Hindawi
Discrete Dynamics in Nature and Society
Volume 2022, Article ID 5604674, 22 pages
https://doi.org/10.1155/2022/5604674

Research Article

An Improved Phase Space Reconstruction Method-Based Hybrid Model for Chaotic Traffic Flow Prediction

Yue Hou, Da Li, Di Zhang, and Zhiyuan Deng

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Correspondence should be addressed to Yue Hou; houyue@mail.lzjtu.cn

Received 16 June 2022; Revised 5 August 2022; Accepted 18 August 2022; Published 23 September 2022

Academic Editor: A. E. Matouk

Traffic flow is chaotic due to nonstationary realistic factors, and revealing the internal nonlinear dynamics of chaotic data and making high-accuracy predictions is the key to traffic control and inducement. Given that high-quality phase space reconstruction is the foundation of predictive modeling. Firstly, an improved C-C method based on the fused norm search domain is proposed to address the issue that the C-C method in the phase space reconstruction algorithm does not meet the Euclidean metric accuracy and reduces the reconstruction quality when the infinite norm metric is used. Secondly, to address the problem of insufficient learning ability of traditional convolutional combinatorial modeling for complex phase space laws of chaotic traffic flow, the high-dimensional phase space features are extracted using the layer-by-layer pretraining mechanism of convolutional deep belief networks (CDBNs), and the temporal features are extracted by combining with long short-term memory (LSTM). Finally, an improved probabilistic dynamic reproduction-based genetic algorithm (PDRGA) is proposed to address the problem of the hybrid model falling into a local optimum when learning the phase space law. Experiments are conducted in three aspects: phase space reconstruction quality analysis, comparison of optimization algorithm convergence, and prediction model performance comparison. The experimentation with two data sets demonstrates that the improved C-C method combines the advantages of the high accuracy metric of the L2 norm with the low operational complexity of the infinite norm, achieving a balance between reconstruction quality and algorithm efficiency. The proposed PDRGA optimization algorithm is a lightweight improvement of the traditional genetic algorithm (GA) and solves the problem that the model tends to fall into a local optimum by optimizing the initial weights of CDBN. Meanwhile, the five error evaluation indexes of the proposed PDRGA-CDBN-LSTM hybrid model are lower than those of the baseline model, providing a new modeling idea for chaotic traffic flow prediction.

1. Introduction

With the expansion of car ownership in cities, the contradiction between the limited urban traffic road network resources and the increasing traffic demand of inhabitants is becoming more pronounced, which can potentially lead to road congestion, poor transportation efficiency, and high fuel consumption. Analyzing the characteristics of traffic flow data and making predictions is one of the most effective ways to mitigate such issues, as it can provide suggestions for route planning and traffic control for travelers and traffic management, and is a crucial step in promoting the development of smart transportation systems in urban areas. Traffic flow data are considered to have three types of terms, namely link flow, path flow, and node pair demand (O/D) [1], with link flow being influenced by road network structure, real-time road conditions, and uncertainty events and exhibiting nonlinear chaotic dynamics [2]. The characterization and prediction of link traffic flow data with chaotic features has been one of the most popular study issues in the field of transportation.

The research on chaotic link traffic flow prediction consists of two major phases: the chaoticity analysis phase and the traffic flow prediction phase. First, the analysis of the chaotic dynamics of traffic flow is an important foundation for traffic flow prediction, and the crucial step is the chaotic phase space reconstruction operation, which can fully expand the chaotic attractors collapsed in the original time series data into a new data organization [3], so as to more intuitively demonstrate the chaotic nature of the data. Then,

following the completion of the chaotic characterization of the data, a traffic flow prediction model suited for the chaotic phase space data feature extraction is created. The prediction of traffic flow should take into account not only the training capability, prediction accuracy, and robustness of the selected prediction model but also the feature extraction capability of the model for chaotic phase space data organization, so that the model can extract rich dynamical information from chaotic data, thereby enhancing its nonlinear fitting performance. Therefore, the research on predicting chaotic traffic flow focuses on two aspects: the chaotic phase space reconstruction operation of traffic flow data and the selection of traffic flow prediction models.

The phase space reconstruction of chaotic traffic flows is based on Takens' theorem [4], which finds the orderly pattern from the seemingly disordered random-like data distribution and reorganizes the data form by the phase space reconstruction method, so that the attractors in the traffic flow are expanded in the high-dimensional space as the phase space, which facilitates the extraction of chaotic features in subsequent prediction models. The embedding dimension and delay time, according to this theorem, are the two key reconstruction parameters for phase space reconstruction of traffic flow data [5]. Currently, there are two types of methods for phase space reconstruction. One type of algorithm calculates the embedding dimension and delay time of phase space reconstruction parameters independently, such as the false nearest neighbor method, G-P method, and minimum description length method [6–8] for calculating the embedding dimension and the mutual information method and autocorrelation method [9] for calculating the delay time. The other type assumes that the two reconstruction parameters do not follow an independent identical distribution [10], and can find the correlation between the two parameters using the embedding window theory to calculate the two reconstruction parameters simultaneously with higher efficiency, such as the C-C method [11]. The operation of traffic flow phase space reconstruction provides a database for the research of prediction models. In this study, the phase space is reconstructed using the C-C method, which has the properties of simple calculation and high efficiency.

The construction of a traffic flow model is the primary technical method for predicting traffic flow, and it is also the embodiment of the application value of traffic flow characteristics analysis. There are three primary phases in the study of models for predicting traffic flow: statistical models, classical machine learning models, and deep learning models. Earlier, statistical models were often used to predict traffic flow time series with simpler characteristics, such as the autoregressive model, the moving average model, the Kalman filter model, and the Autoregressive integrated moving average model (ARIMA) [12, 13], among which the ARIMA model has strong forecasting ability, and the seasonal ARIMA (SARIMA) [14] performs better in the forecasting task with seasonal adjustment to tap into the periodic similarity pattern in the traffic flow time series based on ARIMA. Nonetheless, when dealing with data with more complex characteristics, it is difficult for statistical models to capture effective traffic flow evolution patterns. Hence, machine learning algorithms with stronger nonlinear fitting capabilities, such as linear regression, k-nearest neighbor, and neural networks [15–17], have received more attention. As an important branch of machine learning, neural network models, including artificial neural networks (ANN), fuzzy neural networks (FNN), and radial basis neural networks (RBFNN), are equipped with a multinode network memory function to extract more complex nonlinear feature information from the historical traffic flow. Deep learning models deepen the hierarchical structure of neural networks [18, 19], and this multi-hidden layer network improves the value density of feature information in the process of multilevel parameter transfer, abstracts the low-level feature distribution into high-level feature information, and strengthens the feature representation ability of the model in comparison to traditional neural networks, which can learn the deeper traffic flow evolution laws. Deep belief networks (DBNs) [20] expand the network depth by multilayer stacking based on restricted Boltzmann machines, whose hidden layer units are trained to capture the correlation of higher-order data exhibited at the visual layer, thereby enabling the network to more closely approximate the real system energy state of the data. The LSTM [21] model, on the other hand, utilizes a multiunit depth gating mechanism to improve the model's ability to retain the data's evolving attributes over time. The outstanding fitting capacity of the deep learning model serves as a significant guide for the model's construction.

This study proposes a method for predicting connection traffic flow data with chaotic characteristics. First, an improved phase space reconstruction C-C method is used to demonstrate the chaotic dynamics of the data. Second, a hybrid deep learning model architecture applicable to chaotic phase space data is used to extract the data features from the phase space data, and an improved intelligent algorithm component is installed on the model to optimize its parameters. The remainder of the paper is structured as follows: Section 2 describes the current state of the research field and the paper's contribution; Section 3 introduces the data chaos analysis process with an emphasis on the improved phase space reconstruction principle; Section 4 describes the main framework of the hybrid deep learning model; In Section 5, experiments are conducted and analyzed in terms of phase space reconstruction quality analysis, comparison of convergence performance of optimization algorithms, and comparison of prediction model performance; Finally, conclusions are drawn in Section 6.

2. State of Art

Chaotic link traffic flow prediction necessitates consideration of the problems that exist in each of the two phases, specifically the limited computational accuracy and efficiency of the phase space reconstruction algorithm used in the chaos analysis phase and the poor learning ability of the model constructed in the traffic flow prediction phase for the evolutionary laws of chaotic data. Consequently, the research follows two approaches: the reconstruction

performance enhancement problem of the phase space reconstruction algorithm and the deep learning prediction model construction problem of chaotic traffic flow.

In the first approach, a succession of improvement schemes for enhancing the reconstruction quality and reducing the algorithm's computational complexity have been developed. Among the methods to calculate the embedding dimension and delay time of reconstruction parameters independently, Zhang et al. [22] incorporated the saturation judgment parameter into the Cao method and applied the complementary criterion to jointly compute the saturation reference, thereby reducing the subjectivity of embedding dimension assessment. Li et al. [23] applied the density clustering algorithm and the particle swarm algorithm to the G-P algorithm to establish an optimization model for the automatic identification of scale-free zones, which compensates for the deficiency that the definition of scale-free zones used for embedding dimension determination is overly dependent on human experience. Jiang et al. [24] utilized the probability distribution matrix to determine the number of distribution points in each edge region, which simplifies the calculation of delay time in response to the complex criterion for determining the number of intervals of equal edge distribution in the mutual information method. The C-C method based on the embedding window theory can calculate the two reconstruction parameters concurrently with higher efficiency. However, the $L2$ norm measure used in the C-C method to measure the high-dimensional phase point spacing conforms to the definition of Euclidean distance without metric error, but its calculation logic is more complicated and generates more redundancy. Wang and Chen [25] substituted the infinite norm phase point spacing metric for the $L2$ norm while calculating the phase space reconstruction parameters based on the topological equivalence theory. In comparison to the $L2$ norm, the infinite norm metric is easier to calculate and can effectively improve computational efficiency. However, due to the difference between its metric principle and the definition of Euclidean distance, there are inevitable metric errors that affect the algorithm's accuracy. Lu et al. [26] introduced adjustable parameters to trade-off speed and accuracy based on an infinite norm in order to increase the accuracy of delay time selection while decreasing the operational cost. However, the problem of metric bias of the infinite norm in the Euclidean phase point space still exists, and the spacing of high-dimensional phase points cannot be correctly described, indicating that the efficiency of the infinite norm is enhanced at the expense of their accuracy. In view of this, to achieve the goal of high-quality phase space reconstruction, the idea of improving the reconstruction parameters of the C-C method with accuracy tending toward the $L2$ norm and efficiency tending toward the infinite norm needs to be investigated. In light of this, the concept of enhancing the reconstruction parameters of the C-C method with accuracy tending toward the $L2$ norm and efficiency tending toward the infinite norm needs to be investigated in order to achieve high-quality phase space reconstruction.

In the second approach, numerous deep learning algorithms have been applied to traffic flow prediction tasks in recent years. Zhang et al. [27] estimated the data time lag and data volume using a feature selection method and a convolutional neural network (CNN) to discover the characteristics of traffic flow. Fu et al. [28] used a multiunit LSTM and gated recurrent unit (GRU) for short-term traffic flow prediction and compared it to an ARIMA model to validate the deep learning model's nonlinear fitting performance. Despite the fact that the hierarchical structure of the deep learning model can explore the deeper feature patterns of the data to learn the true representation of the data, due to the chaotic nature of the link traffic flow, the deep learning model trained by gradient descent is prone to be sparse when extracting the complex arrangement of the phase points of chaotic data, and the update parameters containing feature information are easily sparse during the long distance between layers, which leads to the phenomenon of gradient dispersion. Moussa and Owais [29] used pretrained deep convolution neural networks (DCNNs) to perform the prediction task and let the model adopt the previous training results through transfer learning. This method of cumulative learning avoids the time-consuming process of learning data from scratch and the potential training issues it may generate. Subsequently, Moussa et al. [30, 31] used the model feature of the deep residual neural network (DRNN) that can bypass some network layers to reduce the transmission distance of feature information, and the model shortcut connection approach substantially diminishes the training difficulties caused by the excessive number of network layers. Huang et al. [32] introduced a multitask regression layer to DBN to predict traffic flow, and the layer-by-layer greedy pre-training mechanism employed in DBN provides complete network parameter adjustments at each layer to prevent the long-distance transfer of feature information. Due to the complexity of the features contained in the chaotic link traffic flow data, not only does the overall data exhibit chaotic dynamics but also the time dependence in the data evolution process. Therefore, the single modeling strategy is inadequate for traffic flow feature extraction. Zheng et al. [33] developed a multimodel combination construction idea, utilizing deep embedding components to capture and identify feature information, CNN to learn traffic flow data characteristics, and LSTM to maintain long-term time reliance on historical data. Huang et al. [34] predicted chaotic systems in the form of a combined CNN and LSTM architecture, and the two models were utilized to learn the chaotic and temporal features of the system, which enabled a more comprehensive understanding of the system dynamics. In light of this, the selection of deep learning models must take into account the applicability of complex chaotic phase space data trained by multilevel neural networks, while the multifeature extraction perspective of the hybrid model building concept can extract the complex data characteristics of chaotic link traffic flow more thoroughly, which can be used as an important modeling reference.

