Applied Mathematics and Computation 291 (2016) 266–278

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Forecasting traffic time series with multivariate predicting method

Yi Yin, Pengjian Shang*

Department of Mathematics, Beijing Jiaotong University, Beijing 100044, PR China

ARTICLE INFO

Keywords:
Multivariate predicting method
Univariate predicting method
K-nearest neighbor (KNN) nonparametric regression model
Forecast accuracy measure
Traffic time series

ABSTRACT

Scalar time series considered in most studies may be not sufficient to reconstruct the dynamics, while using multivariate time series may demonstrate great advantages over scalar time series if they are available. Multivariate time series are available in the traffic system and we intend to examine the issue for the real data in the traffic system. In this paper, we propose the multivariate predicting method and discuss the prediction performance of multivariate time series by comparison with univariate time series and K-nearest neighbor (KNN) nonparametric regression model. The three kinds of forecast accuracy measure for multivariate predicting method are smaller than those for the other two methods in all cases, which suggest the predicting results for traffic time series by multivariate predicting method are better and more accurate than those based on univariate time series and KNN model. It demonstrates that the proposed multivariate predicting method is more successful in predicting the traffic time series than univariate predicting method and KNN method. The multivariate predicting method has a broad application prospect on prediction because of its advantage on recovering the dynamics of nonlinear system.

1. Introduction

The study engaged in the traffic systems has developed rapidly recently. A great number of empirical studies have shown that the traffic system is a typically complex system in terms of system engineering [1,2]. Traffic flow often exhibits irregular and complex behavior [3–5]. The traffic states are closely related to daily life and get more and more concerns from people. Thus, there is a major requirement for providing dynamic traffic control to make and continuously update predictions of traffic variables and link times for several minutes into the future using real-time data [6]. Time series analysis such as classical time series analysis [7–9] and nonlinear time series analysis has drawn more attention and been applied successfully [10]. There have been successful attempts by many researchers when analyzing or forecasting traffic signals in recent decades from different aspects [11–23].

Multivariate time series data are common in experimental and industrial systems. But most of studies concern only scalar time series. In principle, scalar time series are generically sufficient to reconstruct the dynamics of the underlying systems according to the embedding theorem [24]. However in practice there may be large gains in using all of the measurements. Hence, for real data we cannot sure that any given scalar time series is sufficient to reconstruct the dynamics. On the other hand, using multivariate time series may demonstrate great advantages over scalar time series if they are available,

* Corresponding author.
E-mail address: pjshangdu@163.com (P. Shang).

http://dx.doi.org/10.1016/j.amc.2016.07.017
0096-3003/

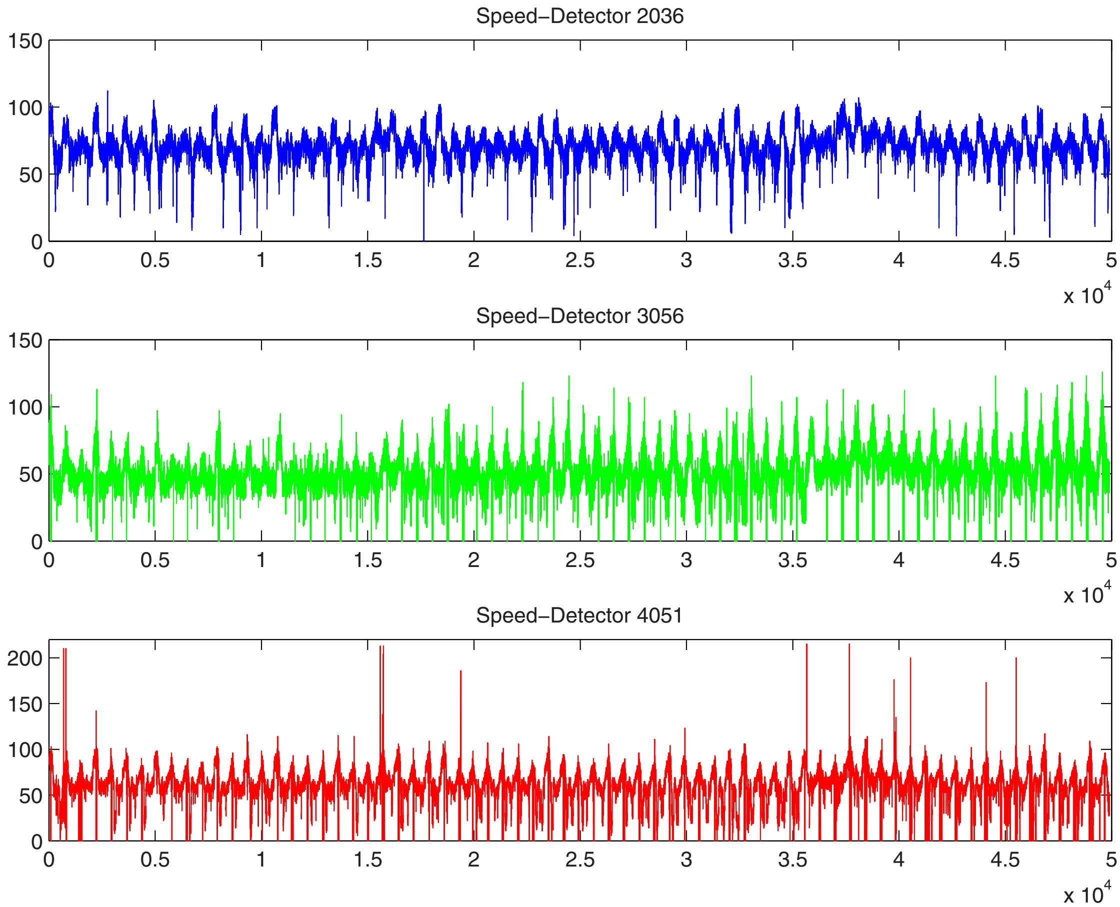

Fig. 1. The speed time series recorded by detectors 2036, 3056, 4051 including the first 49,900 data points.

especially if the system is noisy. Multivariate time series are available in the traffic system and are necessary for understanding the property of the traffic systems. Thus, forecasting traffic variables from multivariate time series arouses our great interest. In this paper, we propose the multivariate predicting method and discuss the prediction performance of multivariate time series by comparison with univariate time series and K-nearest neighbor (KNN) nonparametric regression model.

The structure of this paper is organized as follows. We briefly introduce the procedure of multivariate predicting method in Section 2. In Section 3, we present the traffic time series we used in this paper and show the predictions for univariate time series and multivariate time series which reveal the advantage of multivariate predicting method on recovering the dynamics of nonlinear traffic system. Finally, the conclusions are drawn.

2. Multivariate predicting method

We consider an M-dimensional time series $X_1, X_2, \ldots, X_N$ in a nonlinear system, where $X_i = (x_{1,i}, x_{2,i}, \ldots, x_{M,i})$, $i = 1, 2, \ldots, N$. When $M = 1$, the case changes into scalar time series. First, we make a time-delay reconstruction:

$$\begin{aligned} V_n = (&x_{1,n}, x_{1,n-\tau_1}, \ldots, x_{1,n-(d_1-1)\tau_1}, \\ &x_{2,n}, x_{2,n-\tau_2}, \ldots, x_{2,n-(d_2-1)\tau_2}, \\ &\quad \ldots \quad \ldots, \\ &x_{M,n}, x_{M,n-\tau_M}, \ldots, x_{M,n-(d_M-1)\tau_M}) \end{aligned} \tag{1}$$

where τ_i, d_i, $i = 1, \ldots, M$, denote the time-delays and the embedding dimensions, respectively. In general, following the embedding theorem [24,25] there is a function $F: \mathbb{R}^d \to \mathbb{R}^d (d = \sum_{i=1}^{M} d_i)$ such that

$$V_{n+1} = F(V_n) \tag{2}$$

The equivalent form of (2) can also be expressed as:

$$x_{1,n+1} = F_1(V_n)$$

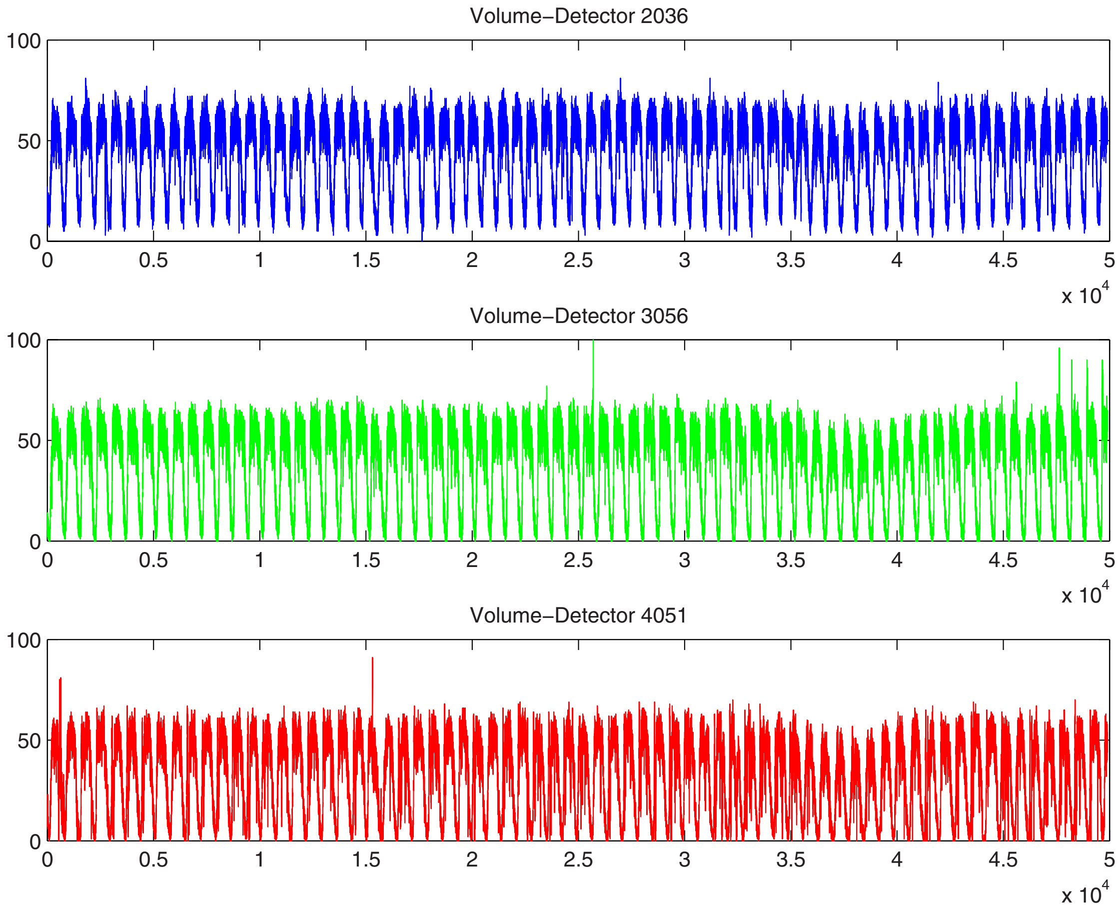

Fig. 2. The volume time series recorded by detectors 2036, 3056, 4051 including the first 49,900 data points.

$$x_{2,n+1} = F_2(V_n)$$
$$\vdots \quad \vdots$$
$$x_{M,n+1} = F_M(V_n) \tag{3}$$

Then we need to consider the selection of the time delays τ_i and embedding dimensions d_i, $i = 1, \ldots, M$ which makes Eq. (2) or Eq. (3) hold. It is known that various methods, such as mutual information [26] and autocorrelation [27] have been proposed to choose the time delay for a scalar time series. Thus, the time-delay τ_i for each scalar time series $x_{i,1}, x_{i,2}, \ldots, x_{i,N}$ can be calculated by mutual information in this paper. After choosing the good time delays for multivariate time series, we focus on the issue that how to choose embedding dimensions from multivariate time series data.

For the purpose of finding the optimal embedding dimensions for the multivariate time series, we consider each equation in (3) separately, because different F_i may need different minimum embedding dimensions. Thus, we consider the selection of embedding dimension for F_1 in (3), while the procedure of finding the embedding dimension for the other function $F_2, \ldots, F_M$ is the similar to that for F_1.

For any given set of dimensions $d_1, \ldots, d_M$, we obtain a series of delay vectors V_n defined in (1), where $n = max_{1 \le i \le M}(d_i - 1)\tau_i + 1, \ldots, N$. We find the nearest neighbor $V_{\eta(n)}$ for each V_n i.e.,

$$\eta(n) = argmin\|V_n - V_j\| : j = max_{1 \le i \le M}(d_i - 1)\tau_i + 1, \ldots, N, \ j \neq n \tag{4}$$

where $\|a\| = \|(a_1, a_2, \ldots, a_d)\| = (\sum_{i=1}^{d} a_i^2)^{1/2}$. Next, we calculate the mean multi-step prediction error which can be obtained from a simple nearest neighbor predictor. This very simple predictor is used to estimate the error because it is necessary that its error be small if the functions F_i are to exist and be continuous. The error is defined as

$$E(d_1, d_2, \ldots, d_M) = \frac{1}{N - J_0} \sum_{n=J_0}^{N-T} |x_{1,n+T} - x_{1,\eta(n)+T}|, J_0 = max_{1 \le i \le M}(d_i - 1)\tau_i + 1 \tag{5}$$

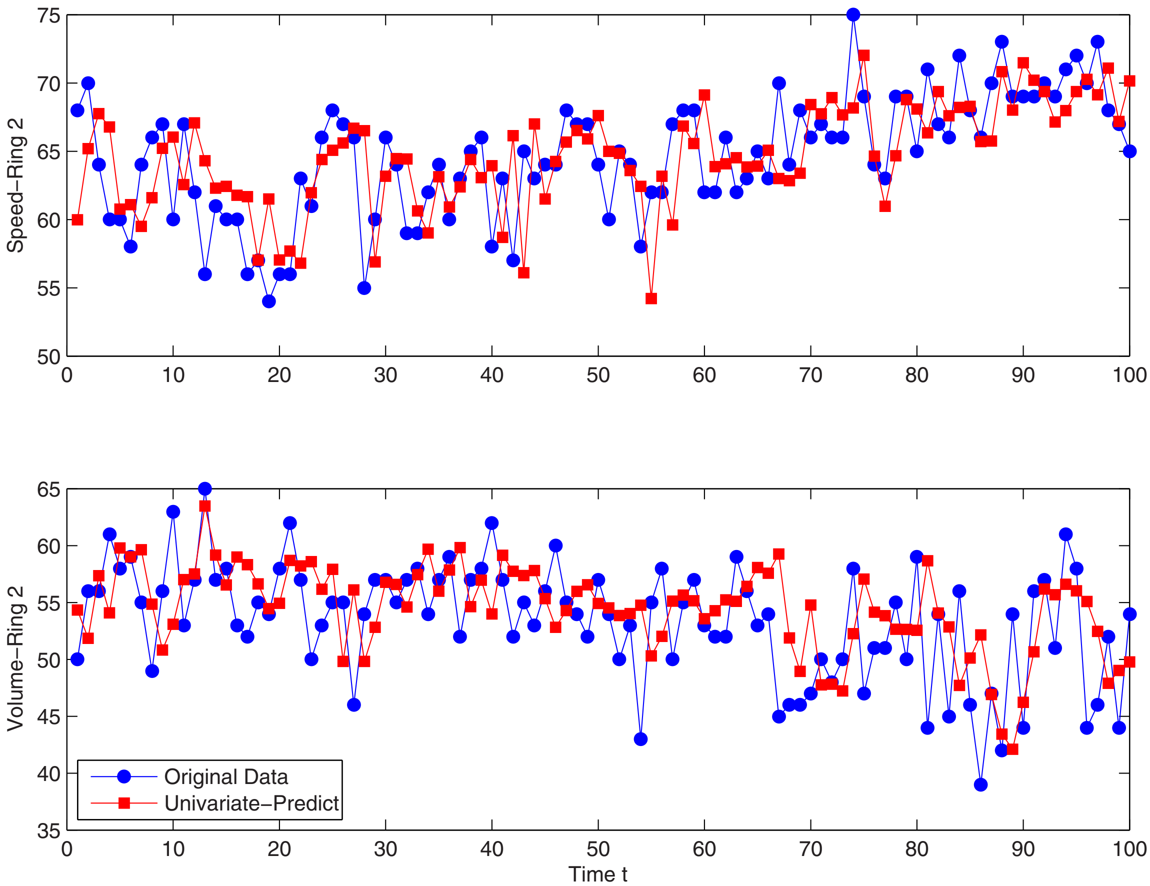

Fig. 3. Predicted values by univariate time series and actual values of speed (a) and volume (b) for detector 2036 in Ring 2. The predicted values are obtained by using the first 49,900 data points of univariate time series to fit the local linear predictor and testing 1-step ahead predictions on the last 100 data points.

