A computationally efficient 2-satge method for short-term traffic prediction on urban roads

Miss Fangce Guo

PhD candidate

Imperial College London

Dr. Rajesh Krishnan

Research Associate

Imperial College London

Prof. John Polak

Professor

Imperial College London

Abstract
Short-term traffic prediction plays an important role in intelligent transport systems. With accurate predictions of the traffic state, transport network managers can develop more sophisticated strategies to mitigate traffic problems before they occur. A number of statistical and data mining methods for traffic prediction are available in the academic literature, including time series analysis, Kalman filter and non-parametric methods. Most traffic prediction methods, however, need the process of training or parameter optimisation, which increases the difficulty of their implementation and limits their robustness. This paper presents a novel 2-stage prediction structure using the technique of Singular Spectrum Analysis (SSA) as a data pre-processing step to improve the prediction accuracy. Moreover, a new time series prediction method named Grey system Model (GM) is introduced to reduce the dependency on method training and parameter optimisation. To demonstrate these improvements, this paper compares the prediction results of SSA structure model with that of a non-SSA method. Another time series method, Seasonal Auto-Regressive Integrated Moving Average (SARIMA) approach was chosen as the prediction function in these models. Prediction models using GM and SARIMA methods were calibrated and evaluated using real traffic flow data from a corridor in Central London under both normal and incident traffic conditions. The prediction accuracy comparisons show that the data pre-processing before the procedure of prediction using traditional time series methods can improve the final prediction accuracy. In addition, the results indicate that GM method outperforms SARIMA under both normal and incident traffic conditions on urban roads.

1 Introduction

Since the use of Inductive Loop Detectors (ILDs) in Intelligent Transportation Systems (ITS), transport network managers have access to large amounts of real-time or near real-time traffic data, such as traffic flow and travel time. These traffic variables can be used by researchers to forecast short-term traffic conditions. An accurate and robust traffic prediction model can help traffic authorities provide services in a proactive manner. However, the traffic prediction model is becoming increasingly more complicated, especially using data-driven methods. In addition, more training data and computational time are required in order to obtain more accurate, robust and reliable prediction results.

The main objective in this paper is to use a novel 2-stage prediction structure to improve traffic prediction accuracy. This paper is to compare the prediction accuracy of GM and

SARIMA based prediction models with SSA and non-SSA step on urban under both normal and incident traffic condition.

2 Background

2.1 Previous traffic prediction

A large number of statistical and data mining methods for traffic prediction have been published over the last three decades. There are many alternative ways of classifying these methods, such as dividing them based on the type of machine learning tools: e.g. parametric and non-parametric techniques (Vlahogianni et al.,2004), classifying based on modeling approaches, namely data driven, model-based and instantaneous approaches (van Lint, 2004), dividing into two categories based on the traffic conditions: prediction methods under normal and abnormal traffic conditions (Guo et al., 2010), or categorising based on the type of information used to model the recurrent traffic process (Krishnan & Polak, 2008). A brief classification of prediction methods used within transport engineering is summarised as follows:

- *Traffic model based method*: uses the simulation of the traffic system itself including the traffic flow, road network and signal control plan. This method considers the detailed simulation of the activities and decision making of drivers on the road network. Microscopic traffic models focus on the prediction of individual vehicle trajectories based on assumptions of driver-behaviour (Ben-Akiva et al., 1998). Macroscopic traffic prediction models centre on the prediction of a stream of traffic based on analogies of vehicular traffic flow with fluid and gas-dynamic (van Lint, 2004). A lot of cities take this approach to predict traffic variables, such as New York, Madrid, Beijing and Southampton that are trying to use this approach to predict traffic parameters.
- *Statistical method:* uses the statistical relationship in the training dataset to select the optimal parameters in the fitting procedure. Future data can be predicted based on the model built using training dataset. A lot of statistical methods on the accurate short-term prediction of traffic variables have been proposed, such as linear model (Sun et al., 2003) and the ARIMA model (Ahmed & Cook, 1979; Hamed et al., 1995; Williams & Hoel, 2003).
- *Data mining based method:* is defined as “an essential process where intelligent methods are applied in order to extract data patterns”(Han & Kamber, 2006). These methods search a set of historical observations from records similar to the current conditions and use these to estimate the future state of the system. Numerous data mining based models have been developed for short-term traffic prediction, such as k-Nearest Neighbours (kNN) (Clark, 2003; Davis & Nihan, 1991; Guo et al., 2010; Krishnan & Polak, 2008; Oswald et al., 2001; Smith & Demetsky, 1997), kernel method (El Faouzi, 1996), Support Vector Regression (Wu et al., 2004) and Neural Network(Park & Rilett, 1999).

