Nonlinear Dyn (2016) 85:179–194
DOI 10.1007/s11071-016-2677-5

ORIGINAL PAPER

Multiple measures-based chaotic time series for traffic flow prediction based on Bayesian theory

Yongfu Li · Xiao Jiang · Hao Zhu · Xiaozheng He · Srinivas Peeta · Taixiong Zheng · Yinguo Li

Received: 1 December 2015 / Accepted: 10 February 2016 / Published online: 25 February 2016

Abstract Considering the chaotic characteristics of traffic flow, this study proposes a Bayesian theory-based multiple measures chaotic time series prediction algorithm. In particular, a time series of three traffic measures, i.e., speed, occupancy, and flow, obtained from different sources is used to reconstruct the phase space using the phase space reconstruction theory. Then, data from the multiple sources are combined using Bayesian estimation theory to identify the chaotic characteristics of traffic flow. In addition, a radial basis function (RBF) neural network is designed to predict the traffic flow. Compared to the consideration of a single source, results from numerical experiments demonstrate the improved effectiveness of the proposed multi-measure method in terms of accuracy and timeliness for the short-term traffic flow prediction.

Keywords Traffic flow · Chaos theory · Phase space reconstruction · Bayesian estimation · Multi-measure time series · RBF neural network

Yongfu Li (✉) · X. Jiang · H. Zhu · T. Zheng · Yinguo Li
Chongqing Collaborative Innovation Center for Information Communication Technology, College of Automation and Center for Automotive Electronics and Embedded System, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
e-mail: laf1212@163.com; liyongfu@cqupt.edu.cn

Yongfu Li · X. He · S. Peeta
School of Civil Engineering and The NEXTRANS Center, Purdue University, West Lafayette, IN 47906, USA

1 Introduction

Due to advances in information and communication technologies, traffic flow modeling [1–8], traffic control, and operations based on intelligent transportation systems are being implemented to mitigate the ever-worsening traffic jams and the pressure imposed on transportation systems [9–14]. The effectiveness and efficiency of the traffic control and operations rely heavily on the accurate prediction of the traffic state that involves multiple measures such as speed, density, and flow. However, based on the fundamental traffic relationships, flow is related to speed and density. Hence, there is a need to predict traffic flow accurately to capture the characteristics of the traffic state evolution. As traffic conditions change dynamically, the prediction of traffic flow is typically modeled as a time series.

Although an accurate prediction of the traffic flow is important for effective traffic control and operations, the inherent uncertainty, nonlinearity, and complexity of traffic dynamics impede the real-time prediction accuracy of the traffic flow time series, especially when the prediction period allowed is short to support quick responses to traffic events. However, recent studies indicate that the short-term traffic flow time series has nonlinear chaotic phenomena [15]; namely, small differences in the initial condition would yield large difference in the prediction results. This chaotic feature precludes the use of traditional models such as the linear regression model to achieve the prediction accuracy and timeliness required for traffic control and

operations. It motivates the need to explore nonlinear dynamical system-based approaches, especially chaos theory, for the accurate high-frequency prediction of traffic flow in real time [16].

In the literature, time series models based on chaotic nonlinear dynamical systems have been developed and applied to enhance the reliability and accuracy of prediction by dynamically updating the prediction model to help it converge to the actual physical process [17]. The time series prediction methods for these chaotic systems can be categorized into two types: local-region methods and whole-region methods. Widely used local-region methods include maximum Lyapunov exponent-based local-region method [18], zero-order local-region method [19], one-rank local-region method [20], and high-order local-region method [18]. Ma et al. [16] applied the Lyapunov exponent-based local-region method to predict the traffic flow and concluded that using multiple traffic measures produces better results compared to using a single traffic measure. They also found that the prediction accuracy significantly depends on the length of prediction time step. The merits of local-region methods lie in their simplicity and low computational requirements. However, they cannot predict traffic flows that do not exist in the historical data. Whole-region methods are able to address this limitation of local-region methods in predicting states that are not present in the historical data.

In the transportation literature, whole-region methods include the polynomial model [21] and the neural network model [22]. In the category of whole-region methods, the neural network model has the potential to identify a highly nonlinear system due to its ability to approximate complex nonlinear systems after the data preprocessing step [23]. Hence, many studies have applied neural network models to predict the time series of traffic flow. Karlaftis and Vlahogianni [22] summarized the types of neural networks used in the transportation context, including back propagation (BP) neural network, multilayer perception (MLP) neural network, and radial basis function (RBF) neural network. Jayawardena and Fernando [24] applied the RBF neural network for hydrologic modeling and runoff simulation in a small catchment and reported that it is computationally more efficient than the BP neural network. Park et al. [25] compared the BP neural network with the RBF neural network and concluded that the RBF neural network requires lesser training time and offers higher accuracy than the BP neural network. Chen and Grant-Muller [26] found that the RBF neural network outperforms the MLP neural network in terms of the prediction performance and a simpler structure with only one hidden layer. Celikoglu [27] demonstrated the better prediction accuracy and faster convergence rate of RBF neural network model compared to other models for a chaotic system. Zhang et al. [28] investigated the short-term prediction of traffic flow using a nonlinear time series and proposed a RBF neural network prediction method based on chaos theory. Their simulation experiments illustrated the strength of the RBF neural network in prediction accuracy. In summary, the RBF neural network can overcome the limitations of other neural networks such as the determination of local optima and slow convergence rate. Hence, it has been widely used for short-term traffic flow prediction.

The neural network models specified heretofore focus on the time series prediction of a single traffic measure. Note that traffic system state is nonlinear and dynamic, and can be characterized by three measures: speed, density, and flow. The fundamental relation of traffic flow states that traffic flow is the product of traffic speed and traffic density. This inherent relationship among these three measures can be leveraged to improve the prediction accuracy of a single measure by using available time series information on all three measures. Therefore, it is needed to construct a model for time series prediction which applies information on all traffic measures to accurately characterize the change of traffic flow.

In the literature on time series prediction, several studies proposed fusion models based on Bayesian estimation theory to combine multiple time series into a single time series. Based on Takens' embedding theorem [29] and Bayesian estimation theory [30,31], Cong et al. [17] proposed a model to fuse information from multiple sources into a single phase space. Their study resolves the problems inherent in single and multiple phase space reconstruction and verifies that a time series based on multiple measures (e.g., traffic speed, density and flow in the context of this study) can provide more comprehensive information than one based on a single measure. Krese and Govekar [32] reconstructed phase space for a time series and proposed multi-step adaptive predictive method based on chaos theory. Their numerical results showed that a chaotic time series prediction based on reconstruction of phase

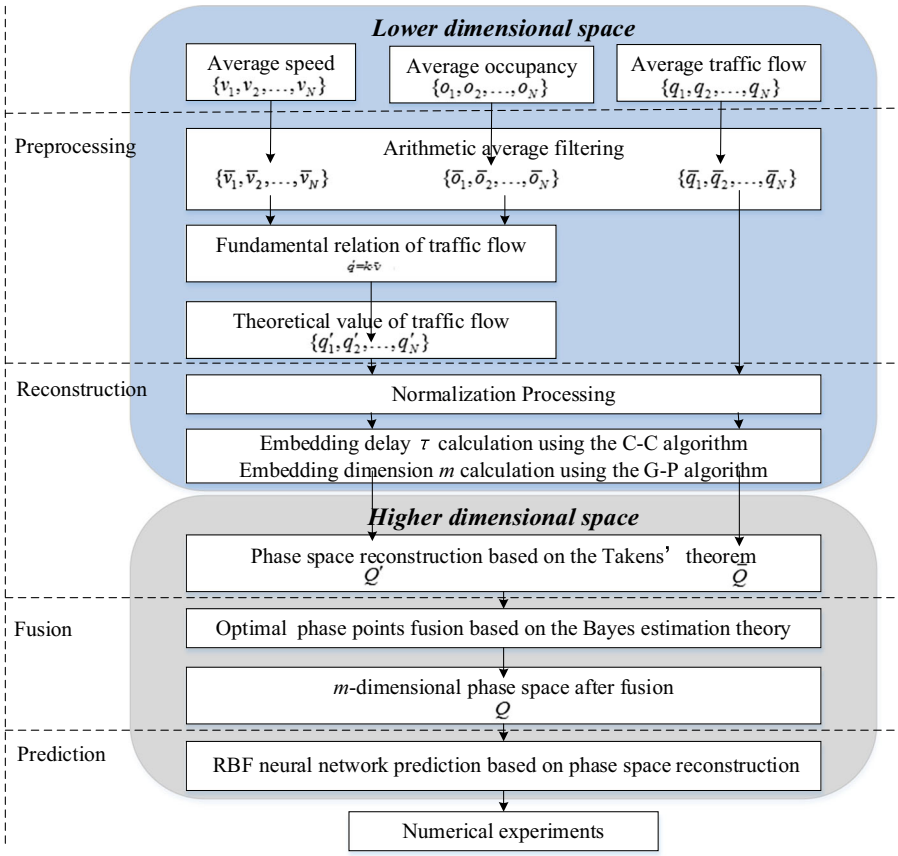


Fig. 1 Proposed framework

space has a higher accuracy than other prediction methods when performing a short-term prediction of the time series.

