

# **RANDOM PROCESS MODEL FOR TRAFFIC FLOW USING A WAVELET - BAYESIAN HIERARCHICAL TECHNIQUE**

Bidisha Ghosh, Biswajit Basu, and Margaret O'Mahony

Department of Civil, Structural and Environmental Engineering

Trinity College, Dublin

Dublin 2

## **ABSTRACT** (148 words)

The paper proposes a random process traffic volume model based on wavelet analysis (WA) and Bayesian Hierarchical methodology (BHM). The model has potential applications at junctions where a regular traffic condition data collection system (such as a loop-detector) is not established or is malfunctioning for a considerable period of time. Unlike typical short-term forecasting algorithms, large traffic flow datasets including information on current traffic scenarios are not required for the proposed model. A non-functional daily trend of urban traffic flow observations is established incorporating discrete WA where the variations in the traffic flow over the trend are modelled using BHM. The time-varying variability (within day) of the urban traffic flow observations from an intersection has been established to be an intrinsic character of the data – it is not clear what you are trying to say here – if you mean what it states, this is nothing new. A case study has been performed at two busy junctions in the centre of Dublin to validate the effectiveness of the strategy – which strategy?.

## 1 INTRODUCTION

Continuous traffic condition data collection from a large number of junctions in an urban transport network is important for use in the Urban Traffic Control (UTC) systems of today. Although UTC systems are equipped with sophisticated computational and mechanical tools, it is often possible to have several junctions in an urban transport network where traffic conditions (e.g. volume, speed or density) data is not collected regularly (minor intersections). In addition, there can be cases in a network where the data collection system (loop-detector or video imaging system) is not working or is out of service for a considerable period of time. A possible solution to such problems is the development of a traffic volume model which is not directly dependent on recent past and real-time information on traffic flow observations.

This paper develops a random process traffic volume simulation model using wavelet analysis (WA) and Bayesian Hierarchical methodology (BHM). The potential application of the proposed simulation model depends on a novel and unique concept it is not clear to me on the basis of what you have written here that it is unique – can you be more explicit about why it is unique and different to the work of others please of modelling the daily trend underlying the traffic flow over a day in a non-functional manner using discrete WA and subsequently capturing the time-varying variations of the traffic flow observations over the trend implementing a BHM.

The proposed discrete WA based non-analytic trend modelling captures the long-term change of the mean level of traffic flow observations in an efficient and flexible manner. In transport modelling WA was introduced primarily for developing automatic incident detection algorithms [Samant and Adeli (2000), Adeli and Samant

(2000), Adeli and Karim (2000), Karim and Adeli (2002), Teng and Qi (2003) and Karim and Adeli (2005)]. It is only recently that researchers in Intelligent Transport Systems (ITS) have shown interest in using this technique for extracting useful information from the existing archived data related to traffic conditions. The traffic flow time-series observations (non-stationary data) are required to be transformed to a stationary process for being modelled using conventional time-series techniques (Chatfield, 2004). The advantage of the application of wavelet analysis is in the inherent capability of accounting for non-stationary time-series. In wavelet analysis based traffic flow modelling, the multi-resolution analysis (MRA) technique (Mallat, 1989) has the potential to be used. This technique, with the advantage of a fast computational algorithm, has been used in feature or pattern extraction at different scales in data/image analysis and can be used as well for multi-scale forecasting of traffic flow. The technique of de-noising using wavelet pre-processing (Sun et al. 2004) has been used by a few researchers for increasing the efficiency of existing short-term traffic flow forecasting models. Chen et al. (2004) combined a wavelet transform with the Markov model to forecast traffic volume. In traffic pattern modelling, varied studies on optimized aggregation level (Qiao et al. 2003); data reduction (Venkatanarayana et al. 2006) and mesoscopic-wavelet model (Ghosh-dastidar and Adeli 2003) have been carried out. Due to very limited research on wavelet based traffic flow modelling in highly congested metropolitan areas there is scope for further exploration in this field.

The use of Bayesian statistics is quite recent in the field of traffic flow modelling and forecasting. Some work has been done in short-term traffic flow forecasting using Bayesian networks (Zhang et al., 2004), Hierarchical regression models (Tebaldi et

al., 2002), Bayesian SARIMA models (Ghosh et al. 2007). No literature is available so far on the application of wavelet analysis and subsequent analysis of the randomness using BHM to model traffic flow.

## **2 TRAFFIC FLOW OBSERVATIONS**

The aim of the study in this paper is to develop a random process traffic volume model to simulate univariate traffic flow time-series data for urban signalised intersections. The proposed model is applied to several intersections in the centre of Dublin to describe and evaluate the methodology. The study at two representative junctions (TCS 183 and TCS 439) is presented in this paper. The univariate traffic flow time-series data used for modelling are obtained from the inductive loop-detectors embedded in the streets of both junctions as part of the urban traffic control (UTC) data collection system used in Dublin. A map of the junctions is given in figure 1. Junction TCS 183 is four-armed with one-way traffic on two approaches. Tara Street (one-way) has four lanes, with traffic flowing from south to north. The traffic volume passing through Tara Street, measured on the loop-detectors numbered 1, 2, 3, 4 are continuously recorded. The junction TCS 439 is a four-armed junction with one-way traffic on both – refers to two – may need to explain why this is not four approaches. The traffic volume passing through Townsend Street, measured on the loop-detectors numbered 1, 2, 3 is used in describing and evaluating the proposed traffic volume model.

The data aggregation interval is unique to the data collection system of the existing urban traffic control system of any city. The data interval can vary from a few seconds to one hour. In short-term traffic flow forecasting models, data aggregate intervals

from 3 min. to 30 min. are used based on the forecasting algorithms (Xie et al. 2007). In view of the wavelet analysis techniques to be applied on the traffic volume observation a short time interval of 5 min. is chosen for this paper. In the case of an urban transport network, the weekend travel dynamics is inherently different from the travel dynamics during weekdays. In this study, the modeling is essentially carried out on the data observed during weekdays. A plot of the traffic flow data from both junctions in vehicles per hour (vph) against time is shown in figure 2.

### 3 NON-FUNCTIONAL TREND MODEL

#### 3.1 Multi-Resolution Wavelet Analysis (MRWA)

The wavelet transform provides a time-frequency representation of any signal/time-series data. The basis of wavelet analysis is decomposing a signal into shifted and scaled versions of the original (or *mother*) wavelet. MRWA uses techniques by which different frequencies of a signal are analyzed with different resolutions by using an efficient numerical algorithm.

The discrete wavelet transform (DWT) (Mallat, 1989) of a signal  $X(t)$  consists of the collection of coefficients

$$\left\{ \begin{array}{l} c_j(k) = \langle X, \varphi_{j,k}(t) \rangle \\ d_j(k) = \langle X, \psi_{j,k}(t) \rangle \end{array} \right\} j, k \in Z \quad \text{where, } Z=1,2,3\dots \quad (1)$$

where  $\langle *, * \rangle$  denotes inner product,  $t$  is time,  $\{d_j(k)\}$  are the detail coefficients at level  $j$  ( $j = 1,2,\dots,J$ ) and  $\{c_j(k)\}$  are the approximate coefficients at level  $J$ . The signal  $X$  to be analyzed is integral-transformed using a set of basis functions

$$\psi_{jk}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (2)$$

which is constructed from the mother-wavelet  $\psi(t)$  by a time-shift operation ( $k$ ) and a dilation operation ( $j$ ). The function  $\varphi_{jk}(t)$  is a time shifted version of the mother-wavelet scaling function  $\varphi_j(t)$ :  $\varphi_{jk}(t) = \varphi_j(t - k)$ .  $\varphi_j(t)$  is a low-pass function which can separate the low frequency component of the signal. Thus DWT decomposes a signal into a large timescale (low frequency) approximation and a collection of details at different smaller timescales (higher frequencies).

The original signal can be reconstructed back from the decomposed approximation and the detail components. Thus, the original signal can be represented as,

$$X(t) = A_J(t) + \sum_{j \leq J} D_j(t) \quad (3)$$

where,  $A_J(t)$  is the reconstruction of the approximation coefficients  $c_j$  at level  $J$  and  $D_j(t)$  is the reconstruction obtained from the detail coefficients  $d_j$  at level  $j$ . In the reconstructed approximation ( $A_J$ ) and in the reconstructed details obtained at each stage/level ( $D_1, D_2, D_3 \dots D_J$ ) the numbers of data points remain the same as the original dataset.

### 3.2 Trend Model

The ‘trend’ of a time-series data can loosely be defined as the ‘long-term’ change of the mean level of the data (Chatfield, 2001). In the daily trend model of the traffic flow observations from an urban signalised intersection, the word ‘long-term’ indicates stability over time on a daily basis. In this study, MRWA is used to develop a representative daily trend model underlying the traffic flow observations over a day (Ghosh et al., 2006).

