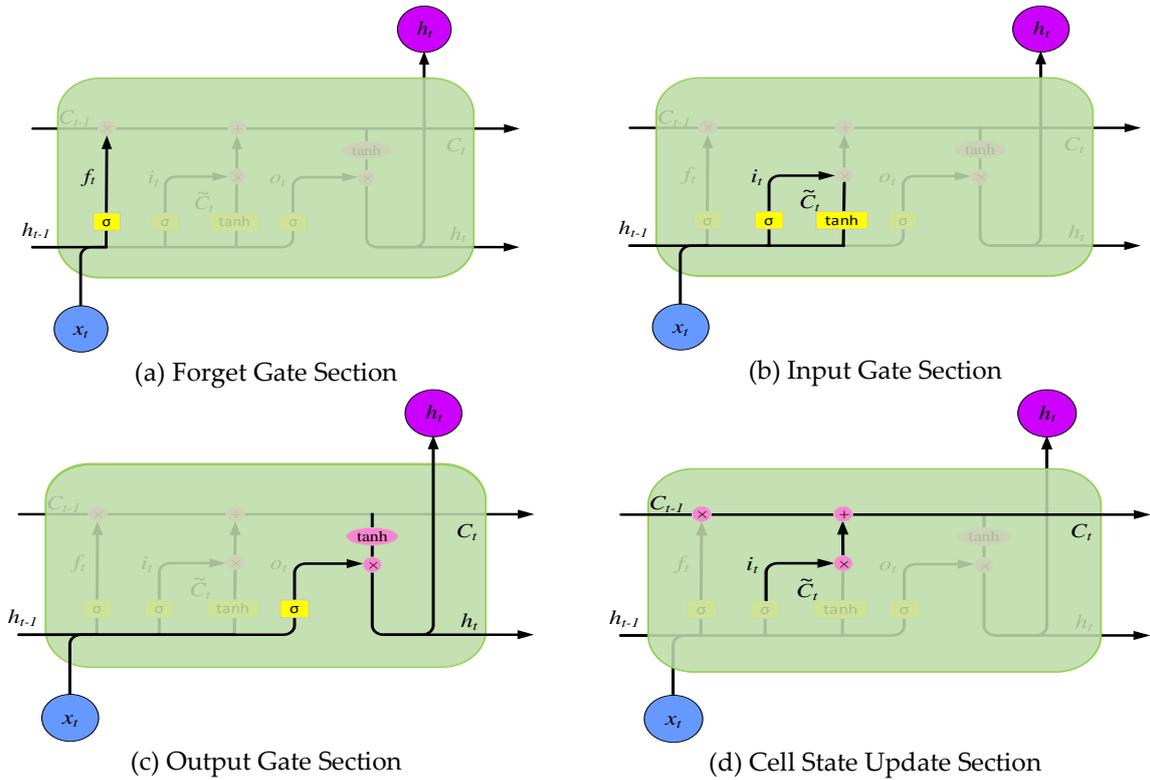


Fig. 4. The principle of LSTM

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (13)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (14)$$

Where  $W_f$  and  $b_f$  represent the weights and biases of the forget gate, respectively.  $[h_{t-1}, x_t]$  represents the concatenation of the previous hidden state and the current input at the current time step. The specific process is shown in the following diagram.

(2) The input gate determines the proportion of the current input  $x_t$  that is saved in the current cell state  $C_t$ , while employing the tanh function to generate a candidate cell state, as shown in equations (15)-(16).

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (15)$$

$$\begin{cases} i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \end{cases} \quad (16)$$

Where  $W_i$  and  $b_i$  represent the weights and biases of the input gate, respectively. The specific process is shown in the following diagram.

(3) The output gate determines the amount of the original output information that is multiplied pointwise with the cell state passed through the tanh layer, to obtain the final model output, as shown in equations (17).

$$\begin{cases} o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t \times \tanh(C_t) \end{cases} \quad (17)$$

Where  $W_o$  and  $b_o$  represent the weights and biases of the output gate, respectively. The specific process is shown in the following diagram.

(4) State update, including cell state  $C_t$  and hidden state  $h_t$

update, as shown in equations (18)-(19).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (18)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (19)$$

The LSTM principles described above are illustrated in Fig.4.

The process of updating the cell state, which is a key component of the LSTM model and is depicted in the figure above. Numerous factors influence the LSTM model's prediction accuracy, including data quality and quantity, as well as model parameters such as the number of hidden layer neurons, learning rate, and time step. This study employs the CEEMDAN algorithm to effectively eliminate noise from the data to improve data quality. Furthermore, the CSSA algorithm is used to find the best hyperparameters of the LSTM model, allowing for accurate short-term passenger flow prediction. It should be noted that these techniques are critical for achieving high-precision predictions in the LSTM model, and their application has great potential for furthering research in this domain. The main process of this study is shown in Fig.5.

#### IV. RESULT AND DISCUSSIONS

##### A. Clustering Results

The POI kernel density within the 1000 m radius of the station is shown in the Fig.6. It clearly displays the density of various POIs near the routes, providing detailed insights into the land use characteristics within the station's influence area. URT stations can be classified more effectively based on this, in conjunction with factors such as route density and passenger flow attributes [42].