2.2 Singular Spectrum Analysis (SSA)

Singular Spectrum Analysis (SSA) was firstly proposed as an adaptive noise-reduction algorithm based on Karhunen-Loeve transform (Sivapragasam et al., 2001). SSA is also used as a pre-processing data method, because it can decompose an original time series to a smoothed trend curve and a noise series based on the spectrum analysis of the input original time series. It is widely used in the rainfall and runoff prediction, not transportation. Simões et al. (2011) used SSA method to extract smoothed components in the study of rainfall time series and applied support Vector Machine (SVM) technique to prediction. Sivapragasam et al. (2001) used SSA as a data pre-processing step in runoff forecasting.

The basic SSA algorithm has two main stages: decomposition and reconstruction. The decomposition stage includes the embedding step and singular value decomposition (SVD). The step of embedding is to construct the trajectory matrix. Singular value decomposition (SVD) of the trajectory matrix turns into the decomposed trajectory matrices based on their singular values. The stage of reconstruction generates subgroups of the decomposed trajectory matrices based on the results of SVD and calculates the diagonal averaging to reconstruct a new time series as an additive component of the initial series. A detailed explanation of the SSA method can be found in Chapter 2 of Hassani (2007).The SSA method can be summarised in the following four steps:

Step1: This step is an embedding step that transfers the original time series to the trajectory matrix. Let $X = \{x_i, i = 1,2, \dots, N\}$ be the initial daily data. Time series X is mapped into L lagged vectors, where the value of L is the embedding dimension. The trajectory matrix T_X is written as

$$T_X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_{K+L-1} \end{pmatrix} \quad (1)$$

where $K = N - L + 1$.

Step 2: This step uses Singular Value Decomposition (SVD) to change the trajectory matrix generated in the Step 1 into a decomposed trajectory matrix. Applying SVD to the trajectory matrix, the matrix T_X is decomposed into $T_X = UDV^T$, where U and V are the left and right eigenvectors and $D = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_L)$ is a diagonal matrix. The corresponding singular values is $\{\sigma_i = \sqrt{\lambda_i}, \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L > 0, i = 1,2, \dots, L\}$, where λ_i is the eigenvalue. Hence, the trajectory matrix T_X can be written as

$$T_X = \sum_{i=1}^{L} u_i \sigma_i {v_i}^T. \quad (2)$$

Step 3: The decomposed trajectory matrix will be reconstructed in this step. This step is the grouping step and corresponds to splitting the matrices, computed at the SVD step, into several groups and summing the matrices within each group. The grouping turns a partition of the set $\{1,2, \dots, L\}$ into the collection of m disjoined subsets of $I = \{I_1, I_2, \dots, I_m\}$. Thus,

$$\begin{aligned} T_X &= f_{SVD}(T_1, T_2, \dots, T_L) \\ &= f_{Grouping}(T_{I_1}, T_{I_2}, \dots, T_{I_m}) \end{aligned} \quad (3)$$

Assume that there are only two groups of the trajectory matrix, namely T_L and T_R, where $L \cup R = I$ and I is the entire set. Hence, $T_L + T_R = T_I$. $T_I = \sum_{i \in I} u_i \sigma_i {v_i}^T$ and $T_L = \sum_{i \in L} u_i \sigma_i {v_i}^T$.

Step 4: A new time series is transformed by the grouped matrices obtained in Step 3. The corresponding operation is called diagonal averaging. It is a linear operation and maps the trajectory matrix of the initial series into the initial series itself. In this way, a decomposition of the initial series into several additive components can be obtained.

2.3 Grey System Model (GM)

The GM based method predicts the future values of a time series based only on a set of the recently observed data depending on the window size of the predictor (Kayacan et al., 2010). It is assumed that all data values to be used in grey models are positive, and the sampling frequency of the time series is fixed. From the simplest point of view, grey models which will be formulated below can be viewed as curve fitting approaches. Grey system based method is an alternative approach in time series prediction. However, this method is widely used in financial domain not in transportation. Kayacan et al. (2010) summarised the theory and applications of grey system-based models and applied it to the prediction of the foreign currency exchange rates. Chang & Tsai (2008) used a grey system model trained by Support Vector Regression (SVR) method to predict equity volume index.

GM(1,1) type of grey model is the most widely used in the literature, as “Grey Model First Order One Variable”. This model is a time series forecasting model. The differential equations of the GM(1,1) model have time-varying coefficients. In other words, the model is renewed as the new data become available to the prediction model. When the order of the difference equation $n = 1$, a grey system model can easily obtain the predicted value and the result of this first-order grey differential equation is an exponential curve.

