Short-term traffic flow forecast of highway network based on chaos time series method

Hao Zhang*

Department industrial engineering, Tsinghua University, Beijing, China, 100084

Received 1 August 2014, www.cmnt.lv

Abstract

With the construction of the highway network and the growing traffic flows, demand for real-time control and guiding service have become increasingly prominent. As for short-term forecast of highway network, it is not only the basis and foundation of the real-time control and guiding service for traffic flows, and the precise forecast result will have a magnificent impact on improving the traffic capacity and service levels. This paper builds a short-term traffic flows forecast model for highway network based on chaotic time series analysis and prediction theory. The forecast of the traffic flows in given areas can be calculated. Results show that this model is feasible and has a high accuracy.

Keywords: highway network, traffic flow, chaos theory, time series, prediction

1 Introduction

With the increasing completed highway network and traffic flow, the demand for real time control and guidance services of highway traffic flow becomes increasingly prominent. Short-term traffic flow forecast of highway network not only is the foundation and basis of real time control and guidance services of highway traffic flow, and the accuracy of prediction results is important to improve traffic capacity and level of service of highway network.

According to the traffic flow prediction for different purposes, the focus and requirements of the prediction are also different. On the one hand, for long-term traffic flow forecasting, the demand for prediction accuracy is not high. On the other hand, in order to make network traffic control scheme, especially to make traffic control and guidance service, the prediction period is greatly reduced. Known for short-term traffic flow forecast, it requires high precision. At present, short-term traffic flow forecast mainly serves for real-time control and guidance services for city traffic so it requires high real-time control. The maximum period of traffic control is 2.5-3min generally and traffic guidance is 5min. Thus, the period of short-term traffic flow forecast of city traffic is generally no more than 15min. However, with the development of highway network in China, traffic congestion is increasingly frequent and serious in the local highway. Therefore, the demand for real time control and guidance services of highway traffic flow is increasingly urgent. Due to the difference between the highway traffic flow characteristics and city road network, it is necessary to study short-term traffic flow forecast of regional highway network.

2 Choosing short-term traffic flow forecast method of highway network

Traffic flow forecast is to speculate the future traffic flow according to the present traffic flow data, namely to make real-time prediction of traffic flow for the next time $t+\Delta T$ and subsequent several times at the moment t. Usually, the traffic flow within 15min is called the short-term traffic flow, so to predict the traffic flow between t and $t+\Delta t$ which does not exceed 15min is short-term traffic flow forecast.

There are many prediction methods for short-term traffic flow forecast [1-3,5,7], but generally can be divided into two categories: one is based on traditional statistics methods, such as regression prediction, time series prediction, Kalman Filtering model and so on; the other category is the forecast method based on neural network, fuzzy mathematics and nonlinear theory, which does not pursue mathematical derivation strictly and clear physical meaning of the object but pay more attention to the fitting effect of the real traffic flow phenomena. Among the above prediction methods, historical mean method and regression method use least square estimation to compute parameters, which is a simple calculation but the uncertainty and nonlinear of the traffic flow are difficult to reflect and the influence of random factors are difficult to overcome; neural network forecasting method, including BP network, fuzzy neural network and the high order neural network trains parameters complicatedly and needs long computing time; Kalman Filtering method and nonlinear time series method are characterized as high real-time and accuracy. Due to the complex feature of highway traffic flow whose change process is nonlinear and uncertain, and it is difficult to find a mathematical model to reflect the traffic flow characteristics in practice, this paper adopts second category of forecasting methods,

**Corresponding author's* e-mail: statistics.zhang@gmail.com

namely short-term traffic flow forecast of highway network based on chaos theory.

To predict short-term traffic flow of highway network based on chaos theory, it is primarily to distinguish whether the road network traffic flow system is a chaotic system. If it is a chaotic system, there can be short-term prediction, but not for long-term forecast; secondly, to determine the embedding dimension and time delay parameters and to find out the last known points in phase space, we need to use phase space reconstruction technique to make phase space reconstruction for traffic flow time series data; then we find out several adjacent points in the phase space and determine the predictive value of fitting function and separation in a known point as the center.

