Scientific Research and Essays Vol. 5 (4), pp. 389-394, 18 February, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 © 2010 Academic Journals

Full Length Research Paper

Dynamic forecasting of traffic volume based on quantificational dynamics: A nearness perspective

Yan-hong Tang* and Bao Xi

School of Management, Harbin Institute of Technology, 150001, Harbin, China.

Accepted 21 January, 2010

Accurate and timely forecasting of traffic volume has long been regarded as a key point in transportation, planning and management. In order to realize effective and efficient traffic forecasting, this paper investigates quantificational method of dynamic factors from the perspective of nearness. The dynamic modeling method based on quantificational dynamics (that is, quantificational disposal of dynamic factors according to nearness) is proposed and this method can significantly improve the forecast effectiveness and efficiency. Swarm simulation is adopted as a new tool with regard to the field of traffic forecasting for analysis and verification. The testing results show that the proposed method outperforms traditional ones in choosing training samples and constituting forecasting models. This work contributes to the consideration and evaluation of dynamic factors in scientific forecasting and may bring some enlightenment to relevant scientific researchers and engineers.

Key words: Traffic forecasting, dynamic factors, quantificational dynamics, nearness, swarm simulation.

INTRODUCTION

Forecasting of traffic volume plays a key role in reducing traffic congestion, enhancing control performance of transportation infrastructure and improving traffic safety. Forecasting results of traffic volume can be applied to control of intersections, freeways, tunnels and even parking lots. Considering period, forecasting techniques of traffic volume can be divided into the categories of long term, intermediate term and short term (Chen, 1997).

The steady increased traffic volumes in both rural and urban roads in recent years have resulted in congestions in many road traffic systems. Intelligent transportation systems (ITS) provide solutions for alleviating the increasing congestion problems. Accurate and timely forecasting of traffic volume is critical for effective control of traffic congestion in ITS environment (Jiang and Adeli, 2005). Short-term traffic volume forecasting has long been regarded as a critical concern for the development and application of ITS. In particular, such traffic flow forecasting supports 1. the development of proactive traffic control strategies in advanced traffic management systems (ATMSs), 2. real-time route guidance in advanced traveler information systems (ATISs) and 3. evaluation of these dynamic traffic control and guidance strategies as well (Zheng et al., 2006). In an early report on the architecture of intelligent transportation systems (Cheslow et al., 1992), it was clearly indicated that the ability to make continuous predictions of traffic flows and link travel times for several minutes into the future, using real-time traffic data, is a major requirement for providing dynamic traffic control and guidance. In the last two decades, the growing need for short-term prediction of traffic parameters embedded in a real-time intelligent transportation systems environment has led to the development of a vast number of forecasting algorithms (Vlahogianni et al., 2006). For information purposes, the highway capacity manual (Transportation Research Board, 2000) suggests using a 15 min traffic flow rate. In this study, traffic volume of a 15 min interval is taken for as the study object.

Literature review

The short-term forecasting of traffic conditions has had an active but somewhat unsatisfying research history Davis and Nihan, 1991). A variety of methods have been applied to short-term traffic volume forecasting, including the multivariate time-series model (Williams et al., 1998), (the Kalman filtering method (Okutani and Stephanedes,

*Corresponding author. E-mail: tangyanhong@yahoo.cn, tyh119@126.com.

1984; Xie et al., 2007) and the nonparametric regression model (Davis and Nihan, 1991; Smith et al., 2002). Over the past decade, a number of papers have been published on the application of neural network models for forecasting traffic flow taking advantage of their ability to capture the uncertain and complex nonlinearity of time series.

Smith and Demetsky (1994, 1997) compared the back propagation (BP) neural network model with the ARIMA model for predicting short-term traffic flow. They concluded that the BP neural network model was superior to the linear statistical ARIMA model because the former was more sensitive to the dynamics of traffic flow than the latter and did not experience the over forecast characteristics of the ARIMA model.