In addition, intelligent algorithmic components are frequently applied in the field of transportation. Owais et al. [35] used a genetic algorithm (GA) with good adaptability and extensibility to solve the bus network design problem. In 2018, this author further considered bus station design [36]

and selected GA as a complete hierarchical solution for the bus network design problem, while multiobjective nature problems are also solved in this approach. The GA algorithm is also frequently employed in traffic flow model optimization problems. Zhang et al. [37] used GA to optimize the filter weights and parameters of the temporal convolutional neural network (TCN) to find the optimal adaptation of traffic flow prediction models, thereby enhancing their accuracy. Using GA, Zhou et al. [38] optimized critical parameters such as penalty parameters of support vector regression in the hybrid traffic flow prediction model to improve the merit-seeking capability during model training. The application of GA components in model parameter optimization strategy is very informative, but the restricted convergence efficiency of GA will additionally increase the training cost of the combined model, and the convergence accuracy of the algorithm also needs to be enhanced.

The main contributions of this paper, which are further explored in light of the aforementioned literature, are summarized as follows: (1). At the stage of chaotic analysis of link traffic flow data, the idea of phase space reconstruction based on the C-C method with an improved fusion norm search domain is proposed to ensure the accuracy of reconstruction parameter calculation while reducing the complexity of the operation, and a game balance is achieved between the quality of the phase space reconstruction and the efficiency of the algorithm. (2). In the stage of prediction model construction, the PDRGA-CDBN-LSTM hybrid deep learning model adaptable to complex feature processing of chaotic link traffic flow is constructed. The combination of CDBN and LSTM can learn the chaotic dynamics and time-dependent relationship of link traffic flow from multiple perspectives, allowing for more comprehensive data feature extraction and enhanced model prediction performance. (3). On the basis of the hybrid model, an intelligent algorithm parameter optimization strategy is further incorporated to realize the lightweight improvement of GA through the probabilistic dynamic reproduction operation (i.e., the PDRGA method), and it is used to assist the optimization of the initial weight parameters of the hybrid model to solve the problem that the initial solution space of the model is randomly distributed, resulting in a tendency to fall into a local optimum during training.

3. Chaotic Phase Space Reconstruction of Link Traffic Flow

Due to the interference of external uncertainties, the actual measured traffic flow data will invariably exhibit various types of noise. It thus affects the quality of the training data and even causes the chaotic nature to degenerate into randomness [39], causing the subsequent prediction process to fail. Figure 1 depicts the processing of chaotic traffic flow data, in which the chaotic traffic flow is first smoothed by the wavelet transform denoising method to remove noise while retaining the original data's local key features. The denoised high-quality measured traffic flow sequence is the input vector for the improved C-C method based on the fused norm metric, which is reconstructed in phase space [40] to expand the attractor structure in high-dimensional space to better show the internal nonlinear law of chaotic data for the chaotic feature extraction task of the hybrid model, and finally, the small data sets method [41, 42] is used to validate the data chaos and normalize it [43] in order to generate the chaotic phase space tensor T and the traffic flow time series S required for the hybrid model.

3.1. Wavelet Transform Denoising Preprocessing. Commonly employed in signal processing, the wavelet transform denoising method has high operational efficiency and outstanding smoothing performance, and its adaptive time-frequency window permits time-frequency localization, which keeps the data's essential characteristics while reducing noise. Given that the essence of the chaotic signal wavelet transform is the projection process of a phase space singular attractor to filter vector space [44, 45], its transformation is comparable to phase space reconstruction topology [46] without compromising the data's chaotic nature rule.

The specific algorithm flow of wavelet transform denoising is depicted in Figure 2. Initially, wavelet decomposition is performed on the original traffic flow data to separate the high-value density traffic flow time series signal from the noise signal. Subsequently, the wavelet coefficients are obtained after selecting the Haar wavelet basis and determining the number of decomposition layers. Finally, the reconstruction of wavelet inversion is carried out to produce the smoothed traffic flow time series.

3.2. The Fundamentals of Phase Space Reconstruction C-C Method. The majority of the research on the internal chaos of denoised traffic flow time series is conducted in high-dimensional space by reconstructing the phase space to enable the system to more precisely depict the structure of chaotic attractors [47, 48]. Due to its low computing complexity, robustness, and ability to get two types of phase space reconstruction parameters concurrently, the C-C method of the reconstruction algorithm is widely used, and it is more suitable for the volume of the observed traffic flow sequence in this research.

The C-C method, which is based on Takens' reconstruction theory, first calculates the correlation dimension [49, 50] as follows:

$$C(m,N,r,t) = \frac{2}{M(M-1)} \sum_{1 \le i < j \le M} \theta\left(r - d_{ij(P)}\right),$$
$$d_{ij(P)} = \left\| X_i - X_j \right\|_{(P)}, \tag{1}$$

where m and t represent the phase space embedding dimensions and time delays, X and M represent the phase points and the number of phase points, r and $d_{ij(P)}$ represent the distance between phase points under the search radius and P norm metrics, respectively, $\theta(x)$ is the Heaviside function, and P is the adopted norm. To reduce the pseudotemporal correlation, the test statistic is defined using an average chunking strategy.

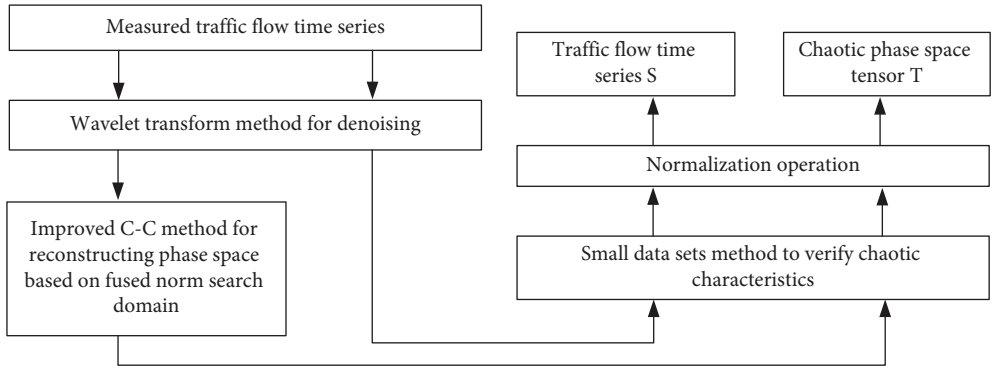


Figure 1: Flowchart of chaotic traffic flow preprocessing.

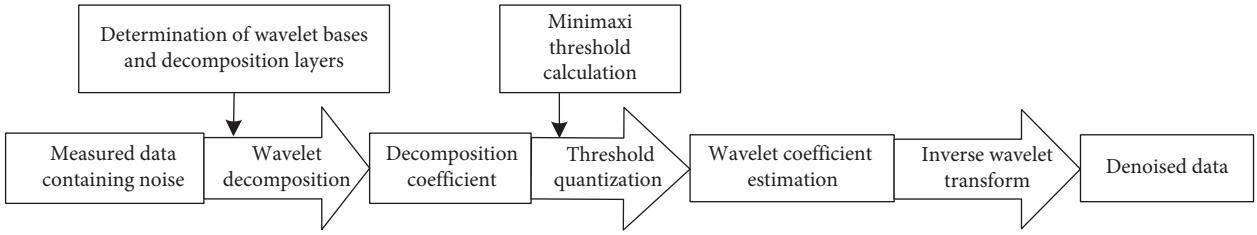


Figure 2: Flowchart of the wavelet denoising algorithm.

$$S(m,N,r,t)=\frac{1}{t}\sum_{i=1}^{t}\left[C_i\left(m,\frac{N}{t},r,t\right)-C_i^m\left(1,\frac{N}{t},r,t\right)\right]. \quad (2)$$

When $n=1$, the definition of the extreme difference statistic is as follows: $N\longrightarrow\infty$

$$\Delta S(m,t)=\max\{S(m,r,t)\}-\min\{S(m,r,t)\}. \quad (3)$$

Together with the BDS statistical test and reasonable estimates of m and r, calculate the statistic as follows:

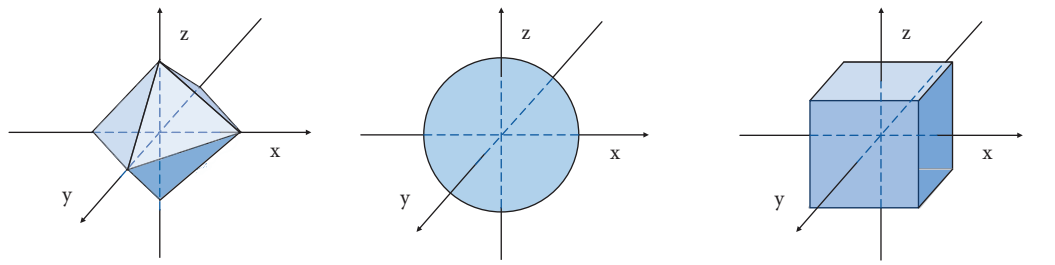
$$\begin{aligned}\overline{S}(t)&=\frac{1}{16}\sum_{m=2}^{5}\sum_{j=1}^{4}S\left(m,r_j,t\right),\\ \Delta\overline{S}(t)&=\frac{1}{4}\sum_{m=2}^{5}\Delta S(m,t),\\ S_{cor}(t)&=\Delta\overline{S}(t)+|\overline{S}(t)|.\end{aligned} \quad (4)$$

The optimal time delay τ_d is determined by the initial local minimum of $\Delta\overline{S}(t)$. According to the embedding window theory, the embedding dimension m can be derived from the global minimum of $S_{cor}(t)$.

3.3. Improved C-C Method Based on Fused Norm Search Domain. In the traditional C-C method, accuracy and computing complexity are in conflict. In calculating the correlation dimension, the algorithm employs a non-repetitive traversal of neighboring sites, wherein the phase spacing is frequently measured using the $L2$ norm and the infinite norm. The $L2$ norm has the advantage of meeting the requirements of the Euclidean distance metric and precisely describing the distance between two phases, but its calculations are more complex and less efficient. The infinite norm offers the benefits of simple calculation and minimal spatial and temporal complexity, but its metric principle does not comply with the Euclidean law and contains certain inaccuracies, which degrade the quality of the reconstruction. No matter which of the two-distance metrics is chosen, it is impossible to achieve a compromise between metric accuracy and computational efficiency. The three norms are given as follows:

$$\begin{aligned}L1\,norm:\ d_{ij(1)}&=d_{(1)}\left(X_i,X_j\right)=\sum_{l=0}^{m-1}\left|x_{i+lt}-x_{j+lt}\right|,\\ L2\,norm:\ d_{ij(2)}&=d_{(2)}\left(X_i,X_j\right)=\sqrt{\sum_{l=0}^{m-1}\left(x_{i+lt}-x_{j+lt}\right)^2},\\ Infinite\,norm:\ d_{ij(\infty)}&=d_{(\infty)}\left(X_i,X_j\right)\\ &=\max\left\{\left|x_{i+lt}-x_{j+lt}\right|\right\},\end{aligned} \quad (5)$$

where X_i、X_j is the phase point, $0\le l\le m-1$. In order to provide a more intuitive description of the search domain of the three types of norms, the metric distances of the three norms are unified in the three-dimensional case, and the search domain function of each norm and its geometric mapping are obtained as shown in Figure 3. Where $\Delta x,\Delta y,\Delta z$ is the difference in relative coordinates between the two picture locations. From the distance definition,



$L1\ norm: r_{L1(X_i,X_j)} = |\Delta x + \Delta y + \Delta z|$ $L2\ norm: r_{L2(X_i,X_j)} = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$ $Infinite\ norm: r_{\inf(X_i,X_j)} = \max\{\Delta x, \Delta y, \Delta z\}$

(a) (b) (c)

Figure 3: Search domain functions and geometric mappings of each paradigm: (a) the l1 norm metric; (b) the l2 norm metric; (c) the infinite norm metric.

search domain, and geometric mapping of the aforementioned three types of norms, it can be determined that in distance calculation, the $L1$ norm and infinite norm only use addition and taking extremes, which are less complex than the exponential operations of the $L2$ norm; in terms of metric accuracy, only the search radius from the origin of the sphere domain to the sphere boundary of the $L2$ norm is required. Depending on the relative angle, the distance from the origin to the search border of the $L1$ norm and the infinite norm regular multicell domains varies. On this basis, the formula for calculating distance in the fused norm search domain is defined as follows:

$$\begin{aligned} d_{ij} &= \min\left\{\omega_a \cdot d_{ij(1)}, \omega_b \cdot d_{ij(\infty)}\right\} \\ &= \min\left\{\omega_a \cdot d_{(1)}\left(X_i, X_j\right), \omega_b \cdot d_{\infty}\left(X_i, X_j\right)\right\} \\ &= \min\left\{\omega_a \cdot \max\left\{\left|x_{i+lt} - x_{j+lt}\right|\right\}, \omega_b \cdot \sum_{l=0}^{m-1}\left|x_{i+lt} - x_{j+lt}\right|\right\}, \\ &\quad 0 \le l \le m-1, \end{aligned} \tag{6}$$

where ω_a、 ω_b are the weight parameters, and the values of the other two variables must satisfy two requirements.