At present, there are many achievements of chaos theory to predict short-term traffic flow in domestic and foreign application [4,6,13,18], but most of these studies are used for traffic flow forecasting on a section in the city road network, while there needs further research for multiple sectional short-term traffic flow forecasting methods and models for highway network. The following will discuss the use of multidimensional chaotic time sequence to construct multiple sectional short-term traffic flow forecast model of highway network.

3 Short-term traffic flow forecast model of highway network based on chaos theory

3.1 BUILDING MULTIDIMENSIONAL CHAOTIC TIME SERIES PREDICTION MODEL

First, construction of multidimensional time series matrix.

Assumption (1): choosing traffic flow statistical data of *M* sections, corresponding *M* time series is obtained. X_i, $i = 1, 2, \ldots, M$.

Assumption (2): the length of each traffic flow time series sample is *N*. Thus, a multidimensional time series matrix as follows:

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,1} & \cdots & x_{2,N} \\ \vdots & \vdots & \cdots & \vdots \\ x_{N,1} & x_{N,1} & \cdots & x_{N,N} \end{bmatrix}. \quad (1)$$

Second, identification of chaotic time series. To make chaotic time series analysis of the traffic flow time series X_i respectively above, and select time delay σ_i and embedding dimension d_i to make phase space reconstruction, calculating Lyapunov index of each time series. If the index is greater than zero, it indicates that the time series is a chaotic time series, where the related theory and the method of chaotic time series can be used for analysis and modeling.

Third, reconstruction of multidimensional phase space. Based on the chaotic time series, according to multidimensional phase space reconstruction method proposed by Liangyue Caoetal (1998), the reconstruction of phase space is obtained:

$$\begin{aligned} V_n = (&x_{1,n}, x_{1,n-r_1}, \ldots, x_{1,n-(d_1-1)r_1}, \\ &x_{2,n}, x_{2,n-r_2}, \ldots, x_{1,n-(d_1-1)r_2}, \\ &\ldots, \\ &x_{M,n}, x_{M,n-r_m}, \ldots, x_{M,n-(d_1-1)r_m}), \end{aligned} \quad (2)$$

$$n = \max_{1SiSM}(d_i - 1)\tau_i + 1, \ldots, N, \quad (3)$$

Forth, construction of prediction model. According to the embedding theorem, there is a smooth mapping function in a *D* dimensional space like $F\left(d = \sum_{i=1}^{M} d_i\right)$: $R^{\mathrm{d}} \to R^{\mathrm{d}}$, If d or di is sufficiently large to make $V_{\mathrm{n}+1} = F(V_{\mathrm{n}})$, then the model can be described as:

$$V_{1,n+1} = F_1(V_{\mathrm{n}}), V_{2,n+1} = F_2(V_{\mathrm{n}}), \ldots V_{M,n+1} = F_M(V_{\mathrm{n}}). \quad (4)$$

After establishing the prediction model, there comes out a phase space after reconstruction of multi-dimensional chaotic time series. According to the embedding theorems of chaos theory, there is a smooth mapping function *F* which can reflect the law of phase space, and the smooth mapping function F can establish a dynamic mathematical model to fit it to predict the motion trace of phase space, which can get the forecasting results of highway network traffic flow. After finishing the reconstruction of phase space through multidimensional time series, we adopt prediction methods like global method, local method and maximum Lyapunov exponent to establish the mathematical model to approximate the mapping function F. In this paper, we use the weighted zero order local forecasting method to make calculation, using V_n as the center point to predict V_{n+1}.