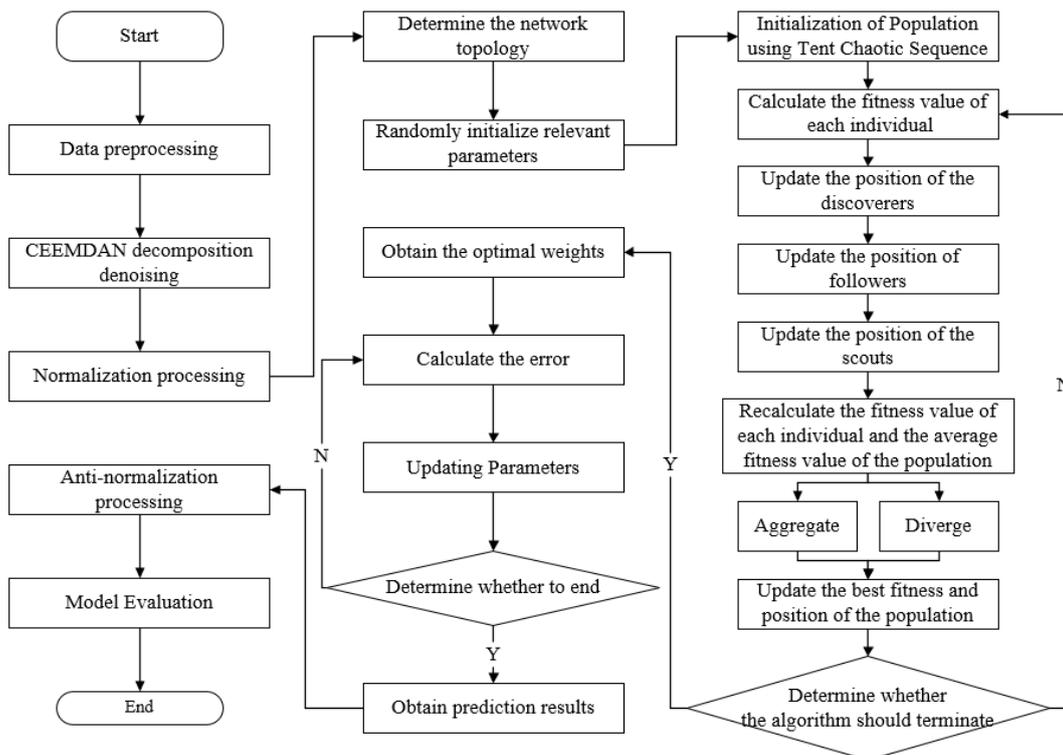


Fig. 5. Research Process

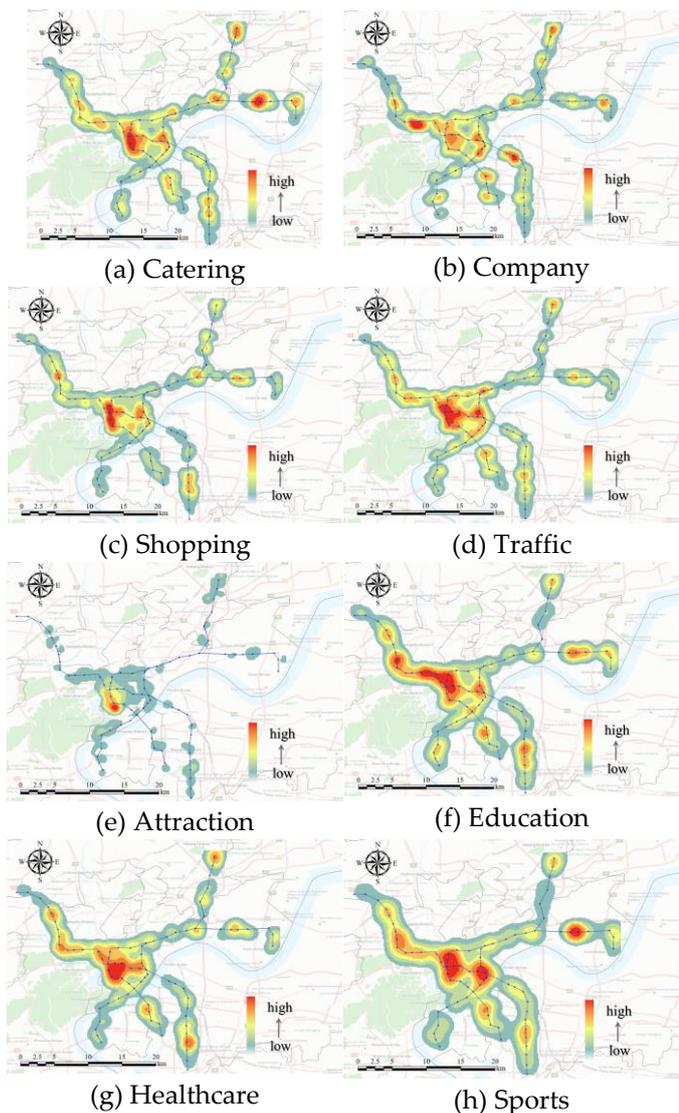


Fig. 6. POI Density Heatmap

Based on previous research experience, this study classifies URT stations into four categories, and the statistical results are shown in Table II.

TABLE II  
RESULTS OF STATIONS CATEGORIZATION

Category	Station Name	Number
Category 1	Xianghu Station, Jinshahu Station, Yunshui Station, etc.	42
Category 2	Xixing Station, Jiubao Station, People's Square Station, etc.	19
Category 3	Ding'an Road Station, Wulin Square Station, Nanxingqiao Station, etc.	13
Category 4	Hangzhou East Railway Station, Passenger Transport Center Station, etc.	6

Category 1: Residential stations. These stations have a wide coverage of residential areas and offer services oriented towards daily life, such as restaurants, shopping facilities, and educational institutions. The proportion of passengers entering the station during the morning peak period is higher than that during the evening peak period, as shown in Fig.7.

Category 2: Comprehensive stations. These stations have a variety of POIs near the station, including both life-oriented and production-oriented services. The proportion of passengers entering and exiting the station during the morning and evening peak periods is relatively balanced.

Category 3: Commercial stations. These stations have a mix of life-oriented and job-oriented services, with a higher volume of passenger flow during the evening peak period compared to the morning peak period.

Category 4: Transportation hub stations. This category includes URT interchange stations and railway hub stations.

The surrounding land use of these stations does not have any distinct features, but they have a high volume of passenger traffic, which is evenly distributed throughout the day.