(1) In the high-dimensional situation, the relative relationship between ω_a and ω_b should cause the geometric mapping of the fused norm search domain to have regular covariance in order to approach the high-dimensional spherical domain of the $L2$ norm. In this study, we define its geometric mapping as an Archimedean body and obtain the equation (12).

$$\begin{cases} l_1 = \dfrac{\omega_a - \omega_b}{\cos(\pi/4)}, \\ l_2 = \dfrac{\omega_a}{\cos(\pi/4)} - \dfrac{\omega_a - \omega_b}{\sin(\pi/4)}. \end{cases} \tag{7}$$

(2) The entire point-seeking volume of the fused norm search domain must be kept within an acceptable range. In this study, the search volume of the advanced search domain and the $L2$ norm are held constant in order to obtain the equation (13).

$$\begin{cases} V_{improve} = \dfrac{256}{675}\sqrt{5}\,\omega_a^3, \\ V_{l_2} = \displaystyle\int_0^{2\pi} d\theta \int_0^{\pi} \sin\phi d\phi \int_0^{r(\phi,\theta)} r^2 dr, \\ V_{improve} = V_{l_2}. \end{cases} \tag{8}$$

Obtaining the parameter weights ω_a, ω_b and constructing the fused norm search domain, whose search domain function and 3D geometric mapping are depicted in Figure 4.

It is straightforward to determine that the fusion norm metric employs simple polarization and addition and subtraction operations in the calculating distance, which has less spatial and temporal complexity and more computational efficiency than the $L2$ norm. The Archimedean body mapping domain of the fusion norm search domain is more similar to the sphere mapping domain of the $L2$ norm in terms of metric accuracy, and the metric error is less than that of the $L1$ norm and infinite norm. The algorithmic flow of the improved C-C method based on the fused norm search domain is summarized in algorithm 1.

To further analyze from a theoretical perspective the performance of point finding accuracy in the fused norm search domain, the theoretical relative polarity evaluation index is defined as follows:

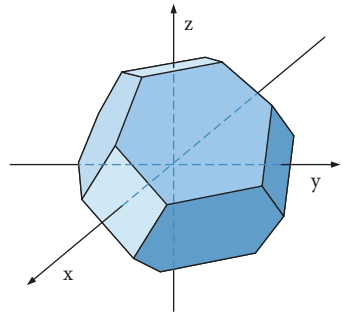
$$S_{(P)} = \frac{\max d_{(P)} - \min d_{(P)}}{r_{(P)}}, \tag{9}$$

where P represents the accepted norm metric, r is the unit length, and $\max d_{(P)}$, $\min d_{(P)}$ are, respectively, the greatest and lowest Euclidean distances for each norm search radius. The $L2$ norm and Euclidean distance metrics have been unified, so the theoretical relative range is 0.

The theoretical relative range of the enhanced search domain is lowered by 80.9% relative to the $L1$ norm and by 88.8% relative to the infinite norm, as shown in Table 1. The preceding analysis demonstrates that the proposed fused

Input: Traffic flow time series after denoising $X = (x_1, x_2, \ldots, x_N)$
Initialization: Maximum chunking volume of average chunking strategy $d_{\max} = 80$;
BDS statistical test conclusion $m = 2, 3, 4, 5, r_i = j \cdot \sigma/2, j = 1, 2, 3, 4$;
Variance of the time series $\sigma = \sum_{i=1}^{N} (x_i - \overline{x})^2/N$;
Number of phase points $M = N - (m-1)\tau$;
for $t = 1, 2, \ldots, d_{\max}$ **do**
The statistic $S(m, N, r, t)$ is sequentially cut into mutually disjoint t subseries: $\{S_1(m, N/1, r, 1), S_2(m, N/2, r, 2), \ldots, S_t(m, N/t, r, t)\}$;
for $m = 2, 3, 4, 5$**do**
for $j = 1, 2, 3, 4$**do**
Reconstructs the subsequence as a phase space: $X = \{X_i \mid X_i = (x_i, x_{i+t}, \ldots, x_{i+(m-1)t}), i = 1, 2, \ldots, N - (m-1)\tau\}$
for $i = 1, 2, \ldots, M$**do**
The distance of phase points in phase space is measured using the fused norm search domain:
$d_{i,\widehat{i}} = \min\{\omega_a \cdot \max\{|x_{i+l\tau} - x_{\widehat{i}+l\tau}|\}, \omega_b \cdot \sum_{l=0}^{m-1} |x_{i+l\tau} - x_{\widehat{i}+l\tau}|\}, 0 \le l \le m-1$;
Statistical domain phase points with θ function and solving the correlation integral $C_t(m, N/t, r, t), C_t^m(1, N/t, r, t)$;
Calculate the mean statistic and the extreme difference statistic $S(m, r, t), \Delta S(m, t)$;
end
end
end
end
Calculate the statistic $\overline{S}(t), \Delta\overline{S}(t), S_{cor}(t)$ and plot its variation with time delay t;
Determine the optimal time delay τ_d and embedding window τ_w based on the curve pattern;
Calculating the embedding dimension $m = (\tau_w/\tau_d) + 1$ ◦
Output: the optimal time delay τ_d and the embedding dimension m

Algorithm 1: Improved C-C method based on a fused norm search domain.



Fusion norm : $r_{fusion\ (Xi,\ Xj)} =$
$\min\{\omega_a \cdot \max\{\Delta x + \Delta y + \Delta z\}, \omega_b \cdot |\Delta x + \Delta y + \Delta z|\}$

Figure 4: Search domain function and geometric mapping of the fused norm.

norm search domain improves the metric accuracy compared to the $L1$ norm and infinite norm while ensuring the simplicity of the operation. The computational efficiency and metric accuracy of the fused norm search domain will be evaluated experimentally.

3.4. Discrimination of Chaotic Characteristic. On the basis of the phase space reconstruction of the measured traffic flow, the small data sets method is utilized to validate the chaotic character of the measured link traffic flow sequences and to offer support for future chaotic time series prediction modeling. First, each phase point in the phase space is traversed, and the transient separation is constrained.

$$d_j(0) = \min\left\|X_i - X_j\right\|, |i - j| > P, \tag{10}$$

where $d_i(0)$ is the minimum distance between the reference phase point X_i and the evolving phase point X_j , and P is the sequence's average time period. Sato et al. propose that the maximum Lyapunov exponent λ can be determined from the average divergence rate of the nearest neighbors of each point in the phase point orbit.

$$\ln d_i(k) = \ln C_i + \lambda(k\Delta t), i = 1, 2, \ldots, M, \tag{11}$$

where M represents the phase point count. Using the least squares method, the regression line $y(k)$ of the preceding equation was fitted.

$$y(k) = \frac{1}{q \cdot \Delta t} \sum_{i=1}^{q} \ln d_i(k), \tag{12}$$

where q is the number of nonzero $d_i(k)$ and the slope of the regression line is the maximum Lyapunov exponent λ, When $\lambda > 0$, the chaotic dynamics of the measured link traffic flow data can be proved.

4. Hybrid Chaotic Traffic Flow Prediction Model

In this section, the hybrid prediction model is built from the perspective of chaotic traffic flow characteristics. Considering that the spatial distribution of phase points in the phase space tensor of traffic flow follows the law of attractor evolution and that the phase space tensor is reconstructed from the traffic flow temporal vectors, which have both temporal laws [51, 52], the chaotic feature extraction module and the temporal feature extraction module are constructed to evolve the chaotic and temporal characteristics of traffic

Table 1: Theoretical relative ranges of each norm metric.

	$L1$ norm metric (%)	Infinite norm metric (%)	Fused norm metric (%)
Theoretical relative range S (%)	42.3	73.2	8.1

flow, respectively. The complex nonlinear rules of the chaotic phase space and the sensitivity to beginning values make it easy for the model to reach a local optimum during training, resulting in an issue with prediction performance decline. In this study, the model optimization module was developed to increase the prediction accuracy and learning capability of the model by lowering the parameter discovery error.

Figure 5 illustrates the construction of the hybrid PDRGA-CDBN-LSTM traffic flow prediction model developed in this paper, whose primary structures are as follows:

(1) Data preprocessing: initially, wavelet denoising is performed on chaotic traffic flow to improve the value density of the traffic flow data, followed by the phase space reconstruction to fully expand the attractor structure, and finally, the chaotic characteristics of traffic flow are distinguished and normalized to obtain the phase space tensor T and the traffic flow time series S.

(2) Chaotic feature extraction module: the spatial characteristics of the phase space tensor T are extracted using the layer-by-layer pretraining process of CDBN and fed into the connection layer for fusion with the traffic flow time series S.

(3) Temporal feature extraction module: The tensor generated from the fusion of the connection layer is fed into the LSTM to extract temporal features based on the spatial features of the phase space [53] in order to acquire more comprehensive spatio-temporal aspects of the chaotic traffic flow.

(4) Intelligent algorithm optimization module: the initial weights of CDBN are optimized using a lightweight modified PDRGA in order to prevent the CDBN from falling into a local optimum when processing a complex high-dimensional phase space tensor.

4.1. Chaotic Feature Extraction Module. When the model extracts the chaotic phase space features, the phase point dimension is typically high, resulting in extremely complex evolutionary features. A neural network with a few layers has a limited ability to fit its complex nonlinear dynamics features. Therefore, a deep network structure is necessary to extract the high-level distribution patterns. In gradient descent, the deep convolutional model learns features through gradient feedback, and its gradient values containing feature information are weakened during back propagation with long-distance transmission between network layers, resulting in insufficient extraction of chaotic features of the traffic flow by the model, which affects the accuracy of the prediction.

The system energy function of CDBN is given as follows:

CDBN is a restricted Boltzmann machine (CRBM) stacked multiconvolutional deep learning model. CDBN only updates one layer of parameters in each training until the next layer is trained, and the parameter updating process is not affected by the depth of the network, so it can learn the chaotic nonlinear law within the measured traffic flow more thoroughly and precisely and extract the high-dimensional phase point space features of the chaotic phase space tensor more efficiently. Figure 6 illustrates the layer structure of CDBN, in which the CRBM layers are indirectly coupled and each layer is subject to convolution and probabilistic maximum pooling operations.

$$E(v,h) = -\sum_{k=1}^{K}\sum_{i,j=1}^{N_H}\sum_{r,s=1}^{N_W} h_{ij}^{k} w_{rs}^{k} v_{i+r-1,j+s-1} - \sum_{k=1}^{K} b_k - c\sum_{i,j=1}^{N_v} v_{i,j} = -\sum_{1}^{K} h^{k}\cdot\left(\widetilde{W}^{k} * v\right) - \sum_{k=1}^{K} b_k \sum_{i,j}^{N_H} h_{ij}^{k} - c\sum_{i,j}^{N_v} v_{ij}, \tag{13}$$

where v and h are the values of the explicit and implicit layer cells, w is the value of the convolution kernel weights, b and c are the bias values of the hidden layer cells, h and the visible layer cells v, and $\widetilde{W}^{k}$ is the summation operation of the K sets of convolution kernels over the elements in the data matrix.

On the basis of the energy function, the joint probability distribution of the explicit hidden layer can be determined as follows:

$$P(v,h) = \frac{\exp\left(-E(v,h)\right)}{Z}, \tag{14}$$

where Z is the function of collocation. The conditional probability distribution function for the activation of explicit hidden layer cells is expressed as follows:

$$\begin{aligned} P\left(h_{ij}^{k} = 1 \mid v\right) &= \frac{1}{1+\exp\left(-\left(\left(\widetilde{W}^{k} * v\right)_{ij} + b_k\right)\right)}, \\ P\left(v_{ij} = 1 \mid h\right) &= \frac{1}{1+\exp\left(-\left(\left(\sum_k W^{k} * h^{k}\right)_{ij} + c\right)\right)}. \end{aligned} \tag{15}$$

The posterior probabilities of node weights and back propagation of the next layer can be derived based on the

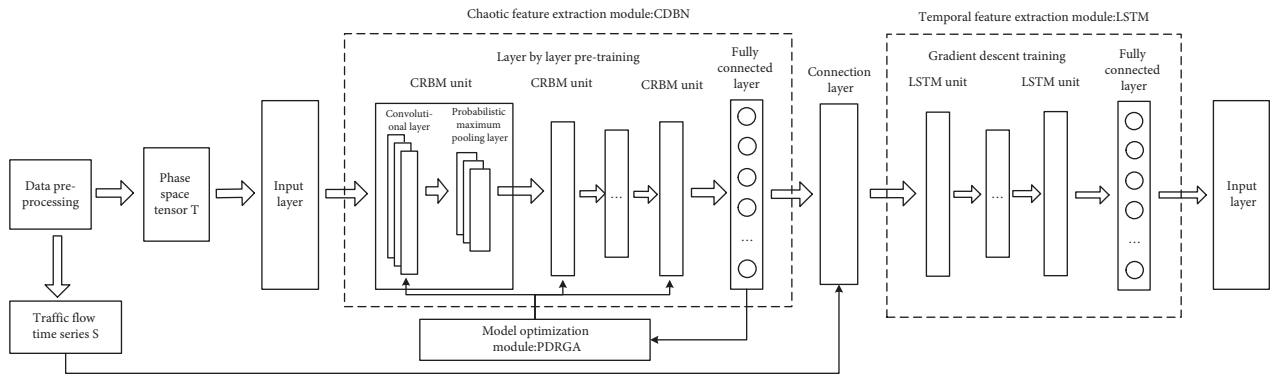


Figure 5: Structure diagram of the PDRGA-CDBN-LSTM hybrid model.