3.2 STEPS OF WEIGHTED ZERO ORDER LOCAL PREDICTION CALCULATION

Steps of weighted zero order local prediction calculation are as following based on reconstruction of phase space:

First, calculating K adjacent points $V_{n,j}^{\min}$, which is the minimum Euclidean distance with V_n in the phase space after reconstruction, the distance of each point are respectively R_j, $j = 1, 2, \ldots, K$, and let $R_{\min}$ be the minimum value. Thus, each adjacent point's weight can be defined as:

$$P_j = e^{-l(R_j - R_{\min})} / \sum_{j=1}^{K} e^{-l(R_j - R_{\min})}. \quad (5)$$

Secondly, V_{n+1} is calculated according to the following equation:

$$V_{n+1} = \sum_{j=1}^{K} P_j V_{n,j}^{\min} . \quad (6)$$

Lastly, according to:

$$\begin{aligned} V_{n+1} = (& x_{1,n+1}, x_{1,n+1-r_1}, \cdots, x_{1,n+1-(d_1-1)r_1}, \\ & x_{2,n+1}, x_{2,n+1-r_2}, \cdots, x_{1,n+1-(d_1-1)r_2}, \\ & \ldots, \\ & x_{M,n+1}, x_{M,n+1-r_m}, \cdots, x_{M,n+1-(d_1-1)r_m}), \end{aligned} \quad (7)$$

The V_{n+1} can be isolated from the final traffic flow prediction results as:

$$\left(x_{1,n+1}, x_{2,n+1}, \ldots, x_{M,n+1}\right). \quad (8)$$

There are two parts of the parameters in the above model. One part is the time σ_i and embedding dimension d_i of phase space reconstruction. The other part is the neighbor number K and the calculation of adjacent weight parameter l in the weighted zero order local forecasting method. In the theory of chaotic time series analysis, there are many ways to determine the time delay σ_i and embedding dimension d_i, such as the auto-correlation method, mutual information method and C-C method. Methods to determine the embedding dimension are G-P algorithm, the false neighbor method C -C method and Cao method. This paper use C-C method to determine the time delay and embedded dimension. The parameters K represents the number of neighboring points. Too few adjacent points leads to low prediction accuracy, while too many neighbor points is unnecessary because it may make the computational complexity increase and affect the prediction effect; the parameter l represents the weights of adjacent points, its weight is determined according to the proximity to the center point of the Euclidean distance in the forecast. In certain Euclidean distance, less l means greater proportion. In this paper use the particle swarm optimization algorithm to select the parameters of K and l, because the particle swarm algorithm is simple and effective and the solving process is easier.

4 Empirical studies

4.1 SOURCES OF DATA

In order to test the model and the method, the traffic flow data comes from Jinggangao Highway, Hurong Highway, Huyu Highway and Yinfu Highway Yuenan Station in the Hubei highway network from 00:00 in September 15 to 23:59 in September 15th, 2013. Through the correlation data pre-processing and completed analysis, the traffic flow of Section 1, Section 2,… Section 6 are sources of data. The flow curves are shown in Figure 1 to 3 (Section 1 for example). The time interval Δ_t is respectively 5 min, 10min and 15min. Then the prediction results and the measured results are compared based on established model, parameters optimization and prediction.

4.2 STEPS OF THE EMPIRICAL STUDIES

Firstly, to make pre-treatment of the selected section traffic flow data and make correlation analysis with the completed data.

Secondly, calculating the time delay embedding dimension and Lyapunov index of each time series using the C-C method shown in Table 1. The lyapnuvo index is positive shown in Table 1, which suggests that the 6 time series are all chaotic time series so we can use the chaos theory to analyze and predict.