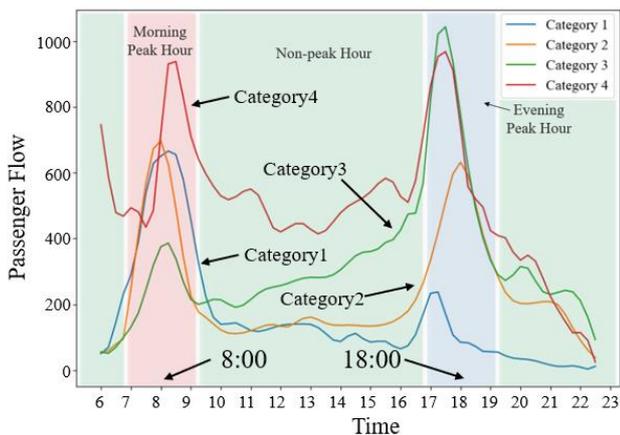


Fig. 7. Different Types of Passenger Flow in URT Stations

The graph above depicts how the passenger flow for various types of URT stations varies. Some stations have unimodal traffic patterns, while others have bimodal or multimodal traffic patterns. Furthermore, peak passenger flow for different types of stations can occur in the morning or evening. As a result, analyzing only one or two URT stations is insufficient for predicting short-term passenger flow. In this study, we classified Hangzhou's URT stations into different types and chose two representative stations from each type to predict passenger flow. We hoped to achieve more accurate results by conducting horizontal comparisons of prediction performance across these stations, thereby improving the reliability and applicability of our findings.

To validate the effectiveness of site categorization, we ran a similarity analysis of the historical passenger flows of the target sites and calculated averages for each site category, as shown in Fig.8. According to the graph, passenger flows for the first three categories of sites remain relatively stable during the weekdays, which could be attributed to commuting and school-related activities. The fourth category of sites, on the other hand, has significantly lower passenger flow similarity, which could be because these sites serve as transportation hubs with high passenger volumes and greater mobility, making predictions more difficult.

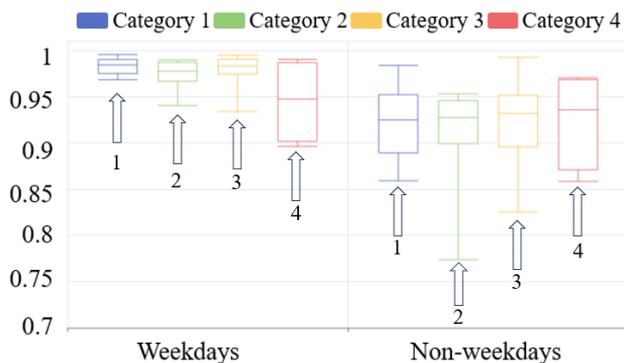


Fig. 8. Distribution of Site Similarity among Different Site Types

The passenger flow similarity for the first three categories decreases slightly during non-working days, while the

similarity for the fourth category remains relatively unchanged. This implies that the shift between working and non-working days has less of an impact on the fourth category of sites. Overall, the first category of sites has the most stable passenger flow similarity, followed by the second and third categories, and the fourth category has the least stability.

**B. CEEMDAN Decomposition Results**

Based on the above-mentioned station classification results, this study divides the stations into four types and builds different prediction models for each type, which are then cross-validated. Cross-validation is the process of training a model on data from one station and then testing its performance on data from other stations to determine the accuracy of the model's predictions for those stations. Because passenger flow patterns differ between station types, cross-validation effectively evaluates the model's generalization ability. Table III shows the selection of eight representative stations for short-term passenger flow prediction. For convenience in subsequent discussions, the first column of Table III is labeled 1–4 for the four stations, and the second column is labeled a–d for the four stations, where station 1 represents Xianghu Station and station b represents Qingling Station.

TABLE III  
RESEARCH SUBJECT

Urban Rail Transit Stations		
Category 1	Xianghu Station	Feihong Road Station
Category 2	Xueyuan Road Station	Qingling Station
Category 3	Ding'an Road Station	Wulin Square Station
Category 4	Cheng Station	Hangzhou East Railway Station

To denoise the signal at Xianghu Station, the CEEMDAN decomposition was used. The Pearson correlation coefficient was used to determine the correlation between the decomposed signal and the original signal, which indicates whether the signal information was lost during the decomposition process, to ensure the integrity of the original signal.

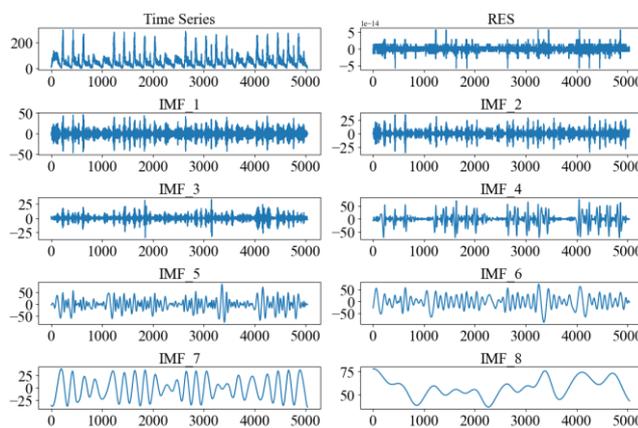


Fig. 9. The Decomposition Results of CEEMDAN

The CEEMDAN decomposition results are shown in Fig.9, which includes ten IMF components and one residual

component. The IMF components are arranged in decreasing frequency order from 1 to 10, with high-frequency noise removed (representing unstable passenger flow in URT) and low-frequency signals retained (representing long-term stable passenger flow, which is easier to investigate for passenger flow changes). The Pearson correlation coefficient was used to ensure the integrity of the original signal after removing high-frequency noise, with a coefficient of 0.9689 and p-value of 0. Thus, after removing high-frequency noise, the signal retained much of the original signal information and had smoother data with higher research value, as illustrated in detail in Fig.10. The same procedure was used for the remaining other URT stations.