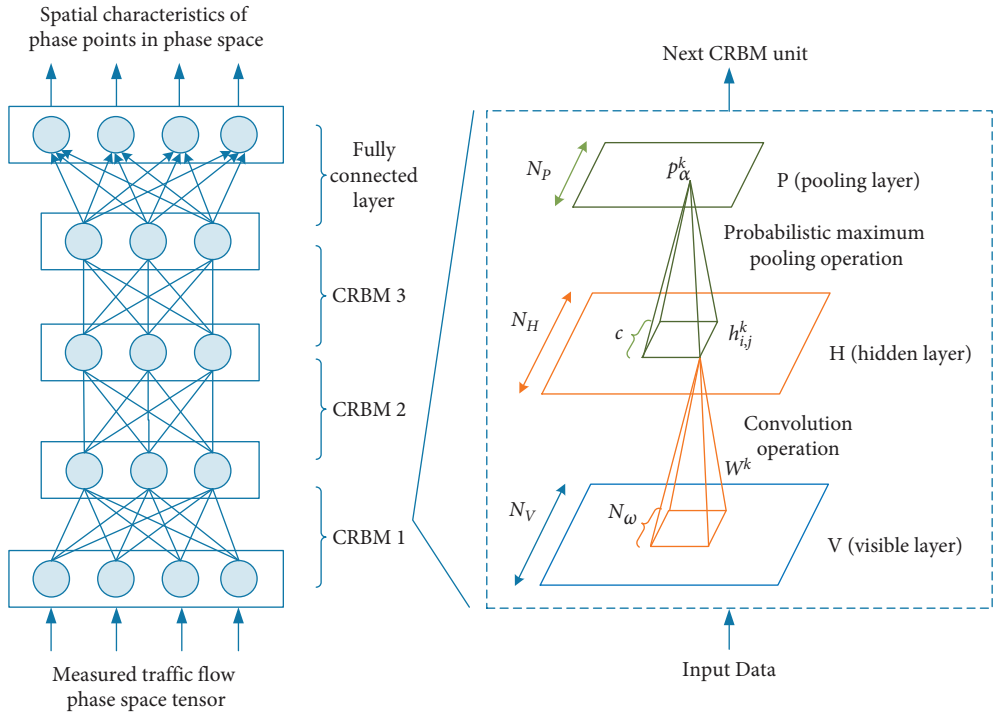


Figure 6: Hierarchy diagram of the CDBN network.

bidirectional transportability and conditional probability distribution of CDBN, and the parameters of the explicit and implicit layers can be updated by continuous round-trip sampling between the explicit and implicit layers using Gibbs block sampling. The probabilistic maximum pooling operation maximizes the shared distribution of high-dimensional phase space data and the Gibbs sampling outcomes of lower-level cells. The conditional probability of pooling units' activation is expressed as follows:

$$P\left(h_{ij}^k = 1|v\right) = \frac{\exp\left(I\left(h_{i,j}^k\right)\right)}{1 + \sum_{(i,j)\in B_\alpha}\exp\left(I\left(h_{i,j}^k\right)\right)}, s.t. \sum (i,j) \in B_\alpha h_{i,j}^k \le 1, \forall k, a,$$
$$P\left(p_\alpha^k = 0|v\right) = \frac{1}{1 + \sum_{(i,j)\in B_\alpha}\exp\left(I\left(h_{i,j}^k\right)\right)}, s.t. \sum (i,j) \in B_\alpha h_{i,j}^k \le 1, \forall k, a, \tag{16}$$

where $I(h_{i,j}^k)$ is the activation probability of the hidden layer unit's information parameter and B_α is the pooling layer mapping block. In order to complete the update of CDBN parameters, the network is then trained using the k-steps contrastive divergence (CD-k) algorithm.

$$\Delta w_{i,j} = \sum_{i=1}^{N_V}\sum_{j=1}^{N_H}\left[P\left(h_i = 1|v^0\right)v_j^0 - P\left(h_i = 1|v^k\right)v_j^k\right],$$
$$\Delta b_j = \sum_{j=1}^{N_H}\left[v_j^0 - v_j^k\right], \tag{17}$$
$$\Delta c_i = \sum_{i=1}^{N_V}\left[P\left(h_i = 1|v^0\right) - P\left(h_i = 1|v^k\right)\right],$$

where $\Delta w_{i,j}, \Delta b_j, \Delta c_i$ represents the amount of parameter adjustment during each iteration for network weight value $w_{i,j}$ and bias values b_j, c_i for the visible and hidden layers.

4.2. Temporal Feature Extraction Module. In the chaotic phase space tensor, not only does the spatial distribution of high-dimensional phase points adhere to the law of attractor change, but so does the temporal evolution of phase points. Therefore, further mining the time-dependent relationship based on the spatial characteristics of phase points can provide a more comprehensive understanding of the macro and microvariations of traffic flow. Commonly used RNN networks for temporal feature extraction are trained with gradient values that are susceptible to gradient disappearance or explosion due to multiple cumulative multiplication operations during the propagation process, making it difficult for the model to learn the long-time dependence of the data [54]. As a special RNN, the LSTM cell state may be propagated between layers using only linear summation [55], and gradient values do not degrade during the propagation process, which enhances the model's capacity to extract long-term features and improves its accuracy performance compared to RNN. The LSTM consists of three fundamental elements: the forgetting gate, the input gate, and the output gate. Figure 7 depicts the cell structure of the LSTM, and the present cell state of the LSTM is also depicted.

As indicated in the picture, σ and tanh represent the sigmoid activation function and the hyperbolic tangent activation function, respectively. The forget gate f_t selects the memory of the previous moment's information based on the input values x_t and h_{t-1}, while the input gate i_t controls the information update of the current cell state S_t based on the cell state update value $\widetilde{S}_t$, and the output gate o_t controls the information content of the next moment's cell output based on the current cell state S_t. The following is how the LSTM is computed:

$$\begin{aligned} f_t &= \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right), \\ i_t &= \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right), \\ o_t &= \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right), \\ \widetilde{C}_t &= \tan h\left(W_C\left[h_{t-1}, x_t\right] + b_C\right), \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t, \\ h_t &= o_t \tan h\left(C_t\right), \end{aligned} \tag{18}$$

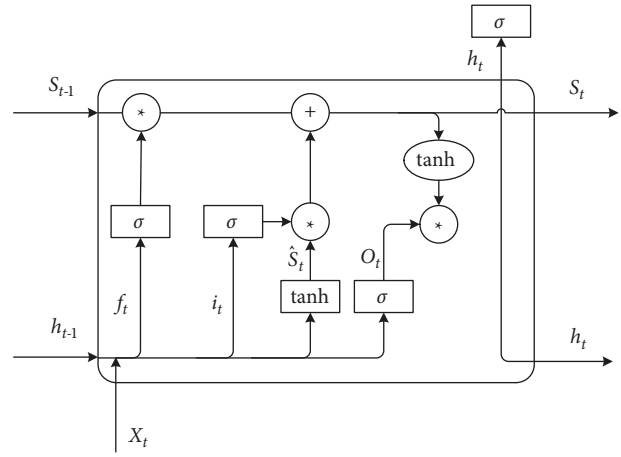


Figure 7: Structure diagram of the LSTM unit.

where W and b are the weights and bias values, respectively, of each gating unit.

4.3. Model Optimization Module. GA [56] is a model for optimizing population viability by computer simulations of natural selection and genetic reproduction mechanisms of biological evolution, and its three core steps of genetic manipulation [57] achieve iterative population renewal by eliminating backward individuals and recombining chromosomal coding and gene mutation. Sometimes, the highly adaptive chromosomal gene fragments of superior individuals are broken down at the breakpoint, and their offspring fail to inherit high-quality genes, resulting in low individual evaluation. Meanwhile, the fixed population variation rate lacks the differentiated consideration of the richness requirements for seeking superiority at different iterations. All of these factors reduce the convergence efficiency of GA, resulting in additional iterations and higher training costs in the pursuit of precision.

To reduce the training cost of the algorithm, this paper proposes an improved PDRGA based on probabilistic dynamic reproduction, which allows the selection of breakpoints in the crossover operation based on the probability of the fitness of one parent in the total fitness of the two parents, so as to retain the high fitness gene fragments of the best individuals and accelerate convergence speed, while the variation rate is dynamically adjusted with the number of individuals. The variation rate is dynamically altered based on the number of iterations to accommodate the varying requirements of global and local search abilities at various iteration stages. Figure 8 depicts the detailed algorithmic procedure.

Let each iteration's population size be n, after the algorithm completes the selection operation in the iteration process, the remaining individuals are sorted by fitness value, and the fitness threshold condition f_t is set to further divide the population into a high fitness population A and a low fitness population B. The threshold condition $f_t f_t$ is defined as follows:

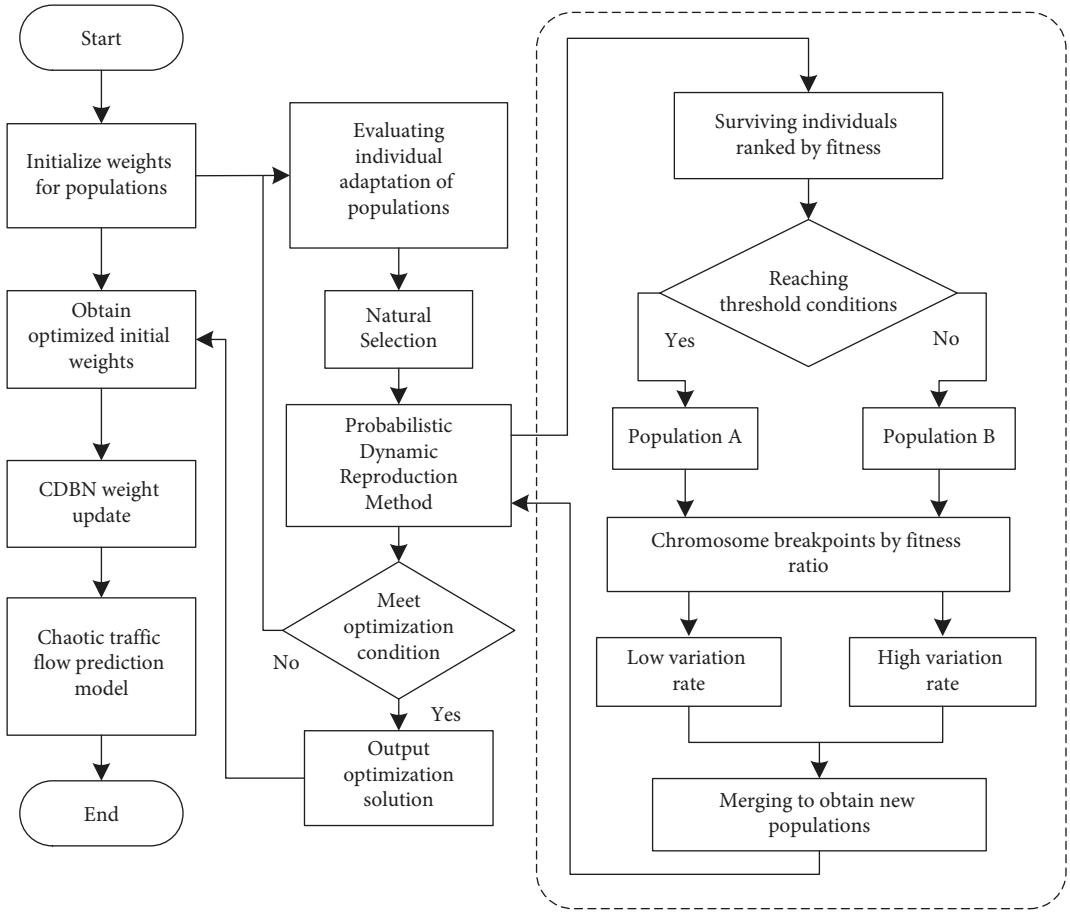


Figure 8: Flowchart of the PDRGA optimization algorithm.

$$f_t = \frac{\sum_{i=1}^{n} f_i}{n}, \tag{19}$$

where f_i represents the individual's fitness value for the population crossover operation, the probability p of one parent's fitness relative to the overall fitness of both parents is calculated and used to determine the position of the gene breakpoint c_{break}.

$$p = \frac{f_i}{f_i + f_i}, i \in [1, n], j \in (i, n],$$
$$c_{\text{break}} = p \cdot n_{\text{gene}}, \tag{20}$$

where f_i and f_j are, respectively, the fitness values of the two parents, and n_{gene} is the number of genes on one chromosome. This breakpoint strategy produces progeny groups A and B containing, respectively, more and less adaptable gene segments. The variation rates v_{Ai} and v_{Bi} for offspring are set as follows:

$$v_i = \left(\frac{gi_{\max}}{g_{\max}}\right),$$
$$v_{Ai} = 0.7v_i, \tag{21}$$
$$v_{Bi} = 1.3v_i,$$

where v_i represents the baseline variation rate, $g_{\max}$ represents the maximum number of iterations, and g_i represents the current number of iterations. v_i controls the total level of variation among the children, which can be reduced dynamically as the number of repetitions increases. A larger overall variation rate at the start of the iteration can accelerate convergence, whereas a lower overall variation rate at the conclusion of the iteration can increase convergence accuracy and prevent the algorithm from degrading into a random search process. The variation rates of populations A and B are specified differently based on v_i. The former variation rate can retain more high fitness gene fragments to guide the general direction of the optimization search, whereas the latter variation rate can increase population randomness to avoid falling into the local optimum and continue the new round of iterations until the optimization is complete after obtaining the new population.