where T is the number of prediction steps to be used. Eq. (5) is used in direct predictions (where we obtain a T-step predictor by fitting a model directly to the T-step-ahead data). This paper uses $T = 1$ to find the optimum embedding dimensions for one-step ahead prediction. The maximum error (for low-noise data only) or the root-mean-square error can also be considered. The error measure E depends on the dimensions $d_1, d_2, \ldots, d_M$. In this paper, we choose the embedding dimensions $d_1, d_2, \ldots, d_M$ by minimizing the error E, i.e.,

$$(d_1, d_2, \ldots, d_M) = argmin E(d_1, d_2, \ldots, d_M) : (d_1, d_2, \ldots, d_M) \in Z^M, \sum_{i=1}^{M} d_i \neq 0 \tag{6}$$

where $Z^M = \prod_{i=1}^{M} Z_i$, and Z_i denotes all non-negative integers. There are minimum values for $d_1, d_2, \ldots, d_M$ which give the multivariate time series a proper embedding. Before $d_1, d_2, \ldots, d_M$ reach the minimum values the error E decreases, while $d_1, d_2, \ldots, d_M$ are increased significantly beyond them, E increases which is caused by using the data in the far past which has lost correlation with the present to make a prediction. To speed up the computation of finding the minimum embedding dimensions, we usually set some maximum dimension D_{max} in advance instead of performing an exhaustive search in (6), which means that

$$(d_1, d_2, \ldots, d_M) = argmin E(d_1, d_2, \ldots, d_M) : 0 \leq d_i \leq D_{max} \forall i, \sum_{i=1}^{M} d_i \neq 0 \tag{7}$$

Besides, when $M = 1$, this approach of selection the minimum embedding dimension is applicable to scalar time series.

After having already chosen the time delays τ_i and embedding dimensions $d_i, i = 1, \ldots, M$, we decide to use local linear prediction for final prediction. The local linear method has been used in scalar chaotic time series predictions extensively [28]. First, we make a time-delay reconstruction based on the chosen time delays τ_i and embedding dimensions d_i and obtain the delay vectors V_n defined in (1). Because we adopt one-step ahead prediction in this paper, a functional relationship

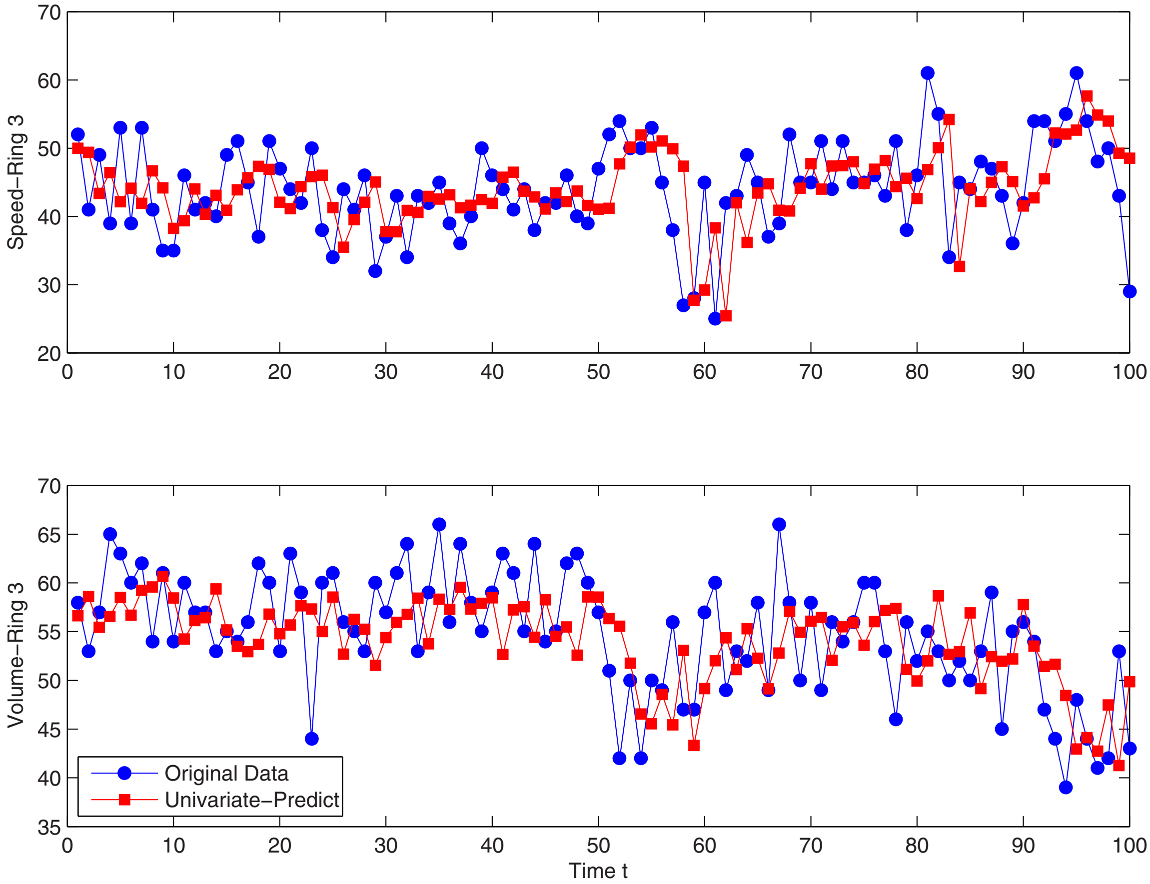

Fig. 4. Predicted values by univariate time series and actual values of speed (a) and volume (b) for detector 3056 in Ring 3. The predicted values are obtained by using the first 49,900 data points of univariate time series to fit the local linear predictor and testing 1-step ahead predictions on the last 100 data points.

between the current state vector V_n and the future state vector V_{n+1} has been assumed in Eqs. (2) and (3). The next step is to find a predictor f_i which approximates F_i in (3).

Without loss of generality, we consider the problem of finding the predictor f_1 for F_1 in (3). We utilize the Euclidean norm, denoted by $\|\|$, to find the k nearest neighbors of V_n, i.e., the k state vectors V_j with $j < n$ that minimize $\|V_n - V_j\|$. We then construct a local predictor, regarding each neighbor V_j as a vector in the domain and $x_{1,j+1}$ as the corresponding point in the range. The simplest approach to construct a local predictor is approximation by nearest neighbor, or zeroth-order approximation, i.e., $k = 1$ and $x_{1,n+1pred} = x_{1,j+1}$. A superior approach is the first order, or linear, approximation, with our taking k greater than d_i, and fitting a linear polynomial to the pairs $(V_j, x_{1,j+1})$. Likewise, the predictor f_i for F_i can also be obtained by fitting a linear polynomial to the pairs $(V_j, x_{i,j+1})$. The predicted values $x_{1,n+1}, \ldots, x_{M,n+1}$ can be calculated by the current vector V_n and the predictor f_i.

3. Results and analysis

3.1. Traffic data description

There are numerous parameters in traffic systems that can be measured, among which, traffic speed, volume, and occupancy are mostly studied. In this study, we use the data observed from detectors 2036, 3056, 4051 which lie on the 2nd, 3rd and 4th Ring Road (Beijing, China) respectively over a period of about 11 weeks, from August 11 to October 26, 2012. The reason why we choose the data from these detectors is that the traffic data from these important areas of the economy, culture, and entertainment of Beijing include the basic traffic situation in Beijing and the traffic states in these areas are closely related to daily life and get more concerns from the people in Beijing. In this paper, we investigate the data including traffic speed time series and traffic volume time series by detectors 2036, 3056, 4051. The raw data are collected every 20 s at each detector location, which are located approximately every half a kilometer. Raw data are screened for errors and then aggregated into two-minute data for average. The one-hour recorded traffic time series are about 30 data points, and