This study proposes a new short-term traffic flow prediction method which fuses multiple traffic measures into a multi-dimensional space using chaos theory and Bayesian estimation theory. The proposed method applies RBF neural network to predict the variation trend of the time series of traffic flow. In particular, the chaotic characteristics of the time series of traffic measures are investigated using saturation correlation dimension and maximum Lyapunov exponent method. Then, for fusion purpose, multiple time series of traffic measures are mapped to a high-dimensional space through the selection of the embedded dimensions of reconstructed space and delay time. Finally, a multi-measure chaos prediction framework is developed for the prediction of traffic flow. The study simulation results illustrate that the proposed method is effective in terms of accuracy and timeliness for short-term traffic flow prediction.

2 Methodology

Considering that the average density of traffic flow can be directly calculated by the measured occupancy [see Eq. (25)], hence, in this study, a traffic state is characterized by three traffic measures: speed, occupancy, and flow. The time series of these three measures can be integrated into a single time series of the fused traffic flow through the multi-source integration method [16,17] so that the predicted traffic flow can be dynamically updated more accurately due to the correlation between these three measures. Hence, to improve the prediction accuracy, a multi-source fusion method is proposed based on the phase space reconstruction that will be explain hereafter.

2.1 Framework

We propose a two-level framework to predict the traffic flow as shown in Fig. 1. The proposed framework includes two components. The first compo-

nent processes and normalizes the data in a lower-dimensional space, and the second component fuses the data and predicts traffic flow in a higher-dimensional space.

Following Fig. 1, the framework is summarized as follows:

Step 1: Preprocessing. To reduce the effect of measurement noises, the average filtering method proposed in [33] is used to process the raw data. Then, the theoretical traffic flow is calculated in terms of the average occupancy and speed.

Step 2: Reconstruction. To enable the fusing of traffic flow, occupancy, and speed, these measures are normalized to the range [0, 1]. Then, the C-C algorithm [34] and G-P algorithm [35] are used to compute the embedding delay τ and dimension m for phase space reconstruction, respectively. In addition, to guarantee that the time series of all measures can be completely unfolded and without distortion, we choose the largest embedding dimension m and the minimum delay τ according to the phase space reconstruction theory.

Step 3: Fusion. To obtain the optimal phase point, the Bayesian estimation approach is used in the phase space to fuse the k-th phase points of the time series of traffic flow, occupancy, and speed.

Step 4: Prediction. To predict the traffic flow based on the fusion results, the RBF neural network is designed.

2.2 Phase space reconstruction

Phase space reconstruction is regarded as the basis of chaotic time series analysis and widely used in nonlinear system analysis [36]. Takens [29] proposed the delay coordinates method of phase space reconstruction for chaotic time series analysis. This study applies the delay coordinates method to reconstruct the multi-measure phase space.

Suppose a time series represented by $[X_1, X_2, \ldots, X_M]^{\mathrm{T}}$ contains M measures. Each element X_i can be viewed as an individual time series consisting of N data points, i.e., $X_i = \{x_i(t) \mid t = 1, 2, \ldots, N\}$. Denote $x_{i,t} = x_i(t)$ for all $i = 1, 2, \ldots, M$ and $t = 1, 2, ..., N$. Then,

$$X = [X_1, X_2, \ldots, X_M]^{\mathrm{T}} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M,1} & x_{M,2} & \cdots & x_{M,N} \end{bmatrix}. \quad (1)$$

Due to the inconsistent dimensions among the measured time series of the three traffic measures, a normalization process is needed for each time series according to Eq. (2).

$$y(n) = \frac{x(n) - \min(x(n))}{\max(x(n) - \min(x(n)))}. \quad (2)$$

Equation (3) indicates the time series after the normalization:

$$Y = [Y_1, Y_2, \ldots, Y_M]^{\mathrm{T}} = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,N} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{M,1} & y_{M,2} & \cdots & y_{M,N} \end{bmatrix}. \quad (3)$$

According to Takens' embedding theorem [29], a chaotic nonlinear time series can be reconstructed as a recovery phase space of the original dynamical system by selecting the embedding dimension m and delay τ. More specifically, for a chaotic time series with multiple measures, the C-C algorithm [34] can be applied to capture the correlation between the measures to improve the prediction accuracy. For each time series Y_i of measure $i = 1, 2, \ldots, M$, its optimal embedding dimension m_i and delay time τ_i can be determined using the G-P algorithm [35]. To capture all attributes of the M measures in a single recovery phase space, the largest embedding dimension m and the minimum delay time τ are used in the phase space reconstruction process. Denote

$$m = \max_i (m_i), \text{ and } \tau = \min_i (\tau_i). \quad (4)$$

For each measure $i = 1, 2, \ldots, M$, the corresponding time series $Y_i = \begin{bmatrix} y_{i,1} & y_{i,2} & \ldots & y_{i,N} \end{bmatrix}$ can be reconstructed into a phase space represented by:

$$Q_i = \begin{bmatrix} y_{i,1} & y_{i,1+\tau} & \cdots & y_{i,1+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ y_{i,k} & y_{i,k+\tau} & \cdots & y_{i,k+(m-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i,K} & y_{i,K+\tau} & \cdots & y_{i,K+(m-1)\tau} \end{bmatrix}, \quad (5)$$

where the number of phase points K is determined by $N-(m-1)\tau$. The k-th row of Q_i, $\left\{y_{i,k}, y_{i,k+\tau}, \ldots, y_{i,k+(m-1)\tau}\right\}$, represents the k-th time series data, and all rows together represent the m-dimensional phase space.

2.3 Phase point fusion in phase space based on Bayesian estimation

This section introduces a multivariate data fusion method to fuse the information on multiple dimensions into a single time series to improve the short-term prediction accuracy. In general, the reconstructed phase space of a single measure can approximately prescribe the characteristic of the attractor of the original dynamical system [37]. However, due to data incompleteness, a single-measure time series is unable to comprehensively reveal the characteristics of the system in other dimensions. Multivariate data fusion methods fuse the multiple-measure time series into a single one, which integrates the system characteristics associated with the different dimensions.

This study proposes a multivariate data fusion method based on Bayesian estimation. Bayesian estimation-based data fusion methods make full use of the prior information of the measures and sample information. Therefore, Bayesian estimates in parameter estimation usually have smaller mean square error [38]. By applying the Bayesian estimation-based multivariate data fusion method, the phase points of multiple measures can be fused into a single one such that the characteristics provided by the phase points can complement each other and mitigate the measurement errors in a single measure. By doing so, the estimation results can better approximate the true state of the system.

According to the phase space reconstruction in the previous section, each measure $i=1,2,\ldots,M$ has the same number of phase points K. Denote the phase point $k=1,2,\ldots,K$ of the reconstructed phase spaces of measures $i=1,2,\ldots,M$ as:

$$\mathbf{y}_k^i=\left[y_{i,k}, y_{i,k+\tau}, \ldots, y_{i,k+(m-1)\tau}\right] \quad \forall i=1,2,\ldots,M; k=1,2,\ldots,K. \tag{6}$$

Then, for a phase point k, collecting the specific phase points from each of the reconstructed phase spaces of measures forms a phase collection point that is denoted by:

$$D_k=\left[\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right] \quad \forall k=1,2,\ldots,K. \tag{7}$$

Denote $z_k (k=1,2,\ldots,K)$ as the fused phase point based on the given reconstructed phase spaces. According to the Bayesian estimation, the probability of z_k can be presented by:

$$p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)=\frac{p\left(z_k; \mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)}{p\left(\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)}. \tag{8}$$

Assume that $z_k \sim N\left(z_0, \sigma_0^2\right)$ and $D_k \sim N\left(z_k, \sigma_h^2\right)$, where z_0 and σ_0^2, respectively, are the average and variance of z_k, and σ_h^2 is the covariance matrix of D_k. Denote $\alpha=\frac{1}{p\left(\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)}$. The posterior probability of the z_k is determined by:

$$p\left(z_k|D_k\right)=\frac{p\left(D_k|z_k\right)p\left(z_k\right)}{p\left(D_k\right)}=\alpha p\left(D_k|z_k\right)p\left(z_k\right). \tag{9}$$