Yun et al. (1998) investigated the performance of a BP neural network model, a finite impulse response model (a linear filtering method) and a time-delayed recurrent model (a dynamic BP neural network) for forecasting the traffic volume. They used three different traffic flow data sets collected from interstate highways, intercity highways and urban intersections with very different characteristics in terms of volatility, period and fluctuations. Their study showed that the time-delayed recurrent BP neural network model outperformed other models in forecasting very randomly moving traffic flow. In contrast, the FIR model demonstrated better forecasting accuracy than the time-delayed recurrent network for relatively regular periodic data. However, the BP model had its inherent shortcomings such as lack of an efficient constructive model (for example, requiring arbitrary selection of the number of hidden nodes), slow convergence rate resulting in excessive computation time and entrapment in a local minimum.

In order to overcome the aforementioned shortcomings of the BP neural network model, some scholars (Park, 2002) proposed a fuzzy-neural network approach for forecast-ting the short-term traffic flow (Park et al., 2002; Yin et al., 2002). They concluded that the composite method required less computing time and provided better forecasting accuracy than BP models.

Until now, no single predictor had yet been developed that presented it to be universally accepted as the best. As traffic flow itself is a complicated process influenced by many dynamic factors, it is found that using composite models to describe and forecast the effect of different factors on the traffic flow is appropriate (Zheng et al., 2006). Besides, how to consider dynamic factors in traffic forecasting model is still a problem that has not been properly solved.

The dynamic modeling

Traffic flow characterized as intense fluctuations in micro-cosmic level (Vlahogianni et al., 2006) presents some general law in macroscopical level. The fluctuation law of traffic flow falls into two categories divided by workday and holiday. In addition, the fluctuation is influenced by weather. Although traffic engineers have been aware of these phenomena for many years, relevant quantificational description of these phenomena are difficult to obtain. First, the quantitative forecasting should base on historical data analysis; therefore, random factors and disturbances cannot be exactly calculated. Second, some complicated factors, such as weather variations, although their influences have been known, are difficult to be quantificationally depicted. Based on these considerations, the improvement of dynamic forecasting model should focus on the following aspects: (1) The periodical variation of traffic flow along with different date types and different characteristic times; (2) The spatio-temporal law of traffic flow and (3) The external influence such as temperature, sunlight and visibility and so on.

The dynamic forecasting of traffic volume can be regarded as a large-scale multi-mapping problem and the forecasting model is uneasy to be identified. Since artificial neural network (ANN) has the strength of approaching discretionary nonlinear function and simulating multi-variable problem wonderfully without knowing the function relations between independent variable and dependent variable (Isik, 2009; Sancak, 2009), the ANN model can conveniently consider some factors such as visibility, sunlight, date type, etc. Theoretically, ANN models suit for application in dynamic forecasting of traffic volume. However, the traditional method for training a multi-layer feed-forward ANN is the BP algorithm. Although it has been successfully applied in many domains and has been improved on and on, the BP algorithm still has some weaknesses, which have not been overcome thoroughly. For example, the slow training and converging BP-ANN algorithm is difficult to be applied to select dynamic samples. While proper sample selection can increase training speed and forecastting precision in dynamic forecasting, so training time can be reduced significantly and unnecessary interference from irrelevant samples can be avoided if the selected samples have similar character with the forecast. The extended kalman filter (EKF) (Tsai et al., 2005) training algorithm gauges the weights as per the principle of minimum root mean squared covariance, which needs much less iteration than the BP algorithm. Moreover, free from calculation of convergent parameters makes it convenient to be applied in rapid calculation and selecting dynamic samples (Wang and Papageorgiou, 2005). Therefore, the EKF-ANN algorithm is applied to dynamic forecasting of traffic volume in this paper.

The nearness between the prospective date and the historical date can be expressed as follows (Miao and Xi, 2008):

$$S(k)=\alpha[(T_{\max}-T_{\max}^{K})^2+(T_{\min}-T_{\min}^{K})^2]+\beta(V-V^k)^2+\gamma(W-W^K)^2+\delta\Delta D^k \quad \text{.........(1)}$$

Where $T_{\max}$ - the highest temperature in the prospective

Table 1. Quantificational disposal of meteorological factors.