The enhanced PDRGA can select the breakpoint location probabilistically and dynamically adjust the variance rate based on the number of iterations and individual evaluation, which improves the convergence efficiency and the optimization accuracy, achieves higher optimization accuracy with fewer iterations, and realizes the lightweight enhancement of GA. In this paper, the initial weight values of convolutional kernels in CDBN are optimized by using PDRGA.

Table 2: Attributes of traffic flow data in the Hefei demonstration zone ITS dataset and PeMSD8 dataset.

	Hefei demonstration zone ITS dataset	PeMSD8 dataset
Data collection locations	High-tech industrial demonstration zone in the Hefei, Anhui, China.	Highway network in San Bernardino, California, United States.
Acquisition device	City road surveillance camera	Caltrans performance measurement system (PeMS) device
Number of selected detectors	Htw001	S001
Time range of the collected data	From June 30, 2016, to July 29, 2016	From July 1, 2016, to August 29,2016
Data acquisition time interval (min)	5	5
Length of data	8640	17280

5. Experiment Analysis and Discussion

The code development platform for the experiment is Python 3.6; PyCharm is selected as the integrated development environment; the CPU model of the computing device is Intel Core i7-9750 2.60 GHz; and the running memory is 16 GB. The link traffic flow data are obtained from two measured datasets: the ITS dataset of the Hefei demonstration zone and the PeMSD8 dataset, and the measurement tools are assumed to be accurate, so the error in measurement is not considered.

In the Hefei city dataset, the traffic flow data from the video traffic flow detection subdataset located at the intersection of Huangshan Road and Tianzhi Road are selected, and the subdataset includes 2 road sections and 6 detectors; in the PeMSD8 dataset, the traffic flow measurements collected by Caltrans in the San Bernardino County subdataset are selected, and the subdataset includes 8 road sections and 170 detectors. Both types of data selected are discrete time slices of continuous vehicle flow collected by detectors throughout the day, and the specific attributes of the data are shown in Table 2.

5.1. Phase Space Construction Reconstruction Experiment. To verify the accuracy and efficiency of the fused norm search domain improvement C-C method, phase space reconstruction experiments are conducted using the $L2$ norm metric, the infinite norm metric, and the fused norm search domain metric. The optimal time delay τ_d is determined as the first local minimum of $\Delta\bar{S}(t)$, whereas the embedding window τ_w is determined as the global minimum of $S_{\text{cor}}(t)$. Figures 9 and 10 depict the experimental results of the phase space reconstruction for the Hefei demonstration zone ITS dataset and PeMSD8 dataset, respectively.

Due to the different principles of each norm metric, the range of the curve values and the overall trend under each mode vary significantly. However, these differences have no bearing on the choice of reconstruction parameters, which must be determined based on the relative change law under this type of mode. Given that the $L2$ norm is theoretically free of Euclidean distance metric error and that the reconstruction parameters are reliably obtained, it can be utilized as the accuracy measure of other norms.

Figure 9(a) depicts the statistical curve with the $L2$ norm metric for the Hefei city dataset. The embedding dimension m is calculated to be 7 and the optimal time delay τ_d is 8. These parameters are the standard of reference for subsequent norms. In Figure 9(b), the embedding dimension m of the infinite norm metric is 6, differing by 1 from the reference standard, and the optimal delay τ_d is 7, differing by 3 from the reference standard. In Figure 9(c), the evolution trend of the improved metric is closer to the $L2$ norm, which reflects the geometric similarity of the topological relationship between the two search domains; consequently, the embedding dimension m of 7 and the optimal time delay τ_d of 8 are identical to the reference standard without deviation. Similarly, in the PeMSD8 dataset, Figure 10(a) obtains the reference standard by taking the $L2$ norm metric with an embedding dimension m of 5 and an optimal delay τ_d of 5. The computation of Figure 10(b) using an infinite norm metric provides an embedding dimension m of 5, but the optimal delay τ_d is 7, which deviates by 2 from the reference standard. Figure10(c) demonstrates that the computation of the two reconstruction parameters of the improved metric is identical to the reference standard.

As shown in Table 3, the method using the $L2$ norm metric has the longest running time, but the number of search domain seek points is theoretically error-free, therefore the program's output serves as the accuracy reference standard. Thus, it is clear that the $L2$ norm compromises computing efficiency for precision. The algorithm with an infinite norm metric has the shortest running time; however, there is a discrepancy between the number of search domain seeks and the reference standard, 17.3% for the Hefei demonstration zone ITS dataset and 17.8% for the PeMSD8 dataset. This indicates that there is a metric bias in the infinite norm, which affects the quality of phase space reconstruction, and the computational efficiency is enhanced at the expense of metric accuracy. The running time of the proposed improved algorithm and the infinite norm remain at an approximate level, compared with the $L2$ norm. The operation time of the Hefei demonstration zone ITS dataset is reduced by 937.82 seconds, the efficiency is increased by 20.3%, the search domain point finding error is only 3.1% compared with the reference standard, and the point finding error is reduced by 14.2% compared with the

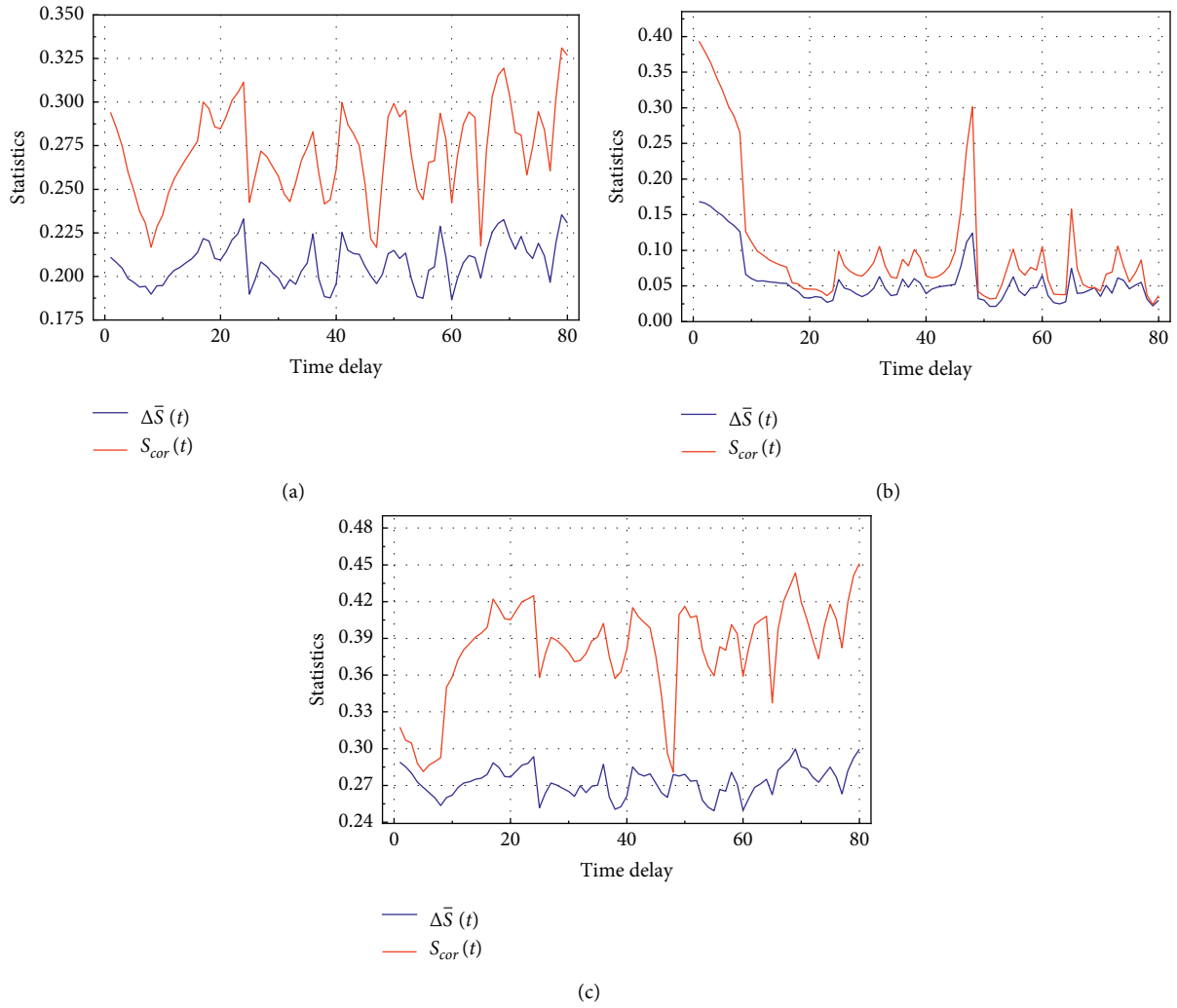


Figure 9: Experimental results of the C-C algorithm under each measurement method: (a) the $L2$ norm measurement; (b) the infinite norm measurement method; and (c) the fused norm measurement.

infinite norm search domain. The time of the PeMSD8 data set is reduced by 855.99 seconds, the efficiency is increased by 17.0%, the point finding error in the search domain is only 3.3% compared with the reference standard, and the point finding error is reduced by 14.5% compared with the infinite norm search domain. The comprehensive analysis demonstrates that the improved algorithm combines the advantages of high efficiency of infinite norm calculation and high accuracy of the $L2$ norm metric and reduces the time complexity of the algorithm while ensuring the improvement of accuracy, achieving a balance between accuracy and efficiency.

The chaotic discriminant algorithm is required to verify the chaotic characteristic of the original link traffic flow data after the phase space reconstruction, as well as to confirm the reasonableness of the phase space reconstruction parameters calculation, which cannot generate the incorrect parameters to destroy the chaotic nature of the system. Using the small data set method, the chaotic discriminant parameter Lyapunov exponent is determined, and the results are presented in Figure 11.

The red dashed line in the graph is the fitted regression line, and its slope value represents the greatest Lyapunov exponent. The exponents of the Hefei demonstration zone ITS dataset and the PeMSD8 dataset are 0.0032 and 0.0095, respectively, both of which are higher than 0. This study demonstrates that the measured link traffic flow data of the two datasets contain chaotic dynamics and demonstrates that the phase space reconstruction does not damage the chaotic characteristics of the data.

5.2. Convergence Comparison Experiment of Optimization Algorithm. In this study, the improved PDRGA algorithm is utilized to optimize the initial weight parameters of CDBN to prevent the model from sliding into a local optimum when dealing with complex, high-dimensional phase space tensors. The algorithm optimization experiment is designed to intuitively evaluate the advantages of PDRGA in terms of convergence efficiency and precision. The experiment brings the population vector X closer to the target vector Z, yielding the algorithm's optimal solution, which is the vector

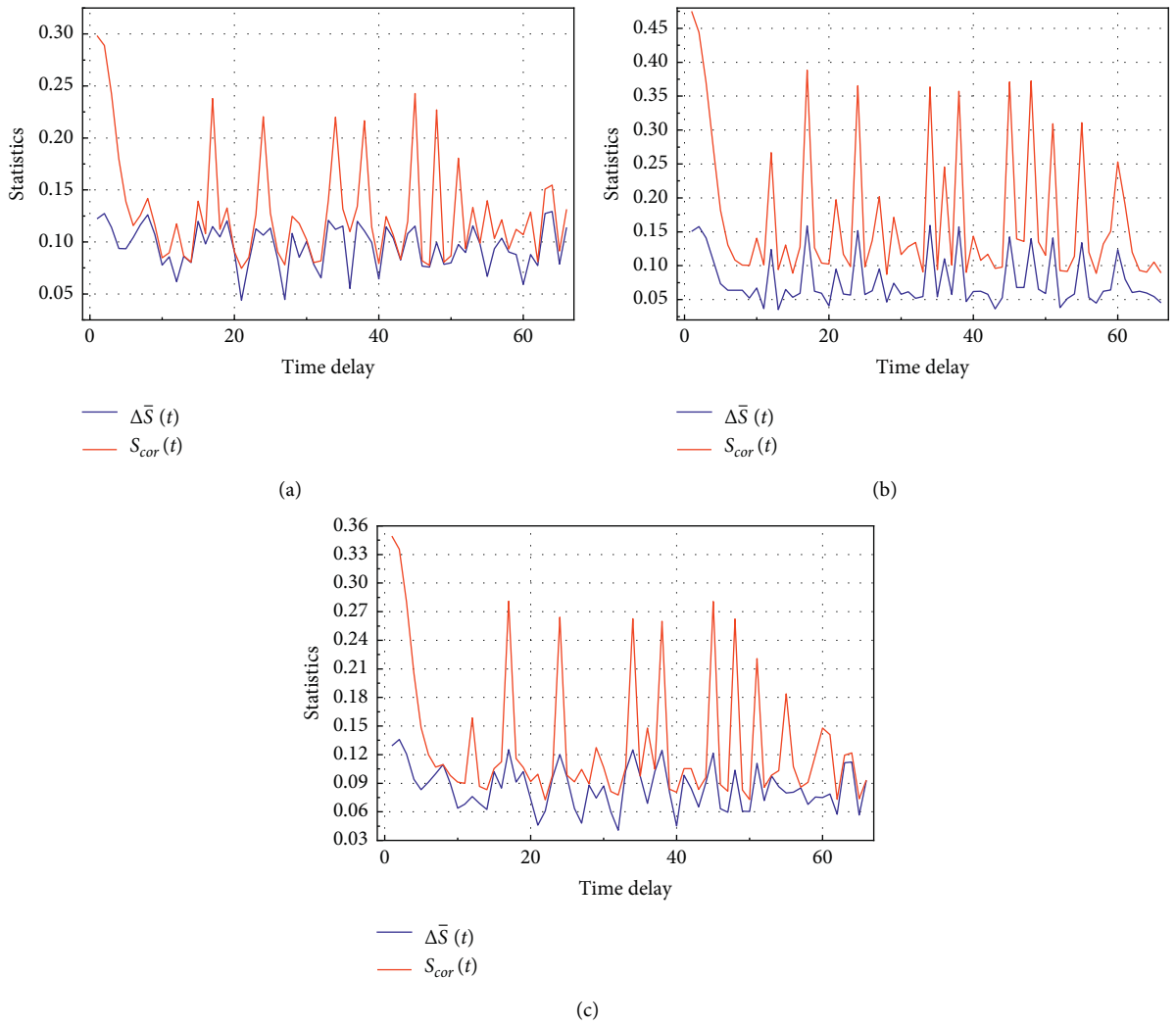


Figure 10: Experimental results of the C-C algorithm using the PeMSD8 dataset: (a) the $L2$ norm measurement; (b) the infinite norm measurement method; and (c) the combined norm measurement.

