Journal of Sustainable Built Environment

https://doi.org/10.70731/yk3fpz22

Short-term Traffic Flow Prediction: A Method of MEA-LSTM Model Based on Chaotic Characteristics Analysis

Zixuan Zhang [a], Dongliang Zhao [b], Yan Li [c], Zhe Huang [d,*]

[a] School of Transportation Engineering, Chang'an University, Xi'an 710064, China
[b] Shandong Provincial Communications Planning and Design Institute Group Co., LTD, Jinan 250002, China
[c] School of Civil Aviation, Northwestern Polytechnical University, Xi'an 710072, China
[d] Hebei Xiong'an New Area Management Committee Construction and Transportation Management Bureau, Xiong'an 070001, China

KEYWORDS

short-term traffic flow prediction,
intelligent transportation,
chaotic characteristics,
machine learning

ABSTRACT

To effectively improve the accuracy of short-term traffic flow prediction, an improved Long Short-Term Memory (LSTM) method is proposed using the Mind Evolution Algorithm (MEA). Firstly, to address the issues of abnormal and missing traffic flow data, a Neighborhood Stacked Denoising AutoEncoder (NSDAE) is used for data repair. Then, the maximum Lyapunov exponent is used to determine the chaotic characteristics. Meanwhile, based on Bayesian estimation theory, the features of three-parameter sequences are fused in high-dimensional space using phase space reconstruction technique to obtain reconstructed multi-parameter fused traffic flow data. Finally, by taking advantage of the fact that MEA can divide the data into several subpopulations for optimal search separately, a prediction model based on MEA to improve LSTM is proposed. The results show that compared to the other two traditional data restoration methods, the NSDAE has higher accuracy, with the lowest average values of RMSE, MAE, and MAPE. Through the phase space reconstruction technique, the feature fusion of three parameters of traffic flow is realized in high-dimensional space, which makes up for the insufficiency of a single time-series data that cannot comprehensively levy the characteristics of traffic flow. The MEA-LSTM model outperforms the LSTM model in terms of prediction accuracy, computational efficiency, and generalization ability, and its RMSE, MEA, and MAPE are reduced by 24.3%, 28.9%, and 30.1%, respectively.

1. Introduction

Under the background of intelligent transportation, the use of traffic big data to predict future traffic conditions and reasonably guide residents to travel according to the prediction results is an effective way to alleviate traffic congestion [1]. In traffic state prediction, short-time traffic flow prediction (less than 15 min) is a very important branch, which has high practical application value in travel path optimization, traffic diversion, dynamic signal control, and other aspects [2]. At present, due to the sufficient access to traffic big data and the relative maturity of deep learning theory, the adoption of deep neural networks as the core methodology of short-term traffic flow prediction has become a hot-

* Corresponding author at: Hebei Xiong'an New Area Management Committee Construction and Transportation Management Bureau, Xiong'an 070001, China.
E-mail address: huangzhe@xiongan.cn (Zhe Huang)

Received 20 November 2024; Accepted 28 November 2024
Available online 30 November 2024

spot and mainstream of research [3]. AutoEncoders (AE), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Gated Recurrent Unit (GRU), Graph Neural Networks (GNN), and various composite deep neural network models have been used in the field of short-term traffic flow prediction [4-8].

As a kind of RNN, the Long Short-Term Memory (LSTM) shows its powerful time series processing ability in short-term traffic flow prediction, which is especially suitable for dealing with such complex and time-series data as traffic flow. LSTM is designed to combine the short-term and long-term temporal information and exhibits superior time-series prediction performance [9]. Yang et al. proposed a short-term traffic flow prediction method, LSTM+, that can sense both long short-term memory and remarkably long distances. This method can effectively improve the problem of the LSTM extremely long-term memory shortage [10]. Wei et al. constructed an AE-LSTM prediction method. AE obtained the internal relationship of traffic flow by extracting the features of upstream and downstream traffic flow data, and the LSTM network utilized the obtained feature data and historical data to predict complex linear traffic flow data [11]. To promote the forecast accuracy, Zhao et al. proposed a novel traffic forecast model based on LSTM network that considered temporal–spatial correlation in traffic system via a two-dimensional network which was composed of many memory units [12].

However, data completeness and validity are the basis for traffic flow prediction. Traffic flow data comes from multiple sources and the patterns in the data are multimodal, which results in less accurate predictions if only a single variable is considered [13]. Although LSTM networks have achieved good results in traffic flow prediction, insufficient data continuity, lack of integrity, and incomplete inclusion of information will greatly reduce the prediction performance of LSTM [14]. For this reason, many scholars based on the chaotic characteristics of traffic flow parameters, traffic flow phase space reconstruction in order to obtain complete information, comprehensive content of high-quality traffic flow data [15-17]. In addition, the prediction performance of LSTM networks is affected by the hyperparameters (number of nodes in the hidden layer, number of iterative cycles, initial learning rate) [18]. Optimizing the hyperparameters can make the model fit the training data better, improve the prediction accuracy and enhance the generalization ability [19].

Given the above research deficiencies, this paper firstly proposes a Neighborhood Stacked Denoising AutoEncoder (NSDAE) method used for traffic data repair, and highlights its effect by comparing with many other data repair methods. Then, the repaired data is subjected to chaotic system determination, while phase space reconstruction is carried out, and the chaotic characteristics of the three parameters of traffic flow are analyzed. Aiming at the problem that a single traffic flow parameter cannot characterize all the features of the traffic system, the Bayesian estimation theory is introduced to fuse multiple traffic flow parameters in the phase space, and the phase space reconstruction sequence containing multiple traffic flow feature information is obtained, which provides an effective data basis for traffic flow prediction. Finally, by combining the respective advantages of the Mind Evolution Algorithm (MEA) and LSTM model, an improved LSTM prediction method based on the MEA is proposed, and a comparative analysis of the prediction accuracy of the MEA-LSTM model is carried out in terms of prediction accuracy, prediction efficiency, and generalization ability to validate the accuracy of the model.