Weather condition	Sunny	Cloudy	Windy	Foggy
Quantificational value	1	2	3	4
Weather condition	Rainy	Thunder storm	Snowy	Snow storm
Quantificational value	5	6	7	8

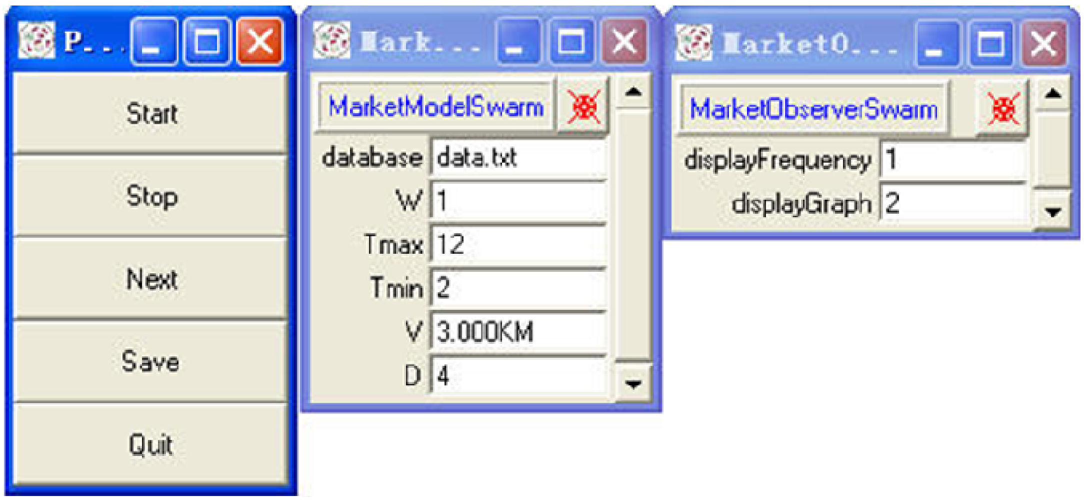

Figure 1. Parameters of the swarm simulation.

date;

$T_{\max}^{K}$ - the highest temperature in the date K;
$T_{\min}$ - the lowest temperature in the prospective date;
$T_{\min}^{K}$ - the lowest temperature in the date K;
V -the visibility in the prospective date;
V^{K} - the visibility in the date K;
W - the weather type in the prospective date;
W^{K} - the weather type in the date K;
ΔD^{k} - the date interval between the prospective date and date K;
α, β, γ and δ are the coefficients that reflect the influence of temperature, visibility, weather type and date interval on traffic flow. The meteorological factors W and W^{K} can be quantificationally disposed as Table 1.

If the $S(k)$ is smaller, the nearness will be higher. Based on the meteorological data of the prospective date, the nearness between the historical date and the prospective date can be obtained. The samples with smaller $S(k)$ are selected as the dynamic training sample set.

The resultant samples can be expressed along time series as formula (2):

$$X = [X_1 \quad X_2 \quad X_3 \quad \cdots \quad X_M] \qquad (2)$$

Where, X is the sample set, X_i $(1 \le i \le M)$ is a vector in the set, the components in it are corresponding historical data. The M is the total number of the resultant samples.

Generally, $(I \times O)^{1/2}$ is chose as the number of hidden layer as per experience, where, I and O respectively denote the unit number in the input layer and output layer. The historical data in the resultant samples, the weather of the prospective date and the date type are taken as the input; the output is the predicted traffic volume.

Swarm-aided simulation example

The Swarm platform designed by the Santa Fe Institute supports a sort of tools that can simulate and validate the above models. It is a collection of software libraries that can support the ANN simulation program (Johnson, 2004). The Swarm needs fewer external initial variables and is good at coping with short-term traffic flow influenced by dynamic factors. Figure 1 shows the parameters of the swarm simulation on dynamic forecasting of short-term traffic volume.