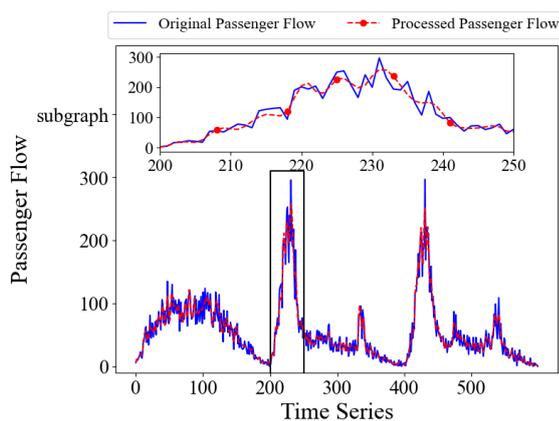
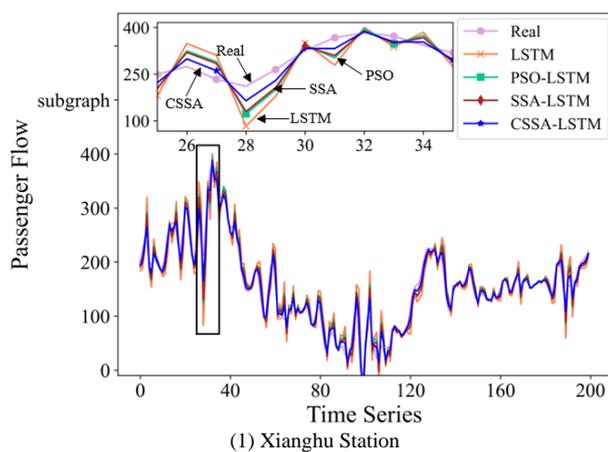


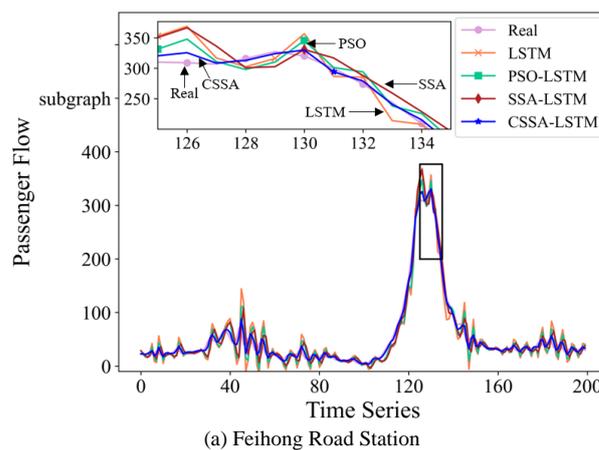
Fig. 10. Comparison before and after Decomposition

C. CSSA-LSTM Prediction Results

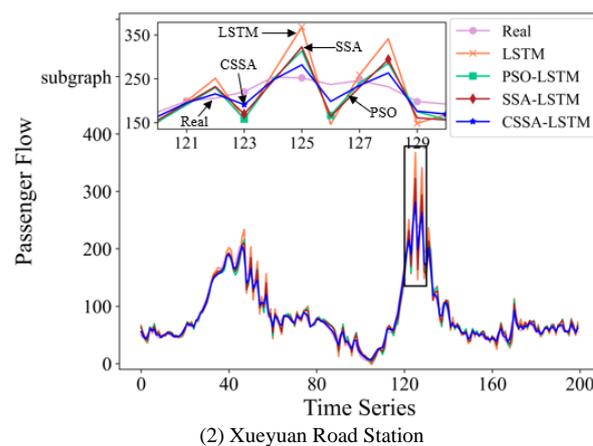
To meet scientific writing standards, we conducted a comparative analysis of the CSSA-LSTM model's effectiveness in comparison to three other models: LSTM, PSO-LSTM, and SSA-LSTM. These models were trained and tested using data samples collected from each station over a 24-day period with a granularity of 5 minutes, yielding approximately 4800 data points per station. The training and testing data were split in a 9:1 ratio. The LSTM model included 50 hidden neurons, a 10-step time step, a learning rate of 0.01, and 50 epochs. The hidden neurons in the CSSA-LSTM, SSA-LSTM, and PSO-LSTM models were determined by their respective optimization algorithms, while the other parameters remained constant. The prediction results for different categories of stations for each model are presented in Fig. 11.