2. Dataset description

2.1. Data source

This paper used traffic volumes, average location speeds, and lane occupancy rate to characterize traffic flows. The data were obtained from the Performance Measurement System (PeMS), a freeway performance evaluation system developed by the California Department of Transportation in conjunction with the University of California, Berke-

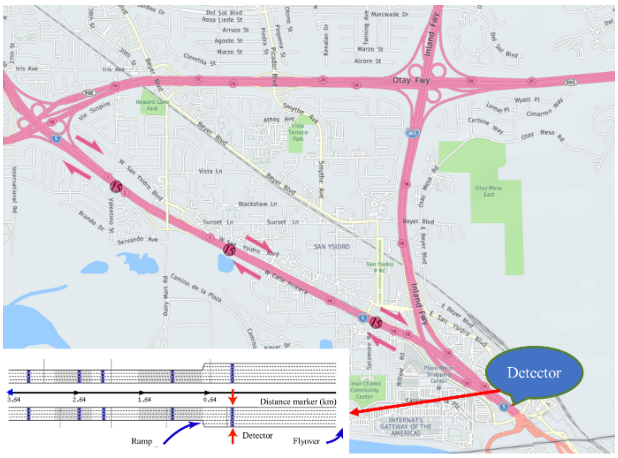


Fig. 1. Target detector location

Table 1. The examples of the raw data

5 Minutes	Flow (Veh/5 Minutes)	Occupancy (%)	# Lane Points	% Observed	Speed (mph)
5/4/2023 0:00	6	0.20	4	75.00	67.10
5/4/2023 0:05	7	0.20	4	75.00	67.00
5/4/2023 0:10	6	0.20	4	75.00	69.60
5/4/2023 0:15	9	0.20	4	75.00	67.80
5/4/2023 0:20	10	0.30	4	75.00	64.50
5/4/2023 0:25	7	0.20	4	75.00	69.10
5/4/2023 0:30	9	0.20	4	75.00	68.80
5/4/2023 0:35	4	0.10	4	75.00	69.00
5/4/2023 0:40	4	0.10	4	75.00	70.10
…	…	…	…	…	…

ley. As shown in Fig. 1, the data detected by detector number 1118326 in the I5-S freeway was selected as the base dataset. Further, the traffic data collected by this detector from May 4, 2023, to May 30, 2023, a total of 27 days of actual roadway measurements, with a data recording interval of 5 min, which meets the requirement of short-term traffic flow prediction duration. The data recording interval of the detector shows that 288 groups of data can be obtained in a single day and a total of 7775 groups of traffic data in the selected period. The examples of the raw data are shown in Table 1.

2.2. *Data patching*

Due to the internal failure of the detector, external changes, and other reasons, the detector data acquisition process results in erroneous data and missing data. Therefore, before the establishment of the traffic flow prediction model, it is necessary to deal with data redundancy, temporal drift, error, loss, and other phenomena occurring in the process of data acquisition [20].

This paper proposed the Neighborhood Stacked Denoising AutoEncoder (NSDAE) model to fill the traffic flow data. Neighborhood is a commonly used method in deep learning, the core of which is to fill in the missing data using the data within the neighborhood of the missing data [21]. Due to the spatio-temporal similarity characteristics of traffic flow data, its missing data are more appropriately handled by the neighborhood correlation filling method. At the same time, combined with Stacked Denoising AutoEncoder (SDAE), different neighborhoods are selected in different times, and the features of the neighborhoods at the missing moments are fully extracted, and then the data are filled in, so as to increase the robustness of the filler

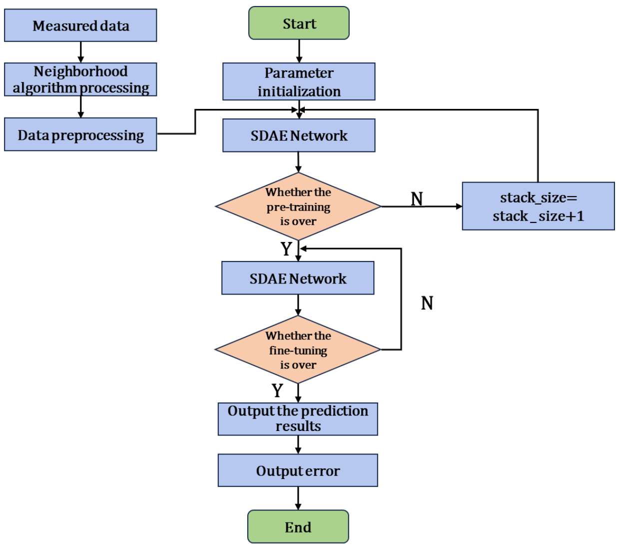


Fig. 2. NSDAE model structure diagram

model, and the data can be obtained closer to the real data [22, 23]. The data patching process for NSDAE is shown in Fig. 2.

NSDAE model is written in Python, where the neighborhood part is written using the NumPy scientific database and the AE part is built with the deep learning framework Keras. The parameter combinations for the NSDAE model are: the missing rate is 10%, the number of SDAEs is 2, each SDAE has three hidden layers, the number of nodes in each SDAE hidden layer is {128, 64, 128}, *window_n=10*, *batch_size=288*, and the number of iterations is 200.

3. Methodology

3.1. *Theory of chaotic properties of traffic flow*

3.1.1. *Lyapunov exponent*

Lyaponuv exponent represents a numerical characteristic of the average exponential divergence of neighboring trajectories in phase space and is often used to determine the chaotic nature of a system [24]. When the dynamical system becomes an n-dimensional discrete system xn+1=F(xn), there exists n Lyapunov exponents $\gamma_1, \gamma_2, \ldots \gamma_n$, Each of γ_i exhibits the motion characteristics of the corresponding orbit. In determining the trajectory of an n-dimensional discrete system, n Lyapunov exponents are generated. The largest of these is called the maximal Lyapunov exponent. If the maximum Lyapunov exponent is positive in a high-dimensional dynamical system, it indicates that the system exhibits chaotic characteristics [25].

3.1.2. Correlation dimension

Correlation dimension D(m) is a measure of the amount of characteristic information contained in chaotic attractors [26]. In chaotic systems, D(m) tends to saturate as the embedding dimension m of the time series increases, so the correlation dimension can also be used as a metric for the determination of chaotic properties. Genetic programming (G-P) algorithm can obtain the time series m by solving the association function [27]. When m increases, if D(m) gradually tends to saturation, then the system satisfies the chaotic property, and D(m) corresponding to the saturation state is the correlation dimension of the attractor of the time series. If the system does not have chaotic properties, D(m) will not tend to saturation, but will grow to positive infinity as m grows [28].

3.1.3. Phase space reconstruction of multiparameter time series

Since the traffic flow data is a one-dimensional time series, it cannot represent the existence of complex motion characteristics within the system. Therefore, the concept of phase space reconstruction needs to be introduced in order to analyze the chaotic characteristics of traffic flow data. Takens points out two parameters that need to be determined in phase space reconstruction: the embedding dimension m and the delay time τ. When selecting the values of m and τ, the connection between them usually needs to be considered [29, 30]. In addition, there is another parameter that in phase space reconstruction: the embedding window width $\tau_W=(m-1)\tau$. In this paper, the C-C algorithm is used to solve τ and τ_W.

After phase space reconstruction, it is difficult to express the complex information inside the chaotic system of traffic flow by considering only a single variable. Therefore, in order to obtain effective data with more complete representation information, this paper fused the three-parameter time series matrices reconstructed from the phase space with phase points in the high-dimensional space based on the internal connection between the three parameters of the traffic flow and based on the Bayesian estimation theory [31]. Based on the fused traffic flow data, a model is built for prediction in order to obtain higher prediction accuracy.