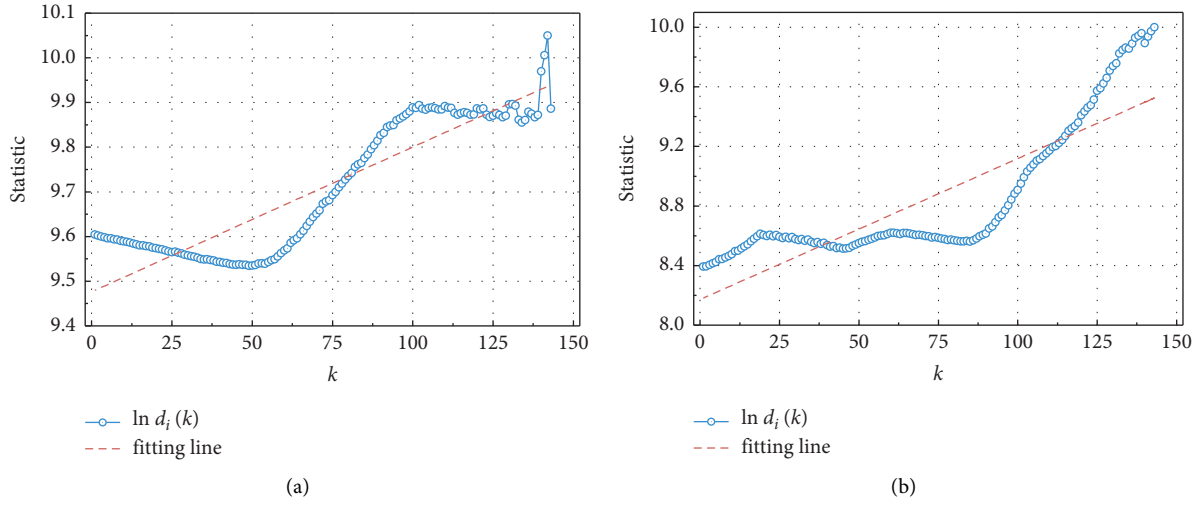


Figure 11: Evolution diagram of $\ln d_i(k)$ with k: (a) the Hefei demonstration zone ITS dataset; (b) the PeMSD8 dataset.

Table 3: Running time and search points of the C-C method under each measurement method.

	Hefei demonstration zone ITS		PeMSD8	
	Algorithm runtime (s)	Searched points (n)	Algorithm runtime (s)	Searched points (n)
L2 norm metric	5118.2432	193167121	5026.7428	159985266
Infinite norm metric	4180.4203	226507968	4208.8855	188496293
Fused norm metric	4125.2575	199227006	4170.7513	165313922

X that is closest to the target vector Z. Where the fitness value is the error evaluation of the population of individuals with respect to the element of the target vector Z, and the optimization aim is to identify the optimal solution vector for the element of the target vector Z , the number of populations is set to 10, the value of each element of the goal vector Z is set to 10, and the termination condition of the algorithm is set to 1000 iterations. The trajectory of the algorithm's convergence error over time is depicted in Figure 12, and Table 4 lists the running times of GA and PDRGA when the GA reaches its final convergence accuracy.

As demonstrated in the graph, compared to the traditional GA, PDRGA exhibits a faster convergence tendency at the beginning of the iteration, and the error in the 100th generation is 0.627, which is significantly less than the GA error of 4,955. The convergence error of PDRGA is always less than that of the traditional GA as the number of iterations increases. When the number of iterations is 235, the convergence trend of PDRGA tends to level off and the error decreases to 0.297, whereas the GA is still in the convergence state and the error is 2.335. PDRGA reaches the optimal convergence error of 0.233 after 403 iterations, while GA reaches this error level after 764 iterations and PDRGA reaches this error level after 91 iterations. Table 3 reveals that the GA requires 3.047 seconds to reach the final convergence error, whereas the PDRGA requires only 0.310 seconds to reach the same error level after 91 iterations, indicating that the lightweight improved PDRGA has a significant reduction in computation time compared to the traditional GA. Experiments demonstrate that the PDRGA increases the convergence speed and optimization accuracy of the method, produces more search accuracy with reduced training costs, effectively realizes the lightweight improvement of the GA, and prevents the CDBN-LSTM model from reaching a local optimum.

5.3. Experiments in Hybrid Model Prediction for Chaotic Traffic Flow. The suggested PDRGA-CDBN-LSTM and other baseline models are utilized in traffic flow prediction tests to evaluate the learning ability of each model with respect to the chaotic law of the observed traffic flow. The normalized findings from the data preprocessing phase are utilized as the research object. To standardize the data scales, 4320 data from the first 15 days were selected from each of the two datasets for the experiment, and the test set and training set are divided into the proportion of 80% to 20%. As model evaluation metrics, the mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE) were chosen, and the following formulas were calculated as follows:

$$\begin{aligned} MSE &= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2, \\ RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2}, \\ MAE &= \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y}_i\right|, \\ MAPE &= \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y}_i - y_i}{y_i}\right|, \\ SMAPE &= \frac{100\%}{n}\sum_{i=1}^{n}\frac{\left|\widehat{y}_i - y_i\right|}{\left(\left|\widehat{y}_i\right| + \left|y_i\right|\right)/2}, \end{aligned} \tag{22}$$

where n represents the predicted length of the data, y_i and $\widehat{y}_i$ represents the predicted and actual values, respectively.

The baseline models are the hybrid CDBN-LSTM model without GA optimization and the single CDBN and LSTM. In addition, the PDRGA-CDBN-LSTM is trained with noisy data in order to examine the impact of noise on model training. Figures 13 and 14 depict the model predictions using the Hefei demonstration zone ITS dataset and the PeMSD8 dataset, respectively.

From the experimental results of the two aforementioned data sets, it can be determined that the calculation results of reconstruction parameters under the infinite norm metric deviate from those under the error-free $L2$ norm metric, indicating that the inaccurate metric will harm the calculation accuracy of reconstruction parameters and thereby diminish the quality of phase space reconstruction. The proposed fused norm search domain metric not only has an evolution trend of the statistical curve that is closer to the reference standard but also the calculation results are error-free in comparison to the reference standard, proving the reliability of the improved C-C algorithm in terms of metric accuracy. The algorithm's running time and number of metric search points are recorded in Table 3 to further validate the accuracy of the improved algorithm and its computational efficiency advantages.

Each model's prediction curves exhibit a peak-and-valley-like cyclical fluctuation pattern with the actual values. From Figures 13(a) to 13(e), utilizing the Hefei demonstration zone ITS dataset, it is evident that CDBN has the

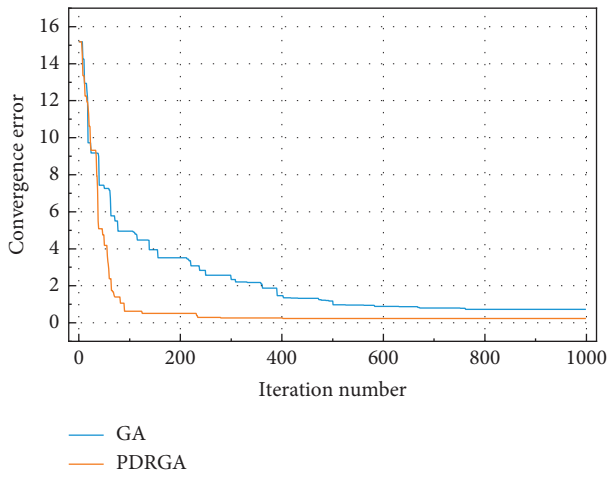


Figure 12: Comparison diagram of the convergence error of the optimization algorithm.

worst fitting performance, with significant fitting bias in peak intervals such as (70, 100) and (180, 210). While LSTM has the second-worst prediction performance, it outperforms CDBN in fitting both peaks and valleys. The fit of the CDBN-LSTM model is significantly better than that of the single model, and the fit of the prediction curves to the true values is better in the peak interval of (70, 210) and the valley interval of (260, 370). The proposed PDRGA-CDBN-LSTM model has the greatest prediction performance among the hybrid models. Compared to CDBN-LSTM, the peaks at time slices 82 and 204 and the valleys at 321 correspond more closely to the true value curve in terms of numerical performance. Nonetheless, the overall performance of the two curves is comparable and requires additional quantitative study. The proposed model trained with noise-containing data exhibits a diminished approximation effect in the high-volatility interval (80, 180). Similarly, in Figures 14(a)–14(e) utilizing the PeMSD8 dataset, CDBN has the worst fit, with fit deviations in the (50, 220) peak interval and the (270, 330) valley interval, whereas LSTM has a relatively better fit performance, but the fit curve in the (50, 190) peak interval is lower than the true value curve. Compared to a single LSTM model, CDBN-LSTM has better fitting precision in the peak interval. The proposed PDRGA-CDBN-LSTM has higher prediction accuracy than the CDBN-LSTM in both the (50, 190) peak interval and the (270, 330) valley interval, and it has the best-fit performance among all models. The PDRGA-CDBN-LSTM trained with noisy data has a more volatile fit curve in the (50, 220) crest interval and slightly lower fit accuracy than the model trained with denoised data. To quantitatively evaluate the performance of each model's fit, additional analysis based on the numerical performance of the error evaluation index is necessary.

Tables 5 and 6 present the error evaluation metrics MSE, RMSE, MAE, MAPE, and SMAPE for the two data sets utilizing various prediction models. The lower the error between the projected and actual values, and the greater the accuracy of the forecast, the smaller the values of the

Table 4: Optimization algorithm running time.

	PDRGA	GA
Algorithm runtime (s)	0.310	3.047

aforementioned metrics. It is evident that the PDRGA-CDBN-LSTM outperforms the other baseline models for all five indexes, and the CDBN-LSTM has the second-best error performance overall. Compared with CDBN-LSTM, PDRGA-CDBN-LSTM reduces the best performing MAPE index by 23.7% and the worst performing RMSE index by 6.7% in the Hefei demonstration zone ITS dataset; in the PeMSD8 dataset, the best performing MAPE index is reduced by 26.0% and the worst performing RMSE index is reduced by 13.1%, indicating that PDRGA enhances the learning ability of the hybrid model for complex high-dimensional chaotic tensors to avoid training into local optimum. In all indexes, the CDBN-LSTM model outperforms the individual CDBN and LSTM models. The best performing MSE index of the CDBN-LSTM is reduced by 45.8% to the LSTM with lower error indexes in the single model, and the worst performing MAPE index is decreased by 18.3% in the Hefei demonstration zone ITS dataset. Similarly, the best performing MSE index was reduced by 19.6%, while the worst performing MAPE index was reduced by 4.9% in the PeMSD8 dataset, demonstrating that the hybrid model can learn more complete data laws from the multiangle feature extraction concept of chaotic dynamics characteristics and time-series evolution laws, thereby achieving the goal of improving the prediction accuracy. In addition, the PDRGA-CDBN-LSTM with denoised data performs better than the noisy data in all indexes, and the largest reduction in the two data sets is 23.4% and 16.6% in SMAPE, indicating that the wavelet denoising reduces the impact of noise on the data value density and improves the model's ability to abstract data features.

Based on the analysis of the fitting curves and error evaluation indexes of each model in both datasets, it can be concluded that the fitting performance of LSTM in the single model is better compared to that of CDBN, indicating that the spatial arrangement of the phase space of chaotic data alone is insufficient for mining the internal patterns of the data. The prediction performance of the hybrid CDBN-LSTM model is superior to that of the individual model, demonstrating that the multiangle feature extraction strategy of the combined model architecture can discover more comprehensive data patterns. The best fitting performance of the proposed model compared to all baseline models demonstrates that the PDRGA prevents the model from falling into local optimum, hence enhancing the model's ability to learn the high-dimensional phase space tensor. The reduced fitting effect of the model trained with noisy data demonstrates that the noise in the measured data can have a substantial effect on the prediction accuracy of the model and that wavelet denoising can effectively increase the data value density.