Expanding Eq. (9) shows that:

$$\begin{aligned}
&p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)\\
&=\alpha\prod_{h=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_h}\exp\left[-\frac{1}{2}\left(\frac{\mathbf{y}_k^h-z_k}{\sigma_h}\right)^2\right]\\
&\quad\times\frac{1}{\sqrt{2\pi}\sigma_0}\exp\left[-\frac{1}{2}\left(\frac{z_k-z_0}{\sigma_0}\right)^2\right]\\
&=\alpha\prod_{h=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_h}\times\frac{1}{\sqrt{2\pi}\sigma_0}\\
&\quad\times\exp\left[-\frac{1}{2}\sum_{h=1}^{M}\left(\frac{\mathbf{y}_k^h-z_k}{\sigma_h}\right)^2-\frac{1}{2}\left(\frac{z_k-z_0}{\sigma_0}\right)^2\right]\\
&=\beta\exp\left[-\frac{1}{2}\sum_{h=1}^{M}\left(\frac{\mathbf{y}_k^h-z_k}{\sigma_h}\right)^2-\frac{1}{2}\left(\frac{z_k-z_0}{\sigma_0}\right)^2\right]\\
&=\beta\exp\left[-\frac{1}{2}\left(\sum_{h=1}^{M}\frac{\left(\mathbf{y}_k^h\right)^2}{\sigma_h^2}+\frac{z_0^2}{\sigma_0^2}\right)\right.\\
&\quad\left.-\frac{1}{2}\left(\left(\sum_{h=1}^{M}\frac{1}{\sigma_h^2}+\frac{1}{\sigma_0^2}\right)z_k^2-2\left(\sum_{h=1}^{M}\frac{\mathbf{y}_k^h}{\sigma_h^2}+\frac{z_0}{\sigma_0^2}\right)z_k\right)\right]\\
&=\gamma\exp\left[-\frac{1}{2}\left(\left(\sum_{h=1}^{M}\frac{1}{\sigma_h^2}+\frac{1}{\sigma_0^2}\right)z_k^2\right.\right.\\
&\quad\left.\left.-2\left(\sum_{h=1}^{M}\frac{\mathbf{y}_k^h}{\sigma_h^2}+\frac{z_0}{\sigma_0^2}\right)z_k\right)\right]
\end{aligned} \tag{10}$$

where $\beta=\alpha\prod_{h=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_h}\times\frac{1}{\sqrt{2\pi}\sigma_0}$ and $\gamma=\beta\exp\left[-\frac{1}{2}\left(\sum_{h=1}^{M}\frac{\left(\mathbf{y}_k^h\right)^2}{\sigma_h^2}+\frac{z_0^2}{\sigma_0^2}\right)\right]$ are independent of z_k.

As can be observed from Eq. (10), the exponent in the equation is a quadratic function of z_k. Therefore, $p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)$ follows a normal distribution. Denote $p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right) \sim N\left(z, \sigma^2\right)$, where z and σ^2, respectively, are the average and variance of $p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right)$. We have:

$$p\left(z_k|\mathbf{y}_k^1, \mathbf{y}_k^2, \ldots, \mathbf{y}_k^M\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{z_k - z}{\sigma}\right)^2\right], \tag{11}$$

Combining Eqs. (10) and (11) yields the following equation:

$$\gamma \exp\left[-\frac{1}{2}\left(\left(\sum_{h=1}^{M}\frac{1}{\sigma_h^2} + \frac{1}{\sigma_0^2}\right)z_k^2 - 2\left(\sum_{h=1}^{M}\frac{\mathbf{y}_k^h}{\sigma_h^2} + \frac{z_0}{\sigma_0^2}\right)z_k\right)\right] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{z_k - z}{\sigma}\right)^2\right]. \tag{12}$$

Eq. (12) implies that:

$$\begin{cases} \frac{1}{\sigma^2} = \sum_{h=1}^{M}\frac{1}{\sigma_h^2} + \frac{1}{\sigma_0^2} \\ \frac{z}{\sigma^2} = \sum_{h=1}^{M}\frac{\mathbf{y}_k^h}{\sigma_h^2} + \frac{z_0}{\sigma_0^2} \end{cases}, \tag{13}$$

Solving Eq. (13) yields:

$$z = \frac{\sum_{h=1}^{M}\frac{\mathbf{y}_k^h}{\sigma_h^2} + \frac{z_0}{\sigma_0^2}}{\sum_{h=1}^{M}\frac{1}{\sigma_h^2} + \frac{1}{\sigma_0^2}}. \tag{14}$$

Therefore, $\hat{z}_k$, the Bayesian estimation value of z_k, can be written as:

$$\hat{z}_k = \int_{\Omega} z_k \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{z_k - z}{\sigma}\right)^2\right] dz_k = z \quad \forall k = 1, 2, \ldots, K. \tag{15}$$

Applying the above procedure to each group of phase points, a new fusion of phase points can be obtained, $Z_k = \hat{z}_k$, and then, a new phase space with m-dimensions can be achieved with the following form

$$Z = [Z_1, \ldots, Z_k, \ldots, Z_K]^{\mathrm{T}}. \tag{16}$$

Each Z_k in Eq. (16) can be expressed as:

$$Z_k = \left(z_k, z_{k+\tau}, \ldots, z_{k+(m-1)\tau}\right) \quad \forall k = 1, 2, \ldots, K \tag{17}$$

where K represents the number of phase space and k represents the arbitrary coordinate point in the time series. Each phase point Z_k includes the main features of every single variable in the new phase space Z and is able to approach the true state of traffic flow. Therefore, the new fusion phase space can accurately characterize the traffic flow.

2.4 RBF neural network based prediction model

In the transportation literature, studies have verified that the traffic state time series can be regarded as a chaotic system [32]. A chaotic system is sensitive to the initial values of measures. A tiny change in the initial values can potentially lead to significant differences in the long-term evolution of the system. This makes the evolution of a chaotic system seem random. Although an accurate long-term prediction for a chaotic system is not possible in general [39], a robust short-term prediction can be feasible. As the short-term traffic state prediction can be regarded as a nonlinear random input/output system, it is feasible to apply an artificial neural network to predict the short-term traffic state sequence based on Kolmogorov continuity theorem [40].

The RBF neural network is a three-layered feed-forward neural network with one radial basis layer. It can uniformly approximate any continuous function with a prospective accuracy. The RBF neural network has local generalization abilities and fast convergence speed [23]. As shown in Fig. 2, the entire network includes three layers: an input layer, a nonlinear hidden layer (radial basis layer), and a linear output layer.

In this study, the typical structure of an m-h-1 RBF neural network has m inputs; this implies embedding dimension m in Eq. (4), h nodes of hidden layer, and 1 output. The vector $z = \left(z_n, z_{n-\tau}, \ldots, z_{n-(m-1)\tau}\right)$ is the input vector of the network, $w = (w_1, \ldots, w_j, \ldots, w_h)$ is the vector of output weight, and z_{n+1} is the output of the network. Denote $\phi(\cdot)$ as the Gaussian radial basis function of the hidden nodes, c_j as the center of radial basis function of hidden node j, and $\|z - c_j\|$ as the Euclidean distance between input vector z and cen-

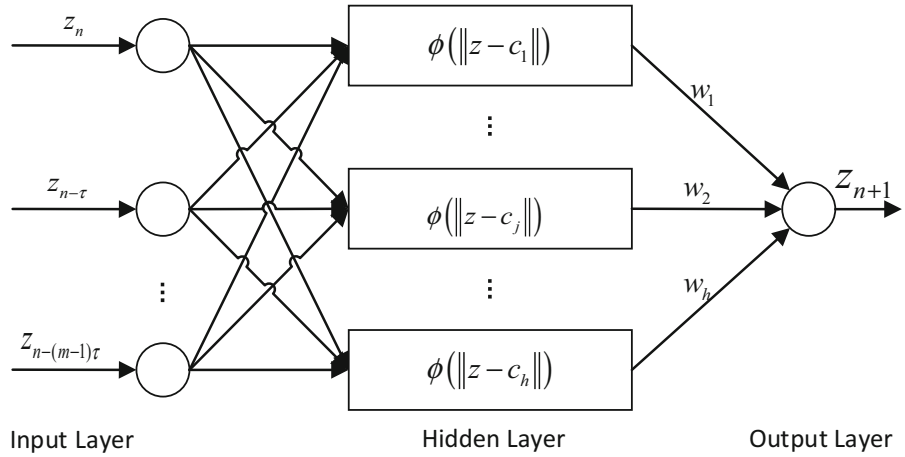


Fig. 2 RBF neural network structure based on phase space reconstruction

ter c_j. The Gaussian radial basis function of the hidden node j is:

$$\phi\left(\|z-c_j\|\right)=\exp\left(-\frac{\left\|z-c_j^2\right\|}{2\sigma_j^2}\right), \tag{18}$$

where σ_j is the j-th neuron of the width parameter of the RBF.

Then, the output of the neural network can be expressed as:

$$z_{n+1}=f(z)=\sum_{j=1}^{h} w_j\phi\left(\|z-c_j\|\right), \tag{19}$$
$$\forall n=(m-1)\tau+1,(m-1)\tau+2,\ldots,N$$

where w_j is the weight imposed on node j and N is the number of time series data points. Assume $n=k+(m-1)\tau$. Eq. (17) is transformed into:

$$z=\left(z_n,z_{n-\tau},\ldots,z_{n-(m-1)\tau}\right) \tag{20}$$
$$\forall n=(m-1)\tau+1,(m-1)\tau+2,\ldots,N,$$

where z is used as the input vectors of the artificial neural network.