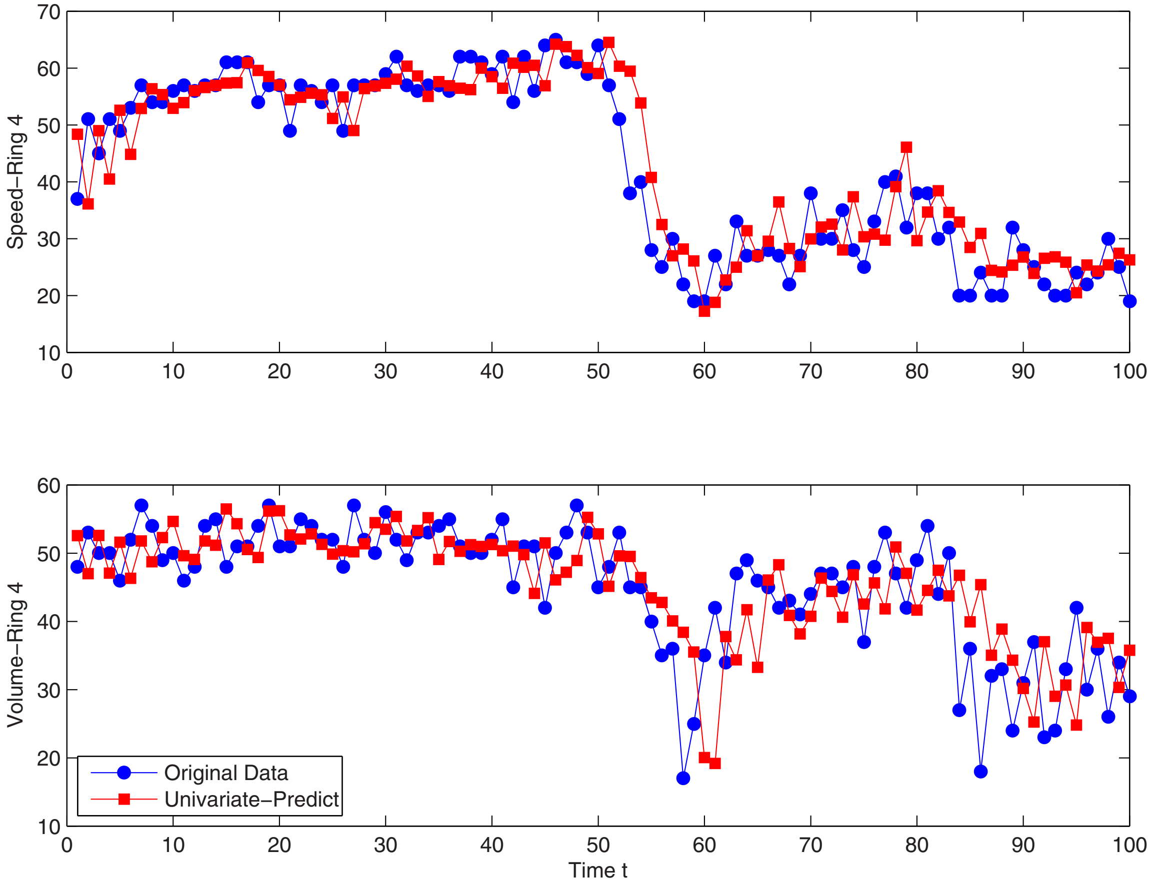

Fig. 5. Predicted values by univariate time series and actual values of speed (a) and volume (b) for detector 4051 in Ring 4. The predicted values are obtained by using the first 49,900 data points of univariate time series to fit the local linear predictor and testing 1-step ahead predictions on the last 100 data points.

the total number of data points for the whole series is about 55,440. For the predictions we use the first 49 900 data points, and then test 1-step ahead predictions on the remaining 100 data points. These speed and volume time series including the first 49,900 data points are presented in Figs. 1 and 2.

3.2. Predicting results from univariate time series

First, we predict the speed and volume time series from univariate time series. We pick a one-dimensional time series, either $y_{1,1}, y_{1,2}, \ldots, y_{1,50000}$ or $y_{2,1}, y_{2,2}, \ldots, y_{2,50000}$, and predict either y_1 or y_2 purely from this one-dimensional time series. To get better predicting results from univariate time series, we adopt one step ahead predictions. We calculate the embedding dimension using the method mentioned in Section 2 taking $M = 1$ because we have a one-dimensional time series. It can be found that the minimum embedding dimension for each of the time series is 3 and the best choice of the time delay τ is 1. Then we use local linear method to test predictions of these two scalar time series. We use the first 49,900 data points to fit the local linear predictor and test 1-step ahead predictions on the last 100 data points for the speed and volume time series by detectors 2036, 3056, 4051. The prediction results are shown in Figs. 3–5. Many criteria are used to evaluate the forecasting performance of different methods in empirical studies. In this paper we employ NMSE (normalized root mean squared error), MASE (mean absolute scaled error) and MAPE (mean absolute percentage) as the forecast accuracy measures. Let denotes the observation of the time series $X(t)$ at time t and $\hat{x}_t$ denotes the forecast of x_t. We define the forecast error $e_t = x_t - \hat{x}_t$ and the percentage error $p_t = 100e_t/x_t$. The scaled error is $q_t = e_t/\frac{1}{n-1}\sum_{i=2}^{n}|x_i - x_{i-1}|$. The NMSE, MASE and MAPE are respectively calculated as:

$$NMSE = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n}(x_t - mean(x))^2}} \tag{8}$$

$$MASE = mean(|q_t|) \tag{9}$$

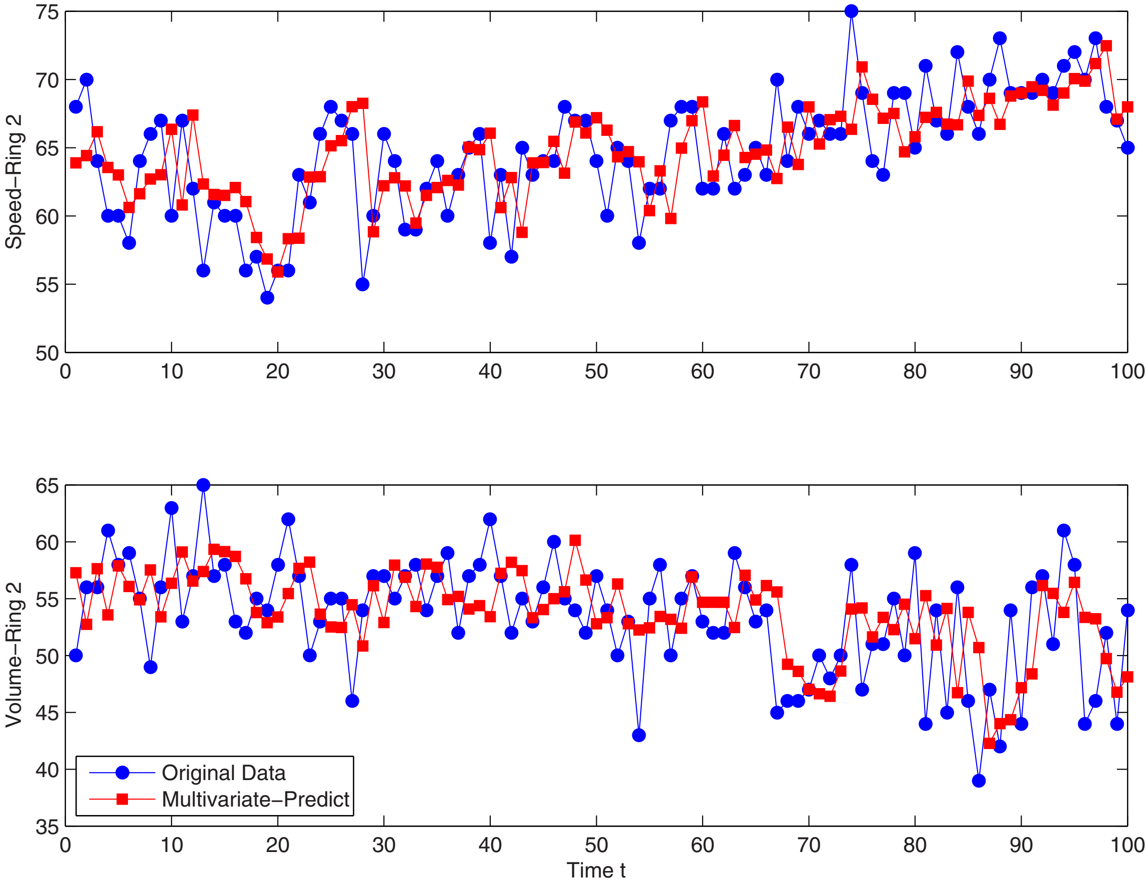

Fig. 6. Predicted values by multivariate time series and actual values of speed (a) and volume (b) for detector 2036 in Ring 2. The predicted values are obtained by using the first 49,900 data points of each series in the two-dimension time series together to fit the local linear predictor and test 1-step ahead predictions on the last 100 data points.

Table 1
Comparison of RMSE, MASE and MAPE values for speed and volume time series by detector 2036.

Detector 2036	$NMSE_{speed}$	$NMSE_{volume}$	$MASE_{speed}$	$MASE_{volume}$	$MAPE_{speed}$	$MAPE_{volume}$
Univariate	0.8837	1.0452	0.9686	0.8395	4.8295	8.0977
Multivariate	0.8215	0.9815	0.8817	0.8109	4.3778	7.7320
KNN	0.8316	1.0322	0.8963	0.8815	4.4603	8.3285

Table 2
Comparison of RMSE, MASE and MAPE values for speed and volume time series by detector 3056.