In this study, DWT associated with the basis *Daubechies' 4* (db4) is used to decompose the signal (time-series traffic flow observations) into different time scales. Initially, the original signal is decomposed into approximation coefficients  $c_1$  (low frequency/fluctuations or variability) and detail coefficients  $d_1$  (high frequency/fluctuations or variability). The approximation coefficients  $c_1$  (relative low frequency components) are again decomposed to approximation coefficients  $c_2$  (low fluctuations) and detail coefficients  $d_2$  (high fluctuations) at the next level. This procedure is repeated for further decomposition. The aim of repeating the decomposition procedure is to find an optimum approximation level for extracting the trend in the data. The optimum approximation level is the one in which the reconstructed approximation coefficients,  $A_m$  ( $m$  is the optimum approximation level), are the optimal smoothed estimate of the traffic flow data which can truly represent the traffic flow pattern on an average day. This is essentially a de-noising technique in signal processing. The local variability in traffic flow observations due to signal control in the urban arterials is considered as the noise (for the mathematical treatment) in this methodology. The traffic flow pattern at any particular approach at any intersection in an urban transport network is similar for weekdays. However, there can be some day-to-day variability due to other factors like, the day of the week, accidents or recurrent congestion in some other part of the transport network etc. These factors are uncontrollable and cannot be modelled as such. So, to obtain a 'regular trend over an average day', the  $A_m$  values over some regular days (approximately 20 days in this study) are to be averaged for a single day. The average trend has a non-analytical functional form and has better flexibility in representing the mean traffic flow over an average day.

In the case study, this methodology of finding a 'regular trend over an average day' is applied to univariate traffic flow observations obtained collectively over each 5 minute interval from the loop detectors at intersections TCS 183 and TCS 439 in Dublin. The date 15<sup>th</sup> June 2005 is chosen arbitrarily as a regular weekday for the modelling. A plot of the traffic flow observations from the chosen sites on that day is given in figure 2, where the traffic flow data series shows a non-stationary nature. Wavelet analysis can handle non-stationarities unlike other time-series modeling techniques where this issue is required to be separately dealt with and transformation of the time-series data is required to be performed (Ghosh et al., 2005). The traffic flow observations over twenty days i.e. four weeks (as weekends are not included), in the month of June-July from the two chosen sites are then decomposed into three levels of resolution using MRA with *Daubechies' 4* wavelet basis function. The three different levels represent three different time scales. The wavelet coefficients for approximations and details are then reconstructed at all three levels. For any modelling purpose the reconstructed values of the approximation and detail coefficients are always used in this study. At each level, during decomposition the high frequency part of the data is separated from the low resolution or the low frequency part. The low frequency part at level three is quite smooth and can be used as a representative of the overall trend over a day in the traffic data. Hence, the third level is considered as the optimum level of decomposition to obtain optimum smoothed estimates of the times-series data. As an example, the approximations at level three, and the details at level one, two and three of the 5 minute aggregate traffic volume from intersection TCS 183 on 15<sup>th</sup> June 2005 are plotted in figure 3. To model a representative trend, level three approximations – do you mean three level approximations? If not, why choose level 3 for the approximation? - explain of the



traffic flow time-series observations over 20 days are taken. An average over 20 days of the level 3 reconstructed approximation coefficients for a single day is used to model the representative daily trend for the two chosen approaches from the two chosen intersections. Explain why? The selection of the average coefficients helps to reduce the effect of certain uncontrollable conditions as described before. In figure 4(A) and (B), the ‘regular trend over an average day’ is plotted over the traffic flow observations on 15th June 2005. From the graphs – which graphs? You will need to be specific as to what the reader should be looking at in the figures because it is not clear it can be observed that the simple trend provides a very good approximation of the traffic volume on any arbitrary day.

A plot of the residuals from junctions TCS 183 and TCS 439 is given in figure 5. The statistical analyses of the residuals are given in Table 1. Mention in the text some of the more important relevant statistics from this Table. The ‘non-functional daily trend model’ forms the skeleton of a background model for simulating traffic flow at an urban intersection. The modelling of the residual obtained after fitting the trend helps to establish a tight simulation interval.

## **4 RESIDUAL MODEL**

### **4.1 Crude Residual Model**

In statistical analysis, residuals with random variability are generally modelled as Gaussian distribution (Chatfield, 2004). For developing a crude model of the residuals, a histogram of the residual data on 15<sup>th</sup> June 2005 is plotted in figure 6(A). From the histogram of the residuals it is evident that the residuals can be approximately modelled as a normal distribution. The normal probability density fit of

the residuals is plotted on the histogram in figure 6(A). The normal density fit matches the histogram only crudely. A confidence interval on the ‘regular trend over an average day’ can be provided based on this residual model. The following equation is used in simulating the confidence limit of this background model based on ‘regular trend over an average day’.

$$\mathbf{y}_{sim} = \mathbf{A}_{trnd} + \boldsymbol{\mu}_{res} + \boldsymbol{\varepsilon}_{res} \quad \text{where, } \boldsymbol{\varepsilon}_{res} \sim N(0, \sigma) \quad (4)$$

where,  $\mathbf{y}_{sim}$  is the simulated traffic volume from the background model;  $\mathbf{A}_{trnd}$  is the average value of level three approximation obtained from the trend model;  $\boldsymbol{\mu}_{res}$  is the mean of the residual dataset and  $\boldsymbol{\varepsilon}_{res}$  is the random part of the residuals.  $\boldsymbol{\varepsilon}_{res}$  values are randomly generated from a normal distribution with mean zero and standard deviation  $\sigma$  which is the sample standard deviation of the residuals. A 95% confidence interval of the residuals is used to form the confidence limit of the background model.

In figure 6(A) and 6(B) the original traffic flow observations on 15th June 2005 are plotted along with simulated confidence intervals from the background model. Most of the observations fall within the confidence limits. This proves that the simplistic background model based on ‘regular trend over an average day’ can be used as an approximate traffic volume simulation model for the intersections where continuous information on traffic conditions is not available.

It is observed from figure 5 that the spread of the residual data points around the mean value is not uniform. The variability of the residual data in off-peak hours of early morning and late night is much less than the variability of the same during peak hours

of the day. The variability is the highest during the evening peak hours. This non-uniform variability signifies that the variance of the residual should not be estimated as a constant parameter for an entire day, but as a variable varying with the time of the day. A bayesian hierarchical model is introduced at this point as a standard statistical method to model the residues with its time varying variance.

## 4.2 Bayesian Hierarchical Residual Model

In simple words, the basic idea behind the Bayesian hierarchical model is to develop a parametric statistical model with parameters which are represented by other parametric statistical models (i.e. it is doubly stochastic in nature). The essence of hierarchical model is that the dependencies among variables in a statistical model can be defined more easily with a tree-like structure. In the case of a Bayesian hierarchical model, while calculating *posterior* density (Lee, 199x), *priors* which themselves depend on other parameters not included in the *likelihood* function can be accounted for.

In this study, the variance of the residual is dependent on time and has to be modelled accordingly using another parametric statistical model. Hence, the residual obtained from the 5 minute aggregate traffic flow observations on 15-06-2005, after fitting the trend model, are further modelled using a Bayesian hierarchical model to account for the time-varying nature of the variance of the residual dataset. If the  $\mathbf{R}$  is the vector of the residual, then in a normal hierarchical model,

$$\mathbf{R}_t \sim N(m, \sigma_t^2) \quad t = 1, 2, \dots, T \quad (5)$$

where,  $m$  is the sample mean of the residual on 15-06-2005 and  $\sigma_t$  is the standard deviation of the residual for each time instant denoted by a subscript  $t$ . As 5 minute aggregate traffic volume is modelled, the vectors  $\boldsymbol{\sigma}$  and  $\mathbf{R}$  are both of dimension

$\{T \times 1\}$  where  $T$  is the number of time intervals or time instants in a day (for 5 minute aggregate traffic flow observations,  $T = 288$ ). The variance  $\sigma^2$ , of the residual dataset  $\mathbf{R}$  changes with the time of the day. To model this time-varying variability of the variance, the following parametric distribution is proposed.

$$\log(\sigma) \sim N(\log(\mathbf{y}), \tau^2) \quad (6a)$$

$$\text{which leads to } \sigma \sim \text{LN}(\log(\mathbf{y}), \tau^2) \quad (6b)$$

As  $\sigma$  is always positive, a lognormal distribution is taken in equation 6 to ensure that all  $\sigma_t$  lie within  $(0, \infty)$ . The lognormal distribution for each  $\sigma_t$  is centred at  $y_t$  with standard deviation of  $\tau$ . The variances of the high resolution components (sum of level 1, 2 and 3 reconstructed detail coefficients) from the 20 day traffic flow observations calculated over each hour of a day are considered as the initial estimates of the standard deviation of the residual ( $y_t$ ) for that hour of the day. The elements of the vector  $\mathbf{y}$  are the same for all time instants within the same hour of the day. The values of the vector  $\mathbf{y}$  of dimension  $\{T \times 1\}$  are calculated from the 20 day dataset and these values are constant for traffic flow observations on an average day at the particular intersection TCS 183 considered in this study. Hence, the Bayesian hierarchical model developed in this study can be considered as applicable to any arbitrary day of the year. The variance  $\tau^2$  of the lognormal distribution of vector  $\sigma$  is assumed to follow a uniform distribution, within a range  $(0, k)$

$$\tau \sim U(0, k) \quad k < \infty \quad (7)$$

where,  $k$  is an arbitrary constant signifying the maximum limit of the values of  $\tau$ . The exact value of  $k$  does not influence the estimation process. In this study the equations 5, 6b and 7 define the Bayesian hierarchical model for the residuals obtained after

subtracting the regular average trend from the traffic flow observations on 15<sup>th</sup> June 2005.