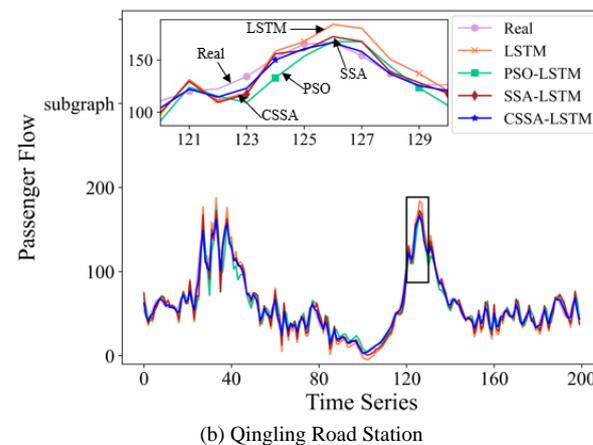
(1) Xianghu Station



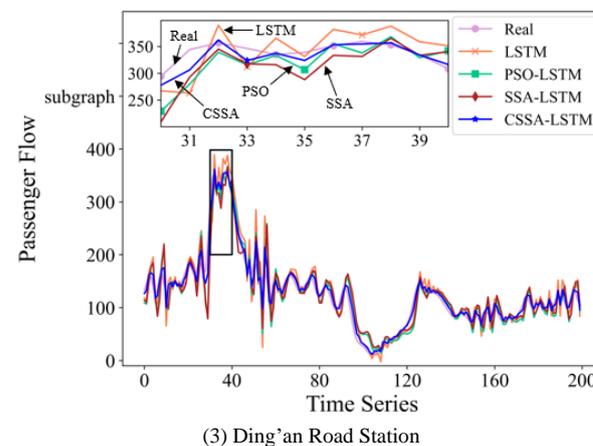
(a) Feihong Road Station



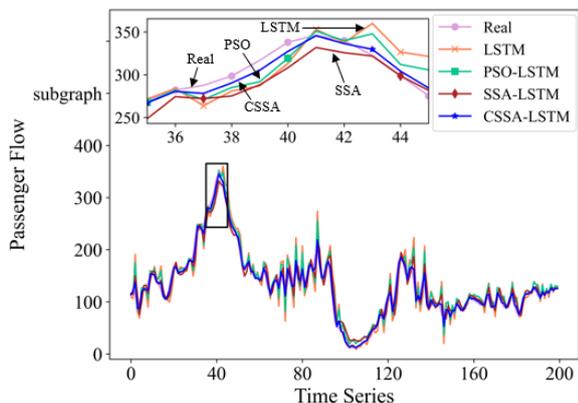
(2) Xueyuan Road Station



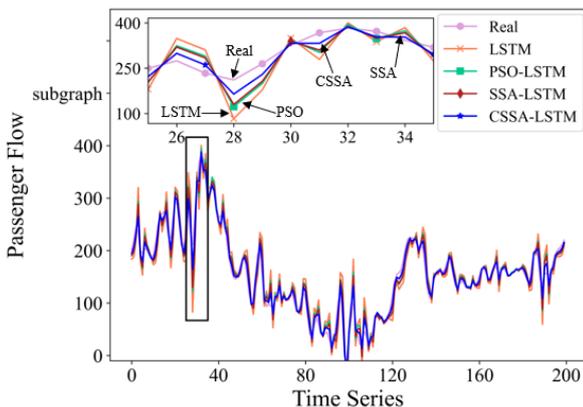
(b) Qingling Road Station



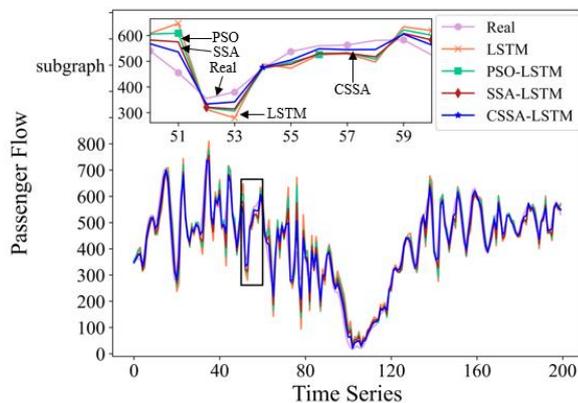
(3) Ding'an Road Station



(c) Wulin Square Station



(4) Cheng Station



(d) Hangzhou East Railway Station

Fig. 11. Comparison of Passenger Flow Predictions among Different URT Stations

Fig. 11 shows the prediction results obtained from various models used to forecast passenger flow at various URT stations. The predicted results were inverse-normalized and compared against the actual passenger flow to allow for a direct comparison of the model's prediction accuracy. The models were evaluated using R-squared, RMSE, MAE, and MAPE metrics. Table IV shows the specific evaluation results.

Fig. 12 shows that as the model's performance advances, so do the projections of passenger flow at different stations. The standalone LSTM model performs the worst, with the greatest RMSE, MAE, and MAPE values and the lowest R-squared value. The level of prediction accuracy varies significantly among station categories. Both the PSO-LSTM and SSA-LSTM models have very erratic forecast accuracy

for various station types, which could not be in line with the day-to-day operational requirements of URT systems. On the other hand, the CSSA-LSTM model exhibits good predictive accuracy for passenger flow at various station categories and is well suited for tackling short-time passenger flow forecasting difficulties at diverse URT station types.

TABLE IV  
MODEL EVALUATION METRICS RESULTS

Station	Model	R2	RMSE	MAE	MAPE
Xianghu Station	LSTM	0.86	19.37	13.11	0.45
	PSO-LSTM	0.91	15.69	10.54	0.53
	SSA-LSTM	0.91	16.04	11.43	0.23
	CSSA-LSTM	0.98	7.31	4.96	0.31
Feihong Road Station	LSTM	0.89	20.34	14.58	0.45
	PSO-LSTM	0.93	16.08	11.69	0.37
	SSA-LSTM	0.94	15.46	11.49	0.36
	CSSA-LSTM	0.99	5.83	4.22	0.13
Xueyuan Road Station	LSTM	0.88	18.11	12.38	0.17
	PSO-LSTM	0.93	14.17	10.15	0.15
	SSA-LSTM	0.94	12.75	8.82	0.13
	CSSA-LSTM	0.98	6.53	4.51	0.06
Qingling Station	LSTM	0.89	12.22	9.23	0.2
	PSO-LSTM	0.91	11.13	8.37	0.17
	SSA-LSTM	0.92	10.21	7.58	0.16
	CSSA-LSTM	0.98	4.89	3.51	0.07
Ding'an Road Station	LSTM	0.84	28.37	19.66	0.16
	PSO-LSTM	0.86	26.61	19.11	0.19
	SSA-LSTM	0.85	27.19	19.73	0.2
	CSSA-LSTM	0.97	11.85	8.16	0.07
Wulin Square Station	LSTM	0.85	25.52	18.71	0.16
	PSO-LSTM	0.91	19.61	14.39	0.14
	SSA-LSTM	0.94	15.92	11.91	0.14
	CSSA-LSTM	0.99	7.65	5.51	0.05
Cheng Station	LSTM	0.82	32.9	24.47	0.32
	PSO-LSTM	0.91	23.03	17.26	0.24
	SSA-LSTM	0.93	20.98	15.55	0.2
	CSSA-LSTM	0.98	12.15	9.03	0.12
Hangzhou East Railway Station	LSTM	0.79	74.65	57.57	0.28
	PSO-LSTM	0.85	62.14	47.69	0.22
	SSA-LSTM	0.9	50.12	38.82	0.19
	CSSA-LSTM	0.96	32.38	24.96	0.12

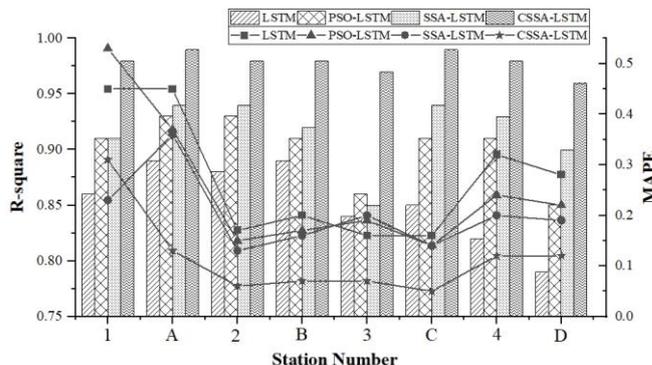


Fig. 12. Comparison of model evaluation metrics (data are from Table IV, selecting R-squared and MAPE, which are not affected by the unit of measurement)

Fig.13 shows that the CSSA-LSTM model has better convergence performance and a faster rate of convergence. As a result, using this technique in real-world applications can increase efficiency.