To use the RBF neural network to predict the chaotic time series, the number of neurons in each layer of the neural network needs to vary according to the chaotic time series. In general, the number of input layer neurons is equal to the chaotic time series in the reconstructed phase space of embedding dimension m, to achieve a better prediction result. The selection of the number of hidden layer nodes does not have an ideal analytical formula and is determined according to experience or experimental analysis. In this study, three indices are used to evaluate the prediction accuracy: mean absolute error (*MAE*), mean absolute relative error (*MARE*), and equal coefficient (*EC*). *MAE* is used to measure the mean absolute value of the error between the predicted and actual values. *MARE* is used to measure the degree of deviation between the actual and estimated values. *EC* is used to illustrate the level of fit between the predicted and actual values and entails an average of more than 0.9 to imply better fit. The performance indices are defined as follows:

$$MAE=\frac{\sum_{h=1}^{N_p}\left|x(h)-\hat{x}(h)\right|}{N_p}, \tag{21}$$

$$MARE=\frac{1}{N_p}\sum_{h=1}^{N_p}\left|\frac{x(h)-\hat{x}(h)}{x(h)}\right|, \tag{22}$$

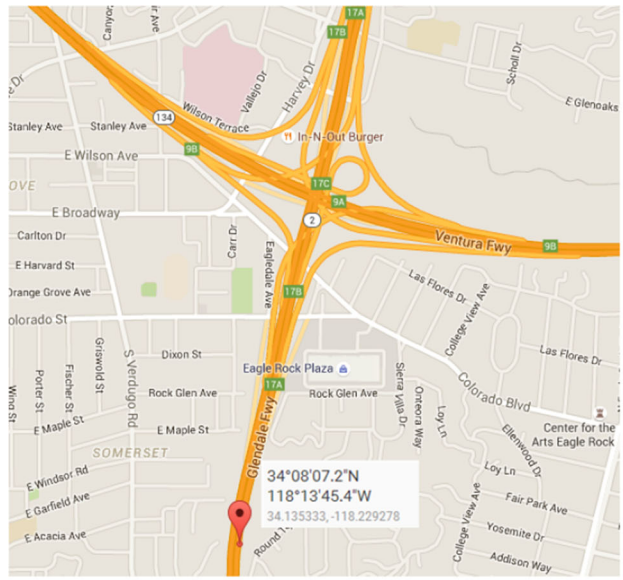
$$EC=1-\frac{\sqrt{\sum_{h=1}^{N_p}\left(x(h)-\hat{x}(h)\right)^2}}{\sqrt{\sum_{h=1}^{N_p}(x(h))^2}+\sqrt{\sum_{h=1}^{N_p}\left(\hat{x}(h)\right)^2}}, \tag{23}$$

where $x(h)$ is the actual value, $\hat{x}(h)$ is the predicted value, and N_p is the length of the predicted time series.

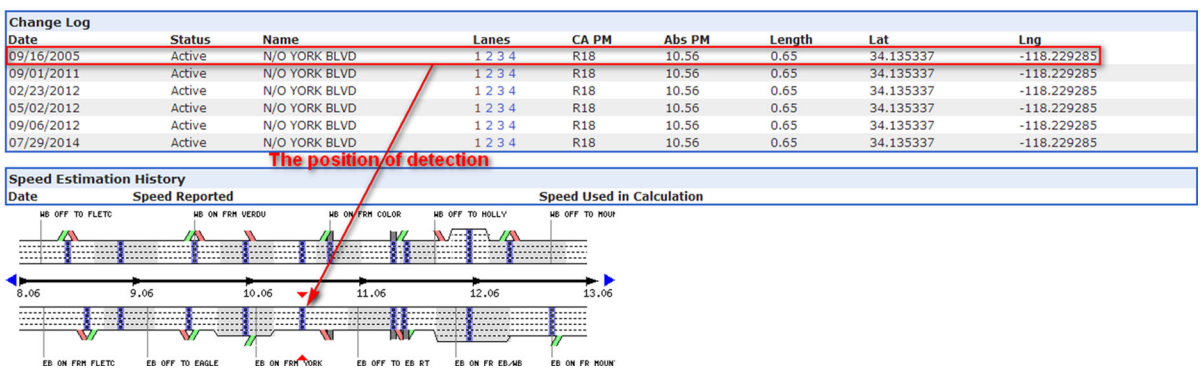
3 Numerical example

To investigate the effectiveness of the proposed short-term traffic state prediction method, a numerical experiment is implemented based on the traffic data collected by the loop detector at SR2-E @ CA PM R18 (Abs PM 10.6) District 7, Los Angeles County, City of Los Angeles, through the California's Freeway Performance Measurement System (PeMS) [41].

(a)



(b)



Change Log

Date	Status	Name	Lanes	CA PM	Abs PM	Length	Lat	Lng
09/16/2005	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285
09/01/2011	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285
02/23/2012	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285
05/02/2012	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285
09/06/2012	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285
07/29/2014	Active	N/O YORK BLVD	1 2 3 4	R18	10.56	0.65	34.135337	-118.229285

Fig. 3 Data collection location: **a** coordinates of data collection location; **b** loop detector location at SR2-E in PeMS

The data set contains information on average speed ($\{v_1, v_2, \ldots, v_N\}$, measured by radar detectors), average occupancy ($\{o_1, o_2, \ldots, o_N\}$, measured by loop detectors), and average traffic flow ($\{q_1, q_2, \ldots, q_N\}$, vehicle count). Each traffic measure is regarded as one measure in the multi-source time series for describing the evolution of traffic state. Data groups are generated every 5 minutes containing the measurements of each traffic measure. Therefore, 288 data groups are generated every day. The data collected from April 20 to April 23, 2015, consisting of 1152 data groups are used in the numerical experiment. Figure 3 shows the data collection location.

To reduce the measurement noises, an arithmetic average filtering [32] is used to process the raw data. Figure 4a–c shows the original time series and the post-processing time series. Figure 4 shows that the evolution patterns of different traffic measures differ from each other. The proposed method leverages the evolution information on all three traffic measures to have a better short-term prediction of traffic flow, through the fusion of multiple time series.

The fundamental relation of traffic flow theory is:

$$q' = k\bar{v} \tag{24}$$

where q' represents average flow, k represents average density, and $\bar{v}$ represents space mean speed. Denote the vehicle length as l and detection range width as d. Then, the relationship between occupancy o and average density k can be expressed as:

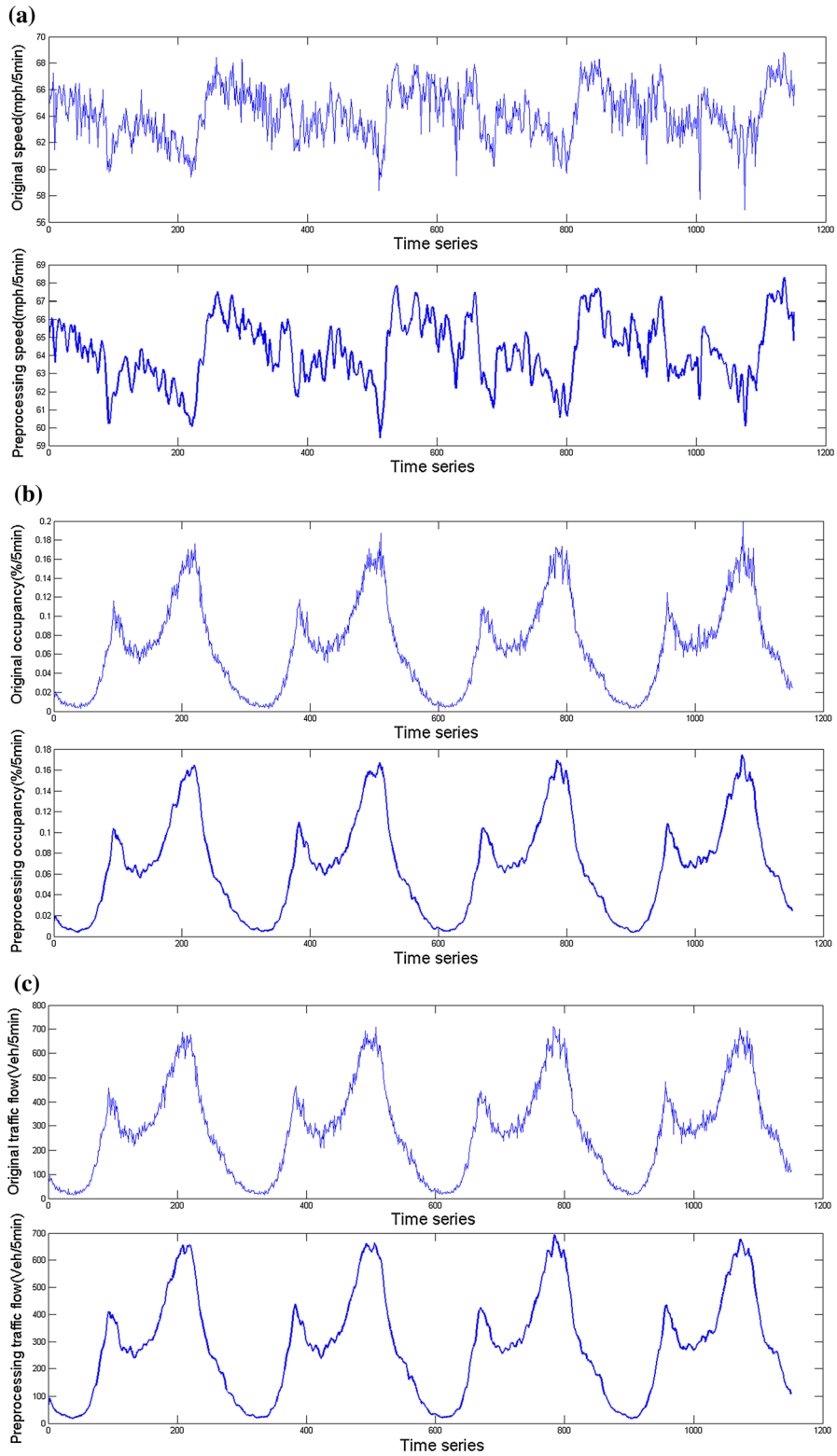


Fig. 4 Time series data: **a** average speed; **b** average occupancy; **c** average traffic flow