In the Bayesian hierarchical model, the unknown parameters to be estimated are  $\sigma(\sigma_1, \sigma_2, \dots, \sigma_{288})$  and  $\tau$ . These unknown parameters are represented by a vector  $\xi = (\tau, \sigma_1, \sigma_2, \dots, \sigma_{288})^T$ . To estimate the vector  $\xi$ , the Bayesian estimation technique (Tanner 1996; Lee 1997) is to be used. For the Bayesian inference, the posterior density of the normal hierarchical model is

$$p(\xi|\mathbf{R}, \mathbf{t}) = p(\tau)L(\sigma|\mathbf{R}, \mathbf{t})L(\tau|\sigma, \mathbf{t}) \quad (8)$$

where,  $p(\xi|\mathbf{R}, \mathbf{t})$  is the posterior density of  $\xi$ ;  $L(\sigma|\mathbf{R}, \mathbf{t})$  is the likelihood function of  $\sigma$  and  $L(\tau|\sigma, \mathbf{t})$  is the likelihood function of  $\tau$ ;  $p(\tau)$  is the prior density of parameter  $\tau$ . According to equation 5,  $\mathbf{R}$  is assumed to follow a normal distribution. Hence, the likelihood function of  $\sigma$  given the vector  $\mathbf{R}$  and the time instant vector  $\mathbf{t}$  (unit time interval = 5 minute) is

$$L(\sigma|\mathbf{R}, \mathbf{t}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right) \quad (9)$$

Similarly according to equation 6b where  $\sigma$  is assumed to follow a likelihood function, the likelihood function of  $\tau$  given  $\sigma$ ,  $\mathbf{y}$  and  $\mathbf{t}$ ,

$$L(\tau|\sigma, \mathbf{t}) = \prod_{t=1}^T \frac{1}{\sigma_t \tau \sqrt{2\pi}} \exp\left[-\frac{(\log \sigma_t - \log y_t)^2}{2\tau^2}\right] \quad (10)$$

$p(\tau)$  is equal to a constant as the prior density of  $\tau$  is assumed as flat on the range  $(0, \infty)$ . Hence, the posterior density from equation 8 is

$$p(\xi|\mathbf{R}, \mathbf{t}) \propto \prod_{t=1}^T \frac{1}{\sigma_t} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right) \prod_{t=1}^T \frac{1}{\sigma_t \tau} \exp\left[-\frac{(\log \sigma_t - \log y_t)^2}{2\tau^2}\right] \quad (11a)$$

which yields,

$$p(\xi|\mathbf{R},\mathbf{t}) \propto \left(\frac{1}{\tau^T}\right) \prod_{t=1}^T \frac{1}{\sigma_t^2} \exp\left(-\frac{R_t^2}{2\sigma_t^2} - \frac{(\log \sigma_t - \log y_t)^2}{2\tau^2}\right) \quad (11)$$

$$\int_0^\infty \dots \int_0^\infty \left(\frac{1}{\tau^T}\right) \prod_{t=1}^T \frac{1}{\sigma_t^2} \exp\left(-\frac{R_t^2}{2\sigma_t^2} - \frac{(\log \sigma_t - \log y_t)^2}{2\tau^2}\right) d\tau d\sigma_1 \dots d\sigma_T \quad (12)$$

By integrating out the other unknown parameters except for the one whose distribution is to be estimated, the ‘marginal distributions’ of the each of the unknown parameters can be found out from the integral in equation 12. The computation of the marginal distributions of the unknown parameters in  $\xi$  involves evaluation of a complex integral with problems of high dimensionality. In Bayesian estimation, numerical integrations are often performed to compute the integrations for which the analytical solution is intractable. However, numerical integration may lead to too many approximations and may even become intractable for large models. In many high dimensional cases of Bayesian analysis, certain refinements of Monte Carlo integration methods such as ??? are often used (Carlin1996). There are different non-iterative and iterative variations of these refinements. *Markov Chain Monte Carlo* (MCMC) is the particular iterative variation of Monte Carlo method in which the simulated values are not in *iid* but are in a Markov chain. In summary, the goal of the MCMC is, given a target distribution  $\Pi(x)$ , a Markov chain  $\{x_n\}$  is required to be constructed whose limiting distribution is  $\Pi(x)$ . There are two popular MCMC algorithms, (i) Gibbs sampler (Geman and Geman 1984) and (ii) Metropolis-Hastings algorithm (Metropolis et al. 1953, Hastings 1970)

In the MCMC method, to simulate the marginal probability distributions for the unknown parameters in the vector  $\xi(\tau, \sigma_1, \sigma_2, \dots, \sigma_T)$ , given an initial condition  $(\tau^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)}, \dots, \sigma_T^{(0)})$  the following 289 steps are to be iterated ( $i$  denotes the number of iteration):

1. Sample  $\tau^{i+1}$  from  $p(\tau^{i+1} | \sigma_1^{(i)}, \sigma_2^{(i)}, \dots, \sigma_T^{(i)}, \mathbf{y}, \mathbf{t})$  using Gibbs sampler technique
2. Sample  $\sigma_1^{i+1}$  using Metropolis Hastings technique
- .
- .
289. Sample  $\sigma_T^{i+1}$  using Metropolis Hastings technique

The initial conditions  $(\tau^{(0)}, \sigma_1^{(0)}, \sigma_2^{(0)}, \dots, \sigma_T^{(0)})$  are as follows,

$$\begin{cases} \sigma^{(0)} = |\mathbf{y} - 0.5| \\ \tau^{(0)} \sim \text{Inverse-gamma} \left[ \frac{T-1}{2}, \frac{\sum_{i=1}^T (\log \sigma_i^{(0)} - \log y_i)^2}{2} \right]^{-1} \end{cases} \quad (13)$$

In step 1, the Gibbs sampler technique is used to simulate the distribution of  $\tau$ . From the posterior density in equation 12, a full conditional distribution for  $\tau$  is as follows,

$$p(\tau | \sigma, \mathbf{R}, \mathbf{t}) \propto \left( \frac{1}{\tau^T} \right) \exp \left( - \frac{\sum_{i=1}^T (\log \sigma_i - \log y_i)^2}{2\tau^2} \right) \quad (14)$$

The full conditional distribution of  $\tau$  can be observed as an inverse gamma distribution with parameters

$$\alpha = T/2 - 0.5 \quad (15a)$$

$$\beta = 0.5 \sum_{i=1}^T (\log \sigma_i - \log y_i)^2 \quad (15b)$$

where the density function of the inverse gamma distribution is as follows,

$$p(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{-(\alpha+1)} \exp\left(-\beta/\tau\right) \quad (15c)$$