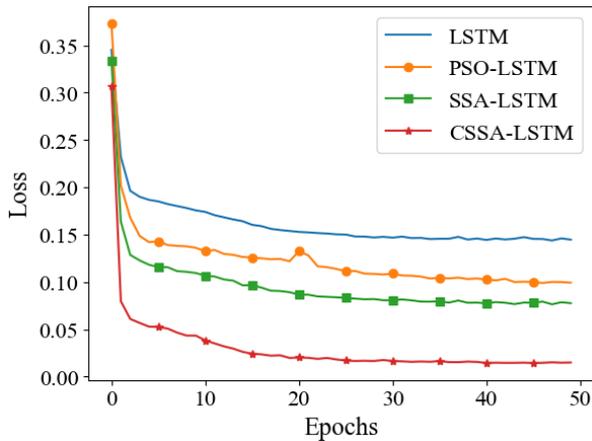
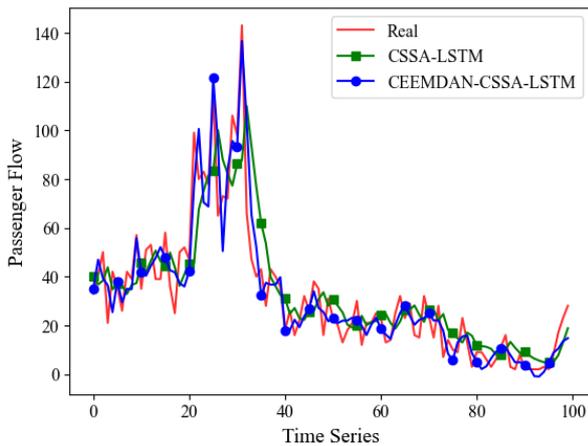
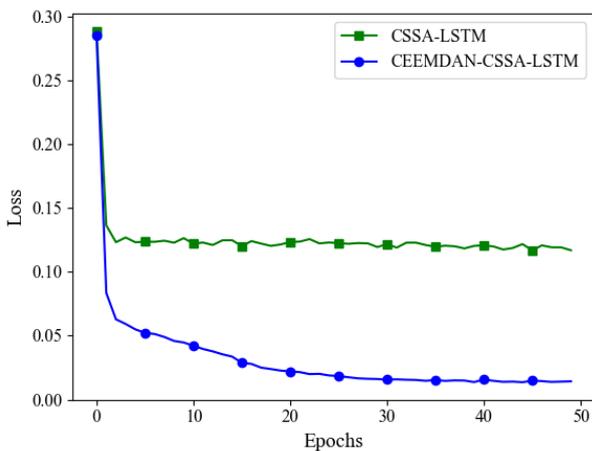


Fig. 13. Loss Function Comparison for Different Algorithms

Using Xianghu Station as an example, we input passenger flow data that has been decomposed and denoised alongside undecomposed and denoised passenger flow data into the CSSA-LSTM model to explore the effect of the CEEMDAN algorithm on model prediction outcomes. Fig. 14 displays the comparison.



(a) Predictive Results Comparison



(b) Convergence Speed Comparison

Fig. 14. The Effects of CEEMDAN Decomposition on CSSA-LSTM's Predictive Performance

According to Fig.15, the R-squared value for predicting passenger flow before denoising is 0.88, but it increases to 0.98 after denoising. This emphasizes how important CEEMDAN preprocessing is for predicting passenger flow because it speeds up model convergence and produces more precise predictions.

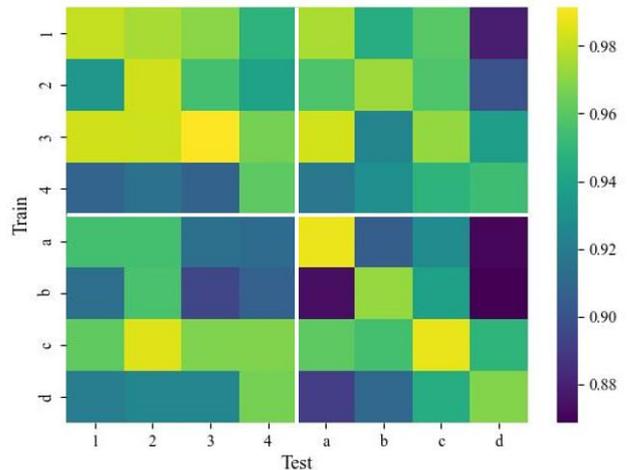


Fig. 15. Cross-validation of different stations

As shown in Fig.15, we performed cross-validation on models for various station categories, and it is clear from the graph that the accuracy of the forecast of the passenger flow is higher for stations of the same kind than it is for stations of different types. The performance is noticeably better than models based on Category 4 data when using passenger flow data from Category 3 stations to train models for predicting passenger flow at other station types. This discrepancy might be explained by the fact that Category 4 stations have a 'peak-hour' pattern with considerable passenger flow variations, which makes it difficult for the model to detect underlying trends.

### V. CONCLUSIONS

This study proposes a CSSA-LSTM neural network for short-time passenger flow prediction in URT systems based on a station classification strategy. First, we evaluate diverse land-use properties within a 1000-meter radius of URT stations to investigate the standards for station classification. We divide stations into four unique classes using the K-Means clustering technique. The original passenger flow data is then broken down and denoised using the CEEMDAN approach to ensure that any unstable passenger flow interference is removed and that the processed data more closely represents the signal's fundamental properties. We then put the CSSA-LSTM model into practice for forecasting short-time passenger flow and evaluate it against alternative algorithms. To evaluate the generalizability of model, we lastly perform cross-validation using pre-trained models for several station categories.