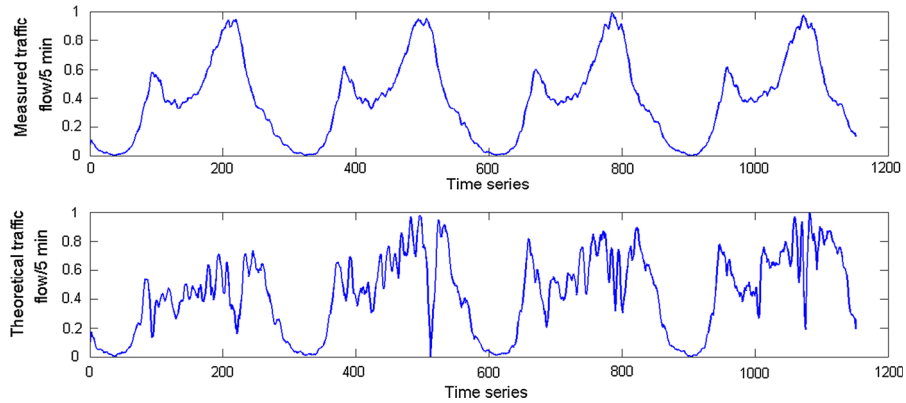


Fig. 5 Measured and theoretical traffic flow time series

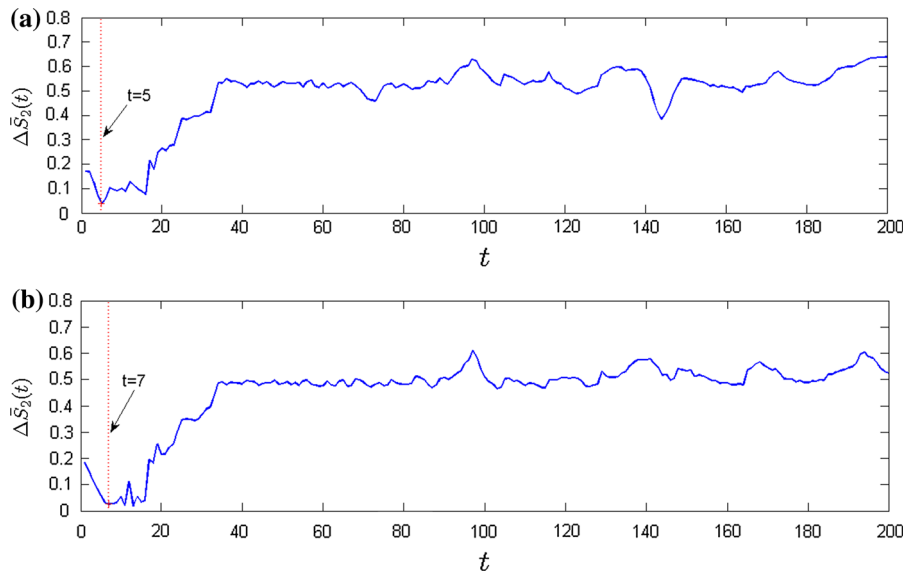


Fig. 6 Delay of time series using C-C algorithm: **a** the measured average traffic flow; **b** the theoretical traffic flow

$$o = \frac{\sum_i (l+d)/v_i}{T} = \frac{l+d}{T}\sum_i \frac{1}{v_i} \tag{25}$$
$$= (l+d)\frac{n}{T}\frac{1}{n}\sum_i \frac{1}{v_i} = (l+d)q'\frac{1}{\bar{v}} = (l+d)k$$

Substituting Eq. (24) into Eq. (25) yields:

$$q' = \frac{o}{l+d}\bar{v}. \tag{26}$$

In this experiment, $l+d$ is assumed to be constant. According to Eq. (26), the theoretical traffic flow q' can be computed in terms of the measured average speed $\bar{v}$ and occupancy o. Figure 5 shows the measured traffic flow q and the theoretical traffic flow q'. From Fig. 5, the overall shape of the measured traffic flow time series is similar to that of the theoretical traffic flow time series; however, the individual features of the measured and theoretical traffic flow also exist.

3.1 Selection of parameters in reconstructing phase space with multiple measures

First, the three measures are normalized to a range of [0–1] according to Eq. (2). Second, the theoretical traffic flow is computed based on the occupancy and speed measurements according to Eq. (26). The embedding delay τ of the measured traffic flow time series and the theoretical traffic flow time series are calculated using the C-C algorithm. The C-C algorithm defines the statistic variable based on the time series correlation integral. Following the C-C algorithm, the time series correlation integral $C(m, N, r, t)$ is defined as [34]

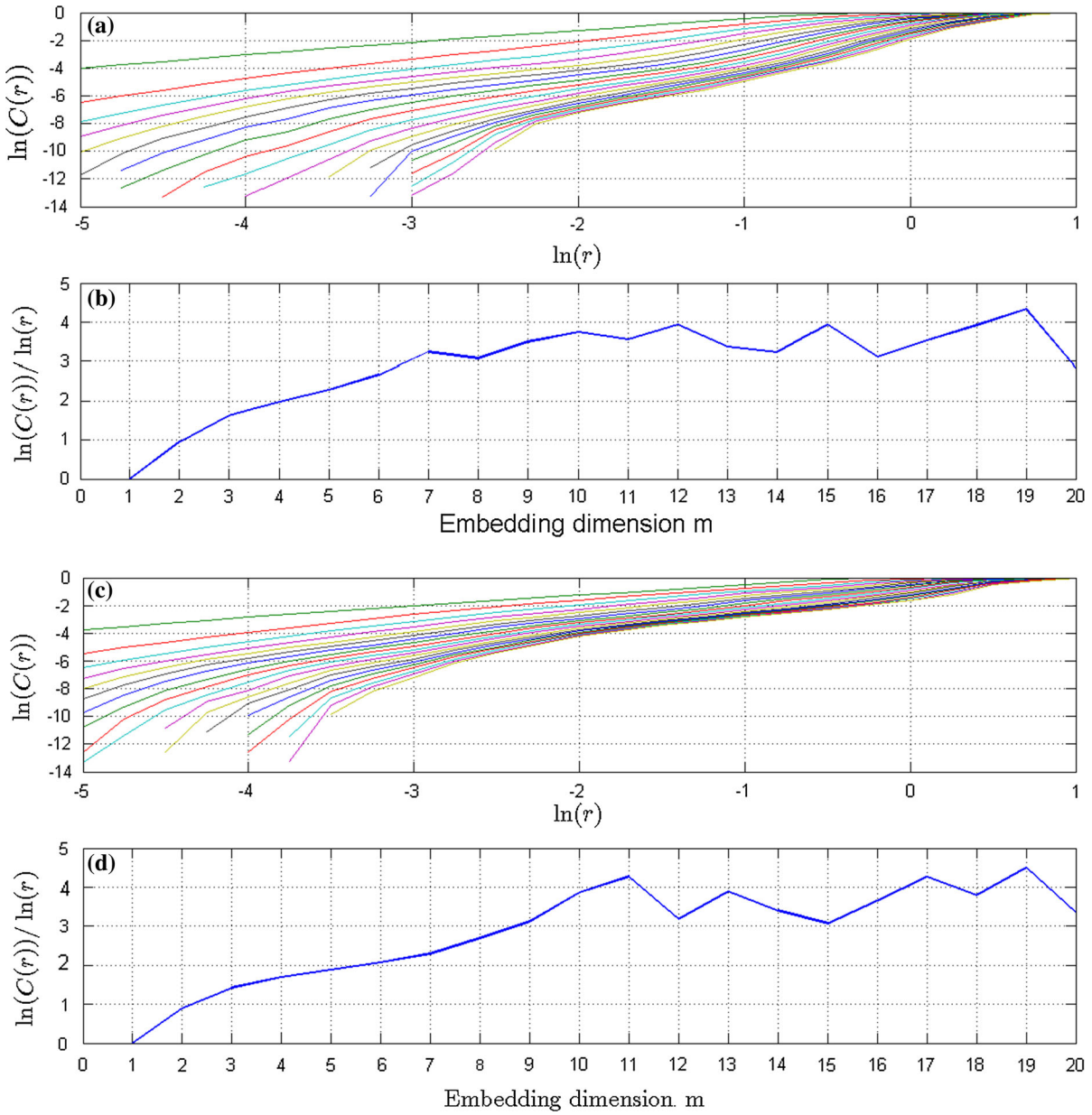


Fig. 7 Correlation dimension using G-P algorithm: **a** correlation dimension of average traffic flow; **b** correlation dimension of average traffic flow based on linear fit; **c** correlation dimension of the theoretical traffic flow; **d** correlation dimension of the theoretical value of traffic flow based on linear fit

$$C(m, N, r, t) = \frac{2}{M(M-1)} \sum_{1 \le i < j \le M} \theta(r - d_{ij}), \quad r > 0 \tag{27}$$

where $M = N - (m-1)t$, $d_{ij} = ||X_i - X_j||$, $\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \ge 0 \end{cases}$.