To determine the precise dynamical properties of a chaotic system, it is necessary to locate the strange attractor and then examine its trajectory in higher dimensional space to uncover the regularity of the chaotic behavior. When a chaotic system involves multiple variables, a portion of its trajectory in multi-dimensional space can be seen as illustrated in Fig. 3.

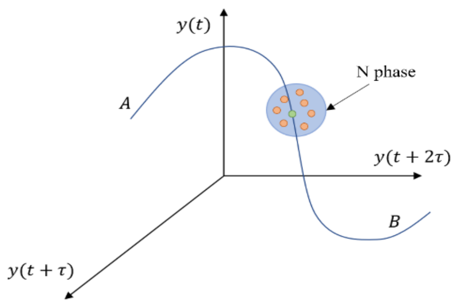


Fig. 3. Phase point distribution in phase space

The real attractor is a phase point in its trajectory, and in principle, if we want to restore its system perfectly, the attractors of the system should all be located in a fixed orbit. However, due to the small information reserve of individual variables in phase space reconstruction, there is a partial deviation between the chaotic attractors generated after phase space reconstruction and the real attractors, which leads to the fact that the chaotic attractors generated by each variable cannot completely present the real characteristics of the original chaotic system. Therefore, these phase points need to be fused to finally obtain an optimal phase point that contains more complete information, has a higher degree of reduction, and is close to the characteristics of the real attractor. By Bayesian estimation theory, given

$$x_k^i = \left[x_{i,k}, x_{i,k+\tau}, \ldots, x_{i,k+(m-1)\tau}\right] \ \forall i = 1,2,\ldots,M;\ k = 1,2,\ldots,K \quad (1)$$

a time series of multiple variables combined, the phase space reconstruction is performed and contains K phase points, which can be expressed by Eq. (1).

$$D_k = \left[x_k^1, x_k^2, \ldots, x_k^M\right] \ \forall k = 1,2,\ldots,K \quad (2)$$

After reconstructing the phase space of the M time series, the k-th phase point of each variable is selected as the set of phase points to be fused, which can be expressed by Eq. (2).

$$P\left(Z_k \mid x_k^1, x_k^2, \ldots, x_k^M\right) = \frac{P\left(Z_k; x_k^1, x_k^2, \ldots, x_k^M\right)}{P\left(x_k^1, x_k^2, \ldots, x_k^M\right)} \quad (3)$$

If $Z_k(k = 1,2,\ldots,K)$ is a phase point after the fusion of the multi-parameter phase space, then the estimated value of $Z_k(k = 1,2,\ldots,K)$ is expressed by Eq. (3).

Assuming $Z_K \sim N(Z_0, \sigma_0^2)$, where Z_0 is the mean value of Z_k and $D_k \sim N(Z_k, \sigma_h^2)$ is the variance,

$$P\left(Z_k \mid x_k^1, x_k^2, \ldots, x_k^M\right) = \gamma \exp\left[-\frac{1}{2}\left(\left(\sum_{h=1}^{M}\frac{1}{\sigma_h^2}+\frac{1}{\sigma_0^2}\right)z_k^2 - 2\left(\sum_{h=1}^{M}\frac{x_h^k}{\sigma_h^2}+\frac{Z_0}{\sigma_0^2}\right)Z_k\right)\right] \quad (4)$$

while, $D_k \sim N(Z_k, \sigma_h^2)$ and σ_h^2 are the covariance

$$\gamma \exp[-\frac{1}{2}((\sum_{h=1}^{M}\frac{1}{\sigma_h^2}+\frac{1}{\sigma_0^2})z_k^2 - 2(\sum_{h=1}^{M}\frac{x_h^k}{\sigma_h^2}+\frac{Z_0}{\sigma_0^2})Z_k)] = \frac{1}{\sqrt{2\pi\sigma}}\exp[-\frac{1}{2}(\frac{Z_k - Z}{\sigma})^2] \quad (5)$$

matrix. Setting the parameter

$$Z = \frac{\sum_{h=1}^{M}\frac{x_h^k}{\sigma_h^2}+\frac{Z_0}{\sigma_0^2}}{\sum_{h=1}^{M}\frac{1}{\sigma_h^2}+\frac{1}{\sigma_0^2}} \quad (6)$$

$\alpha = 1/P(X_k^1, X_k^2, \ldots\ldots, X_k^M)$, the following Eq. (4) is given.

If P obeys a normal distribution, we can obtain

$$\hat{Z}_k = \int Z_k \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{Z_k - Z}{\sigma}\right)^2\right] dZ_k = z; \forall k = 1,2,\ldots,K \quad (7)$$

Eq. (5).

By solving Eq. (5), Eq. (6) can obtain:

Accordingly, the value of the estimated value $\hat{Z}_k$ of the optimal fusion phase point can be obtained by Eq. (7):

Where: K-Total number of phase points in multidimensional space in phase space reconstruction after multi-parameter fusion; k-th phase point in phase space reconstruction after multi-parameter fusion.

3.2. MEA-LSTM model

3.2.1. MEA

Mind Evolution Algorithm (MEA) is a machine learning algorithm that incorporates the two opposing thinking patterns of convergence and divergence found in human cognition. The core of the MEA algorithm is the continuous exploration of the data to find the optimal individual element values in many iterations. In this algorithm, convergence and dissimilation operations are used to continuously generate a new and better subpopulation, and finally the optimal solution is found [32]. The specific steps of the algorithm refer to the Ref. [33]. The optimization process is shown in Fig. 4.

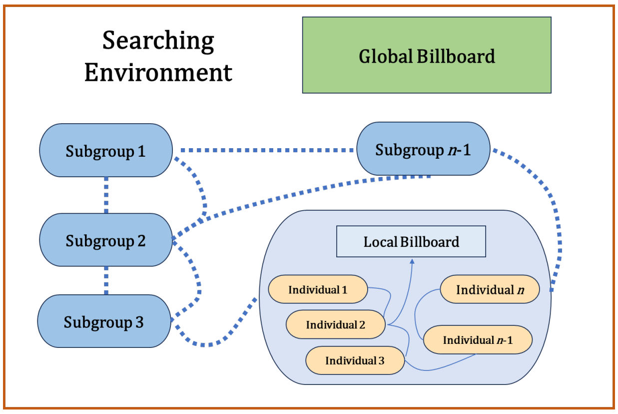


Fig. 4. The mechanism of the MEA algorithm [33]

3.2.2. LSTM

LSTM is a special kind of RNN model, which has both long-time and short-time memory, and improves the gradient dispersion and gradient explosion problems of traditional RNNs. LSTM model has three threshold units, namely, forgetting gate, input gate and output gate, and transfers the information along the temporal sequence through linear operations, which has good memory because the information is re-inputted in a loop each time, and the model achieves the filtering of the input infor-