For simplicity the superscript denoting the number of iterations is not used in equations 14 and 15. Steps 2 to 289 are similar in nature and are used to simulate the values of  $\sigma_1, \sigma_2, \dots, \sigma_T$  in each iteration. The simulation of  $\sigma$  is done using Metropolis-Hastings technique. The candidate values of each of the elements of the vector  $\sigma$  are simulated from the following proposal distribution,

$$\sigma^i \sim \text{LN}\left[\log(\mathbf{y}), (\tau^i)^2\right] \quad (16)$$

The simulated value of  $\sigma^i$  at each iteration is accepted following the Metropolis algorithm. According to this algorithm, each simulated value of the elements of the vector  $\sigma^i$  in each iteration is accepted with a probability

$$\hat{p}(\sigma_t) = \frac{p\left(\sigma_t^i \mid \tau^i, \sigma_1^{i-1}, \sigma_2^{i-1}, \dots, \sigma_{t-1}^{i-1}, \sigma_{t+1}^{i-1}, \dots, \sigma_T^{i-1}, \mathbf{R}, \mathbf{t}\right)}{p\left(\sigma_t^{i-1} \mid \tau^i, \sigma_1^{i-1}, \sigma_2^{i-1}, \dots, \sigma_{t-1}^{i-1}, \sigma_{t+1}^{i-1}, \dots, \sigma_T^{i-1}, \mathbf{R}, \mathbf{t}\right)} \quad (17)$$

or 1 whichever is minimum, with the an ?? acceptance criteria (Ghosh et al. 2007). The steps 1 to 289 are repeated for 10000 times to simulate 10000 values for all the unknown parameters. The simulated values of  $\tau$  for both the junctions are plotted in two graphs in figure 7(A). The simulations show high convergence towards a constant



value of about 0.4 and 23 for junction TCS 183 and TCS 439 respectively. Comment on large difference – why so? In the case of the vector of time varying standard deviation  $\sigma$ , all of the two eighty-eight elements of the vector are simulated separately. Instead of showing a plot of 10000 simulated values of each  $\sigma_i$ , the mean of the simulated values of each  $\sigma_i$  are shown in figure 8(A) and 8(B). The sample standard deviation obtained from the previous Gaussian noise model of the residual is shown as a horizontal line in the same figure. What you are presenting in these figures is not clear – can you be very explicit about what you are trying the reader to interpret from figures 8(a) and 8(b) and explain this in the text. The estimates of  $\sigma$  obtained from the Bayesian hierarchical model change with the time of the day. The estimates during the peak hours are much more than the estimated values  $\sigma_i$  during the rest of the day and the estimates during the early hours in the morning are the lowest of all. You are presenting very large differences here without sufficient explanation – can you expand on what the differences mean please. The nature of the variability of the variance of the residual conforms to the spread of the residual data points in figure 5 – it is not clear how they conform? .

## 5 ANALYSES OF RESULTS

In this section a discussion of the 5 minute aggregate traffic flow simulations for junctions TCS 183 and TCS 439 on 15<sup>th</sup> June 2005 obtained by using the WA-BHM and the WA-Gaussian Noise Models is presented.

In figures 6(A) and 6(B), a 95% confidence interval for traffic flow data from TCS 183 and TCS 439 on 15<sup>th</sup> June 2005 is simulated using the WA-Gaussian Noise Model. From the graph, it can be observed that in both cases, the simulated intervals follow the nature of the traffic but fail to include all the observed data points during

evening and morning peak hours. To illustrate the effectiveness of the WA-BHM model, a 95% confidence interval similar to the Gaussian noise model is constructed on the regular average trend for traffic flow data from intersections TCS 183 and TCS 439 on 15<sup>th</sup> June 2005. The original 5 minute traffic flow observations on 15<sup>th</sup> June 2005 are plotted along with the simulated 95% confidence limit in figure 9(A) and 9(B) respectively. Unlike the WA-Gaussian Noise model, all the traffic flow observations on 15-06-2005 – use the same date format throughout the paper fall within the simulated confidence limits. Being a Bayesian method, the confidence limits adapt according to the variability of the residual data. This adaptation proves most effective during the evening peak hours where most of the observations fall outside the simulated limits from the Gaussian noise model. The number of traffic flow observations obtained from TCS 183 and TCS 439 falling outside the simulated interval for both the models are given in Table 2.

The aim of constructing a simulated traffic volume interval lies in including the extremes of the original traffic volume observations within the simulated limits. Hence, the upper limit of the simulated traffic volume range should consistently account for the maxima points in the traffic flow data while the lower limit should account for the minima points. The maxima and minima points of the traffic flow observations on 15-06-2005 from TCS 439 and TCS 183 are plotted along with the simulated traffic volume ranges from WA-Gaussian Noise model in figure 10(A), 10(B) respectively. The same for the WA-BHM model are plotted in figures 11(A) and 11(B). By comparing the plots, it can be seen that the simulated traffic range from the WA-BHM model can better match the rapid variability of traffic flow at any busy urban signalised intersection. The error estimates of the extreme points over a day and during evening and morning peak hours are given in Table 2. The maxima and

minima point errors over a day are seen to be higher than the same during AM peak and PM peak hours. This justifies that the WA-BHM model is more effective in simulating traffic volume during busy peak hours. The justification you have provided is insufficient – can you expand more on why this model is more useful than using the models of others? It would strengthen it if you could compare the results of this model to that of others and give some examples with numbers of why your model is better.

## 6 CONCLUSIONS

A novel discrete wavelet analysis and Bayesian hierarchical methodology based traffic volume model for signalised urban arterials has been introduced in this paper. Why was this done? What purpose does it serve? How is it different to the models of others? A non-functional wavelet-based ‘regular average daily trend’ is obtained by isolating the low resolution component from the high resolution components of a univariate traffic volume time-series dataset. A Bayesian Hierarchical methodology is further proposed to model the time-varying variability (within day) of the high resolution components. This time-varying variability has been established to be an intrinsic property of the traffic flow time-series dataset – is this a new finding – have others not found this out before?. The potential application of this random process traffic volume model lies in simulating traffic flows at junctions where real-time traffic conditions data collection system are not available or out of service for a considerable period. This methodology can also be used in developing an automatic incident detection algorithm using the concept of variability of variance developed in this paper – you have not done any work in this paper on incident detection so it should not appear in the conclusions. Your conclusions should also include the

benefits in terms of policy – what difference will using your model make in the real world? – again this relates back to the justification for the research.

## **ACKNOWLEDGEMENT**

The research work is funded under the Programme for Research in Third-Level Institutions (PRTLII), administered by the HEA.

**In relation to the figures, you will have to present the data so that it can be clearly interpreted in black and white – some of your figures are relying on colour to differentiate and this is not acceptable for papers submitted to journals.**

## REFERENCES

Adeli, H. and A. Samant. (2000) An Adaptive Conjugate Gradient Neural Network-Wavelet Model For Traffic Incident Detection. *Computer-Aided Civil Infrastructure Engineering*, Vol.15, pp. 251–260.

Adeli, H., and A. Karim. (2000) Fuzzy-Wavelet RBFNN Model for Freeway Incident Detection. *Journal of Transportation Engineering*, ASCE, Vol. 126, pp. 464–471.

Carlin, B. P. and Louis, T. A. (1996) *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall, New York.

Chatfield, C. (2001) *Time-Series Forecasting*. Chapman and Hall/CRC, London.

Chatfield, C. (2004) *The Analysis of Time Series: An Introduction, Sixth Edition*. Chapman and Hall/CRC, London.

Chen, S., Wang, W. and Qu, G.(2004) Combining Wavelet Transform and Markov Model to Forecast Traffic Volume, Proceedings of the Third International Conference on Machine Learning and Cybernetics, Shanghai.

Daubechies, I. (1992) *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 721-741.

Ghosh, B., Basu, B. and O'Mahony, M. M. (2005) Time-Series Modelling for Forecasting Vehicular Traffic Flow in Dublin. *84<sup>th</sup> Annual Meeting of Transportation Research Board (CD-ROM)*, TRB, Washington, D. C.

Ghosh, B., Basu, B. and O'Mahony, M. M. (2006) Analysis of Trend in Traffic Flow. *Proceedings of Institute of Engineering and Technology: Irish Signals and Systems Conference*, pp. 415-419.

Ghosh, B., Basu, B. and O'Mahony, M. M. (2006a) A Bayesian Time-Series Model For Short-Term Traffic Flow Forecasting. In press, *Journal of Transportation Engineering*, ASCE.

Ghosh-Dastidar, S. and Adeli, H. (2003) Wavelet-Clustering-Neural Network Model for Freeway Incident Detection. *Computer-Aided Civil and Infrastructure Engineering*, Vol.18 No.5, pp. 325-338.

Karim, A. and Adeli, H. (2002) Incident Detection Algorithm Using Wavelet Energy Representation of Traffic Patterns. *Journal of Transportation Engineering*, ASCE, Vol. 128, pp. 232-242.

Mallat, S. (1989) A Theory For Multi-Resolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11(7), pp. 674–693.

Mallat, S. (1998) *A Wavelet Tour of Signal Processing (2nd Edition)*. Academic Press, London.