Figure 15 depicts the PDRGA-CDBN-LSTM and the CDBN-LSTM training numbers versus error convergence using two data sets. In experiments using the Hefei

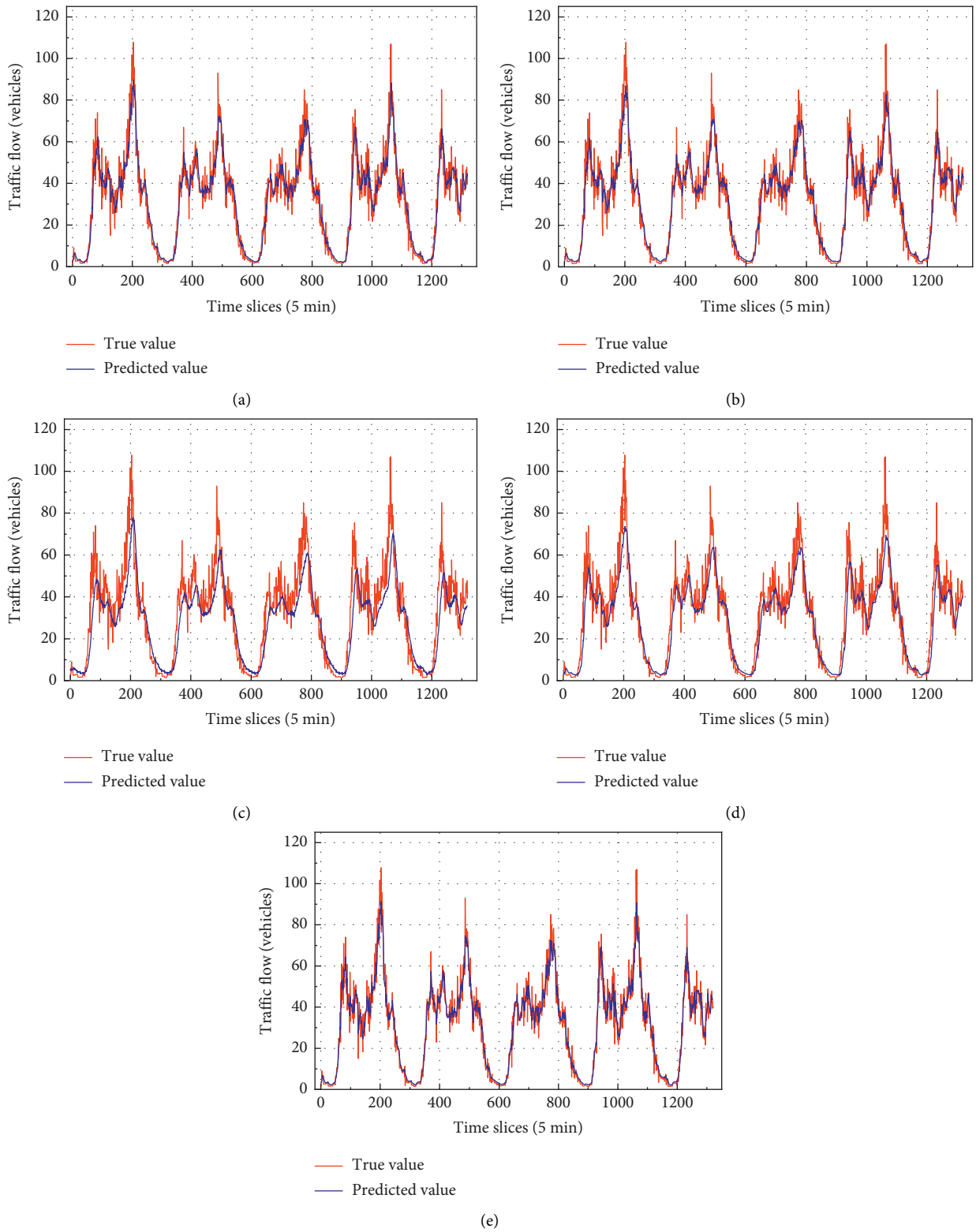


Figure 13: Prediction results of each model using the Hefei demonstration zone ITS dataset: (a) PDRGA-CDBN-LSTM; (b) CDBN-LSTM; (c) CDBN; (d) LSTM; and (e) PDRGA-CDBN-LSTM (noisy data).

demonstration zone ITS dataset, the proposed model converged to a small error range on the 20th iteration, and its curve trend was in a high convergence state at the start of the first 10 iterations of the optimization search. After 25 iterations, the error drops gradually, the convergence state stabilizes, and the training error is always lower than that of

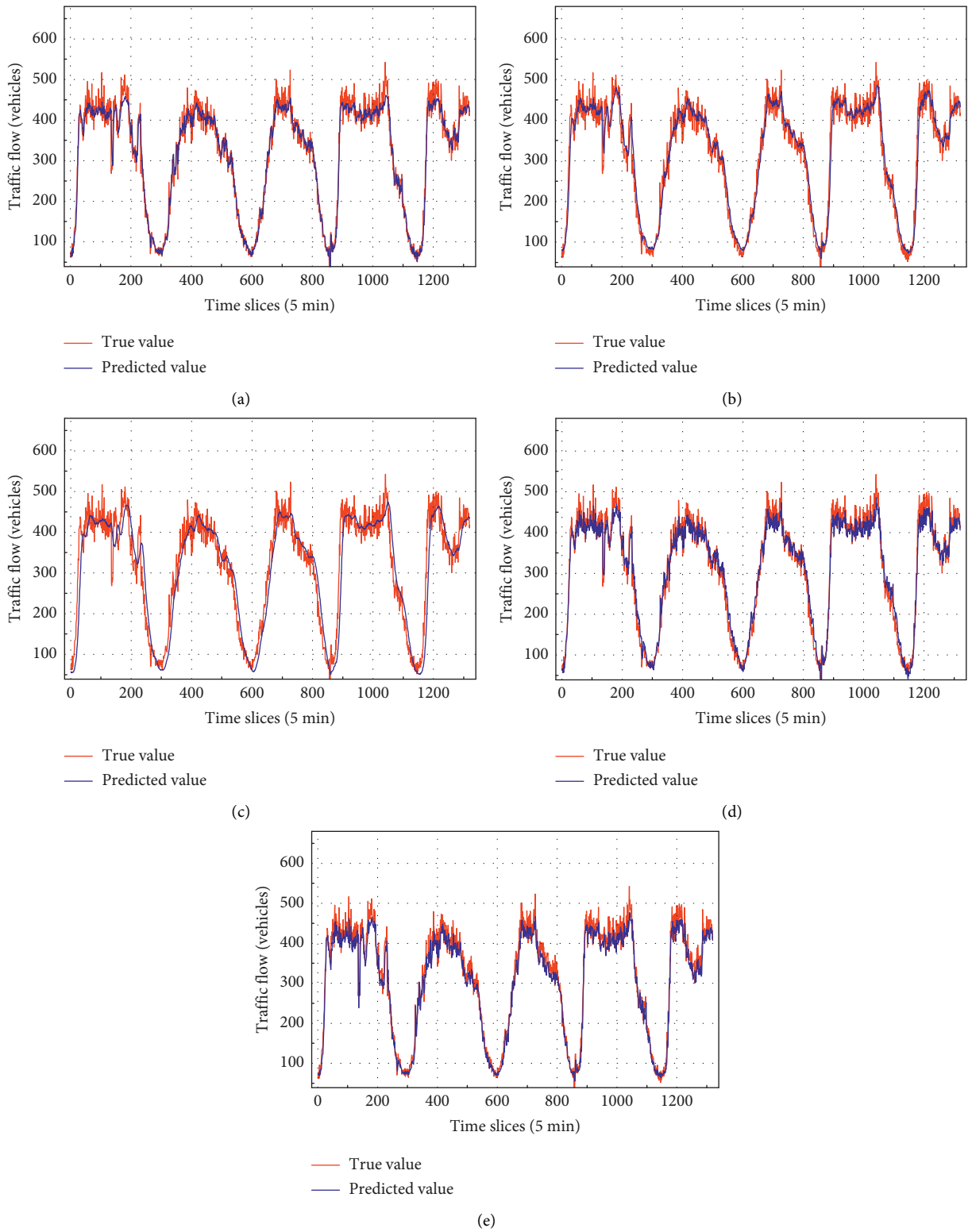


Figure 14: Prediction results of each model using the PeMSD8 dataset: (a) PDRGA-CDBN-LSTM; (b) CDBN-LSTM; (c) CDBN; (d) LSTM; and (e) PDRGA-CDBN-LSTM (noisy data).

the CDBN-LSTM. In the first 90 iterations, CDBN-LSTM converges quickly, but after 100 iterations, the error level stabilizes, the convergence speed slows, and the final convergence error is greater. In the PeMSD8 dataset experiments, the proposed model reached a low convergence error on the 22nd iteration, whereas the first 8 iterations belong to the curve's rapid convergence period. The fast convergence stage of CDBN-LSTM consists of the first 129

Table 5: Error evaluation indicator of each model using the Hefei demonstration zone ITS dataset.

	PDRGA-CDBN-LSTM	CDBN-LSTM	CDBN	LSTM	PDRGA-CDBN-LSTM (containing noise)
MSE	37.7831	43.3622	91.7154	80.0478	39.9624
RMSE	6.1468	6.5850	10.8682	8.9469	6.3216
MAE	4.1546	4.5865	7.8734	6.2781	4.4597
MAPE	0.1743	0.2284	0.3952	0.2798	0.2473
SMAPE	15.8289	19.8725	33.9259	27.5919	20.6687

Table 6: Error evaluation indicator of each model using the PeMSD8 dataset.

	PDRGA-CDBN-LSTM	CDBN-LSTM	CDBN	LSTM	PDRGA-CDBN-LSTM (containing noise)
MSE	694.2836	919.2247	2319.6816	1143.7891	808.6719
RMSE	26.3492	30.3187	48.1630	33.8200	28.4371
MAE	19.0154	22.5588	34.8261	26.2578	21.7602
MAPE	0.0668	0.0903	0.1412	0.0950	0.0777
SMAPE	6.5958	8.6518	14.5699	9.2467	7.9117

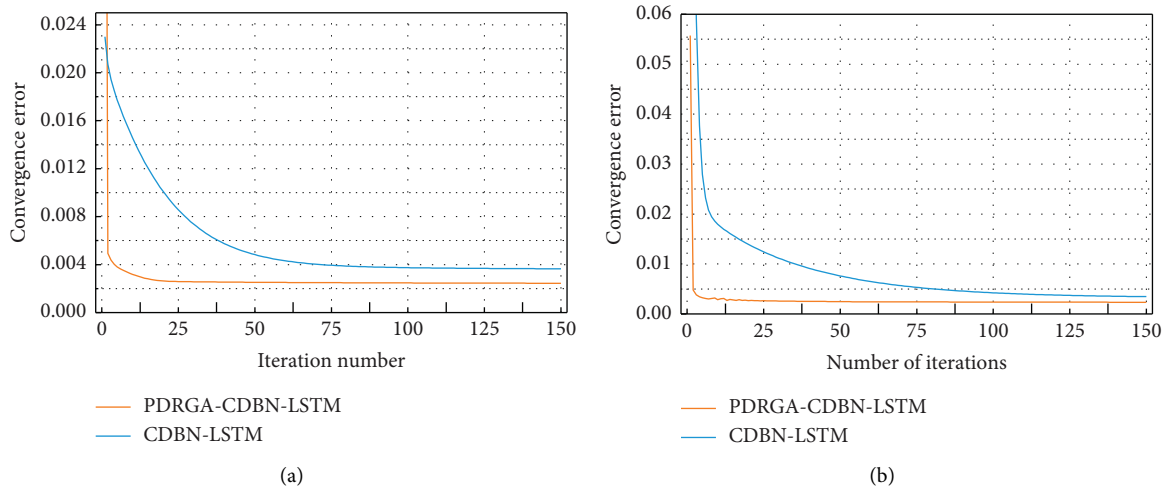


Figure 15: Relationship between iteration number and convergence error: (a) the Hefei demonstration zone ITS dataset; (b) the PeMSD8 dataset.

iterations, and its convergence efficiency and final convergence precision are inferior to those of the proposed model. Experiments on the two data sets reveal that the training error of CDBN-LSTM no longer declines considerably after iterations and remains consistently greater than that of PDRGA-LSTM, indicating that the model cannot discover the global optimal solution and falls into a local optimum. The above analysis indicates that PDRGA optimizes the model's ability to extract features, allowing it to escape the local optimum and reach a lower training error.

The analysis of the integrated model prediction and evaluation metrics reveals that the wavelet denoising method improves the value density of the data and enables the model to learn more effective data patterns during training; the lightweight improved PDRGA can enhance the learning ability of the model and prevent the model from falling into a local optimum; and the proposed PDRGA-CDBN-LSTM has the best performance in the prediction task of the measured chaotic traffic flow under five error metrics compared with all the baseline models.

6. Conclusion

From the perspective of studying the internal chaotic dynamics of the observed traffic flow, this paper focuses on the phase space reconstruction method of the chaotic data and constructs a chaotic-influenced hybrid traffic flow prediction model. In the phase space reconstruction stage, an improved C-C method based on the fusion norm search domain is proposed to reconstruct the phase space in order to assure the accurate and efficient calculation of reconstruction parameters. In order to make the model applicable to the complex nonlinear feature extraction law of chaotic traffic flow, a hybrid PDRGA-CDBN-LSTM prediction model is proposed to extract the spatial features of the chaotic phase space by CDBN and the temporal features jointly with

LSTM, while the initial weight parameters of CDBN are optimized by the improved PDRGA based on probabilistic dynamic multiplication. Experiments on the phase space reconstruction, convergence of the optimization algorithms, and the model prediction performance comparisons demonstrate the following:

(1) The wavelet denoising method may efficiently decrease data noise and enhance data value density, allowing the model to extract more high-frequency critical data features defining the traffic flow change pattern and so improving the prediction accuracy.