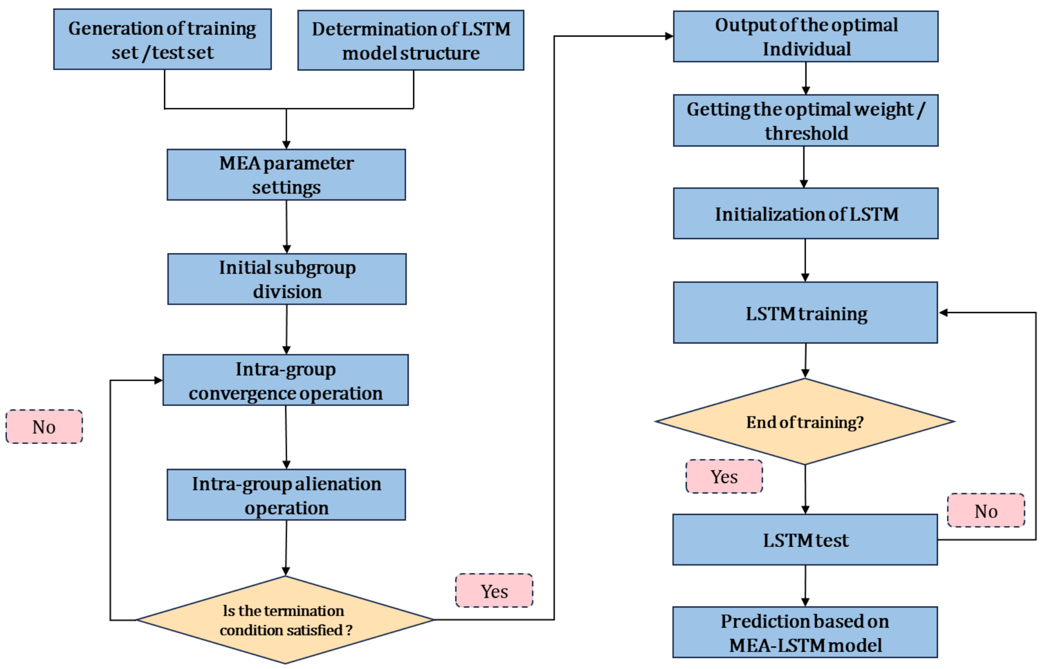


Fig. 5. MEA-LSTM model structure diagram

mation through the ' gate ' structure [34]. The model achieves the filtering of input information through the ' gate ' structure. The specific steps of the algorithm refer to the Ref. [35].

3.2.3. MEA-LSTM model

Aiming at the problems of the LSTM model, this paper improves LSTM model by introducing the MEA, which prompts the output results to reach the optimum quickly. MEA divides the data into several subpopulations for optimization search, alternates between convergence and dissimilation operations, changes the shortcomings of LSTM model parameters that can only be trained one by one, and optimizes the weights and thresholds of the LSTM model quickly. The MEA-LSTM model not only improves the convergence speed and generalization ability of the LSTM model, but also makes the initial weights and thresholds of the model more global, so that the algorithm is more accurate in predicting the traffic flow. Its modelling process is shown in Fig 5.

3.2.4. Parameterization

The traffic flow data after multi-parameter phase space reconstruction is divided into training set and test set. Valid data before 23 May is used as the training set and valid data from 23 May to 30 May is used as the test set.

After searching and trial calculation, the network structure of LSTM model is determined as a single input layer, double hidden layer and single output layer. The excitation functions are sigmoid and tanh functions, the optimization method is Adam, and the number of iterations is set to 500, and the learning rate is set to 0.03. The first 8 periods of the measured traffic flow data are selected as the input values, so the number of neurons in the input layer and the number of neurons in the hidden layer of the LSTM model are 8 and 16 respectively. The main parameters of the MEA model: the population size was set to 800, six winning subpopulations were screened during the run, and six temporary subpopulations were screened; the size of a single subpopulation was 80, and the number of iterations was 20.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \hat{x}_i - x_i^2}{n}} \tag{8}$$

$$MAE = (\frac{\sum_{i=1}^{n} |\hat{x}_i - x_i|}{n}) \tag{9}$$

$$MAPE = (\frac{\sum_{i=1}^{n} |\hat{x}_i - x_i|}{n}) \tag{10}$$

3.2.5. Evaluation indicators

Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) were used to assess the accuracy of the model's prediction results, as shown in Eqs. (8)-(10):

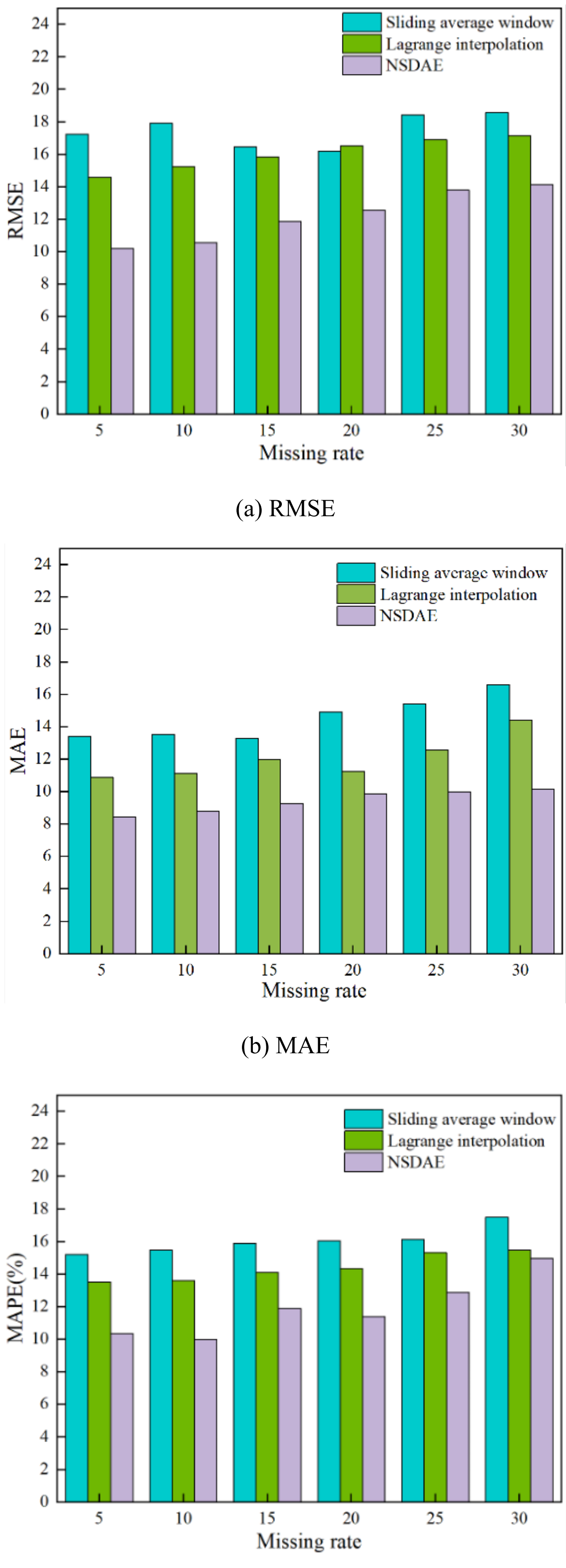
Where: $\hat{x}_i$ - predicted value; x_i - observed value; n - Length of time series data.