The GM(1,1) model can only be used in non-negative data sequences (Deng, 1989). $\{x^{(0)}(k), k = 1,2, \dots, n \text{ and } n \geq 4\}$ is the original positive sampling data. In order to reduce the randomness and improve the regularity, positive data sequences transfer to monotonically increasing sequence using Accumulating Generation Operator (AGO) (Deng, 1989). This method is described as follows

$$x^{(1)}(1) = x^{(0)}(1) \tag{4}$$
$$x^{(1)}(2) = x^{(0)}(1) + x^{(0)}(2) \tag{5}$$
$$\vdots$$
$$x^{(1)}(N) = x^{(0)}(1) + x^{(0)}(2) + \cdots + x^{(0)}(n) \tag{6}$$

or, the above equations can be summarised as

$$x^{(1)}(k) = \sum_{i=0}^{k} x^{(0)}(i), k = 1,2, \dots, n \tag{7}$$

It is clear that the new sequence $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n), n \geq 4\}$ is monotonically increasing, that is

$$x^{(1)}(1) \leq x^{(1)}(2) \leq \cdots \leq x^{(1)}(n) \tag{8}$$

When $n = 1$, the least square estimate sequence of the grey difference equation of GM(1,1) is easily defined as follows (Deng, 1989)

$$x^{(0)}(k) + az^{(1)}(k) = u, k = 2,3, \dots, n \tag{9}$$

where

$$z^{(1)}(k) = c * x^{(1)}(k) + (1 - c) * x^{(1)}(k - 1), 0 \leq c \leq 1 \tag{10}$$

c is the coefficient usually set by 0.5 ; Z is the mean value of adjacent data (Deng, 1989).

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = u \tag{11}$$

$[a, u]^T$ is a sequence of parameters that can be found as follows

$$[a, u]^T = (B^T B)^{-1} B^T Y \tag{12}$$

where $Y = \left[x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), \dots, x^{(0)}(n)\right]^T$

and $B = \left[\left(-z^{(0)}(2), 1\right), \left(-z^{(0)}(3), 1\right), \left(-z^{(0)}(4), 1\right), \dots, \left(-z^{(0)}(n), 1\right)\right]^T$.

The solution of $x^{(1)}(t)$ at time k in the above differential equation is

$$\hat{x}^{(1)}(k + 1) = \left[x^{(0)}(1) - \frac{u}{a}\right] e^{-ak} + \frac{u}{a}. \tag{13}$$

In the beginning, AGO method is used to generate an increasing sequence. Hence, the Inverse Accumulating Generation Operator (IAGO) method is applied to find the prediction valued of original data (Deng, 1989).

$$\begin{aligned} \hat{x}^{(0)}(k + 1) &= \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k) \\ &= \left[x^{(0)}(1) - \frac{u}{a}\right] e^{-ak} - \left[x^{(0)}(1) - \frac{u}{a}\right] e^{-a(k-1)} \\ &= \left[x^{(0)}(1) - \frac{u}{a}\right] e^{-ak}(1 - e^{a}). \end{aligned} \tag{14}$$

2.4 Seasonal Auto-Regressive Integrated Moving Average (SARIMA)

Auto-Regressive Integrated Moving Average (ARIMA) time-series model, which is known as Box-Jenkins approach (Box et al., 1994; Chatfield, 2004), is one of the most commonly used parametric models in time series analysis. This method applies a statistical way to obtain the information from the past and current data of a series. Then it uses this information to predict the future values. The Box-Jenkins approach is generally referred to as an ARIMA model. Given a time series $\{X_t\}$, white noise series $\{e_t\}$, and the backshift operator B the ARIMA(*p, d, q*) structure is defined as

$$\phi_p(B)(1 - B)^d X_t = \theta_q(B) e_t \tag{15}$$

where d is the order of differencing; ϕ_p, $\theta_q(B)$ are polynomials of order p, q respectively, such that

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 -, \dots, \phi_p B^p \tag{16}$$

and

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 -, \dots, \theta_q B^q \tag{17}$$

Exploiting the recurrence of traffic data, a Seasonal ARIMA (SARIMA) model was used byWilliams & Hoel (2003). When seasonal terms are included, an SARIMA*(p, d, q)(P, D, Q)*$_S$ model is defined as

$$\phi_p(B)\Phi_P(B^S)(1-B)^d X_t = \theta_q(B)\Theta_Q(B^S)e_t \tag{18}$$

where Φ, Θ, P, D and Q are the seasonal counterparts of ϕ, θ, p, d and q respectively; S denotes the seasonality. A SARIMA model can provide linear state transition equations which can be applied to recursively produce single and multiple interval predictions.

3 Prediction model

Model 1: is the prediction model without data pre-processing step. This model applies GM(1,1) and SARIMA to the prediction of future traffic variables.

Model 2: uses SSA method as the pre-processing of original traffic data. The original raw traffic series f_t is decomposed into two series: a smoothed series f_s and its residuals f_r. The prediction results of the smoothed series by GM(1,1) or SARIMA prediction method is $\hat{f}_{s+15}$. At the same time, the residual series $\hat{f}_r$ is estimated using the historical average value of the historical residual. The final prediction result is the sum of the smoothed series prediction and the characteristic curves, that is $\hat{f}_{t+15} = \hat{f}_{s+15} + \hat{f}_r$. This traffic prediction model is summarised in the Figure 1.