The statistic variable $S_2(m, N, r, t)$ is defined as, i.e., if $N \to \infty$

$$S_2(m, r, t) = \frac{1}{t} \sum_{s=1}^{t} \left[C_s(m, r, t) - C_s^m(1, r, t) \right] \tag{28}$$

Hence, the deviation of the statistic variable is defined as $\Delta \bar{S}_2(t)$ accordingly. Finally, according to the corre-

lation of $\Delta\bar{S}_2(t)$ and time t, the optimal delay τ can be calculated as it is the time that corresponds to the first minimum value calculated in terms of $\Delta\bar{S}_2(t)$. Figure 6 shows the delays associated with the measured and the theoretical traffic flow time series.

According to the modified C-C algorithm proposed in [42], the values of delay and embedded window can be estimated simultaneously through the correlation integral method. As shown in Fig. 6, the optimal delay is that which corresponds to the first minimum value calculated in terms of $\Delta\bar{S}_2(t)$.

In addition, the correlation dimension algorithm (G-P algorithm) is used to analyze the measured and theoretical traffic flow time series to obtain the correlation integral. For the G-P algorithm, the correlative dimension d will increase with the increase in the embedding dimension m; however, the increase rate is gradually reduced. When the embedding dimension m increases to a certain value, the correlative dimension d will tend to a saturation value. At this point, the minimum value of m corresponding to the saturation value d is the minimum embedding dimension of the time series [35]. Figure 7 shows the correlation integral $C(r)$ and $\ln C(r) - \ln r$ curve with respect to the arbitrary radius r. It indicates that when $\ln(r)$ varies in $[-5, 1]$, the embedding dimension m varies from 2 to 20. Among the families of $\ln C(r) - \ln r$ curves, an optimal fitting straight-line can be obtained. The slope of the optimal fitting straight-line is defined as correlation dimension D. As the dimension m increases, $D(m)$ (where $D(m) = \ln C(r)/\ln r$) gradually converges to a constant saturation value. Consequently, the embedded dimension of the parametric time series is defined as the dimension when $D(m)$ first reaches a stable value. Figure 7 also shows that the parametric time series have chaotic characteristics. In addition, according to chaos theory, the maximum Lyapunov exponent λ is calculated using the small-data method. Table 1 shows the embedded dimension m, delay τ, and the maximum Lyapunov exponents λ of average speed, average occupancy, and traffic flow.

3.2 Fusion and prediction of time series data of multiple measures

According to Eq. (4), we choose the values of phase space reconstruction parameters as $m = 10$ and $\tau = 5$ in Table 1. Then, we reconstruct the phase spaces of the average speed, average occupancy, and average traffic flow according to Eq. (5). Figure 8 shows the reconstruction diagrams and the local characteristics (marked by 1, 2, 3, 4, 5, and 6, respectively.) of the individual parametric time series. From Fig. 8, it can be noted that the local characteristics of average traffic flow and the average occupancy have some similarities, but they differ from the local characteristics of average speed. This indicates that none of the one-dimensional traffic measure time series can fully characterize the traffic system. To capture the traffic system characteristics from different data sources, this study focuses on multiple measures rather than any single measure. Based on Eq. (26), the theoretical traffic flow can be calculated in terms of the average speed and average occupancy. Then, the groups of the measured and theoretical traffic flow time series are embedded into a higher-dimensional space. Consequently, we can obtain an optimal fusion phase space based on the fusion of the phase points of the two groups of measures, according to Eq. (15), in the higher-dimensional space.

Table 1 Parametric values in phase space based on time series data

Measures	m	τ	λ
$v(t)$	10	6	0.0018
$o(t)$	10	5	0.0024
$q(t)$	10	5	0.00059
$q'(t)$	10	7	0.0021

As shown in Fig. 9, Fig. 9a is the phase space reconstruction diagram of the measured average traffic flow time series, and local characteristics 1 and 2 are marked on it. Figure 9b is the phase space reconstruction diagram of the theoretical traffic flow time series, and local characteristics 3, 4, and 5 are marked on it. The local characteristics 1 and 2 of the measured average traffic flow time series are very similar to the characteristics 3 and 4 of the theoretical traffic flow time series, respectively. However, local characteristic 5 has no analog in the measured traffic flow. Figure 9c is the fusion phase space reconstruction diagram of the multi-source time series and is similar to Fig. 9a in general. It reflects local characteristics 1 and 2 of the measured average traffic flow time series and local characteristics 3 and 4 of the theoretical traffic flow time series. In addition, the characteristic on the upper right corner of the dense part reflects local characteristic 5 of the theoretical traffic

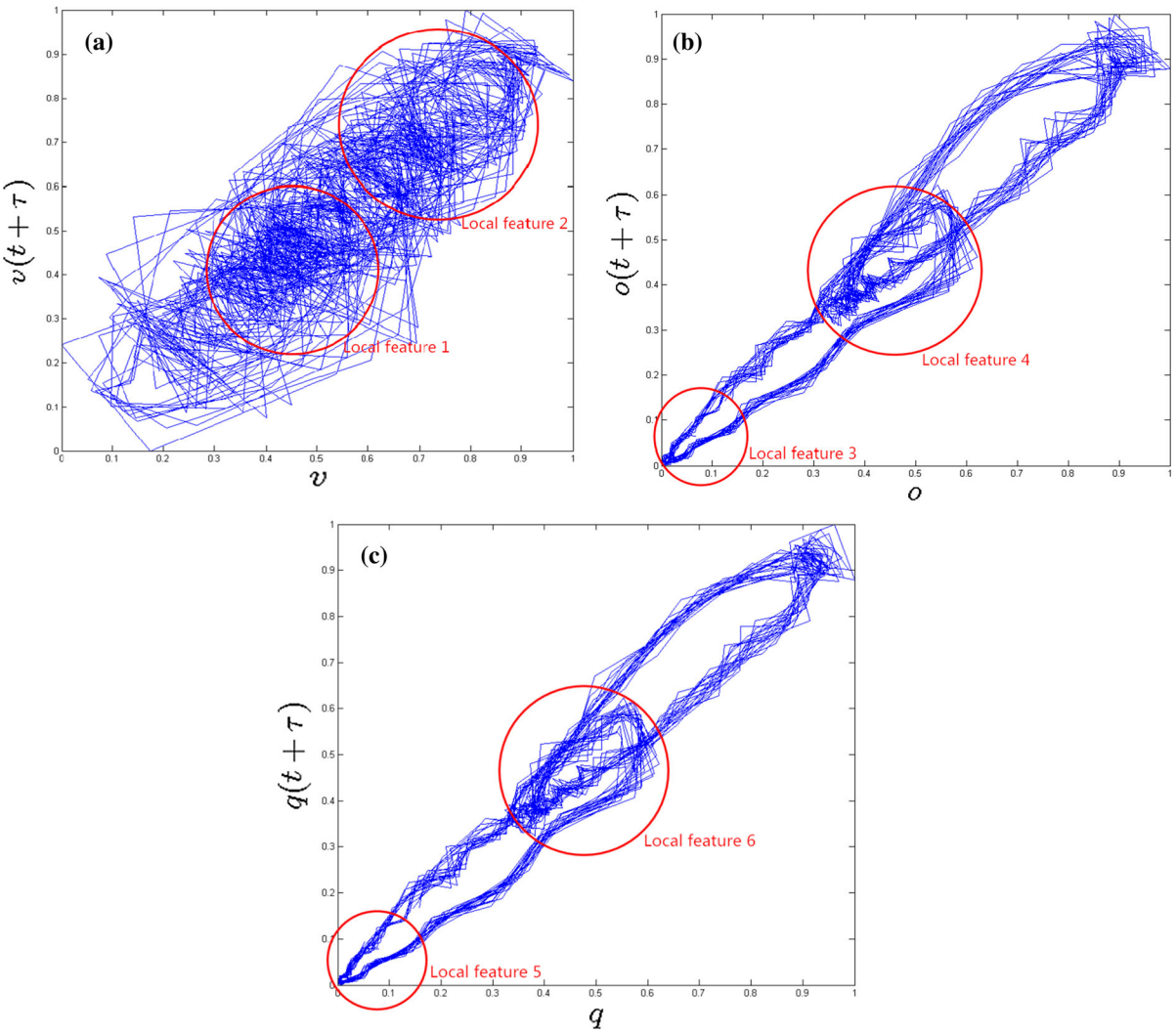


Fig. 8 Individual parametric time series phase space reconstruction diagrams: **a** the measured average speed; **b** the measured average occupancy; **c** the measured average traffic flow

flow as shown in Fig. 9c. Therefore, the multi-source fusion time series phase space contains all the main features of the measured and theoretical traffic flow time series. Based on the above analyses, the proposed method can reflect the characteristics of traffic system effectively. In addition, compared to an individual measure, the proposed method can elicit richer information. Consequently, it can reflect the main characteristics of traffic system accurately.