4. Results and discussions

4.1. Evaluation of the results of data patching

In order to show the performance of NSDAE method, this paper selects two commonly used data patching methods: the sliding average window method [36] and the Lagrange interpolation method [37] for comparative analysis, and evaluates them in combination with the error metrics. Under the setting of different data missing rate, NSDAE, sliding average window method and Lagrange interpolation method are used for data restoration, respectively. The error results of the three methods were obtained when the value interval of the missing data rate was set to [5,30] and the value interval was 5%, as shown in Fig. 6.

As can be seen from Fig. 6, the errors of all three models show an upward trend with the increase of the missing data rate, with the Lagrange interpolation and NSDAE floating more stably, while sliding average window method fluctuates more. As a whole, when the value interval of the missing data rate is [5,30] and the value interval is 5%, the average value of RMSE of NSDAE is 12.174, the average value of MAE is 9.416, and the average value of MAPE is 11.908%, which is the lowest value in all three models. It can be concluded that the error data repair results of NSDAE are significantly better than the other two methods.



(a) RMSE

(b) MAE

(c) MAPE

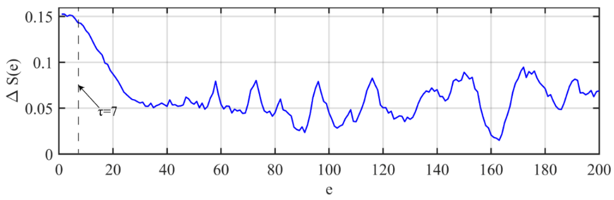
Fig. 6. Comparison of the effectiveness of restoration under different rates of missing data

4.2. Chaotic characterization of traffic flow

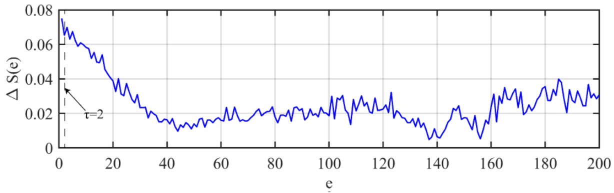
In this paper, valid data (a total of 7,775 sets of traffic data) that have been repaired and processed are used as test data for the determination of chaotic properties and phase space reconstruction.

4.2.1. Three-parameter determination of chaotic properties

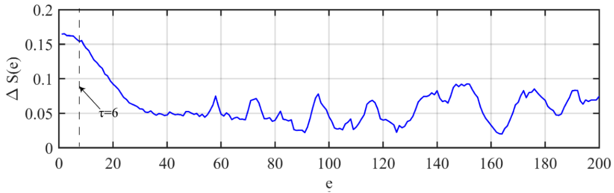
Since the selected valid data contains three variables: traffic volume, average speed and average occupancy rate, the valid data is firstly organized into the time series form, i.e, $\{x(n), n = 1,2,\ldots,7775\}$, for better chaotic characterization. Then, the delay time parameter of the



(a) Traffic volume



(b) Average location speeds



(c) Lane occupancy rate

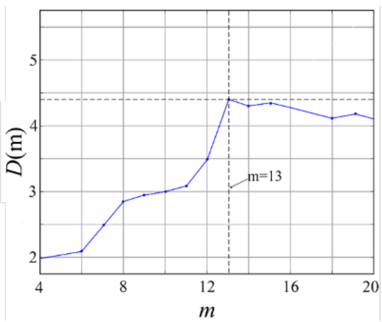
Fig. 7. $\overline{\Delta S}(e)$ curves for three parameters

time series is calculated according to the C-C algorithm, and the optimal delay time is taken as the first local minima of the $\overline{\Delta S}(e)$ curve [38]. The curve for the three parameters is shown in Fig. 7.

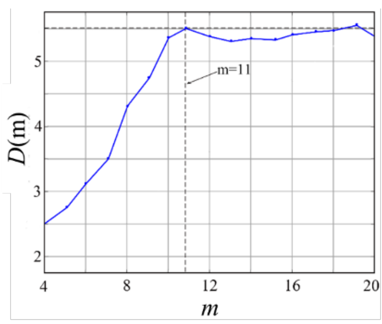
From Fig. 7, the first minima of the three parameters $\overline{\Delta S}(e)$-curves are $\tau = 7$, $\tau = 8$ and $\tau = 9$, respectively, which gives the delay times of the three-parameter time series as $\tau = 7$, $\tau = 8$ and $\tau = 9$, respectively. According to the G-P algorithm, a line graph of the three parameters correlation dimension with the embedding dimension can be obtained, as shown in Fig. 8.

From Fig. 8, it can be seen that D(m) of the three-parameter time series grows with the increase of m. At the moment when m is 13, 11 and 15 respectively, D(m) is close to saturation, so the embedding dimensions are 13, 11 and 15 respectively.

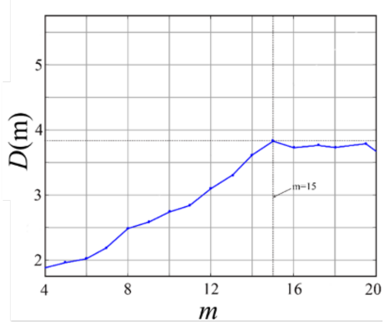
In order to further determine the chaotic properties of the three-parameter time series, the maximum Lyapunov exponent of the series after phase space reconstruction is calculated using the small data volume method [39], and the results are shown in Fig. 9. The red straight line is the regression line fitted by the least squares method, and the slope of this line is the maximum Lyapunov exponent of the



(a) Traffic volume

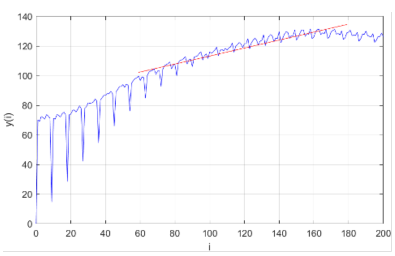


(b) Average location speeds

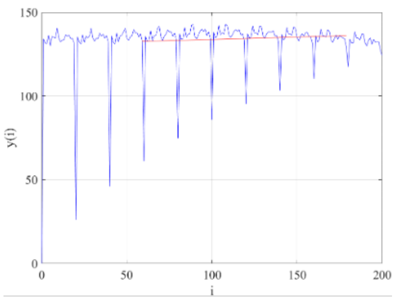


(c) Lane occupancy rate

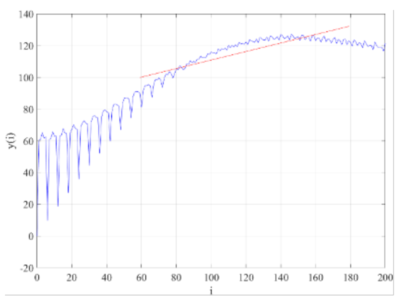
Fig. 8. D(m) curves for three parameters.