Detector 3056	$NMSE_{speed}$	$NMSE_{volume}$	$MASE_{speed}$	$MASE_{volume}$	$MAPE_{speed}$	$MAPE_{volume}$
Univariate	1.0699	0.9094	0.8883	0.8189	13.8964	8.3996
Multivariate	1.0387	0.8971	0.8582	0.8058	13.4775	8.2523
KNN	1.0972	0.9218	0.9396	0.8451	14.6694	8.6384

$$MAPE = mean(|p_t|) \tag{10}$$

The NMSE, MASE and MAPE of predictions from univariate time series for detectors 2036, 3056 and 4051 are shown in the Table 1, 2 and 3, respectively.

3.3. Predicting results from multivariate time series

Next we apply the multivariate predicting method described in Section 2 to forecast the traffic time series. We consider a two-dimensional time series: speed and volume, and use the method of Section 2 to calculate the embedding dimensions

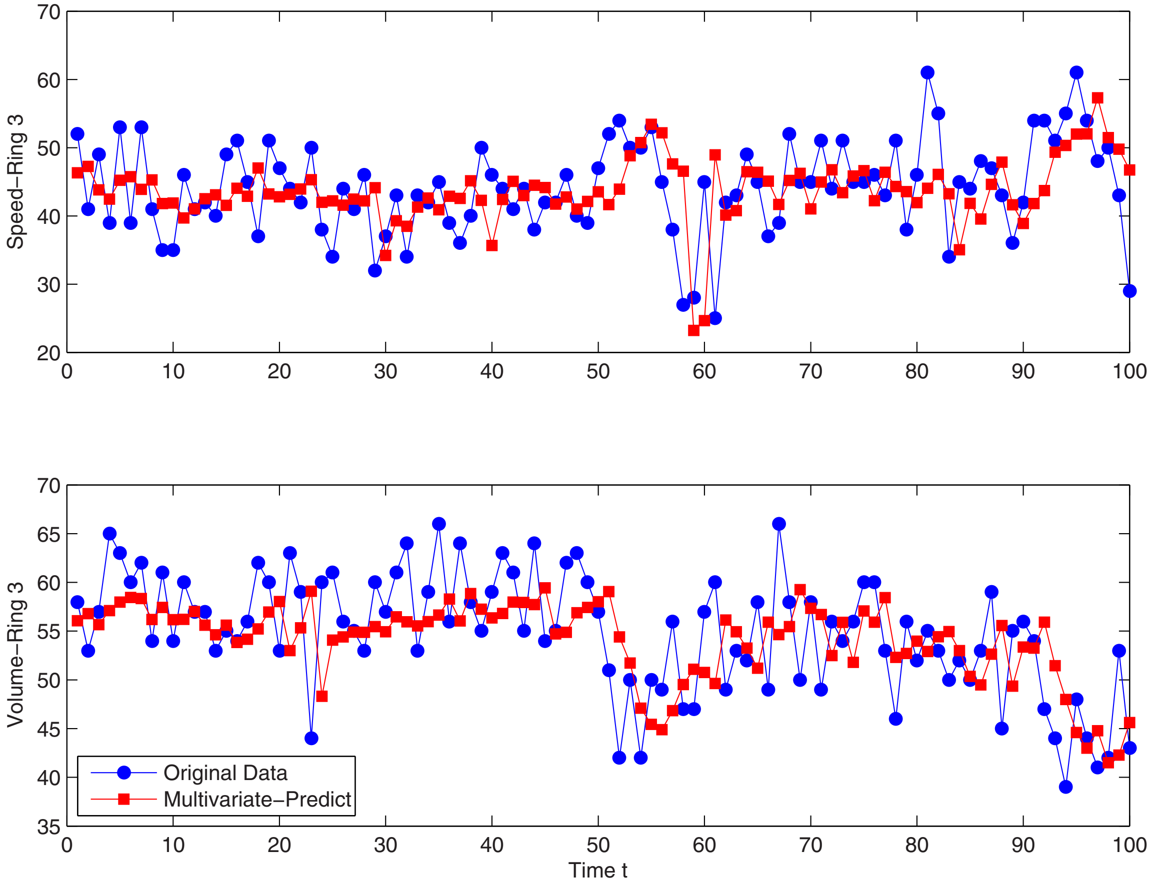

Fig. 7. Predicted values by multivariate time series and actual values of speed (a) and volume (b) for detector 3056 in Ring 3. The predicted values are obtained by using the first 49,900 data points of each series in the two-dimension time series together to fit the local linear predictor and test 1-step ahead predictions on the last 100 data points.

Table 3
Comparison of RMSE, MASE and MAPE values for speed and volume time series by detector 4051.

Detector 4051	$NMSE_{speed}$	$NMSE_{volume}$	$MASE_{speed}$	$MASE_{volume}$	$MAPE_{speed}$	$MAPE_{volume}$
Univariate	0.3974	0.7843	1.1568	1.0486	14.0538	15.2376
Multivariate	0.3525	0.6689	1.0446	0.8688	12.3912	12.7788
KNN	0.3531	0.6772	1.0331	0.8816	12.5174	12.8586

$d_{1_1}, d_{2_1}, d_{1_2}, d_{2_2}$ needed to build the following predictive models:

$$\begin{aligned} y_{1,n+1} &= F_1(y_{1,n}, y_{1,n-1}, \ldots, y_{1,n-(d_{1_1}-1)}, y_{2,n}, y_{2,n-1}, \ldots, y_{2,n-(d_{2_1}-1)}) \\ y_{2,n+1} &= F_2(y_{1,n}, y_{1,n-1}, \ldots, y_{1,n-(d_{1_2}-1)}, y_{2,n}, y_{2,n-1}, \ldots, y_{2,n-(d_{2_2}-1)}) \end{aligned} \tag{11}$$

Minimizing E of Eq. (5), we find the embedding dimensions are $d_{1_1} = d_{2_1} = d_{1_2} = d_{2_2} = 2$. Now we fit a local linear model to the functions F_1, and F_2, respectively, using the first 49 900 data points of each series in the two-dimension time series together, and then test 1-step ahead predictions on the remaining 100 data points. The prediction results are shown in Figs. 6–8. The forecasting performance of multivariate predicting method is presented in Tables 1–3 as well.

For purpose of revealing the performance of multivariate predicting method better, we take the K-nearest neighbor (KNN) nonparametric regression model as a comparison as well. KNN model is extension of nearest neighbor classification [29]. By analyzing the traffic speed data Shang et al. found that chaotic characteristics obviously exist in the traffic system [30]. KNN has been widely applied to forecasting [31–33]. In the KNN model, we first consider a discrete time series $X = x_1, x_2, \ldots, x_n$ where n is the total number of points of the series and x_n is the current state. First, we need to find the nearest value group of near values, also called the nearest neighbors, of the current state x_n in the past data. Then we predict x_{n+1} on basis of those nearest values. For example, if the size of the neighborhood was $k = 1$, and the nearest value was x_j, then we could predict x_{n+1} on basis of x_{j+1}. The definition of the current state of the time series can be extended to include several