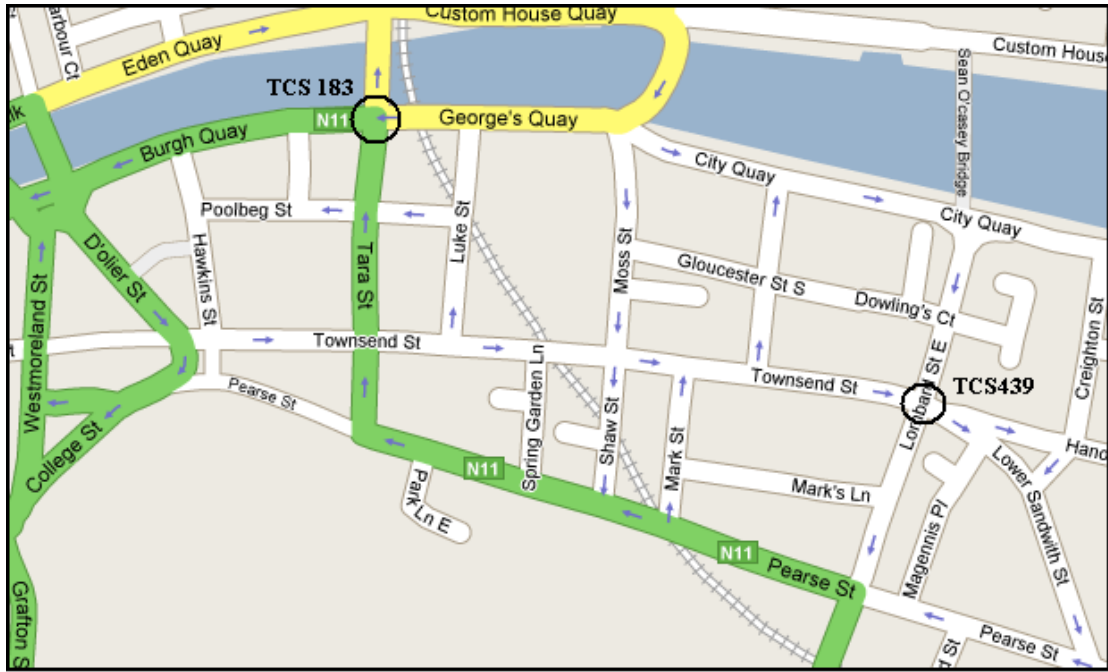
Mathworks. (2000) *Wavelet toolbox for use with MATLAB: User's Guide Version 2*. Mathworks Inc. (<http://www.mathworks.com>)

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953) Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, Vol. 21, pp. 1087-1092.

Qiao, F., Wang, X. and Yu, L. (2003) Optimizing Aggregation Level for ITS Data Based On Wavelet Decomposition. *82<sup>th</sup> Annual Meeting of Transportation Research Board (CD-ROM)*, TRB, Washington, D. C.

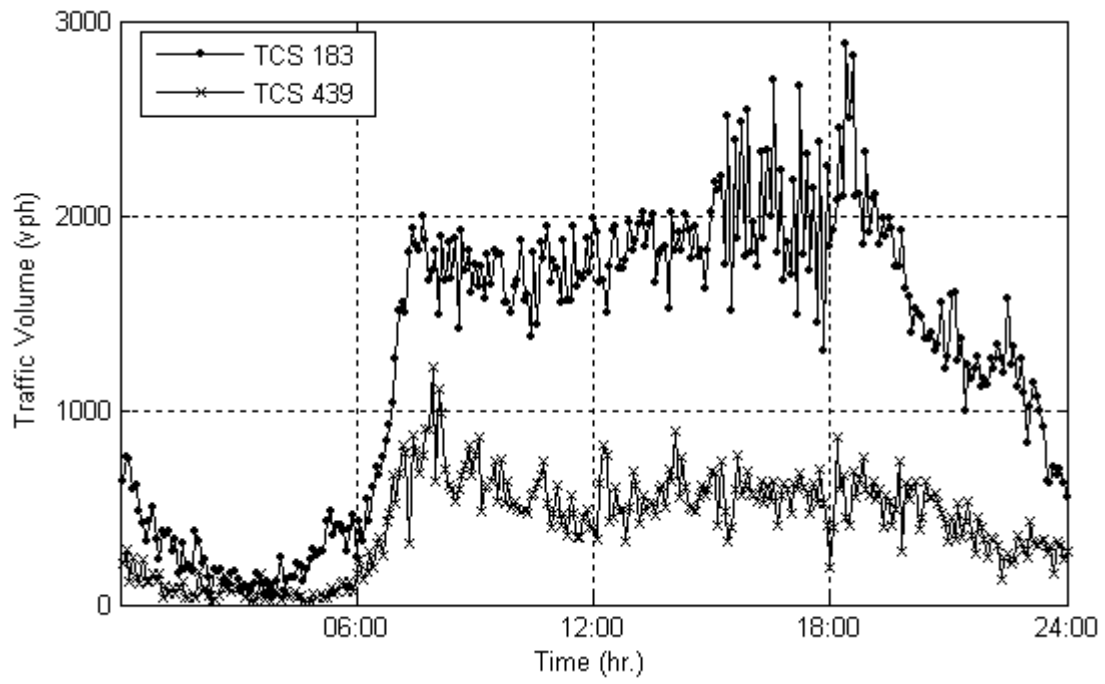
## LIST OF FIGURES

1. Diagram of the Junction TCS 183 and TCS 439.
2. Traffic Flow Observations on 15-06-2005 at Junction TCS 183 and TCS 439
3. Reconstructed Approximation Coefficients at Level 3 and Reconstructed Details Coefficients at Level 1, 2, 3 on 15<sup>th</sup> June 2005
4. (A) Trend and Original Observations on an Arbitrary Day (15-06-2005) from TCS 183
4. (B) Trend and Original Observations on an Arbitrary Day (15-06-2005) from TCS 439
5. Dot Plot of Residual on 15-06-2005
6. (A) Simulated and Original Traffic Volumes on 15-06-2005 from TCS 439.
6. (B) Simulated and Original Traffic Volumes on 15-06-2005 from TCS 183.
7. Simulations of Values of  $\tau$  from TCS 183(upper) and TCS 439 (lower).
8. (A) Simulations of Values of  $\sigma$  for TCS 183.
8. (B) Simulations of Values of  $\sigma$  for TCS 439.
9. (A) Simulations of Values of  $\sigma$  for TCS 183.
9. (B) Simulations of Values of  $\sigma$  for TCS 439.
10. (A) Simulations of Values of  $\sigma$  for TCS 183.
10. (B) Simulations of Values of  $\sigma$  for TCS 439.
11. (A) Simulations of Values of  $\sigma$  for TCS 183.
11. (B) Simulations of Values of  $\sigma$  for TCS 439.