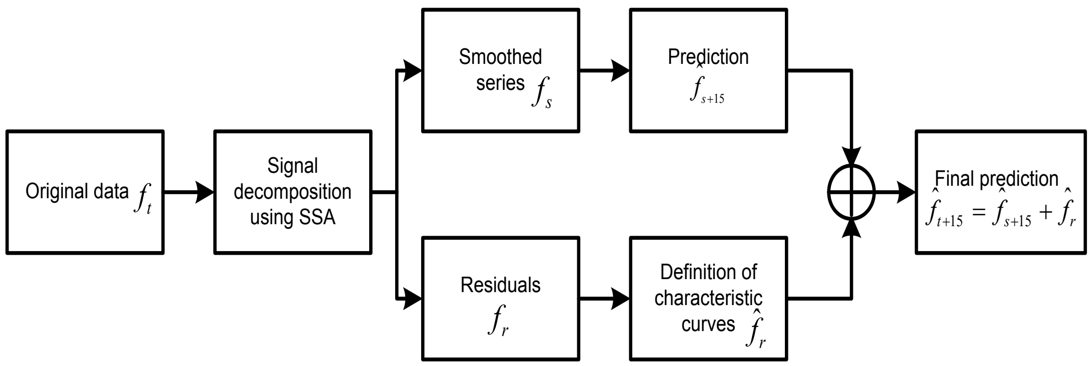

Figure 1. Structure of two-step prediction model using SSA for de-noising

4 Study area and traffic data

All traffic data used in this study are obtained from Inductive Loop Detectors (ILDs) in London as a part of the SCOOT traffic control system (Hunt et al., 1981). The values of 15-minute traffic flow and occupancy are extracted from the ASTRID system (Hounsell & McLeod, 1990) associated with SCOOT. There are over 6000 ILDs in London that provide near real-time traffic data for all the major links. Thus, SCOOT ILD data could be widely used in the applications of traffic estimation and prediction for arterial roads in London due to its comprehensive spatial and temporal coverage (Krishnan, 2008) . Traffic data used in this paper are from the Marylebone Road corridor in the centre of London, as shown in Figure 2.

Information about abnormal traffic conditions used in this paper is obtained from the British Broadcasting Corporation (BBC). BBC provides a traffic information service on the web, where incident information can be broadly classified into planned events and unplanned incidents (Hu et al., 2008). Planned incident information is provided by organisations such as local boroughs, the police, utility, companies and event organisers. Information about unplanned incidents is mainly obtained from Transport for London staff who monitor Closed Circuit television (CCTV) cameras and the police who is informed by the public about accidents and other disruptions(Hu et al., 2008). The traffic information service from the BBC can be used to identify the location, duration and the degree of severity of each incident.

The testing period is from the 5th June to 20th June, 2008. There are 12 days left after filtering out the error data of detector device faults using Daily Statistics Algorithm (DSA) (Chen et al., 2003; Robinson, 2005) and weekend data. In the testing dataset, the first 11 days data are normal, non-incident traffic data. A severe traffic incident happened on the last testing day, 20th June 2008. The accident period is around 18:59:00 to 21:01:21. Accident location is near to the intersection of Macfarren Place and Marylebone Road (the point A in Figure 2).

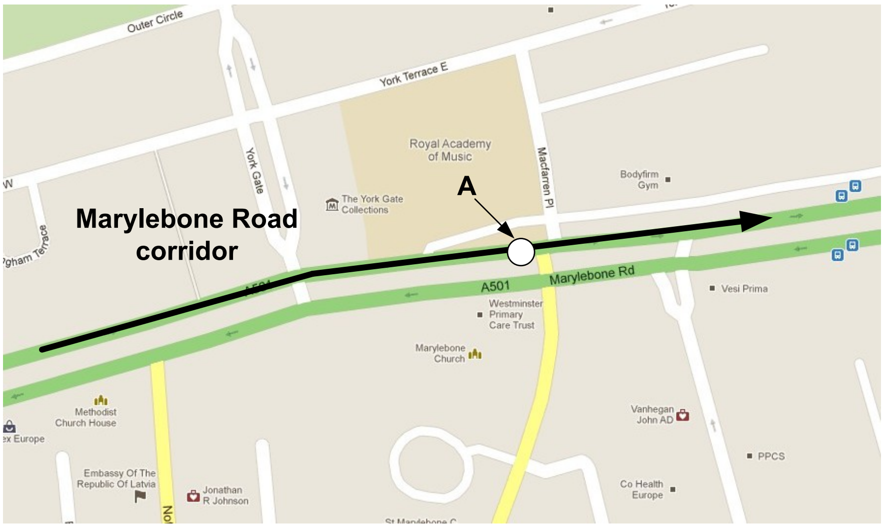

Figure 2. Locations of incident and selected path on Marylebone Road corridor (Source: Google Maps)

5 Prediction results and analysis

5.1 Prediction accuracy measurement

The prediction accuracy is evaluated using three goodness-of-fit measures, namely Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). These measures are employed as follows:

Mean Percentage Error (MPE):

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{f_i-\hat{f_i}}{f_i}\right] \tag{19}$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|f_i-\hat{f_i}|}{f_i}\right) \tag{20}$$

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\sum_{i=1}^{n}\left(\frac{|f_i-\hat{f_i}|^2}{n_i}\right)} \tag{21}$$

here,

f_{i}: the actual traffic flow for observation

$\hat{f}_i$: the predicted traffic flow

n: the number of predictions.