TABLE 1 Result of parameters calculated under 6 time series and three intervals

5min	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6
Time delay	1	1	2	2	1	2
Embedding dimension	3	10	4	4	7	3
Lyapunov	1.1605	1.0095	0.712	0.8286	1.2915	0.2845
10min	Section1	Section2	Section3	Section4	Section5	Section6
Time delay	1	1	3	3	1	2
Embedding dimension	2	4	3	2	5	3
Lyapnuvo	0.3905	0.5632	0.2636	0.2486	0.8283	0.2744
15 min	Section1	Section2	Section3	Section4	Section5	Section6
Time delay	1	1	3	2	1	2
Embedding dimension	1.75	2	3	3	5	3
Lyapnuvo	0.379	0.381	0.31	0.4836	0.4541	0.3756

4.3 COMPUTATIONAL PREDICTION AND ERROR

According to the time delay and embedding dimension, using the multi section model of chaotic prediction and the use of particle swarm optimization algorithm of K and optimization of l parameters, the calculation results of traffic flow of each cross section. The prediction results see Figures 1-3 (in Section 1 as an example), the solid part is the prediction of traffic flow, the dashed part for the measured traffic flow, a unit for every 5 minutes, every 10 minutes and a bus every 15 minutes.

At different time intervals, each section of traffic flow forecast value and the actual value of average error, see Table 2.

TABLE 2 Different time intervals of the cross-section traffic predicted and actual values of average error table

average error	Section1	Section2	Section3	Section4	Section5	Section6
5min	0.026541	0.031395	0.031081	0.018483	0.099386	0.129176
10min	0.01886	0.021353	0.018794	0.006982	0.04686	0.055504
15min	0.018788	0.018246	0.012589	0.004737	0.035359	0.047582

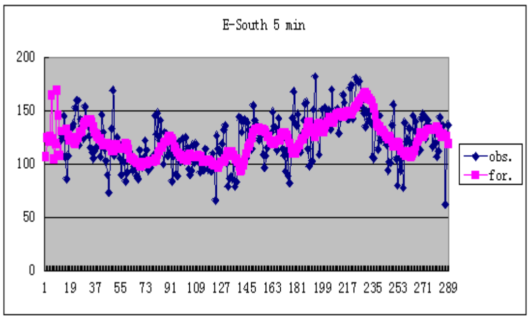

FIGURE 1 Under section 1 in 5min-interval actual and predicted traffic volume comparison chart

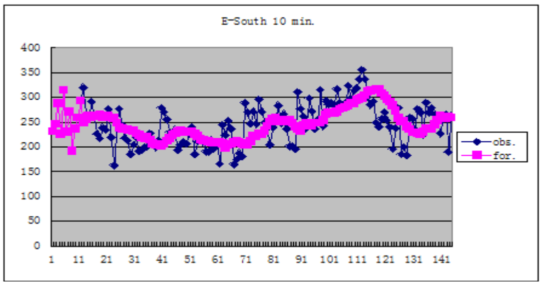

FIGURE 2 Under section 1 in 10min-interval actual and predicted traffic volume comparison chart

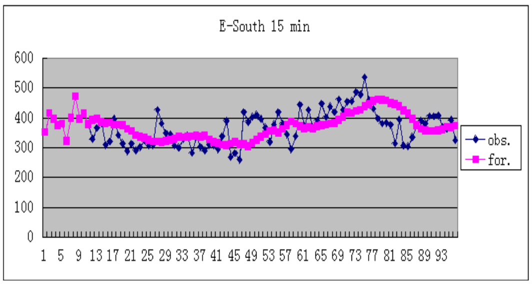

FIGURE 3 Under section 1 in 15min-interval actual and predicted traffic volume comparison chart

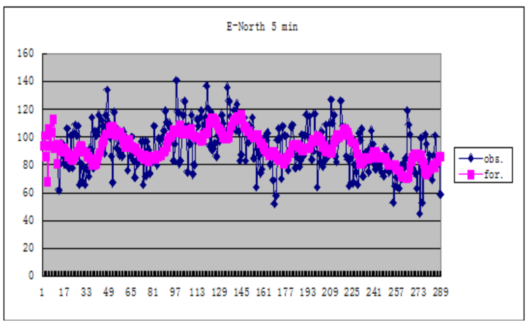

FIGURE 4 Under section 3 in 5min-interval actual and predicted traffic volume comparison chart