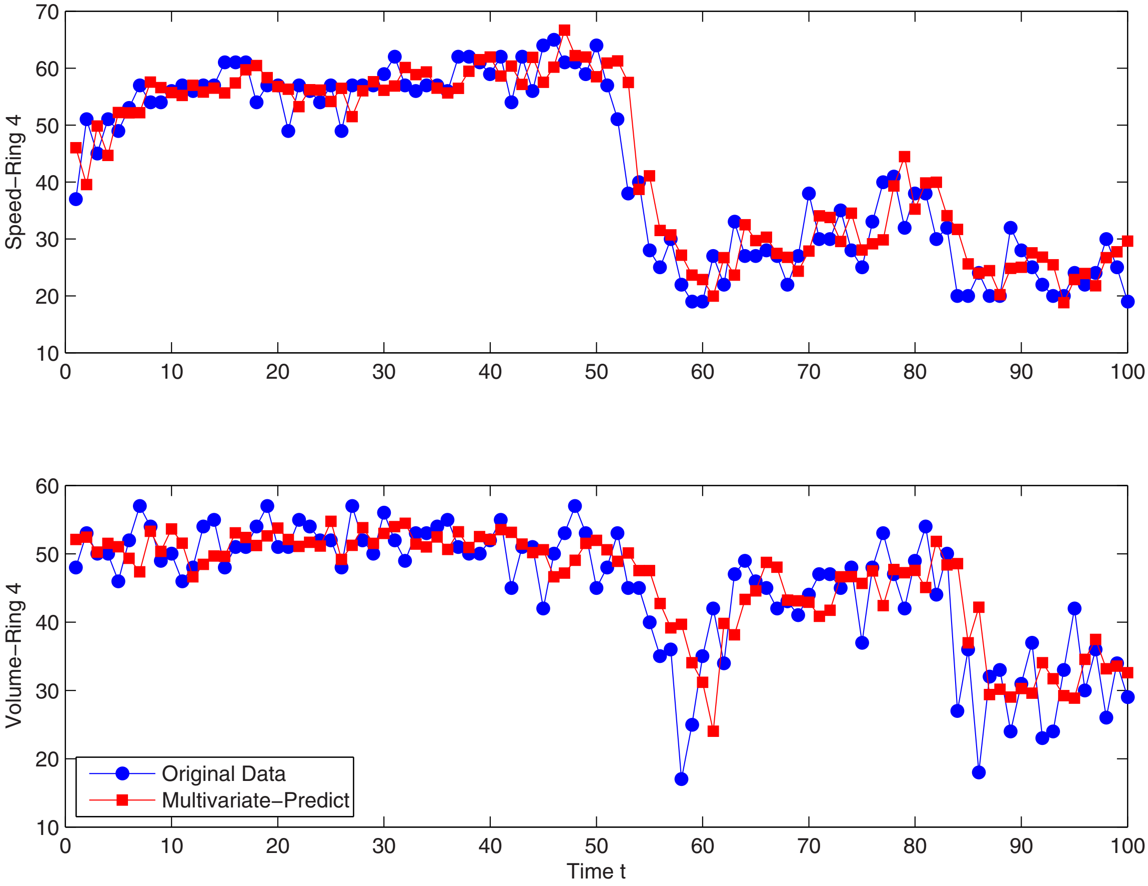

Fig. 8. Predicted values by multivariate time series and actual values of speed (a) and volume (b) for detector 4051 in Ring 4. The predicted values are obtained by using the first 49,900 data points of each series in the two-dimension time series together to fit the local linear predictor and test 1-step ahead predictions on the last 100 data points.

consecutive values $x_{n-l}, x_{n-l+1}, \ldots, x_{n-1}, x_n$ where l is a pattern size such that $1 \le l \le n-1$. Meanwhile, a difference vector $Q = (q_{n-l}, q_{n-l+1}, \ldots, q_{n-1})$ is also defined where $q_i = x - i + 1 - x_i$, $n - l \le i \le n - 1$. The size of pattern should be optimized to obtain the best results because of its important impact on minimizing the error. The procedure of KNN can be shown briefly as follows:

Step 1: Start with a minimal neighborhood size k.

Step 2: Start with a minimal pattern size l.

Step 3: Form the pattern of size l describing the current state, i.e., $(x_{n-l}, x_{n-l+1}, \ldots, x_{n-1}, x_n)$.

Step 4: Search the time series $(x_1, \ldots, x_{n-l-1})$ to find the nearest matches for each nearest match corresponds to an index j, for which the matching pattern is $x_{j-l}, .., x_{j-1}$. The difference vector associated with $x_{j-l}, .., x_{j-1}$ is $Q_j = (q_{j-l}, .., q_{j-1})$, and the final difference associated with match number h is Q_j^h.

Step 5: Estimate the value $\hat{x}_{n+1}$ on the basis of the final differences for all of the nearest neighbors:

$$\hat{x}_{n+1} = x_n + Q_m \tag{12}$$

where $Q_m = \sum_{h=1}^{k} \frac{Q_j^h}{k}$

Step 6: Calculate the root mean squared error (RMSE) between the actual and predicted values, in order to choose neighborhood size k and pattern size l for the entire estimation set:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[x(i) - \hat{x}(i)]^2} \tag{13}$$

Step 7: Repeat Steps 3–6 for pattern sizes $l+1, l+2, \ldots, l_{max}$.

Step 8: Repeat Steps 2–7 for neighborhood sizes $k+1, k+2, \ldots, k_{max}$.

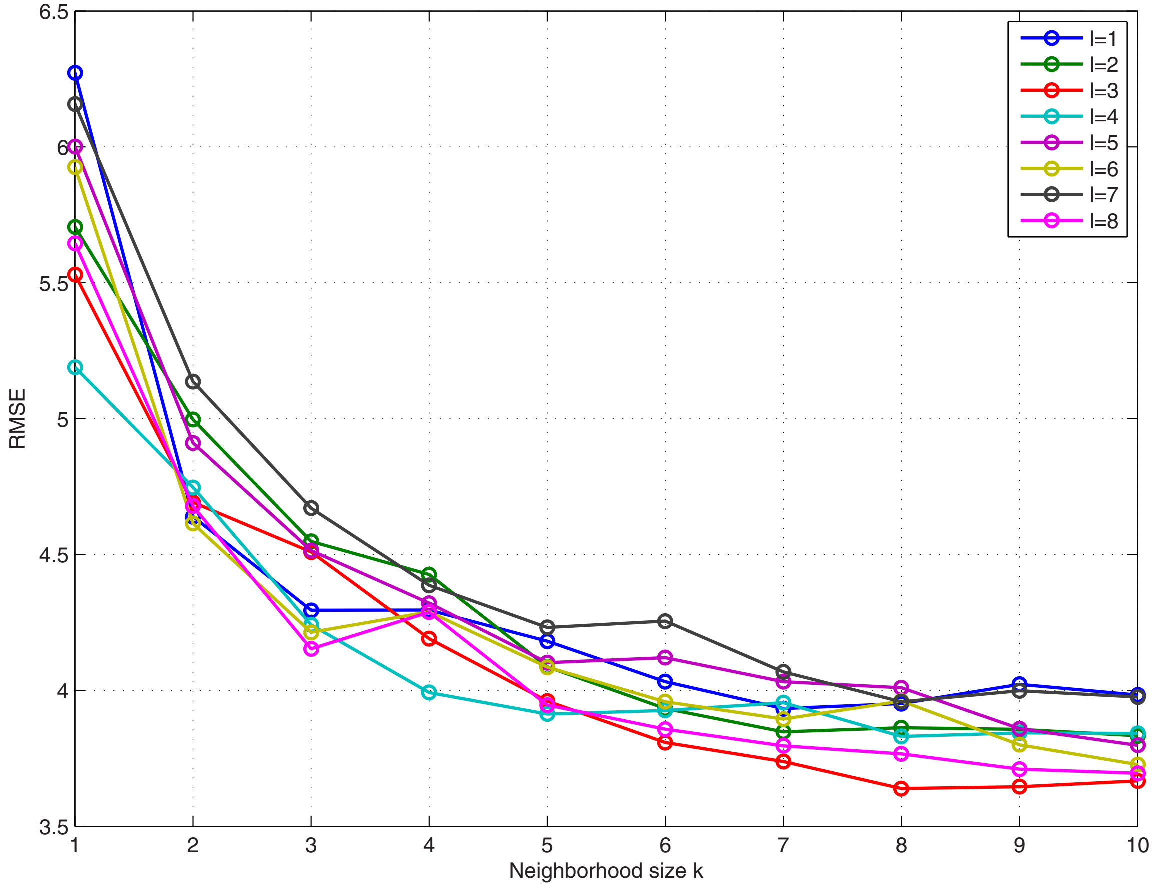

Fig. 9. Choosing the optimal neighborhood size k for the speed time series of detector 2036. It can be found that the optimal neighborhood size k is 3.

Step 9: Choose the optimal pattern recognition model which yields minimal RMSE by optimizing the neighborhood size k and pattern size l.

Thus we use the KNN to predict the speed and volume time series. Figs. 9 and 10 show RMSE performance of different neighborhood size k and pattern size l for the speed time series of detector 2036 as an example. The RMSE performance of different neighborhood size k and pattern size l for the other series are also obtained but not given in this paper for the sake of brevity. It can be concluded from these figures that the optimal neighborhood size and pattern size for each time series is $k = 3$ and $l = 8$. The predicting results by KNN are displayed in Figs. 11–13, while the forecasting performance of multivariate predicting method is given in Tables 1–3.

It can be found from Table 1 that the NMSE, MASE and MAPE values obtained by multivariate predicting method for speed and volume of detector 2036 are smaller than the other two methods, while three kinds of forecast accuracy measure by KNN are smaller than those by the univariate predicting method except for the MASE of volume. Table 1 also presents that the predicting speed shows the smaller NMSE and MAPE values for all the three methods than the predicting volume. As for the speed and volume of detector 3056, the forecasting measures for both speed and volume by KNN are smaller than those by the univariate predicting method, while the forecasting measures by multivariate predicting method are smaller than those by KNN in Table 2. Meanwhile, the predicting volume presents the smaller NMSE and MAPE values for all the three methods than the predicting speed. In terms of the speed and volume of detector 4051, the predicting performance of KNN is better than the univariate predicting method but worse than the multivariate predicting method seen from Table 3. The predicting speed of detector 4051 has the smaller NMSE and MAPE values than the predicting volume for these three methods. As a result, compared with the NMSE, MASE and MAPE values obtained by univariate predicting method, multivariate predicting method and KNN, the NMSE, MASE, MAPE of the predicting speed and volume by multivariate predicting method are smaller than those of the predicting speed and volume by univariate predicting method and KNN model in all cases which demonstrates the proposed multivariate predicting method is more successful in predicting the traffic time series than univariate predicting method and KNN method, while the univariate predicting method and KNN model may have different forecasting performance for different time series. Hence, the multivariate predicting method has a broad application prospect on prediction because of its advantage on recovering the dynamics of nonlinear system.