5.2 Prediction results

Normal traffic condition: The first 11 days of testing dataset on Marylebone road corridor are normal, non-incident traffic data. Both GM and SARIMA were implemented in this scenario. One significant advantage of GM method is that it does not require training process. Using the criteria of Akaike information criterion (AIC) in the fitting process, SARIMA$(1,0,1)(0,1,1)_{96}$ model is selected for traffic prediction in this study. Table 1 shows the prediction accuracy of *Model 1* - without data preprocessing step and *Model 2* - using SSA for two GM and SARIMA using traffic data from Marylebone road corridor under normal traffic conditions. In the model without SSA part, the prediction accuracy of the GM based method is better than that of the SARIMA based method using MAPE and RMSE metric, with the MAPE value of 10.00% and the RMSE value of 134.62veh/h. It also can be seen that the use of data pre-processing improves the prediction accuracy with both GM and SARIMA methods in Model 2. The GM based method still performs better than SARIMA with the data pre-processing structure.

Table 1. Prediction accuracy of two methods under normal traffic conditions

	Methods	MPE(%)	MAPE(%)	RMSE(veh/h)
Model 1: without data pre-processing	GM	3.07	10.00	134.62
	SARIMA	1.72	10.62	146.07
Model 2: with data pre-processing	SSA-GM	3.18	9.57	125.47
	SSA-SARIMA	1.71	10.13	131.92

Incident traffic condition: two models with two prediction methods were tested when an unexpected traffic incident occurred. A serious traffic incident happened in the last day of the testing dataset, 20th June, 2008. Table 2 shows the prediction accuracy of Model 1 and Model 2 under incident traffic conditions. It is clear to see that Model 2 with SSA as the data pre-processing improve the prediction during incident. For GM method, the MAPE value improves from 22.97% to 21.82%; for SARIMA the MAPE value improves from 37.47% to 36.90%. The GM method with SSA has the best prediction accuracy.

Table 2. Prediction accuracy of two methods under incident traffic conditions

	Methods	MPE(%)	MAPE(%)	RMSE(veh/h)
Model 1: without data pre-processing	GM	8.12	22.97	225.62
	SARIMA	25.94	37.47	330.67
Model 2: with data pre-processing	SSA-GM	7.57	21.82	205.27
	SSA-SARIMA	25.50	36.90	319.97

Figure 3 shows the scatter-plot of predicted and observed traffic flows, the auto-correlation plot of predictions, the histogram of error distribution and sample time-series plot between predicted and observed flows for 15-minute ahead prediction in Model 1 using GM method during incident.

Figure 4 presents the prediction performance in Model 1 using SARIMA during traffic incident. Figure 5 is the prediction performance using GM and SSA method during incident. The prediction bias of SSA-GM slightly reduced compared with the model without data pre-processing. Figure 6 shows the prediction results of SSA-SARIMA method. The prediction bias also reduced using the structure of data pre-processing. The prediction bias of GM is much lower than SARIMA.

The incident happened around 18:59:00 that is about the 68th time lag in the last figure of Figure 3, 4 ,5 & 6 and cleared around 21:01:21, the 84th in this figures. It can be seen that traffic flow significantly dropped during this traffic incident period. The SARIMA based method did response on this sudden traffic flow change, because the parameters were chosen using training set without incident. The GM base method can much more quickly detect this drop than the SARIMA method. In summary, the GM based method outperforms SARIMA under incident conditions.