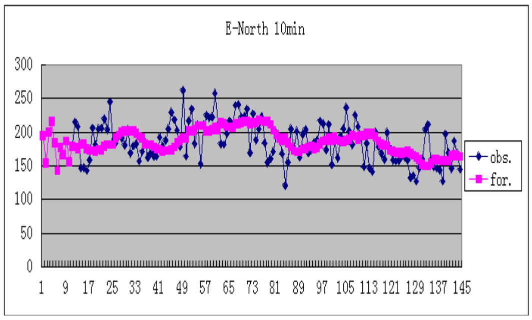

FIGURE 5 Under section 3 in 10 min-interval actual and predicted traffic volume comparison chart

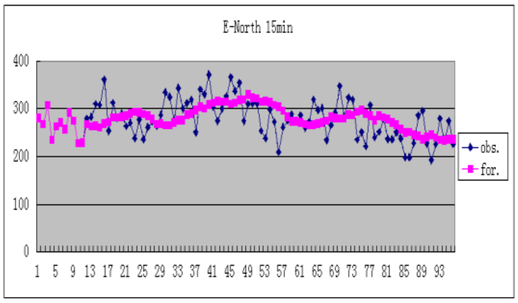

FIGURE 6 Under section 3 in 15 min-interval actual and predicted traffic volume comparison chart

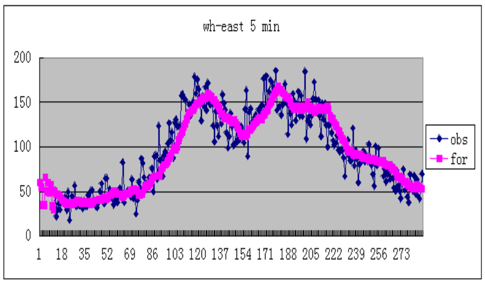

FIGURE 7 Under section 4 in 5min-interval actual and predicted traffic volume comparison chart

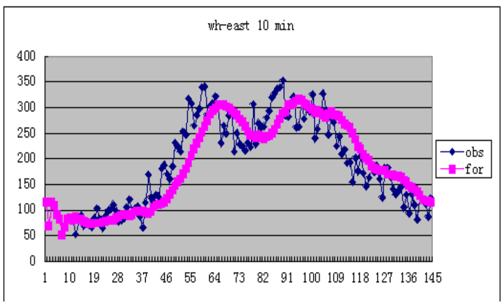

FIGURE 8 Under section 4 in 10 min-interval actual and predicted traffic volume comparison chart

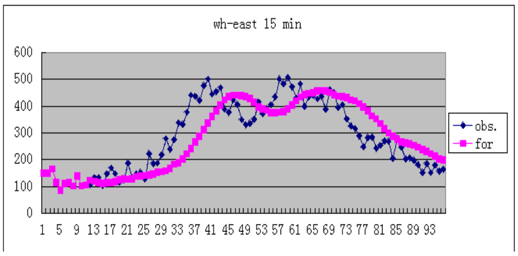

FIGURE 9 Under section 4 in 15 min-interval actual and predicted traffic volume comparison chart

5 Conclusions

Compared the actual and predicted values of traffic flow in section 1 and section3, it shows that the average prediction error tends to decrease with the increasing time intervals. Meanwhile, compared with chaotic time series prediction values in a separate and single section, the prediction accuracy of multiple-sections forecasting method is higher than single-section forecasting methods. Overall, the multiple-sections traffic flow forecasting method in road network based on multi-dimensional chaotic time series prediction is feasible, and the predictive effect is better than of the single-section chaotic time series prediction method. On the contrary to the short-term traffic flow forecasting in urban road network, see Figure 7, Figure 8 and Figure 9 in Section4 located at the intersection of the urban road network and the highway network, if the time intervals relatively increase in short-term traffic flow prediction of highway network, its predictive effect would be better.