Based on the fused traffic flow, the RBF neural network shown in Fig. 2 is designed to predict traffic flow according to Eq. (19). The number of input layer neurons of the RBF neural network is set as the embedded dimension ($m = 10$) of the fusion phase space Z. The inputs are set as the fused traffic flow in Z. Then, we obtain the prediction values of the last 106 samples of the fused traffic flow time series. Figure 10 shows the predicted values of fused traffic flow and measured traffic flow, respectively.

To evaluate the prediction performance, errors associated with average speed, average occupancy, average traffic flow, and fused traffic flow using the proposed framework in terms of *MAE*, *MARE*, and *EC* are given in Table 2. It illustrates that prediction errors based on the fused traffic flow satisfy $MAE < 0.03$, $MARE < 0.04$, and $EC > 0.97$, respectively. It verifies the effec-

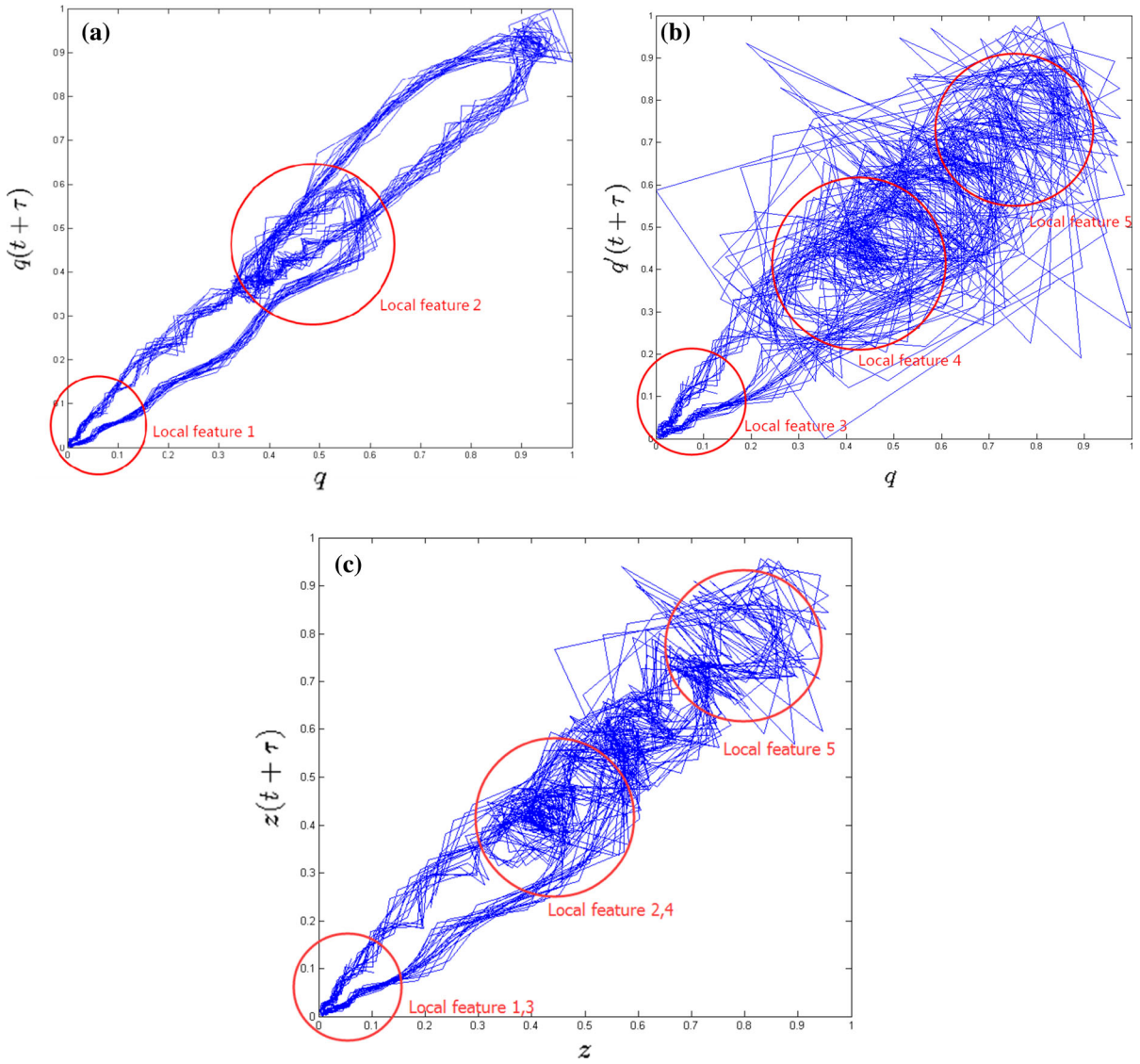


Fig. 9 Time series phase space reconstruction diagram: **a** the measured average traffic flow; **b** the theoretical traffic flow; **c** the fusion of the multi-source time series

tiveness of the proposed RBF-based prediction method for traffic flow. It also illustrates that the proposed framework can predict traffic flow with high reliability and relatively high accuracy in terms of *MAE*, *MARE*, and *EC*.

4 Conclusions

Considering the chaotic characteristics of traffic flow, this study proposes a multi-measure chaotic time series prediction framework to predict traffic flow. To accurately characterize traffic flows, time series of three traffic measures, i.e., traffic speed, occupancy, and traffic flow, are used to reconstruct the phase space using the phase space reconstruction theory. It enables capturing the characteristics of traffic flow in a higher-dimensional space. Then, the optimal fusion phase point of traffic measures is obtained using the Bayesian estimation theory in the phase space. Finally, the RBF neural network is designed to predict traffic flow.

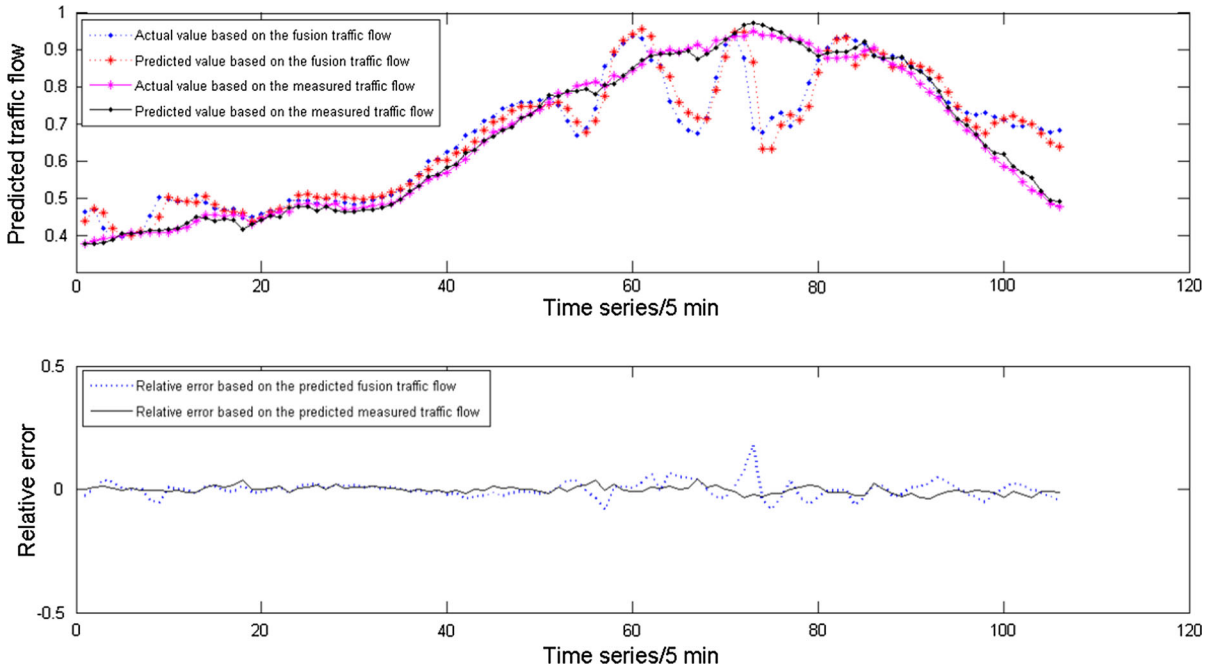


Fig. 10 Prediction of fused traffic flow and measured traffic flow using RBF neural network

Table 2 Prediction performance comparisons

Error index	$V(t)$	$O(t)$	$Q(t)$	$Z(t)$
MAE	0.0286	0.0120	0.0118	0.0235
MARE	0.0386	0.0192	0.0181	0.0348
EC	0.9721	0.9824	0.9891	0.9754

Compared to any single traffic measure, insights from numerical experiments illustrate that the proposed multi-measure-based framework can generate richer information and reflect the main characteristics of traffic flow accurately. Also, it demonstrates the effectiveness of the proposed framework in terms of accuracy and timeliness for short-term traffic flow prediction. Note that this study addresses the relatively uncongested traffic conditions. In the future, we propose to study traffic flow prediction under congestion situations using multiple measures.