(a) Traffic volume

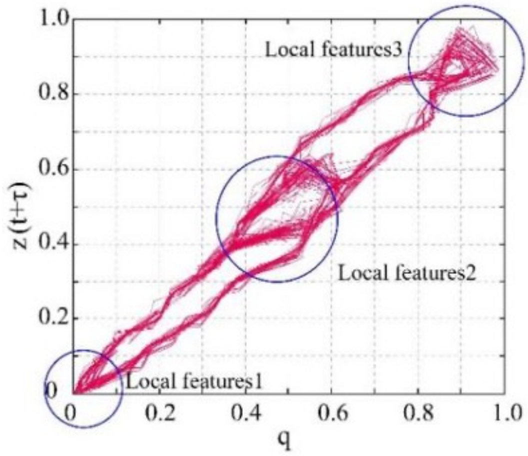


(b) Average location speeds

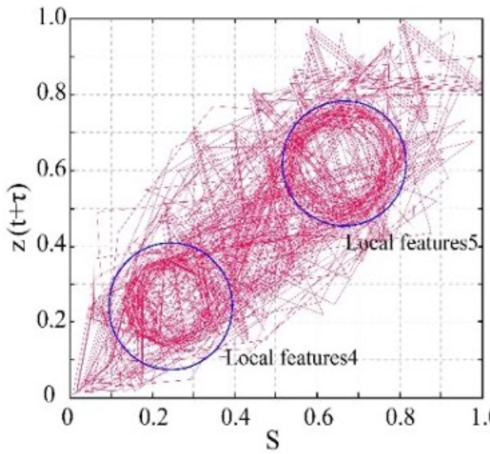


(c) Lane occupancy rate

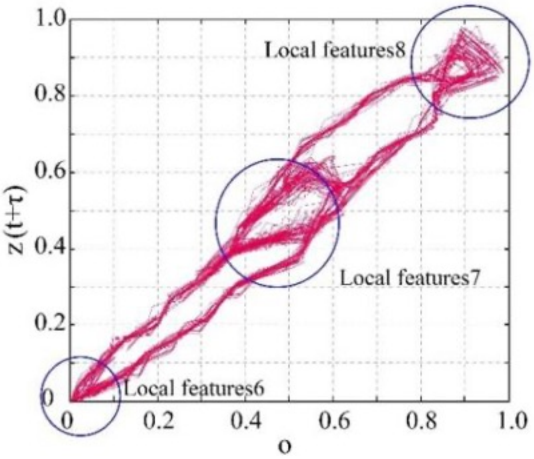
Fig. 9. Maximum Lyapunov exponents curves for three parameters.



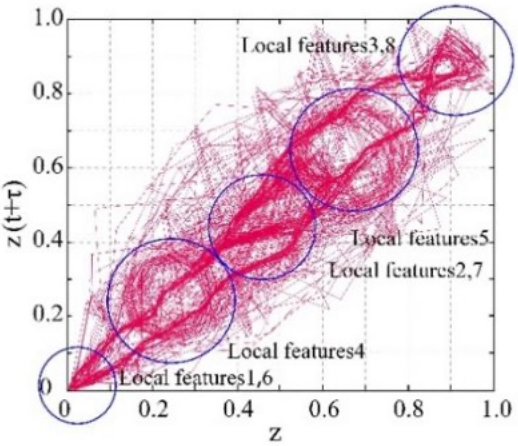
(a) traffic volume



(b) average speed



(c) average occupancy rate



(d) three-parameter fusion

Fig. 10. Phase space reconstruction of chaotic attractors based on multi-parameter fusion.

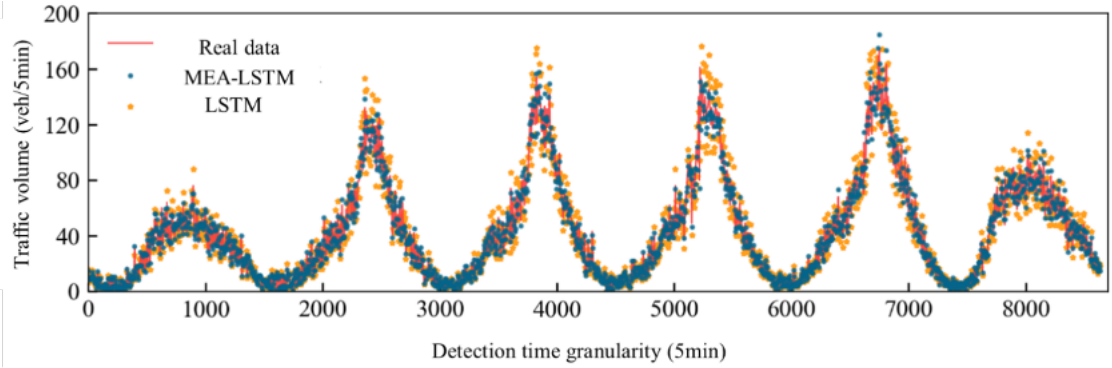
sequence. The maximum Lyapunov exponents of the three-parameter time series after phase space reconstruction are 0.2642, 0.0283 and 0.2615, respectively, and all of them are greater than 0, which proves that the series is a chaotic system.

4.2.2. Three-parameter time series phase space fusion

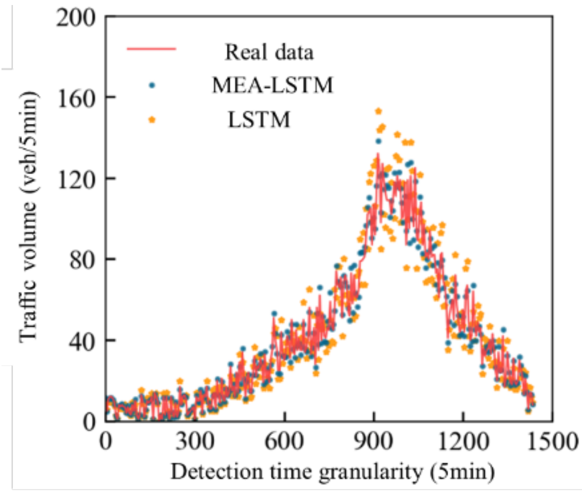
In order to obtain multi-parameter fused traffic flow data, the three-parameter sequences are reconstructed in the same dimension of the phase space using the phase space reconstruction technique. According to Takens' embedding theorem, sequences with chaotic properties can all characterize a more complete phase space structure of the original system when the appropriate delay time and embedding dimension are chosen. Therefore, phase space reconstruction of three-parameter sequences in the same high-dimensional space firstly needs to normalize the three sequences and determine the embedding dimension m. As shown in Fig. 8, all three sequences are essentially saturated at an embedding dimension of 11. Therefore, it is determined that m = 11. Based on the embedding window width solved by the C-C algorithm, the opti-

Table 2. Traffic flow time series phase space reconstruction parameters.