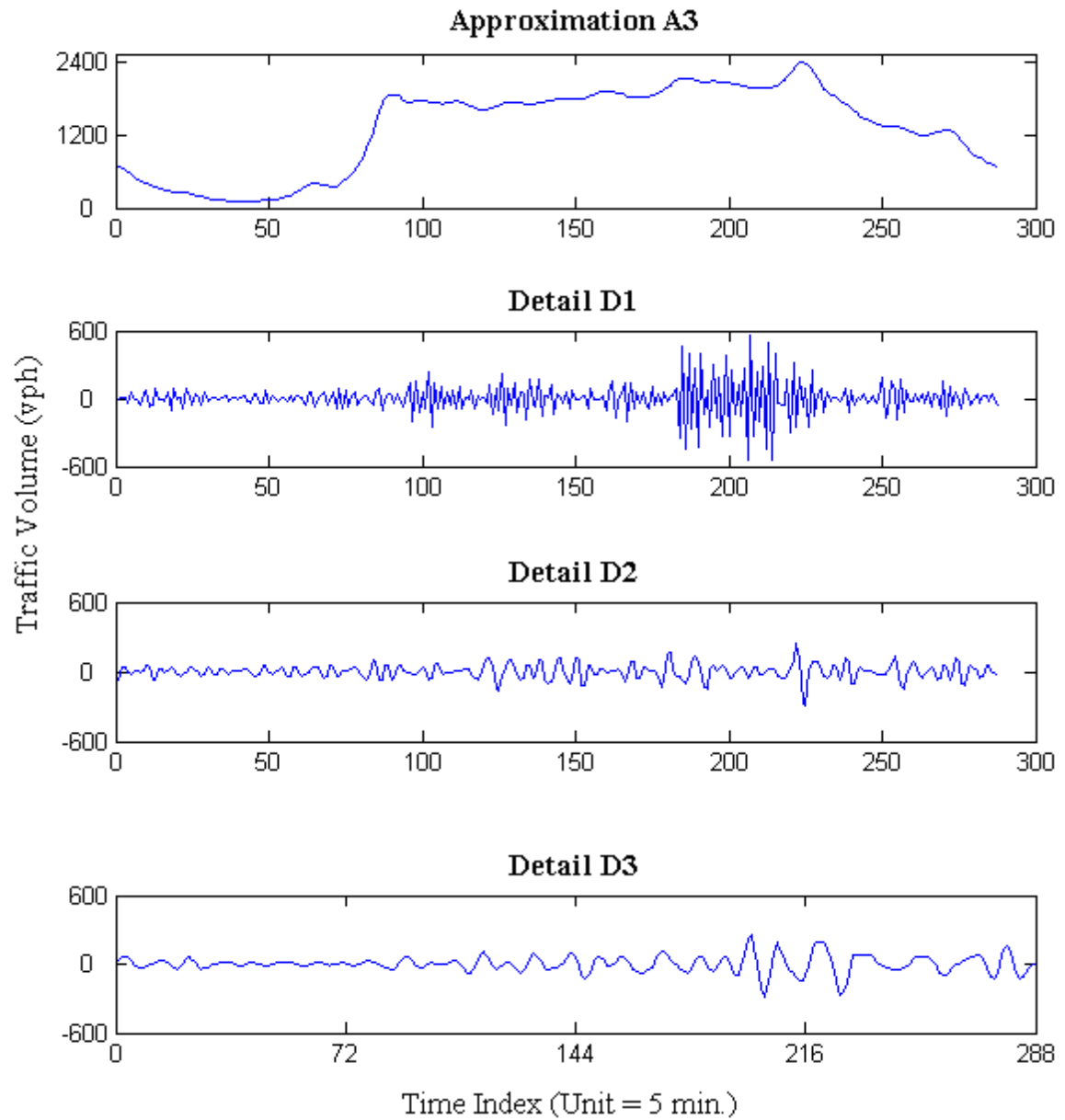


**Figure1**      **Diagram of the Junction TCS 183 and TCS 439.**

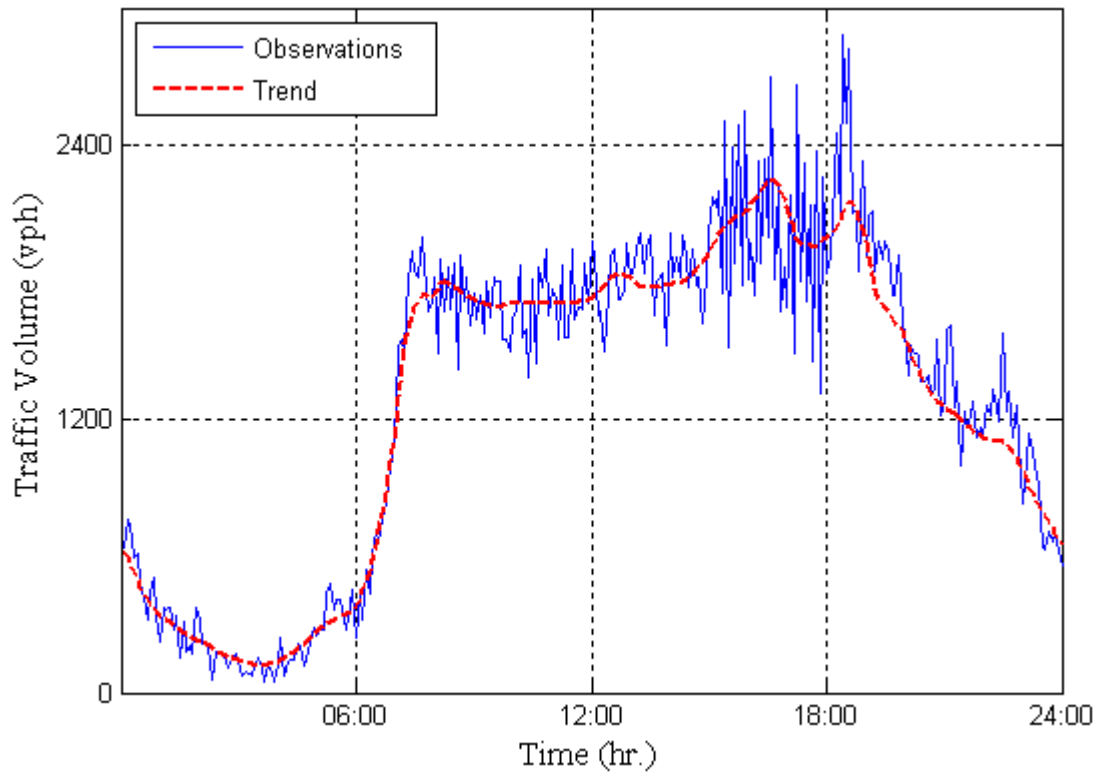




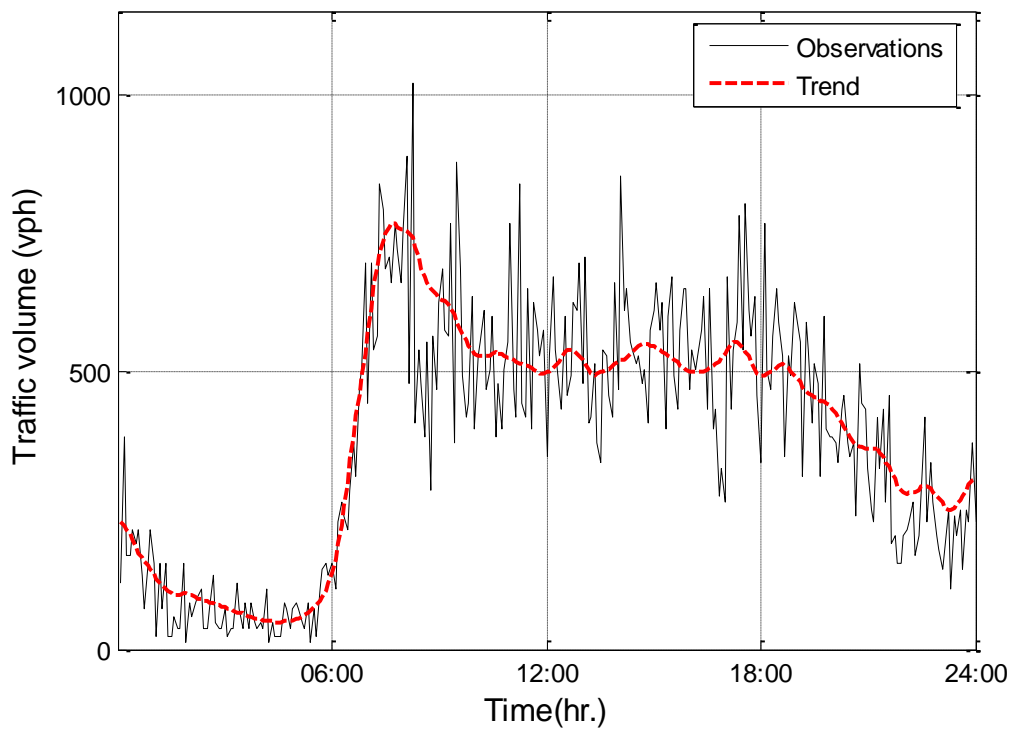
**Figure 2** Traffic Flow Observations on 15-06-2005 at Junction TCS 183 and TCS 439



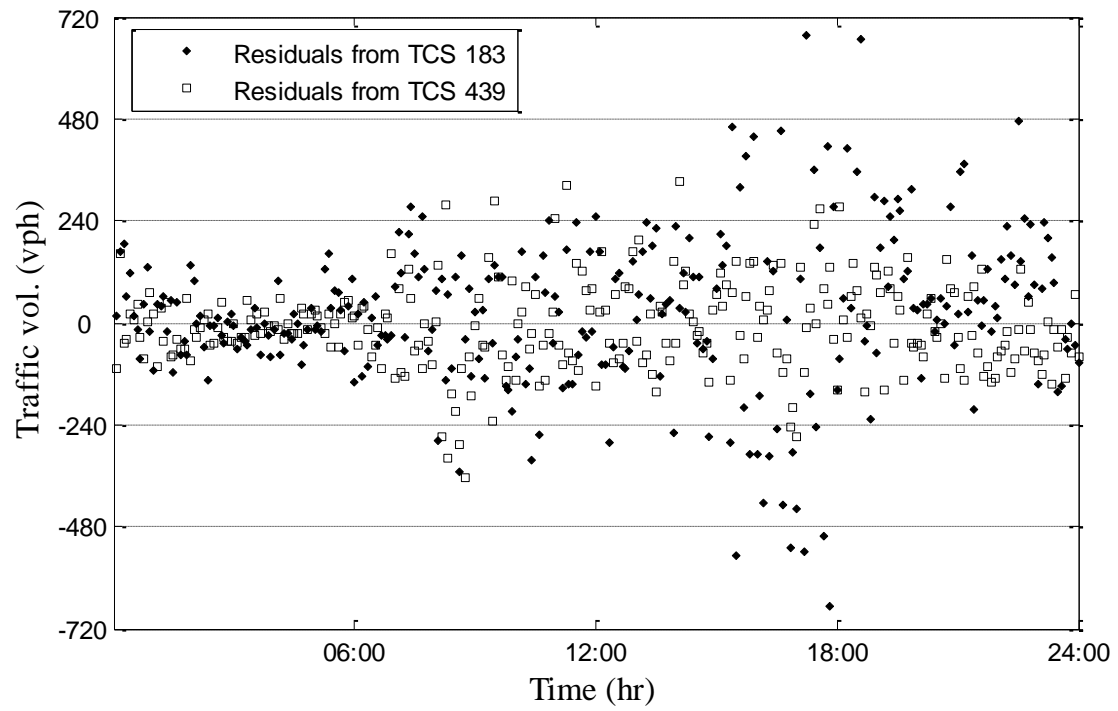
**Figure 3** Reconstructed Approximation Coefficients at Level 3 and Reconstructed Details Coefficients at Level 1, 2, 3 on 15<sup>th</sup> June 2005