The meanings of the parameters in Figure 1 are listed as follows:

(1) Database: the data.txt is the dynamic training sample set, which covers data from the historical as well as from

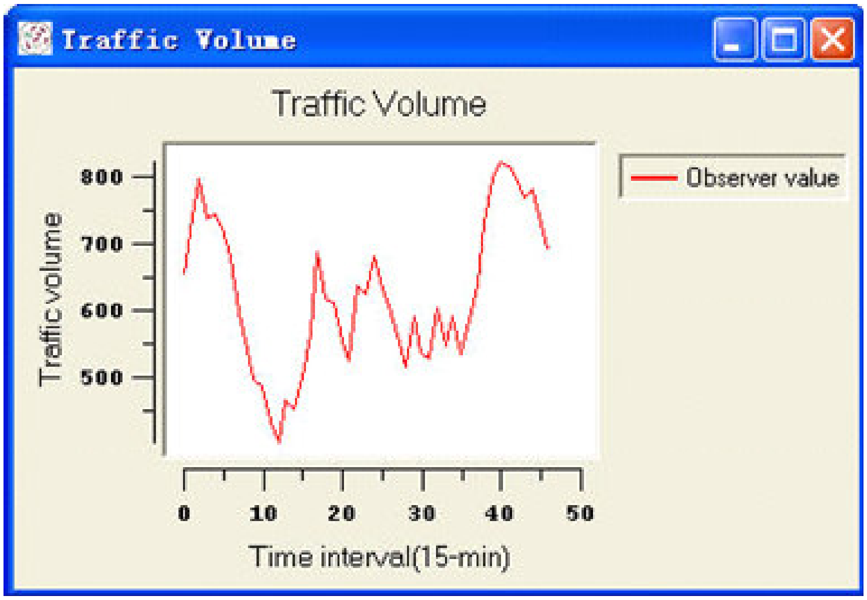

Figure 2. Typical daily traffic volumes at the observation spot.

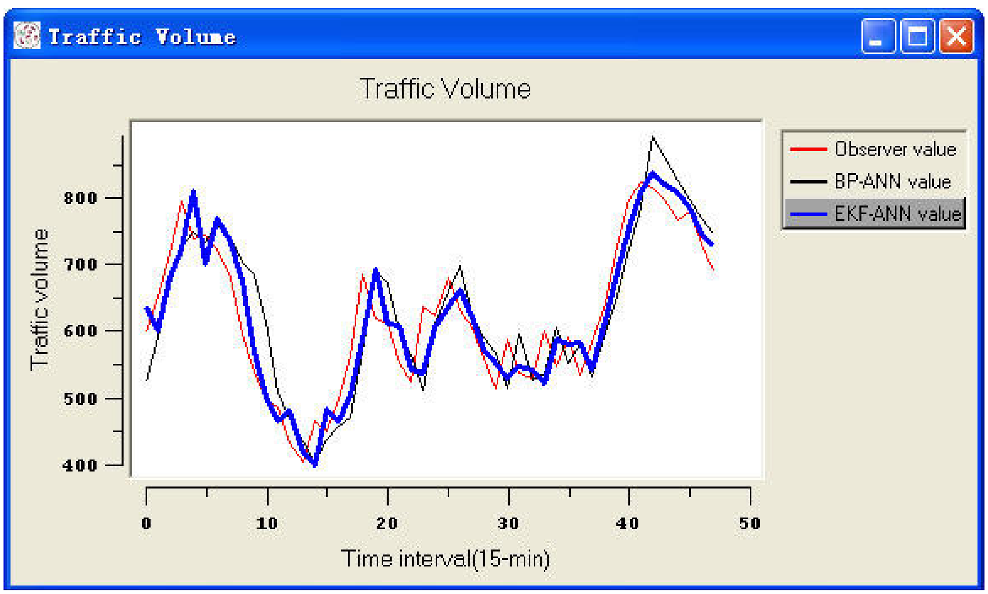

Figure 3. Output of the two predictors.

the recent.
(2) *W*: weather, the number 1 denotes sunny.
(3) *T max*: the highest temperature in the prospective day.
(4) *T min*: the lowest temperature in the prospective day.
(5) *V*: the visibility of the prospective day.
(6) *D*: date type, the number 4 denotes Thursday.

Figure 2 depicts typical daily traffic volumes at the observation spot while Figure 3 presents the output of the two predictors at the observation spot.

RESULT ANALYSIS

In Figure 3, the observed traffic volumes on that day are also presented for comparison. With the exception of a few intervals, the two predictors show a good reflection of the changing trends of traffic flow, while the EKF-ANN predictor gives a better approximation of the actual traffic volume.