Acknowledgments We acknowledge the support from various grant sources: the National Natural Science Foundation of China (Grant No. 61304197), the Scientific and Technological Talents of Chongqing (Grant No. cstc2014kjrc-qnrc30002), the Key Project of Application and Development of Chongqing (Grant No. cstc2014yykfB40001), Wenfeng Talents of Chongqing University of Posts and Telecommunications,"151" Science and Technology Major Project of Chongqing—General Design and Innovative Capability of Full Information Based Traffic Guidance and Control System (Grant No. cstc2013jcsf-zdzxqqX0003)—the Doctoral Start-up Funds of Chongqing University of Posts and Telecommunication (Grant No. A2012-26), and the US Department of Transportation through the NEXTRANS Center, the USDOT Region 5 University Transportation Center.

References

1. Li, Y., Sun, D.: Microscopic car-following model for the traffic flow: the state of the art. J. Control Theory Appl. **10**, 133–143 (2012)
2. Li, Y., Sun, D., Liu, W., Zhang, M., Zhao, M., Liao, X., Tang, L.: Modeling and simulation for microscopic traffic flow based on multiple headway, velocity and acceleration difference. Nonlinear Dyn. **66**, 15–28 (2011)
3. Tang, T., Wang, Y., Yang, X., Wu, Y.: A new car-following model accounting for varying road condition. Nonlinear Dyn. **70**, 1397–1405 (2012)
4. Li, Y., Zhu, H., Cen, M., Li, Y., Li, R., Sun, D.: On the stability analysis of microscopic traffic car-following model: a case study. Nonlinear Dyn. **74**, 335–343 (2013)
5. Tang, T., Shi, W., Shang, H., Wang, Y.: A new car-following model with consideration of inter-vehicle communication. Nonlinear Dyn. **76**, 2017–2023 (2014)

6. Tang, T., Wang, Y., Yang, X., Huang, H.: A multilane traffic flow model accounting for lane width, lane-changing and the number of lanes. Netw. Spat. Econ. **14**, 465–483 (2014)
7. Tang, T., Chen, L., Yang, S., Shang, H.: An extended car-following model with consideration of the electric vehicle's driving range. Phys. A **430**, 148–155 (2015)
8. Li, Y., Zhang, L., Peeta, S., Pan, H., Zheng, T., Li, Y., He, X.: Non-lane-discipline-based car-following model considering the effects of two-sided lateral gaps. Nonlinear Dyn. **80**, 227–238 (2015)
9. Li, Y., Zhang, L., Zheng, T., Li, Y.: Lattice hydrodynamic model based delay feedback control of vehicular traffic flow considering the effects of density change rate difference. Commun. Nonlinear Sci. Numer. Simul. **29**, 224–232 (2015)
10. Li, Y., Yang, B., Zheng, T., Li, Y., Cui, M., Peeta, S.: Extended-state-observer-based double loop integral sliding mode control of electronic throttle valve. IEEE Trans. Intell. Transp. Syst. **16**, 2501–2510 (2015)
11. Li, Y., Zhang, L., Zheng, H., He, X., Peeta, S., Zheng, T., Li, Y.: Evaluating the energy consumption of electric vehicles based on car-following model under non-lane discipline. Nonlinear Dyn. **82**, 629–641 (2015)
12. Hu, J., Zong, C., Song, J., Zhang, Z., Ren, J.: An applicable short-term traffic flow forecasting method based on chaotic theory. Proc. of IEEE 6th Int. Conf. Intell. Transp. Syst **12**(15), 608–613 (2003)
13. Smith, B.L., Williams, B.M., Oswald, R.K.: Comparison of parametric and nonparametric models for traffic flow forecasting. Transp. Res. Part C **10**(4), 303–321 (2002)
14. Vlahogianni, E.I., Karlaftis, M.G.: Comparing traffic flow time-series under fine and adverse weather conditions using recurrence-based complexity measures. Nonlinear Dyn. **69**(4), 1949–1963 (2012)
15. Wang, J., Shi, Q.: Short-term traffic speed forecasting hybrid model based on Chaos-Wavelet Analysis-Support Vector Machine theory. Transp. Res. Part C **27**, 219–232 (2013)
16. Ma, Q., Liu, W., Sun, D.: Multi-parameter fusion applied to road traffic condition forecasting. Acta Phys. Sin. **61**(16), 169501–169509 (2012)
17. Cong, R., Liu, S., Ma, R.: An approach to phase space reconstruction from multivariate data based on data fusion. Acta Phys. Sin. **57**(12), 7487–7493 (2008)
18. Cai, M., Cai, F., Shi, A., Zhou, B., Zhang, Y.: Chaotic time series prediction based on local-region multi-steps forecasting model. Lecture Notes in Comput. Sci. **3174**, 418–423 (2004)
19. Zhao, D., Ruan, J., Cai, Z.: Combination prediction method of chaotic time series. Chin. Sci. Bull. **52**(4), 570–573 (2007)
20. Lv, J., Zhang, S.: Application of adding weight one rank local region method in electric power system short term load forecast. Control Theory Appl. **19**(5), 767–770 (2002)
21. Kumara, K., Paridab, M., Katiyar, V.K.: Short term traffic flow prediction for a non-urban highway using artificial neural network. Procedia Soc. Behav. Sci. **104**, 755–764 (2013)
22. Karlaftis, M.G., Vlahogianni, E.I.: Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transp. Res. Part C **19**(3), 387–399 (2011)
23. Jia, Z., Jin, X., Yuan, Z.: Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections. Transp. Res. Part C **47**, 139–154 (2014)
24. Jayawardena, A.W., Fernando, D.A.K.: Use of radial basis function type artificial neural networks for runoff simulation. Comput. Aided Civil Infrastruct. Eng. **13**(2), 91–99 (1998)
25. Park, B., Carroll, J., Messer, T., Urbanik, I.I.: Short-term freeway traffic volume forecasting using radial basis function neural network. Transp. Res. Rec. **1651**, 39–47 (2007)
26. Chen, H., Grant-Muller, S.: Use of sequential learning for short-term traffic flow forecasting. Transp. Res. Part C **9**(5), 319–336 (2001)
27. Celikoglu, H.B.: Travel time measure specification by functional approximation: application of radial basis function neural networks. Procedia Soc. Behav. Sci. **20**, 613–620 (2011)
28. Zhang, Y., Qu, S., Wen, K.: A short-term traffic flow forecasting method based on chaos and RBF neural network. Syst. Eng. **25**(11), 26–30 (2007)
29. Takens, F.: Detecting strange attractors in turbulence. Lecture Notes in Math. **898**, 366–381 (1981)
30. Jin, S., Wang, D., Qi, H.: Bayesian network method of speed estimation from single-loop outputs. J. Transp. Syst. Eng. Inform. Technol. **10**(1), 54–58 (2010)
31. Wang, J., Deng, W., Guo, Y.: New Bayesian combination method for short-term traffic flow forecasting. Transp. Res. Part C **43**, 79–94 (2014)
32. Krese, B., Govekar, E.: Analysis of traffic dynamics on a ring road-based transportation network by means of 0–1 test for chaos and Lyapunov spectrum. Transp. Res. Part C **36**, 27–34 (2013)
33. Stathopoulos, A., Karlaftis, M.G.: A multivariate state-space approach for urban traffic flow modeling and prediction. Transp. Res. Part C **11**(2), 121–135 (2003)
34. Kim, H.S., Eykholtb, R., Salasc, J.D.: Nonlinear dynamics, delay times and embedding windows. Phys. D **127**, 48–60 (1999)
35. Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. Phys. D **9**(1–2), 189–208 (1983)
36. Ettweina, F., Růžičkab, M., Weber, M.: Existence of steady solutions for micropolar electrorheological fluid flows. Nonlinear Anal. Theor. **125**, 1–29 (2015)
37. Lan, L.W., Sheu, J.B., Huang, Y.S.: Investigation of temporal freeway traffic patterns in reconstructed state spaces. Transp. Res. Part C **16**(1), 116–136 (2008)
38. Yua, J., Goos, P., Vandebroek, M.: A comparison of different Bayesian design criteria for setting up stated preference studies. Transp. Res. Part B **46**(7), 789–807 (2012)
39. Kellert, S.H.: In the Wake of Chaos: Unpredictable Order in Dynamical Systems. University of Chicago Press, Chicago (1994)
40. Ma, T., Zhou, Z., Abdulhai, B.: Nonlinear multivariate time-space threshold vector error correction model for short term traffic state prediction. Transp. Res. Part B **76**, 27–47 (2015)
41. http://pems.dot.ca.gov/
42. Lu, Z., Cai, Z., Jiang, K.: Determination of embedding parameters for phase space reconstruction based on improved C-C method. J. Syst. Simul. **19**(11), 2527–2529 (2007)