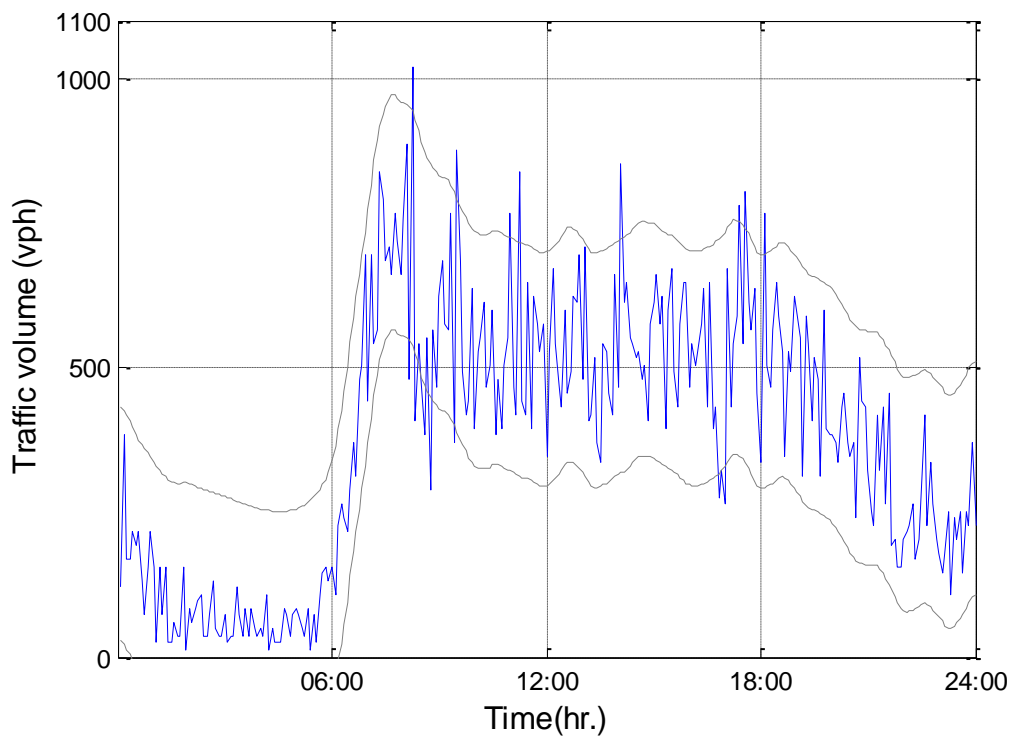
**Figure 4(A) Trend and Original Observations on an Arbitrary Day (15-06-2005) from TCS 183**



**Figure 4(B) Trend and Original Observations on an Arbitrary Day (15-06-2005) from TCS 439**



**Figure 5**      **Dot Plot of Residual on 15-06-2005**

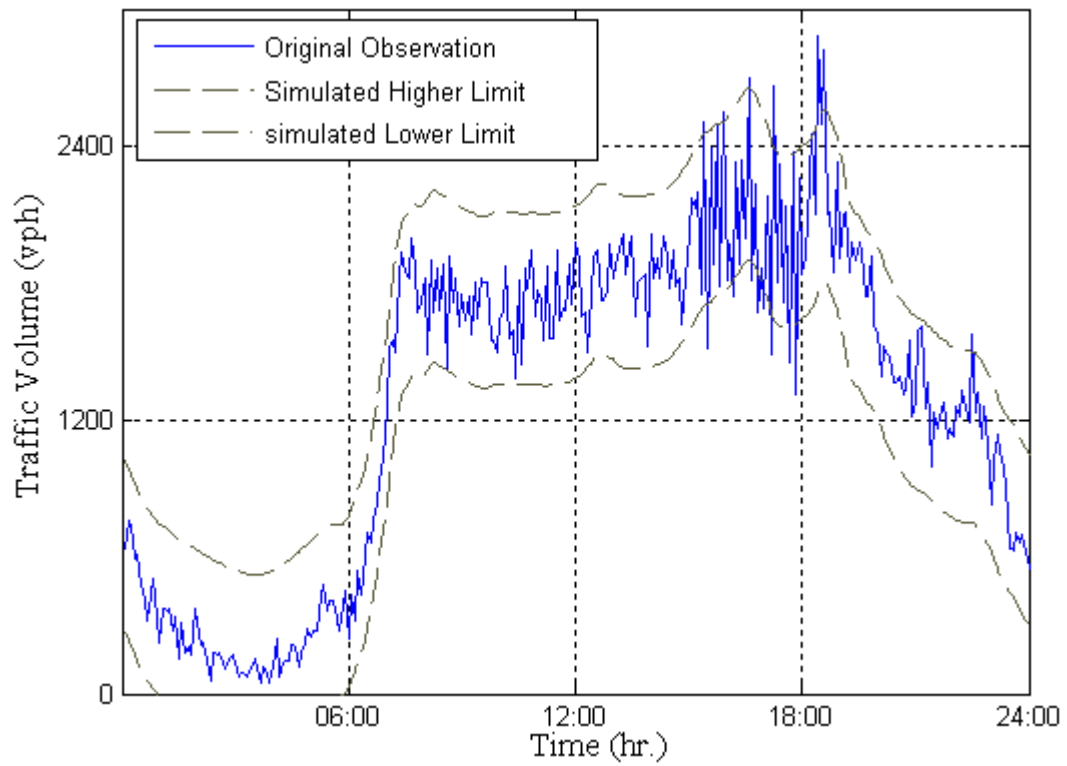


**Figure 6(A) Simulated and Original Traffic Volumes on 15-06-2005 from TCS**

**439.**

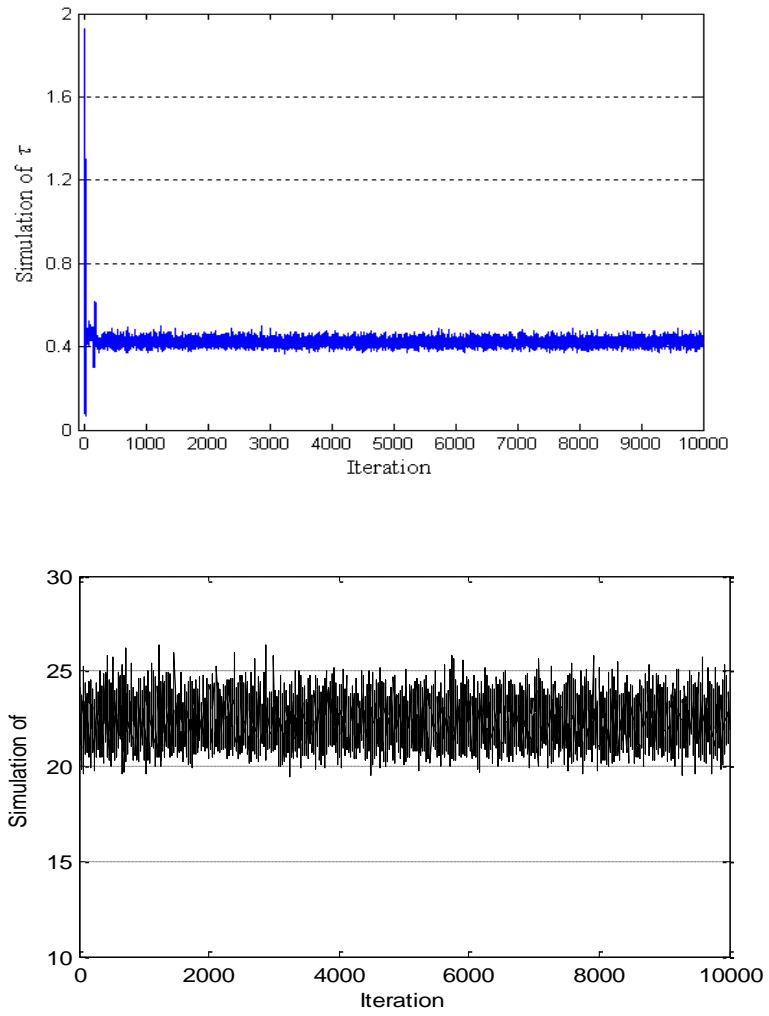
**You will need to make the lines on these figures distinguishable if published in black and white as colour will not be an option.**

The legend on the y axis in Fig 7 is not clear.