Three indices, that is, the mean absolute percentage error (MAPE), root mean squared error (RMSE) and the relative error, are selected and employed to compare the

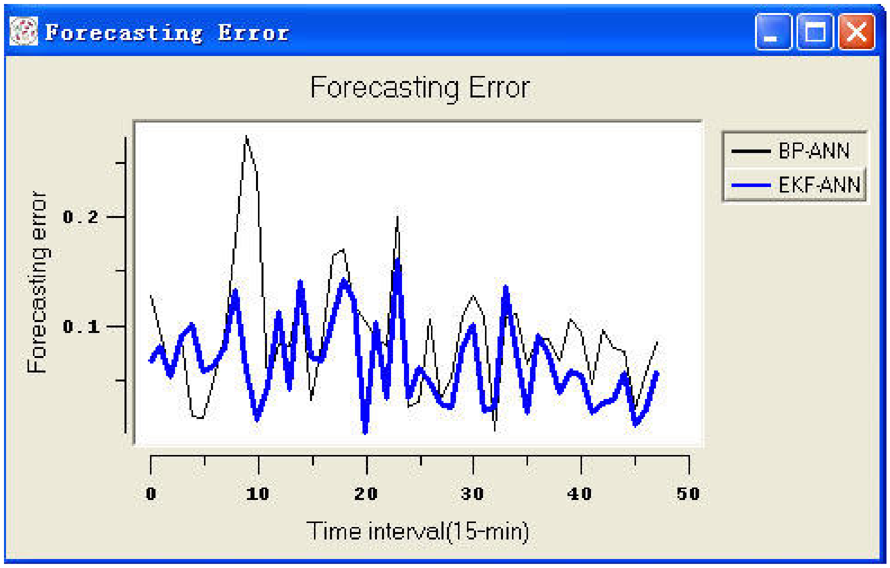

Figure 4. Distribution of the forecasting errors of the two predictors.

Table 2. Performance comparison of the two predictors.

	MAPE	RMSE
BP-ANN	9.07%	10.61%
EKF-ANN	6.43%	7.52%

forecasting performances of the two predictors. The MAPE and RMSE reflect the accuracy and stability of the predictors and the relative error indicates the reliability of the predictors. The MAPE, RMSE and relative error are defined as follows:

$$MAPE = \frac{1}{N} \times \sum_{k=1}^{N} \left| \frac{l_k - l'_k}{l_k} \right| \times 100 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left(\frac{l_k - l'_k}{l_k} \right)^2} \times 100 \quad (4)$$

$$Err(k) = \left| \frac{l_k - l'_k}{l_k} \right| \quad (5)$$

Where l_k denotes the observed traffic volume in time interval *k*; l'_k denotes predicted traffic volume in time interval *k*; *N* denotes the number of time intervals for forecasting, as the observation time is from 7:00 to 19:00 and each time interval is 15 min, therefore, there are 48 intervals in the 12 h (from 7:00 to 19:00) and $N = 48$; $Err(k)$ denotes the relative error between predictive values and observed values in interval *k*.

Equation (3) calculates the average relative error between the predictive values and observed values and it represents the accuracy of the forecasting. Equation (4) represents the sum of the deviations from the average performance during the forecasting in all intervals. It is obvious that a predictor with a large RMSE is not as stable as one with a smaller RMSE. Table 2 shows the performance comparison of the two predictors.

From Table 2, we can see that the EKF-ANN predictor has a better forecasting performance than the BP-ANN predictor in most time of the day in terms of accuracy and stability, which is indicated by the MAPE and RMSE values. It is also found that on the whole day, the EKF-ANN gives a more reliable prediction, as it shows a probability of more than 85% of yielding predictive values with a forecasting error margin of less than 10%. This is higher than that of the other predictor. With such a level of accuracy, the EKF-ANN model can be considered as better for practical application.

Figure 4 shows the distribution of the forecasting errors for the BP-ANN and the EKF-ANN model on the day. As can be seen from the figure, the forecasting results of the EKF-ANN method are ideal and this method is to some extent superior to the BP-ANN method.