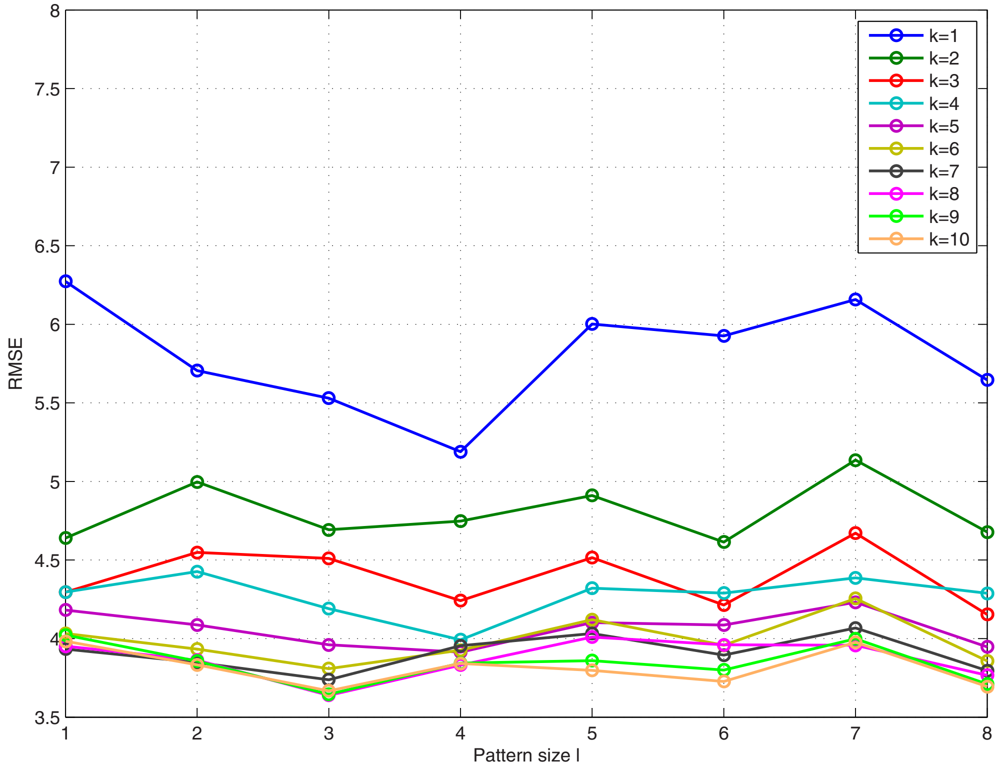

Fig. 10. Choosing the optimal pattern size l for the speed time series of detector 2036. It can be found that the optimal pattern size l is 8.

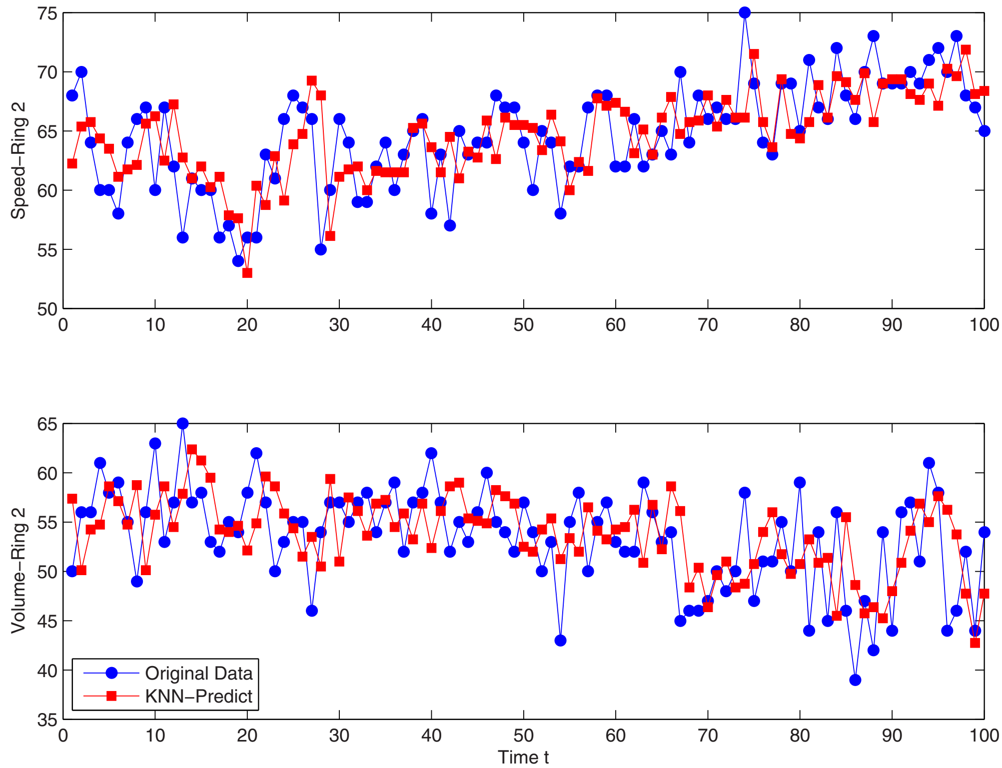

Fig. 11. Predicted values by KNN based on $k = 3$ and $l = 8$ and actual values of speed (a) and volume (b) for detector 2036 in Ring 2.

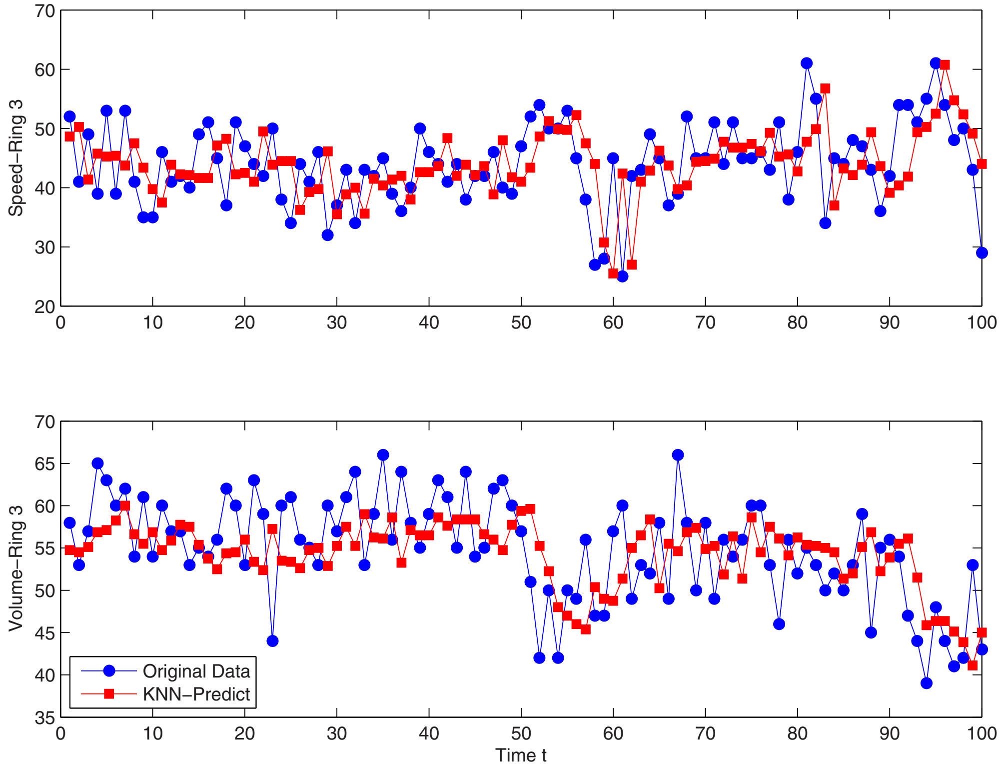

Fig. 12. Predicted values by KNN based on $k = 3$ and $l = 8$ and actual values of speed (a) and volume (b) for detector 3056 in Ring 3.