The research findings include the following: (1) We demonstrated the feasibility and necessity of site categorization by analyzing the degree of similarity among different types of sites. This method allows us to quantify the difficulty of predicting passenger flow at various URT stations. (2) The efficacy of CEEMDAN signal decomposition in greatly boosting the prediction model's

accuracy. (3) Variations in typical URT passenger flow prediction model performance across various station types. For instance, compared to Category 3 and Category 4 stations in this study, the LSTM model had better predictive accuracy for Category 1 and Category 2 stations. The passenger flow attributes, and computational procedures may be the reason for this disparity. Category 1 and Category 2 stations, with their comparatively steady passenger flows, enable better learning of their dynamic properties, resulting in higher prediction performance. The passenger flows at Category 3 and Category 4 stations are relatively complex, and the model training is insufficient, resulting in relatively poor training outcomes. (4) No matter the station category for which passenger flow prediction was done, the CSSA-LSTM model presented in this research showed the best predictive performance. When compared to the other three models, it showed 14.81%, 8.60%, and 6.82% gains in R-squared values, as well as reductions in RMSE values of 61.72%, 53.00%, and 47.48%, MAE values of 61.78%, 53.41%, and 48.25%, and MAPE values of 57.53%, 53.73%, and 64.37%. (5) Extensive experiments showed that this model had faster convergence, shorter training times, and less resource usage. (6) The CSSA-LSTM model demonstrated strong generalization ability based on cross-validation results, however minor variations in predicting accuracy persisted across several station types.

These findings highlight the benefits and contributions of the CSSA-LSTM model as it relates to improving short-time passenger flow forecast in URT systems.

It is important to recognize that the model's performance may be constrained by the cross-validation approach in this study's relatively small sample size. To further improve the predictive power of the model, it is crucial to consider the impact of numerous elements such as different time periods, seasonal fluctuations, and extraordinary occurrences on passenger flow prediction. It is noteworthy that the results of cross-validation differ significantly between the various station categories, with Category 3 stations showing comparably better passenger flow prediction results for stations belonging to other categories, while Category 4 stations show comparably worse results for the same.

#### REFERENCES

- [1] Jhenglong Wu, Mingying Lu, and Chiayun Wang, "Forecasting metro rail transit passenger flow with multiple-attention deep neural networks and surrounding vehicle detection devices," *Applied Intelligence*, vol. 53, no. 15, pp11-16, 2023
- [2] Fuya Yuan, Huijun Sun, Liujiang Kang, Ying Lv, Xin Yang, and Yun Wei, "An integrated optimization approach for passenger flow control strategy and metro train scheduling considering skip-stop patterns in special situations," *Applied Mathematical Modelling*, vol. 118, no. 8, pp412-436, 2023
- [3] Xiao Fu, Yufan Zuo, Jianjun Wu, Yuan Yu, and Sheng Wang, "Short-term prediction of metro passenger flow with multi-source data: A neural network model fusing spatial and temporal features," *Tunnelling and Underground Space Technology incorporating Trenchless Technology Research*, vol. 124, no. 12, 2022
- [4] Zhi Xiong, Jianchun Zheng, Dunjiang Song, Shaobo Zhong, and Quanyi Huang, "Passenger Flow Prediction of Urban Rail Transit Based on Deep Learning Methods," *Smart Cities*, vol. 2, no. 3, pp371-387, 2019
- [5] Kaer Zhu, Ping Xun, Wei Li, Zhen Li, and Ruochong Zhou, "Prediction of Passenger Flow in Urban Rail Transit Based on Big Data Analysis and Deep Learning," *IEEE Access*, vol. 7, no. 8, pp142272-142279, 2019
- [6] Qian Zhang, Xiaoxiao Liu, Sarah Spurgeon, Dingli Yu, "A two-layer modelling framework for predicting passenger flow on trains: A case study of London underground trains," *Transportation Research Part A: Policy and Practice*, vol. 151, no. 9, pp119-139, 2021
- [7] Billy M. Williams, Priya K. Durvasula, and Donald E. Brown, "Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models," *Transportation Research Record*, vol. 1644, no. 1, pp132-141, 1998
- [8] S. Lee, and D. Fambro, "Application of Subset Autoregressive Integrated Moving Average Model for Short-Term Freeway Traffic Volume Forecasting," *Transportation Research Record*, vol. 1678, no. 1, pp179-188, 1999
- [9] Billy M. Williams, and Lester A. Hoel, "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp664-672, 2003
- [10] Miloš Milenković, Libor Švadlenka, Vlastimil Melichar, Nebojša Bojović, and Zoran Avramović, "SARIMA modelling approach for railway passenger flow forecasting," *Transport*, vol. 33, no. 5, pp1-8, 2015
- [11] Yu Wei, and Muchen Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 21, no. 1, pp148-162, 2012
- [12] Young-Seon Jeong, Young-Ji Byon, Manoel Mendonca de Castro-Neto, Said M. Easa, "Supervised Weighting-Online Learning Algorithm for Short-Term Traffic Flow Prediction," *IEEE transactions on intelligent transportation systems*, vol. 14, no. 4, pp1700-1707, 2013
- [13] Pengpeng Jiao, Ruimin Li, Tuo Sun, Zenghao Hou, and Amir Ibrahim, "Three Revised Kalman Filtering Models for Short-Term Rail Transit Passenger Flow Prediction," *Mathematical Problems in Engineering*, vol. 16, no. 3, 2016.
- [14] Yang Li, Xudong Wang, Shuo Sun, Xiaolei Ma, and Guangquan Lu, "Forecasting short-term subway passenger flow under special events scenarios using multiscale radial basis function networks," *Transportation Research Part C: Emerging Technologies*, vol. 77, no. 5, pp306-328, 2017
- [15] Qi Ouyang, Yongbo Lv, Jihui Ma, and Jing Li, "An LSTM-Based Method Considering History and Real-Time Data for Passenger Flow Prediction," *Applied Sciences*, vol. 10, no. 11, 2020
- [16] Hongyan Gao, and Fasheng Liu, "Estimating freeway traffic measures from mobile phone location data," *European Journal of Operational Research*, vol. 229, no. 1, pp252-260, 2013
- [17] Xiaoqing Dai, Lijun Sun, and Yanyan Xu, "Short-Term Origin-Destination Based Metro Flow Prediction with Probabilistic Model Selection Approach," *Journal of Advanced Transportation*, pp1-15, 2018
- [18] Xin Yang, Qiuchi Xue, Meiling Ding, Jianjun Wu, and Ziyou 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, no. 1, 2021
- [19] Feifan Jia, Haiying Li, Xi Jiang, and Xinyue Xu, "Deep learning-based hybrid model for short-term subway passenger flow prediction using automatic fare collection data," *IET Intelligent Transport Systems*, vol. 13, no. 11, pp1708-1716, 2019
- [20] Shahriari Siroos, Ghasri Milad, Sisson S. A., and Rashidi Taha, "Ensemble of ARIMA: combining parametric and bootstrapping technique for traffic flow prediction," *Transportmetrica A: Transport Science*, vol. 16, no. 3, pp1552-1573, 2020
- [21] Lu Zeng, Zinuo Li, Jie Yang, and Xinyue Xu, "CEEMDAN-IPSO-LSTM: A Novel Model for Short-Term Passenger Flow Prediction in Urban Rail Transit Systems," *International Journal of Environmental Research and Public Health*, vol. 19, no. 24, pp16433-16433, 2022
- [22] Cong Xiu, Yichen Sun, Qiyuan Peng, Cheng Chen, and Xunquan Yu, "Learn traffic as a signal: Using ensemble empirical mode decomposition to enhance short-term passenger flow prediction in metro systems," *Journal of Rail Transport Planning & Management*, vol. 22, no. 3, 2022
- [23] Wei Li, Min Zhou, and Hairong Dong, "Classifications of Stations in Urban Rail Transit based on the Two-step Cluster," *Intelligent Automation and Soft Computing*, vol. 26, no. 3, pp531-538, 2020
- [24] Yuhui Sun, Liying Wang, Qingwen Lei, and Qingjiao Cao, "Research on the Vibration Prediction of Hydropower Unit Based on CEEMDAN-IPSO-LSTM," *Yellow River*, vol. 45, no. 5, pp156-162, 2023
- [25] Lianghe Kang, Xiaoqiang Li, Lili Wu, Yue Li, and Xia Zhao, "Predicting Stock Closing Price with Stock Network Public Opinion