(2) In phase space reconstruction experiments utilizing the Hefei demonstration zone ITS dataset and the PeMSD8 dataset, the C-C method based on the improved fused norm search domain improves the operation time by 20.3% and 17.0%, respectively, compared with the $L2$ norm search domain and reduces the point finding error by 14.2% and 14.5%, respectively, compared with the infinite norm search domain. The algorithm is able to strike a compromise between the quality of phase space reconstruction and its efficiency by utilizing the high precision of the $L2$ norm and the convenience of infinite norm computation.

(3) The PDRGA based on lightweight improvement has a lower training cost; whereas the traditional GA requires 764 iterations to achieve final convergence accuracy, the PDRGA can reach the same error level in only 91 iterations, resulting in a significant reduction in computing time and the ability to effectively prevent the model from falling into local optimum.

(4) The proposed PDRGA-CDBN-LSTM model can effectively extract chaotic and temporal features from measured link traffic flow data, and in experiments on the Hefei demonstration zone ITS dataset and the PeMSD8 dataset, the error evaluation index of MAPE decreases by 23.7% and 26.0%, respectively, compared to the best-performing combined CDBN-LSTM model in the baseline. The proposed model outperforms the baseline model in terms of all error indexes, providing a new approach to model construction for link traffic flow prediction.

Data Availability

The traffic flow data used to support the findings of the study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This project was supported by the National Natural Science Foundation of China [Grant no. 62063014], the Natural Science Foundation of Gansu Province, China [Grant no. 20JR5RA407], the Science and Technology Innovation Project of the Gansu Provincial Education Department [Grant nos. 2021CYZC-04].

Supplementary Materials

Along with the article, the traffic flow data used in the Hefei demonstration zone ITS dataset and PeMSD8 dataset and the data descriptions were submitted as supplementary materials. Both datasets were collected at five-minute intervals. The Hefei demonstration zone ITS dataset was collected in the high-tech industrial demonstration zone from June 30, 2016, to July 29, 2016, with 8640 discrete time slices; the PeMSD8 dataset is 17280 discrete time slices collected in the highway network in San Bernardino from July 1, 2016, to August 29, 2016. (*Supplementary Materials*)

References

[1] M. Owais and A. E. Matouk, "A factorization scheme for observability analysis in transportation networks," *Expert Systems with Applications*, vol. 174, p. 114727, 2021.

[2] J. E. Disbro and M. Frame, *Traffic Flow Theory and Chaotic Behavior*, Dept. of Transportation, New York (State), 1989.

[3] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.

[4] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, warwick 1980*, pp. 366–381, Springer, 1981.

[5] H.-g. Ma and C. Z. Han, "Selection of embedding dimension and delay time in phase space reconstruction," *Frontiers of Electrical and Electronic Engineering in China*, vol. 1, pp. 111–114, 2006.

[6] C. J. Cellucci, A. M. Albano, and P. E. Rapp, "Comparative study of embedding methods," *Physical Review A*, vol. 67, no. 6, p. 066210, 2003.

[7] Y. I. Molkov, D. N. Mukhin, E. M. Loskutov, A. M. Feigin, and G. A. Fidelin, "Using the minimum description length principle for global reconstruction of dynamic systems from noisy time series," *Physical Review A*, vol. 80, no. 4, p. 046207, 2009.

[8] M. Wang and L. Tian, "From time series to complex networks: the phase space coarse graining," *Physica A: Statistical Mechanics and Its Applications*, vol. 461, pp. 456–468, 2016.

[9] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1986.

[10] L. Tang and J. Liang, "Cc method to phase space reconstruction based on multivariate time series,"in *Proceedings of the 2011 2nd international conference on intelligent control and information processing*, vol. 1, pp. 438–441, IEEE, Harbin, China, 25-28 July 2011.

[11] H. S. Kim, R. Eykholt, and J. Salas, "Nonlinear dynamics, delay times, and embedding windows," *Physica D: Nonlinear Phenomena*, vol. 127, no. 1-2, pp. 48–60, 1999.

[12] T. Mai, B. Ghosh, and S. Wilson, "Short-term traffic-flow forecasting with auto-regressive moving average models,"*Presented at Proceedings of the Institution of Civil Engineers-Transport*, vol. 167, pp. 232–239, Thomas Telford Ltd, 2014.

[13] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, pp. 1–11, 1984.

[14] X. Luo, L. Niu, and S. Zhang, "An algorithm for traffic flow prediction based on improved sarima and ga," *KSCE Journal of Civil Engineering*, vol. 22, no. 10, pp. 4107–4115, 2018.

[15] M.-b. Pang and X. P. Zhao, "Traffic flow prediction of chaos time series by using subtractive clustering for fuzzy neural network modeling,"in *Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application*, vol. 1, pp. 23–27, IEEE, Shanghai, China, 20-22 December 2008.

[16] L. Zhang, Q. Liu, W. Yang, N. Wei, and D. Dong, "An improved k-nearest neighbor model for short-term traffic flow prediction," *Procedia-Social and Behavioral Sciences*, vol. 96, pp. 653–662, 2013.

[17] L. Rui-Guo, Z. Hong-Li, F. Wen-Hui, and W. Ya, "Hermite orthogonal basis neural network based on improved teaching-learning-based optimization algorithm for chaotic time series prediction," *Acta Physica Sinica*, vol. 64, no. 20, p. 200506, 2015.

[18] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017.

[19] M. Owais, G. S. Moussa, and K. F. Hussain, "Robust deep learning architecture for traffic flow estimation from a subset of link sensors," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 1, p. 04019055, 2020.

[20] H. Tan, X. Xuan, Y. Wu, Z. Zhong, and B. Ran, "A comparison of traffic flow prediction methods based on dbn," *Cictp*, pp. 273–283, 2016.

[21] Z. Zhao, W. Chen, X. Wu, and J. Liu, "Lstm network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.

[22] L.-G. Zhang, S.-Q. Zhang, H.-T. Liu, W. Dong, and S.-S. Song, "Chaotic singular spectrum analysis based on improved phase space reconstruction algorithm and its application," *Acta Metrologica Sinica*, pp. 1299–1306, 2021.

[23] Y. Li, H.-Y. Wang, D.-X. Ren, H.-X. Tian, and H.-Y. Li, "Prediction of the compressor's operating state based on the improved phase-space reconstruction," *Journal of Machine Design*, vol. 37, pp. 84–90, 2020.

[24] A.-H. Jiang, P. Zhou, Y. Zhang, and H.-X. Hua, "Improved mutual information algorithm for phase space reconstruction," *Journal of Vibration and Shock*, vol. 34, pp. 79–84, 2015.

[25] F.-Q. Wang and G.-X. Chen, "Improvement of G-P algorithm and its application," *Chinese Journal of Computational Physics*, vol. 10, no. 3, pp. 345–351, 1993.

[26] Z. B. Lu, Z.-M. Cai, and K.-Y. Jiang, "Determination of embedding parameters for phase space reconstruction based on improved cc method," *Journal of System Simulation*, vol. 19, pp. 2527–2538, 2007.

[27] W. Zhang, Y. Yu, Y. Qi, F. Shu, and Y. Wang, "Short-term traffic flow prediction based on spatio-temporal analysis and cnn deep learning," *Transportmetrica: Transportation Science*, vol. 15, no. 2, pp. 1688–1711, 2019.

[28] R. Fu, Z. Zhang, and L. Li, "Using lstm and gru neural network methods for traffic flow prediction," in *Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 324–328, IEEE, Wuhan, China, 11-13 November 2016.

[29] G. S. Moussa and M. Owais, "Pre-trained deep learning for hot-mix asphalt dynamic modulus prediction with laboratory effort reduction," *Construction and Building Materials*, vol. 265, p. 120239, 2020.

[30] G. S. Moussa and M. Owais, "Modeling hot-mix asphalt dynamic modulus using deep residual neural networks: parametric and sensitivity analysis study," *Construction and Building Materials*, vol. 294, no. 2021, p. 123589, 2021.

[31] G. S. Moussa, M. Owais, and E. Dabbour, "Variance-based global sensitivity analysis for rear-end crash investigation using deep learning," *Accident Analysis & Prevention*, vol. 165, no. 2022, p. 106514, 2022.

[32] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014.

[33] Z. Zheng, Y. Yang, J. Liu, H.-N. Dai, and Y. Zhang, "Deep and embedded learning approach for traffic flow prediction in urban informatics," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3927–3939, 2019.

[34] W. J. Huang, Y.-T. Li, and Y. Huang, "Prediction of chaotic time series using hybrid neural network and attention mechanism," *Acta Physica Sinica*, vol. 70, no. 1, p. 010501, 2021.

[35] M. Owai, M. K. Osman, and G. Moussa, "Multi-objective transit route network design as set covering problem," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 670–679, 2016.

[36] M. Owais and M. K. Osman, "Complete hierarchical multi-objective genetic algorithm for transit network design problem," *Expert Systems with Applications*, vol. 114, pp. 143–154, 2018.

[37] R. Zhang, F. Sun, Z. Song, X. Wang, Y. Du, and S. Dong, "Short-term traffic flow forecasting model based on ga-tcn," *Journal of Advanced Transportation*, vol. 2021, pp. 1–13, 2021.

[38] J. Zhou, H. Chang, X. Cheng, and X. Zhao, "A multiscale and high-precision lstm-gasvr short-term traffic flow prediction model," *Complexity*, pp. 1–17, 2020.

[39] J. Wang, Q. Shi, and H. Lu, "The study of short-term traffic flow forecasting based on theory of chaos," in *Proceedings of the IEEE Proceedings. Intelligent Vehicles Symposium*, pp. 869–874, IEEE, Las Vegas, NV, USA, 06-08 June 2005.

[40] W. Liebert and H. Schuster, "Proper choice of the time delay for the analysis of chaotic time series," *Physics Letters A*, vol. 142, no. 2-3, pp. 107–111, 1989.

[41] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "A practical method for calculating largest lyapunov exponents from small data sets," *Physica D: Nonlinear Phenomena*, vol. 65, no. 1-2, pp. 117–134, 1993.

[42] S. Sato, M. Sano, and Y. Sawada, "Practical methods of measuring the generalized dimension and the largest lyapunov exponent in high dimensional chaotic systems," *Progress of Theoretical Physics*, vol. 77, pp. 1–5, 1987.

[43] J. Sola and J. Sevilla, "Importance of input data normalization for the application of neural networks to complex industrial problems," *IEEE Transactions on Nuclear Science*, vol. 44, no. 3, pp. 1464–1468, 1997.

[44] T. Zhongda, L. Shujiang, W. Yanhong, and S. Yi, "A prediction method based on wavelet transform and multiple models fusion for chaotic time series," *Chaos, Solitons & Fractals*, vol. 98, pp. 158–172, 2017.

[45] A. A. Koronovskii and A. E. Hramov, "Wavelet transform analysis of the chaotic synchronization of dynamical systems," *Journal of Experimental and Theoretical Physics Letters*, vol. 79, no. 7, pp. 316–319, 2004.

[46] M. Han, *Prediction Theory and Method of Chaotic Time Series*, China Water&Power Press, 2007.
[47] N. P. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," *Physical Review Letters*, vol. 45, no. 9, pp. 712–716, 1980.
[48] J. D. Farmer, E. Ott, and J. A. Yorke, "The dimension of chaotic attractors," *Physica D: Nonlinear Phenomena*, vol. 7, no. 1-3, pp. 153–180, 1983.
[49] J. Sprott, "How common is chaos?" *Physics Letters A*, vol. 173, no. 1, pp. 21–24, 1993.
[50] D. Kugiumtzis, "State space reconstruction parameters in the analysis of chaotic time series—the role of the time window length," *Physica D: Nonlinear Phenomena*, vol. 95, no. 1, pp. 13–28, 1996.
[51] D. S. Dendrinos, "Traffic-flow dynamics: a search for chaos," *Chaos, Solitons & Fractals*, vol. 4, pp. 605–617, 1994.
[52] J. Tang, Y. Wang, H. Wang, S. Zhang, and F. Liu, "Dynamic analysis of traffic time series at different temporal scales: a complex networks approach," *Physica A: Statistical Mechanics and Its Applications*, vol. 405, pp. 303–315, 2014.
[53] I. E. Livieris, E. Pintelas, and P. Pintelas, "A cnn–lstm model for gold price time-series forecasting," *Neural Computing & Applications*, vol. 32, no. 23, pp. 17351–17360, 2020.
[54] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," *Presented at International conference on machine learning*, pp. 1310–1318, PMLR, 2013.
[55] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with lstm," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
[56] L. Wang, X. Shu, G. Lu, and F. Gao, "Energetics and structures of hydrogen-vacancy clusters in tungsten based on genetic algorithm," *Science China Physics, Mechanics & Astronomy*, vol. 61, no. 10, p. 107022, 2018.
[57] A. H. Wright, "Genetic algorithms for real parameter optimization," in *Foundations of genetic algorithms*, vol. 1, pp. 205–218, Elsevier, 1991.