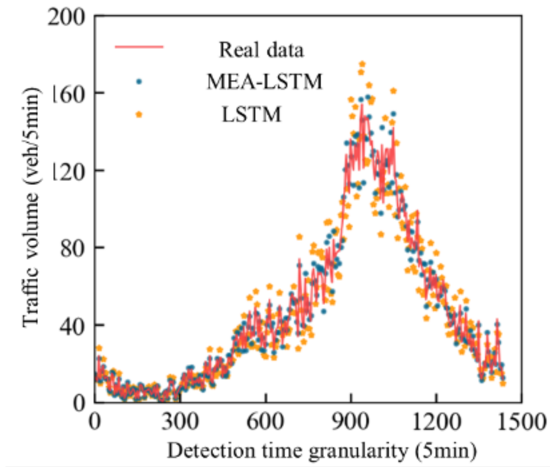
Traffic flow parameter	Delay time	Embedding dimension
Traffic volume	9	11
Average speed	7	11
Average occupancy rate	6	11



(a) Comparison of results from May 25 to May 30



(b) Comparison of results on May 26



(c) Comparison of results on May 27

Fig. 11. Comparison of model prediction results

mal delay time is inverted according to $\tau_w = (m-1)\tau$. After calculation, the traffic flow parameters phase space reconstruction parameters are shown in Table 2.

The three-parameter phase space reconstruction vectors are reconstructed into a high-dimensional phase space of the same dimension, and the Bayesian estimation theory is applied in this high-dimensional space to realize the fusion of the phase points, and a traffic flow data matrix containing more complete feature information is obtained, and the chaotic attractor of this chaotic system is shown in Fig. 10.

As can be seen from Fig. 10, the region of chaotic attractors in the aggregation state is a local feature of the sequence. The chaotic attractor representation maps of the traffic volume and occupancy time series in the traffic flow parameters are to some extent similar, showing three local features, but the chaotic attractor of the average speed time series is different from the first two variables, which also indicates that only using a single variable for phase space reconstruction and traffic flow prediction cannot represent all the characteristics of traffic flow chaotic system, and the prediction accuracy is low. After the fusion of multiple variables in high-dimensional phase space using Bayesian estimation theory, the local characteristics of chaotic attractors of multiple variables are reflected. It shows that after multivariate time series attractor fusion, the fused phase space information contains all the important characteristics of the measured traffic flow data.

The multi-parameter fusion technique proposed in this paper can not only show all the original main features of traffic flow as a dynamical system, but also show the characteristic information of traffic flow parameters in a more comprehensive way than the reconstruction information that can be achieved by a single variable. In the subsequent traffic flow prediction, the application of the data that have undergone multi-parameter phase space fusion to comprehensively consider multiple main features is very favorable to improve the prediction accuracy.

4.3. Traffic flow prediction results based on MEA-LSTM model

The prediction results and prediction errors of the LSTM model and the MEA-LSTM model are shown in Fig. 11 and Table 3.

Combined with Fig. 11 and Table 3, the comparative analysis shows that the fitting degree of the MEA-LSTM model is better than the LSTM model in terms of prediction accuracy, and each evaluation indicator has been optimized, including RMSE index decreased by 2.271, MEA index decreased by 3.096, MAPE index decreased by 3.61 %. It can be seen that the prediction accuracy of the LSTM model optimized by MEA has improved over the single LSTM model, indicating that the MEA-LSTM model proposed in this paper is feasible in the practical application of traffic flow prediction. In terms of prediction speed, the prediction time of the LSTM model is 14.5 min, while that of the MEA-LSTM model is 4.7 min, which is 3.1 times higher. Comprehensive analysis of the experimental results, and the reasons for improving the accuracy and prediction speed of the MEA-LSTM model are as follows: 1) The MEA model can carry out the search process for the optimal solution from multiple populations as well as multiple elements simultaneously, thus improving the search efficiency of the model and giving the model a faster convergence rate. 2) The combination of the MEA model with the LSTM model introduces an optimization-preserving property, where its optimal individuals are preserved after each iteration, ensuring that the run results are all closer to the optimal solution than the previous layer of the model. 3) As the number of iterations increases, the training error of the MEA-LSTM model varies more and more, which indicates that it will always be more efficient in solving the optimal parameters towards the global optimum.

5. Conclusions

The paper first proposes the NSDAE method to repair traffic data and compares it with various other data repair methods to highlight its effectiveness. Meanwhile, an improved LSTM prediction method is constructed based on the MEA by combining the advantages of the MEA and the LSTM model. The main findings are as follows.

(1) In this paper, we propose the NSDAE method to repair the original traffic flow data and compare its performance with other methods to determine the advantages of NSDAE. Through this method, effective data for making short-term traffic flow predictions were obtained.

(2) Considering that it is difficult to comprehensively characterize the traffic flow using univariate data for prediction, this paper firstly determines the chaotic characteristics of the traffic flow using the chaos theory determination method, analyses the similarities and differences among the three parameters, and finds out their important features. Then, according to the determination results, the Bayesian estimation method is selected to reconstruct the phase space of the three parameters, so that the traffic flow parameters are reconstructed into one-dimensional data, and can represent the data characteristics of the three parameters. Reconstructing the multi-dimensional data phase space into one-dimensional data provides a data basis for the establishment of the short-time traffic flow prediction model, and also reduces the computational cost of the model.

(3) In this paper, we take advantage of the fact that the MEA can divide the data into several subpopulations for optimal search, and perform convergence and dissimilation operations alternately, so that the weights and thresholds of the prediction model can be optimized quickly, and then propose the MEA-LSTM model. The MEA-LSTM model proposed in this paper improves the prediction accuracy, computational efficiency, and generalization ability, and can provide a reference basis for related research.

Acknowledgments

This work was supported by Shaanxi Natural Science Foundation for basic research plan under Grant 2024JC-YBMS-373. The authors would like

Table 3. Prediction errors of the LSTM model and the MEA-LSTM model.

Evaluation indicators	LSTM model	MEA-LSTM model
RMSE	9.332	7.061
MAE	10.682	7.586
MAPE	11.842%	8.232%

to express their sincere appreciation to the aforementioned organizations.