Conclusion

This research provides a new insight for consideration of dynamic factors in traffic forecasting and contributes to traffic engineering studies and scientific forecasting research. Quantificational dynamics may help scientific researchers and engineers to dispose fuzzy dynamics from nearness perspective. Through a simulation testing of short-term traffic volume forecasting, this method is demonstrated to be superior in terms of forecast, effectiveness and efficiency. However, this research is preliminary and tentative, further engineering applications should be carried out in future to prove and improve this approach.

ACKNOWLEDGEMENTS

This work is supported by Heilongjiang Provincial Research Base of Management Science and Engineering at Harbin Institute of Technology. The authors would like to thank the anonymous reviewers, Dr. Sivakumar and other relevant editorial board members for their constructive comments and suggestions.

REFERENCES

Chen J (1997). Characterization and implementation of neural network time series models for traffic volume forecasting. Dissertation for the Doctor of Philosophy Degree in Engineering Science, The University of Toledo.

Cheslow M, Hatcher SG, Patel VM (1992). An initial evaluation of alternative intelligent vehicle highway systems architecture, MITRE Report 92w0000063, MITRE Corporation.

Davis GA, Nihan NL (1991). Nonparametric regression and short term freeway traffic forecasting. J. Transport. Eng. 117(2): 178-188.

Isik NS (2009). Estimation of swell index of fine grained soils using regression equations and artificial neural networks. Sci. Res. Essays. 4(10): 1047-1056.

Jiang XM, Adeli H (2005). Dynamic wavelet neural network model for traffic flow forecasting. J. Transport. Eng. 131(10): 771-779.

Johnson P, Lancaster A (2004). Swarm User Guide. Swarm Development Group, www.swarm.org.

Miao X, Xi B (2008). Agile forecasting of dynamic logistics demand. Transport 23(1): 26-30.

Okutani I, Stephanedes YJ (1984). Dynamic prediction of traffic volume through kalman filtering theory. Transport. Res. B-Meth. 18(1): 1-11.

Park B (2002). Composite neuro-fuzzy application in short-term freeway traffic volume forecasting. Transport. Res. Rec. (1802): 190-196.

Park DC, El-Sharkawi MA, Marks II RJ, Atlas LE, Damborg MJ (1991). Electric load forecasting using an artificial neural network. IEEE T Power. Syst. 6(2): 442-449.

Sancak E (2009). Prediction of bond strength of lightweight concretes by using artificial neural networks. Sci. Res. Essays 4(4): 256-266.

Smith BL, Demetsky MJ (1994). Short-term traffic flow prediction: Neural network approach. Transport. Res. Rec. (1453): 98-104.

Smith BL, Demetsky MJ (1997). Traffic flow forecasting: Comparison of modeling approaches. J. Transport. Eng. 123(4): 261-266.

Smith BL, Williams BM, Oswald RK (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. Transport. Res. C-Emer. 10(4): 303–321.

Transportation Research Board (2000). Highway Capacity Manual. National Research Council, Washington, D.C.

Tsai JSH, Yu JM, Canelon JI, Shieh LS (2005). Extended-Kalman-filter-based chaotic communication. Ima. J. Math. Control I. 22(1): 58-79.

Vlahogianni EI, Karlaftis MG, Golias JC (2006). Statistical methods for detecting nonlinearity and non-stationarity in univariate short-term time-series of traffic volume. Transport. Res. C-Emer. 14(5): 351-367.

Wang YB, Papageorgiou M (2005). Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. Transport. Res. B-Meth. 39(2): 141-167.

Williams BM, Durvasula PK, Brown DE (1998). Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models. Transport. Res. Rec. (1644): 132-141.

Xie YC, Zhang YL, Ye ZR (2007). Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. Comput-Aided Civ. Inf. 22(5): 326-334.

Yin H, Wong SC, Xu J, Wong CK (2002). Urban traffic flow prediction using a fuzzy-neural approach. Transport. Res. C-Emer. 10(2): 85-98.

Yun SY, Namkoong S, Rho JH, Shin SW, Choi JU (1998). A performance evaluation of neural network models in traffic volume forecasting. Math. Comput. Model. 27(9-11): 293-310.

Zheng WZ, Lee DH, Shi QX (2006). Short-term freeway traffic flow prediction: Bayesian combined neural network approach. J. Transport. Eng. 132(2): 114-121.