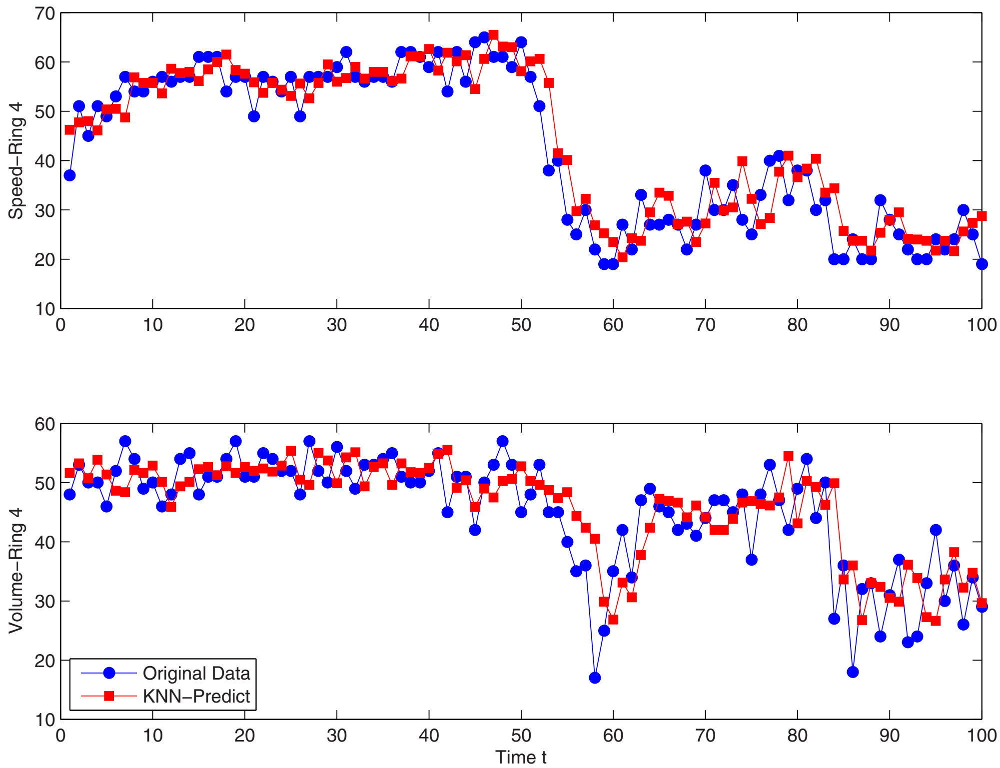

Fig. 13. Predicted values by KNN based on $k = 3$ and $l = 8$ and actual values of speed (a) and volume (b) for detector 4051 in Ring 4.

4. Conclusions

In this paper, a multivariate predicting method is proposed to investigate the issue of how to use multiple data streams for prediction effectively. Multivariate time series are available in the traffic system and are necessary for understanding the property of the traffic systems. Thus, we apply the multivariate predicting method to forecast traffic variables and discuss the prediction performance of multivariate time series by comparison with univariate time series and KNN mode. By compared the predicting results of these methods with the actual data and calculating and comparing the forecast accuracy measure: NMSE, MASE and MAPE, we find that The three kinds of forecast accuracy measure for multivariate predicting method are smaller than those for the other two methods in all cases, which suggests the predicting results for traffic time series by multivariate predicting method are better and more accurate than those based on univariate time series and KNN model. It can be concluded that the proposed multivariate predicting method is more successful in predicting the traffic time series than univariate predicting method and KNN method. The multivariate predicting method has a broad application prospect on prediction and can be applied to nonlinear complex system in various fields because of its advantage on recovering the dynamics of nonlinear system. In the future, the notion of using multivariate time series can combine with the other techniques to reveal the properties and dynamics of nonlinear system from the view of multiple variables.

Acknowledgments

The financial supports from the funds of the National Natural Science Foundation of China (61371130) and the Natural Science Foundation of Beijing (4162047) are gratefully acknowledged.

References

[1] W. Leutzbach, Introduction to the Theory of Traffic Flow, Springer, Berlin, 1988.
[2] B.S. Kerner, The Physics of Traffic, Springer, New York, 2004.
[3] D. Chowdhury, L. Santen, A. Schadschneider, Statistical physics of vehicular traffic and some related systems, Phys. Rep. 329 (2000) 199–329.
[4] D. Helbing, Traffic and related self-driven many-particle systems, Rev. Mod. Phys. 73 (2001) 1067–1141.
[5] B.S. Kerner, The Physics of Traffic, Springer, New York, 2004.
[6] M. Cheslow, S.G. Hatcher, V.M. Patel, An initial evaluation of alternative intelligent vehicle highway systemsarchitectures, Syst. Archit. (1992) Paper number: MTR 92W0000063.
[7] M. Perc, Introducing nonlinear time series analysis in undergraduate courses, Fizika A 15 (2006) 91–112.
[8] S. Kodba, M. Perc, M. Marhl, Detecting chaos from a time series, Eur. J. Phys. 26 (2005) 205–215.
[9] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, Cambridge, 1997.
[10] S. Kostić, M. Perc, N. Vasović, S. Trajković, Predictions of experimentally observed stochastic ground vibrations induced by blasting, PLoS One 8 (2013) 14161–14168.
[11] L.A. Safonov, E. Tomer, V.V. Strygin, Y. Ashkenazy, S. Havlin, Delay-induce chaos with multifractal attractor in a traffic flow model, Europhys. Lett. 57 (2002) 151–158.
[12] K. Daoudi, J.L. Véhel, Signal representation and segmentation based on multifractal stationarity, Signal Process 82 (2002) 2015–2024.
[13] I. Gasser, G. Sirito, B. Werner, Bifurcation analysis of a class of 'car following' traffic models, Physica D 197 (2004) 222–241.
[14] R.E. Wilson, Mechanisms for spatio-temporal pattern formation in highway traffic models, Philos. Trans. R. Soc. A 366 (2008) 2017–2032.
[15] M.Y. Bai, H.b. Zhu, Power law and multiscaling properties of the chinese stock market, Physica A 389 (2010) 1883–1890.
[16] J. Wang, P. Shang, X. Zhao, J. Xia, Multiscale entropy analysis of traffic time series, Int. J. Mod. Phys. C 24 (2013) 1350006.
[17] P. Shang, X. Li, S. Kamae, Nonlinear analysis of traffic time series at different temporal scales, Phys. Lett. A 357 (2006) 314–318.
[18] P. Shang, Y. Lu, S. Kamae, Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis, Chaos Solitons Fractals 36 (2008) 82–90.
[19] Y. Yin, P. Shang, Multiscale multifractal detrended cross-correlation analysis of traffic flow, Nonlinear Dyn. 81 (2015a) 1329–1347.
[20] Y. Yin, P. Shang, Multifractal cross-correlation analysis of traffic time series based on large deviation estimates, Nonlinear Dyn. 81 (2015b) 1779–1794.
[21] Y. Yin, P. Shang, Multiscale detrended cross-correlation analysis of traffic time series based on empirical mode decomposition, Fluct. Noise Lett. 14 (2015c) 1550023.
[22] M. Xu, P. Shang, J. Xia, Traffic signals analysis using qSDiff and qHDiff with surrogate data, Commun. Nonlinear Sci. Numer. Simul. 28 (2015) 98–108.
[23] G. Du, P. Shang, X. Zhao, Multiscale detrended fluctuation analysis of traffic index series, Fluct. Noise Lett. 13 (2014) 315–328.
[24] E. Takens, Detecting strange attractors in turbulence, in: D. Rand, L. Young (Eds.), Dynamical Systems and Turbulence, 898, Springer, Berlin, 1981, pp. 365–381.
[25] T. Saner, J.A. Yorke, M. Casdagli, Embedology, J. Stat. Phys. 65 (1992) 579–616.
[26] A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33 (1986) 1134–1140.
[27] A.M. Albano, A.I. Mees, G.C. deGuzman, P.E. Rapp, Data requirements for reliable estimation of correlation dimensions, in: H. Degn, A. Holden, L. Olsen (Eds.), Chaos in Biological Systems, Plenum, New York, 1987, pp. 207–220.
[28] J.D. Farmer, J.J. Sidorowich, Predicting chaotic time series, Phys. Rev. Lett. 59 (1987) 845–848.
[29] T.M. Cover, P.E. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1967) 21–27.
[30] P. Shang, X. Li, S. Kamae, Chaotic analysis of traffic time series, Chaos Solitons Fractals 25 (2005) 121–128.
[31] J. Wang, P. Shang, X. Zhao, A new traffic speed forecasting method based on bi-pattern recognition, Fluct. Noise Lett. 10 (2012) 59–75.
[32] A. Lin, P. Shang, G. Feng, B. Zhong, Application of empirical mode decomposition combined with k-nearest neighbors approach in financial time series forecasting, ALL Publ. 11 (2012) 590–602.
[33] X.L. Zhang, H.E. Guo-Guang, L.U. Hua-Pu, Short-term traffic flow forecasting based on k-nearest neighbors non-parametric regression, J. Syst. Eng. 24 (2009) 178–183.