- Based on AdaBoost-IWOA-Elman Model and CEEMDAN Algorithm," IAENG International Journal of Computer Science, vol. 49, no.4, pp975-984, 2022
- [26] Gerardo Beni, and Jing Wang, "Swarm Intelligence in Cellular Robotic Systems," Proceedings of NATO Advanced Workshop on Robots and Biological Systems, vol. 102, pp703-712, 1989
- [27] Yuelin Gao, Qinwen Yang, Xiaofeng Wang, Jiahang Li, and Yanjie Song, "Overview of New Swarm Intelligent Optimization Algorithms," Journal of Zhengzhou University (Engineering Science), vol. 43, no. 3, pp21-30, 2022
- [28] J. Kennedy, and R. Eberhart, "Particle swarm optimization," Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp1942-1948, 1995
- [29] Jiankai Xue, and Bo Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," Systems Science & Control Engineering, vol. 8, no. 1, pp22-34, 2020
- [30] Xin Lv, Xiaodong Mu, Jun Zhang, and Zhen Wang, "Chaos sparrow search optimization algorithm," Journal of Beijing University of Aeronautics and Astronautics, vol. 47, no. 8, pp. 1712-1720, 2021.
- [31] Yuan Yuan, Chufu Shao, Zhichao Cao, Wenxin Chen, Anteng Yin, Hao Yue, and Binglei Xie, "Urban Rail Transit Passenger Flow Forecasting Method Based on the Coupling of Artificial Fish Swarm and Improved Particle Swarm Optimization Algorithms," Sustainability, vol. 11, no. 24, pp7230-7241, 2019
- [32] Wenbo Lu, Chaoqun Ma, and Peikun Li, "Research on Sample Selection of Urban Rail Transit Passenger Flow Forecasting Based on SCBP Algorithm," IEEE Access, vol. 8, no. 5, pp. 89425-89438, 2020.
- [33] Tao Zhang, Xiaofeng Yang, Kun Qin, Feifei Li, and Wenshan Luo, "Analysis of path optimization algorithm using seagull theory," Bulletin of Surveying and Mapping, vol. 2022, no. 12, pp110-115, 2022
- [34] Wanli Yang, Xueting Zhou, and Mengna Chen, "New Chaotic Simplified Particle Swarm Optimization Algorithm Based on Logistic Mapping," Computer and Modernization, vol. 12, no. 12, pp15-20, 2019
- [35] Min Han, and Yong He, "Adaptive multi-objective particle swarm optimization with Gaussian chaotic mutation and elite learning," Control and Decision, vol. 31, no. 8, pp1372-1378, 2016
- [36] Lu Zeng, Zinuo Li, Jie Yang, and Xinyue Xu, "Research on short-term passenger flow prediction method of urban rail transit based on CEEMDAN-IPSO-LSTM," Journal of Railway Science and Engineering, no. 12, pp1-14, 2022
- [37] Ei Maazoui Q., Retbi A., and Bennani S., "AUTOMATISATION HYPERPARAMETERS TUNING PROCESS FOR TIMES SERIES FORECASTING: APPLICATION TO PASSENGER'S FLOW PREDICTION ON A RAILWAY NETWORK," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 4, no. 12, pp53-60, 2022
- [38] Xingchen Song, Tongxi Wang, Ziming Xu, and Zhan Shi, "Multistep Forecasting Method of Short-term Power Load Based on VMD-Prophet-Seq2seq," IAENG International Journal of Computer Science, vol. 49, no.3, pp720-727, 2022
- [39] Dan Yang, Kairun Chen, Mengning Yang, and Xiaochao Zhao, "Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features," IET Intelligent Transport Systems, vol. 13, no. 10, pp1475-1482, 2019
- [40] Hongtao Li, Kun Jin, Shaolong Sun, Xiaoyuan Jia, and Yongwu Li, "Metro passenger flow forecasting though multi-source time-series fusion: An ensemble deep learning approach," Applied Soft Computing, no. 2, pp108644-108667, 2022
- [41] Huang Sen, "Time Series Prediction based on Improved Deep Learning," IAENG International Journal of Computer Science, vol. 49, no.4, pp1133-1138, 2022
- [42] Qiong Jiang, "GMM Clustering Based on WOA Optimization and Space-Time Coupled Urban Rail Traffic Flow Prediction by CEEMD-SE-BiGRU-AM," Mobile Information Systems, vol. 2022, no. 3, 2022.