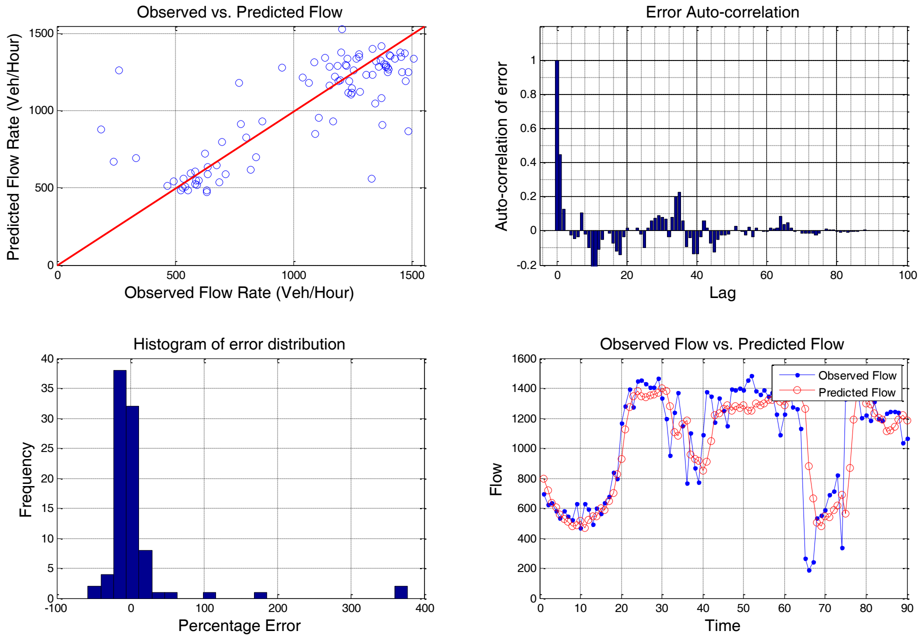

Figure 3. Prediction performance of Model 1 using GM during incident

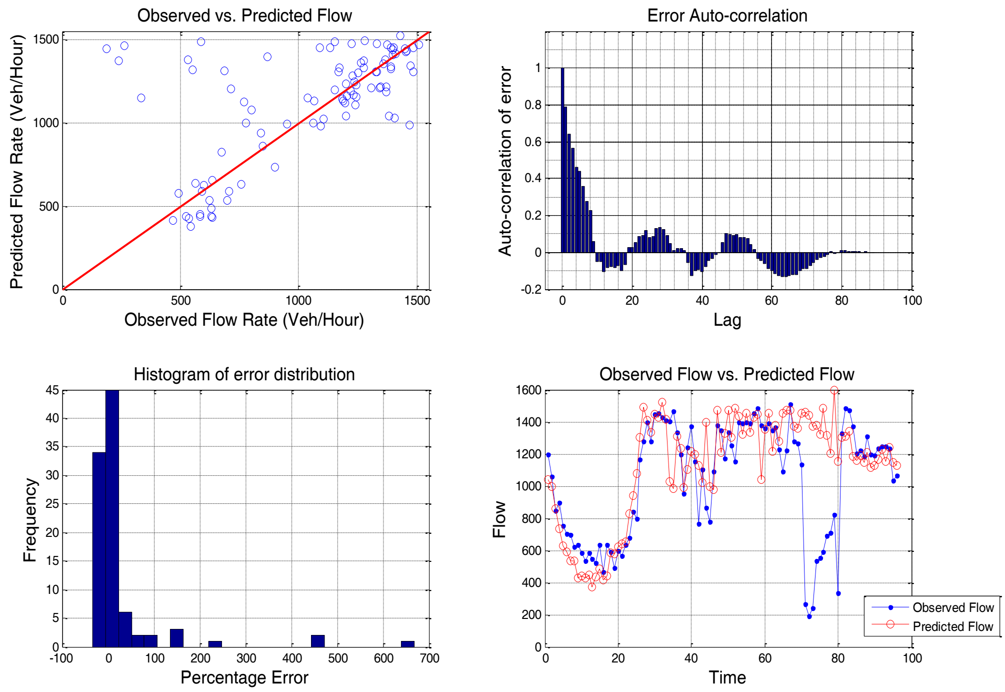

Figure 4. Prediction performance of Model 1 using SARIMA during incident

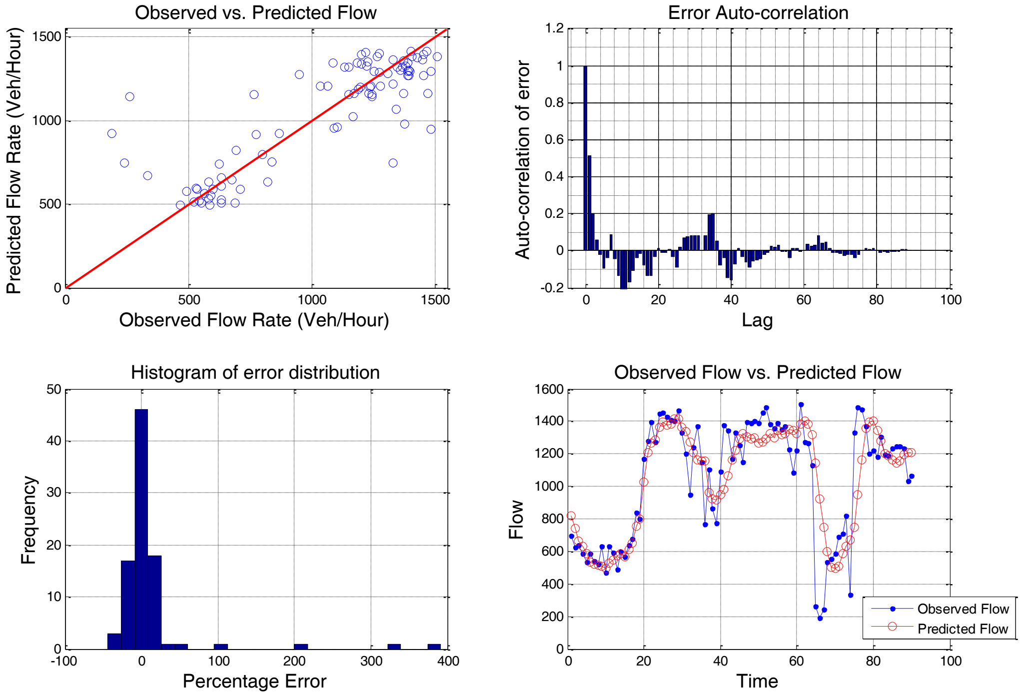

Figure 5. Prediction performance of Model 2 with SSA using GM during incident

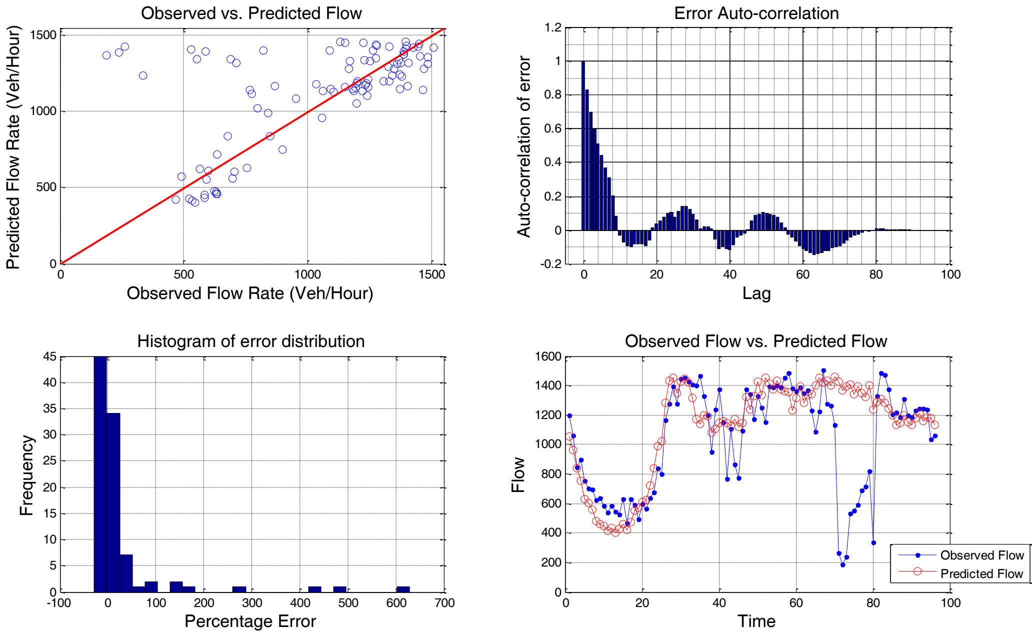

Figure 6. Prediction performance of Model 2 with SSA using SARIMA during incident

6 Conclusion

A new model structure with data pre-processing compared with traditional prediction methods was presented in this paper. Two prediction methods GM and SARIMA with two model structures with and without SSA were tested and their results were compared and evaluated in this paper. The structure of data pre-processing with SSA can improve the prediction accuracy using time series methods during both normal and incident conditions and there is an improvement in traffic prediction by a suitable pre-processing procedure. In addition, traditional GM has method has slightly better prediction accuracy than SARIMA under normal conditions on Marylebone Road corridor. However, the GM method has the better ability to detect and response to the sudden change of traffic patterns which was caused by a traffic incident. In contrast, the SARIMA based method is not very efficient in traffic prediction during abnormal conditions. Moreover, the significant advantage of GM method is that it does not require a complex training and parameter optimisation process. SARIMA needs the fitting process to select optimal parameter sets. Therefore, the GM based prediction model with the structure of data pre-processing with SSA has the best prediction accuracy during incident in this paper.

References

Ahmed, M. & Cook, A. (1979) Analysis of freeway traffic time series data by using Box-Jenkins techniques. *Transportation Research Board,* 722, 1-9.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. & Mishalani, R. (1998) DynaMIT: a simulation-based system for traffic prediction and guidance generation. *TRISTAN III, San Juan, Porto Rico*.

Box, G. E. P., Jenkins, G. M. & Reinsel, G. C. (1994) *Time Series Analysis: Forecasting and Control.* Prentice Hall, Englewood Cliffs, NJ.

Chang, B. R. & Tsai, H. F. (2008) Forecast approach using neural network adaptation to support vector regression grey model and generalized auto-regressive conditional heteroscedasticity. *Expert Systems with Applications,* 34, 925-934.