References

[1] Hao B L 1989 Elementary Symbolic Dynamics and Chaos in Dissipative Systems *World Scientific Publishing Company*

[2] Okotani I, Stephanedes Y J 1984 Dynamic Prediction ofTraffic Volume Through Kalman Filtering Theory *Transportation Research Part B: Methodological* **18**(1) 1-11

[3] Guo M, Lan J, Xiao X, Lu H 2010 Forecasting Short-time Traffic Flow for Beijing 2nd Ring Road Using Chaos Theory *Journal of Transportation Systems Engineering and Information Technology* **10**(2) 106-11

[4] Liu J 2007 Study on Forecasting of Short-term Traffic Flow Based on Fractal Theory *Doctoral Dissertation at Beijing Jiaotong University (in Chinese)*

[5] Zhu X 2007 Research on a Practical Method of Freeway Forecasting Distribution *Central South Highway Engineering* **32**(3) 161-4 *(in Chinese)*

[6] Zhang Y, Qu S, Wen K 2009 Chaotic Property Analysis and Prediction Model Study for Traffic Flow Time Series *China Civil Engineering Journal* **42**(1) 119-23*(in Chinese)*

[7] Yue Y, Han W 2005 Determination of Parameters in the Phase-space Reconstruction of Multivariate Time Series *Control and Decision* **20**(3) 290-3

[8] Ghosh B, Basu B, O'Mahony M 2009 *IEEE Transactions on Intelligent Transportation Systems* **10**(2) 246-54

[9] Tchrakian T T, Basu B, O'Mahony M 2012 *IEEE Transactions on Intelligent Transportation Systems* **13**(2) 519-26

[10] Javad A, Behzad M, Baher A, Ali K S 2012 Forecasting of short-term traffic-flow based on improved neurofuzzy models via emotional temporal difference learning algorithm *Engineering Applications of Artificial Intelligence* **25**(5) 1022-42

[11] Kit Y C, Khadem S, Dillon T S, Palade V 2012 *IEEE Transactions on Industrial Informatics* **8**(2) 255-66

[12] Kit Y C, Dillon T S, Chang E 2013 *IEEE Transactions on Industrial Electronics* **60**(10) 4114-25

[13] Lippi M, Bertini M, Frasconi P 2013 *IEEE Transactions on Intelligent Transportation Systems* **14**(2) 871-82

[14] Kit Y C, Dillon T S, Singh J, Chang E *IEEE Transactions on Intelligent Transportation Systems* **13**(2) 644-54

[15] Zhang X, Onieva E, Perallos A, Osaba E, Lee V C S 2014 Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction *Transportation Research Part C: Emerging Technologies* **43**(1) 127-142

[16] Hosseini S H, Moshiri B, Rahimi-Kian A, Araabi B N 2012 Short-term traffic flow forecasting by mutual information and artificial neural networks *2012 IEEE International Conference on Industrial Technology (ICIT)* 1136-41

[17] Huang Z, Ouyang H, Tian Y 2011 Short-Term Traffic Flow Combined Forecasting Based on Nonparametric Regression Information Technology *International Conference on Computer Engineering and Management Sciences (ICM)* **1** 316-9

[18] Kaya S, Kilic N, Kocak T, Gungor C 2014 From Asia to Europe: Short-term traffic flow prediction between continents *International Conference on Telecommunications (ICT)* 277-82

Author

Hao Zhang, born in December, 1984, Wuhan city, China

Current position, grades: PhD Candidate in Department industrial engineering, Tsinghua University.
University studies: MS, Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, Stony Brook NY USA, 2009. Area of Study: Statistics and Quantitative Finance. BS, National Mathematics Base Class, Wuhan University, Wuhan China, 2007. Area of Study: Mathematics.
Scientific interests: operations research and logistics management.
Publications: 5 papers.