Declaration of Competing Interest

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

References

1. Lv Y, Duan Y, Kang W, et al. Traffic flow prediction with big data: A deep learning approach. IEEE transactions on intelligent transportation systems, 2014, 16(2): 865-873.
2. Shu W, Cai K, Xiong N N. A short-term traffic flow prediction model based on an improved gate recurrent unit neural network. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(9): 16654-16665.
3. Hou Q, Leng J, Ma G, et al. An adaptive hybrid model for short-term urban traffic flow prediction. Physica A: Statistical Mechanics and its Applications, 2019, 527: 121065.
4. Chen C, Liu Z, Wan S, et al. Traffic flow prediction based on deep learning in internet of vehicles. IEEE transactions on intelligent transportation systems, 2020, 22(6): 3776-3789.
5. Miglani A, Kumar N. Deep learning models for traffic flow prediction in autonomous vehicles: A review, solutions, and challenges. Vehicular Communications, 2019, 20: 100184.
6. Tian Y, Pan L. Predicting short-term traffic flow by long short-term memory recurrent neural network// 2015 IEEE international conference on smart city. IEEE, 2015: 153-158.
7. Sun Z, Hu Y, Li W, et al. Prediction model for short-term traffic flow based on a K-means-gated recurrent unit combination. IET Intelligent Transport Systems, 2022, 16(5): 675-690.
8. Peng H, Wang H, Du B, et al. Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. Information Sciences, 2020, 521: 277-290.
9. Lu Z, Lv W, Cao Y, et al. LSTM variants meet graph neural networks for road speed prediction. Neurocomputing, 2020, 400: 34-45.
10. Yang B, Sun S, Li J, et al. Traffic flow prediction using LSTM with feature enhancement. Neurocomputing, 2019, 332: 320-327.
11. Wei W, Wu H, Ma H. An autoencoder and LSTM-based traffic flow prediction method. Sensors, 2019, 19(13): 2946.
12. Zhao Z, Chen W, Wu X, et al. LSTM network: a deep learning approach for short-term traffic forecast. IET intelligent transport systems, 2017, 11(2): 68-75.
13. Abadi A, Rajabioun T, Ioannou P A. Traffic flow prediction for road transportation networks with limited traffic data. IEEE transactions on intelligent transportation systems, 2014, 16(2): 653-662.
14. Xiao Y, Yin Y. Hybrid LSTM neural network for short-term traffic flow prediction. Information, 2019, 10(3): 105.
15. Ma Q, Huang G H, Ullah S. A multi-parameter chaotic fusion approach for traffic flow forecasting. IEEE Access, 2020, 8: 222774-222781.
16. Hong W C. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. Neurocomputing, 2011, 74(12-13): 2096-2107.
17. Cheng A, Jiang X, Li Y, et al. Multiple sources and multiple measures-based traffic flow prediction using the chaos theory and support vector regression method. Physica A: Statistical Mechanics and its Applications, 2017, 466: 422-434.
18. Tang J, Zeng J, Wang Y, et al. Traffic flow prediction on urban road network based on license plate recognition data: combining attention-LSTM with genetic algorithm. Transportmetrica A: Transport Science, 2021, 17(4): 1217-1243.
19. Cai L, Lei M, Zhang S, et al. A noise-immune LSTM network for short-term traffic flow forecasting. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2020, 30(2).
20. Li P, Dong B, Li S, et al. A repair method for missing traffic data based on FCM, optimized by the twice grid optimization and sparrow search algorithms. Sensors, 2022, 22(11): 4304.
21. Li J, Li R, Xu L. Multi-stage deep residual collaboration learning framework for complex spatial–temporal traffic data imputation. Applied Soft Computing, 2023, 147: 110814.
22. Liu X, Zhang H, Niu Y, et al. Modeling of an ultra-supercritical boiler-turbine system with stacked denoising auto-encoder and long short-term memory network. Information Sciences, 2020, 525: 134-152.
23. Mohanty H, Roudsari A H, Lashkari A H. Robust stacking ensemble model for darknet traffic classification under adversarial settings. Computers & Security, 2022, 120: 102830.
24. Young L S. Mathematical theory of Lyapunov exponents. Journal of Physics A: Mathematical and Theoretical, 2013, 46(25): 254001.
25. Sahoo S, Roy B K. Design of multi-wing chaotic systems with higher largest Lyapunov exponent. Chaos, Solitons & Fractals, 2022, 157: 111926.
26. Ghorbani M A, Kisi O, Aalinezhad M. A probe into the chaotic nature of daily streamflow time series by correlation dimension and largest Lyapunov methods. Applied Mathematical Modelling, 2010, 34(12): 4050-4057.
27. Ghorbani M A, Khatibi R, Mehr A D, et al. Chaos-based multigene genetic programming: A new hybrid strategy for river flow forecasting. Journal of hydrology, 2018, 562: 455-467.
28. Jiang J J, Zhang Y, McGilligan C. Chaos in voice, from modeling to measurement. Journal of Voice, 2006, 20(1): 2-17.

29. Takens F. The reconstruction theorem for endomorphisms. Bulletin of the Brazilian Mathematical Society, 2002, 33: 231-262.
30. Noakes L. The Takens embedding theorem. International Journal of Bifurcation and Chaos, 1991, 1(04): 867-872.
31. Bosq D, Ruiz-Medina M D. Bayesian estimation in a high dimensional parameter framework. 2014.
32. Bai H, Cao Q, An S. Mind evolutionary algorithm optimization in the prediction of satellite clock bias using the back propagation neural network. Scientific Reports, 2023, 13(1): 2095.
33. Liu H, Tian H, Liang X, et al. New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks. Renewable Energy, 2015, 83: 1066-1075.
34. Sayed S A, Abdel-Hamid Y, Hefny H A. Artificial intelligence-based traffic flow prediction: a comprehensive review. Journal of Electrical Systems and Information Technology, 2023, 10(1): 13.
35. Tian Y, Zhang K, Li J, et al. LSTM-based traffic flow prediction with missing data. Neurocomputing, 2018, 318: 297-305.
36. Lee C, Ryu T, Kim H, et al. Efficient approach of sliding window-based high average-utility pattern mining with list structures. Knowledge-Based Systems, 2022, 256: 109702.
37. Kai Z, Jinchun S, Ke N, et al. Lagrange interpolation learning particle swarm optimization. PloS one, 2016, 11(4): e0154191.
38. Said A, Abbasi R A, Maqbool O, et al. CC-GA: A clustering coefficient based genetic algorithm for detecting communities in social networks. Applied Soft Computing, 2018, 63: 59-70.
39. Rosenstein M T, Collins J J, De Luca C J. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D: Nonlinear Phenomena, 1993, 65(1-2): 117-134.