Chatfield, C. (1996) *The Analysis of Time Series: An Introduction.* Chapman and Hall, London.

Chen, C., Kwon, J., Rice, J., Skabardonis, A. & Varaiya, P. (2003) Detecting errors and imputing missing data for single-loop surveillance systems. In: *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*, Washington D.C,USA. 160-167.

Clark, S. (2003) Traffic prediction using multivariate nonparametric regression. *Journal of Transportation Engineering-ASCE,* 129, 161-168.

Davis, G. A. & Nihan, N. L. (1991) Nonparametric regression and short-term freeway traffic forecasting. *Journal of Transportation Engineering-ASCE,* 117, 178-188.

Deng, J. L. (1989) Introduction to grey system theory. *The Journal of Grey System,* 1, 1-24.

El Faouzi, N. E. (1996) Nonparametric traffic flow prediction using kernel estimator. In: *Proceedings of Internaional symposium on transportation and traffic theory*. 41-54.

Guo, F., Polak, J. & Krishnan, R. (2010) Comparison of modelling approaches for short term traffic prediction under normal and abnormal conditions. In: *Proceedings of the 13th International IEEE Annual Conference on Intelligent Transportation Systems*, Madeira Island, Portugal.

Hamed, M. M., Almasaeid, H. R. & Said, Z. M. B. (1995) Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering-ASCE,* 121, 249-254.

Han, J. & Kamber, M. (2006) *Data mining: concepts and techniques.* Morgan Kaufmann.

Hassani, H. (2007) Singular spectrum analysis: methodology and comparison. *Journal of Data Science,* 5, 239-257.

Hounsell, N. & McLeod, F. (1990) *ASTRID Automatic SCOOT Traffic Information Database*. Technical Report Contractor Report, Transport and Road Research Laboratory, Department of Transport. 0266-7045.

Hu, J., Krishnan, R. & Bell, M. G. H. (2008) TPEG feed from the BBC: A potential source of ITS data? In: *Proceedings of the 13th International Conference on Road Transport Information and Control*, Manchester, UK. IET.

Hunt, P. B., Robertson, D. I., Bretherton, R. D. & Winton, R. I. (1981) *SCOOT-a traffic responsive method of coordinating signals*. Transport and Road Research Laboratory,

Kayacan, E., Ulutas, B. & Kaynak, O. (2010) Grey system theory-based models in time series prediction. *Expert Systems with Applications,* 37, 1784-1789.

Krishnan, R. (2008) *Travel time estimation and forecasting on urban roads*. PhD Thesis. Centre for Transport Studies, Imperial College London.

Krishnan, R. & Polak, J. W. (2008) Short-term travel time prediction: An overview of methods and recurring themes. In: *Proceedings of the Transportation Planning and Implementation Methodologies for Developing Countries Conference (TPMDC 2008)*, Mumbai, India.

Oswald, R. K., Scherer, W. T. & Smith, B. L. (2001) *Traffic flow forecasting using approximate nearest neighbor nonparametric regression*. UVACTS-15-13-7.

Park, D. & Rilett, L. R. (1999) Forecasting freeway link travel times with a multilayer feedforward neural network. *Computer-Aided Civil and Infrastructure Engineering,* 14, 357-367.

Robinson, S. (2005) *The development and application of an urban link travel time model using data derived from inductive loop detectors*. PhD thesis. Centre for Transport Studies, Imperial College London.

Simões, N., Wang, L., Ochoa, S., Leitão, J. P., Pina, R., Onof, C., Sá Marques, A. & Maksimović, Č. (2011) A coupled SSA-SVM technique for stochastic short-term rainfall forecasting. In: *12th International Conference on Urban Drainage*, Porto Alegre,Brazil.

Sivapragasam, C., Liong, S. Y. & Pasha, M. (2001) Rainfall and runoff forecasting with SSA-SVM approach. *Journal of Hydroinformatics,* 3, 141-152.

Smith, B. L. & Demetsky, M. J. (1997) Traffic flow forecasting: Comparison of modeling approaches. *Journal of Transportation Engineering-ASCE,* 123, 261-266.

Sun, H., Liu, H. X., Xiao, H., He, R. R. & Ran, B. (2003) In *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*Washington D.C, USA.

van Lint, J. W. C. (2004) *Reliable travel time prediction for freeways*. PhD Thesis. Delft University of Technology, Delft, The Netherlands.

Vlahogianni, E. I., Golias, J. C. & Karlaftis, M. G. (2004) Short-term traffic forecasting: Overview of objectives and methods. *Transport reviews,* 24, 533-557.

Williams, B. M. & Hoel, L. A. (2003) Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *Journal of Transportation Engineering-ASCE,* 129, 664-672.

Wu, C. H., Ho, J. M. & Lee, D. T. (2004) Travel-time prediction with support vector regression. *IEEE Transactions on Intelligent Transportation Systems,* 5, 276-281.