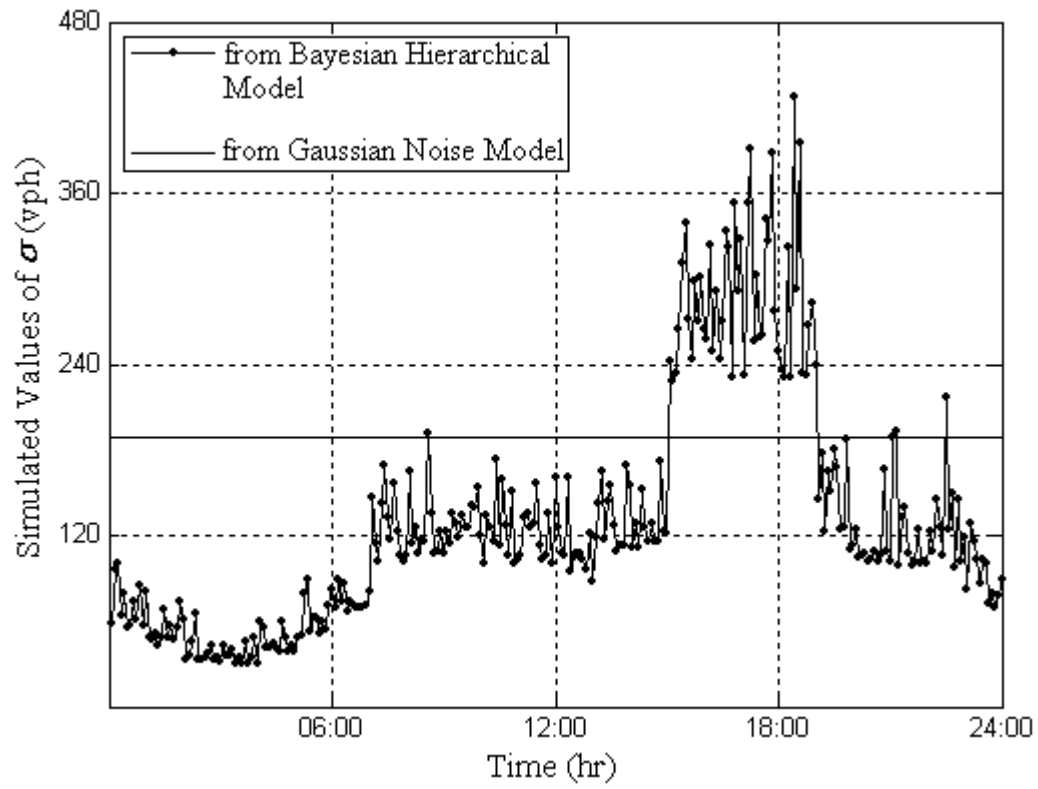


**Figure 6(B) Simulated and Original Traffic Volumes on 15-06-2005 from TCS**

**183.**

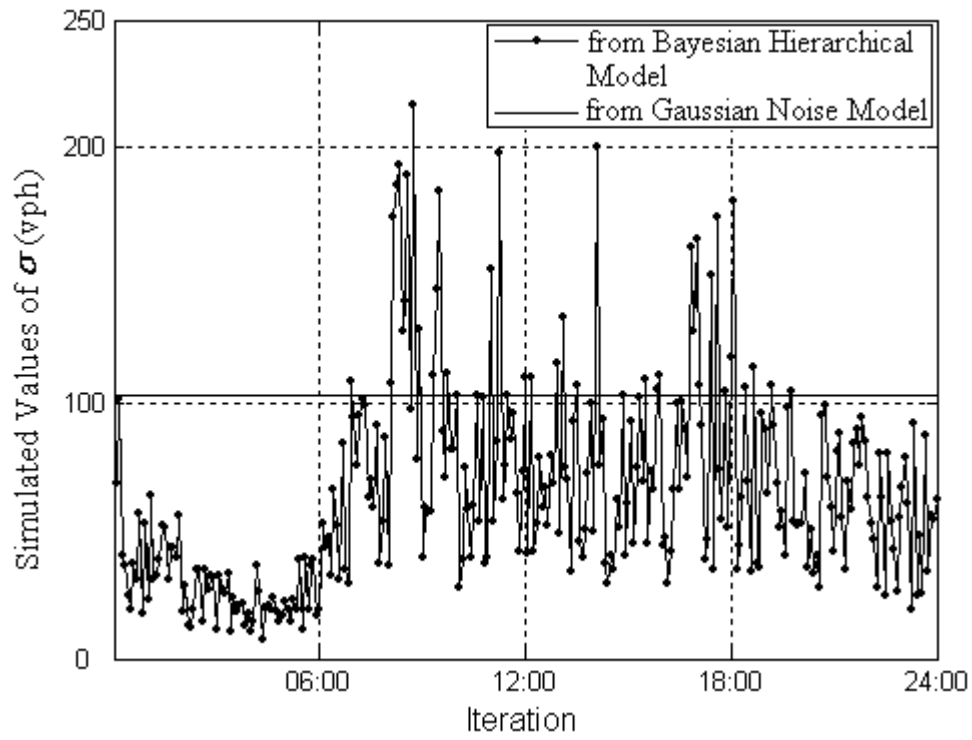


**Figure 7(A) Simulations of Values of  $\tau$  from TCS 183(upper) and TCS 439 (lower).**

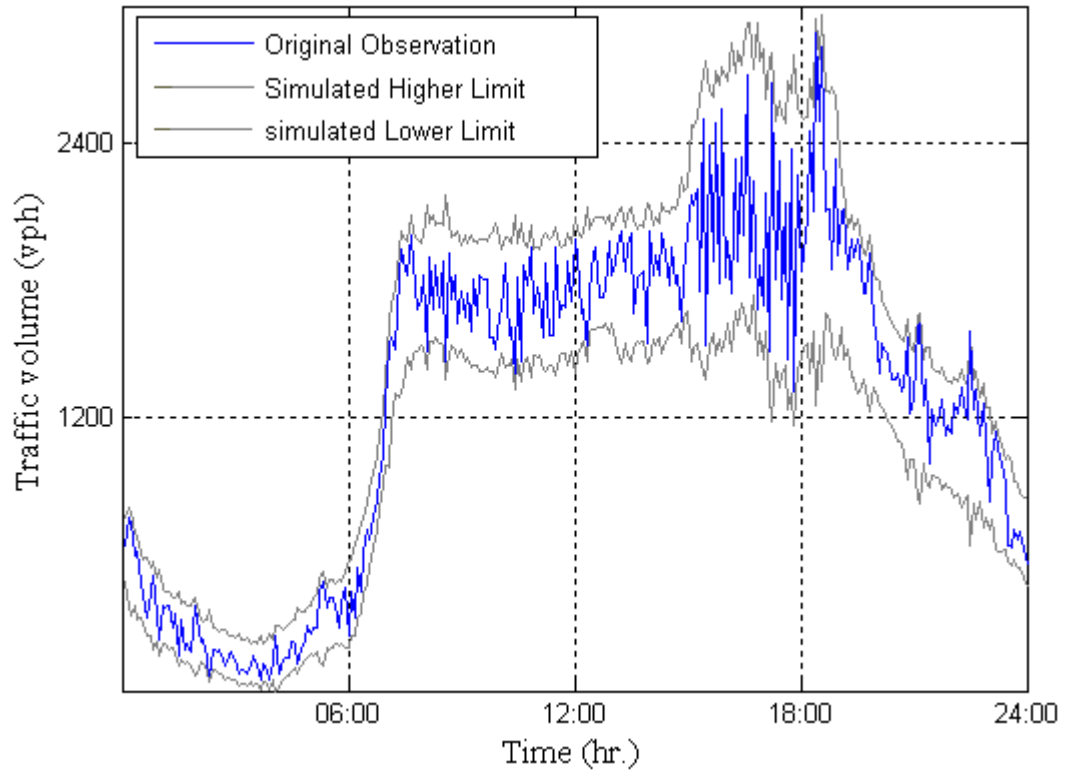


**Figure 8(A) Simulations of Values of  $\sigma$  for TCS 183.**

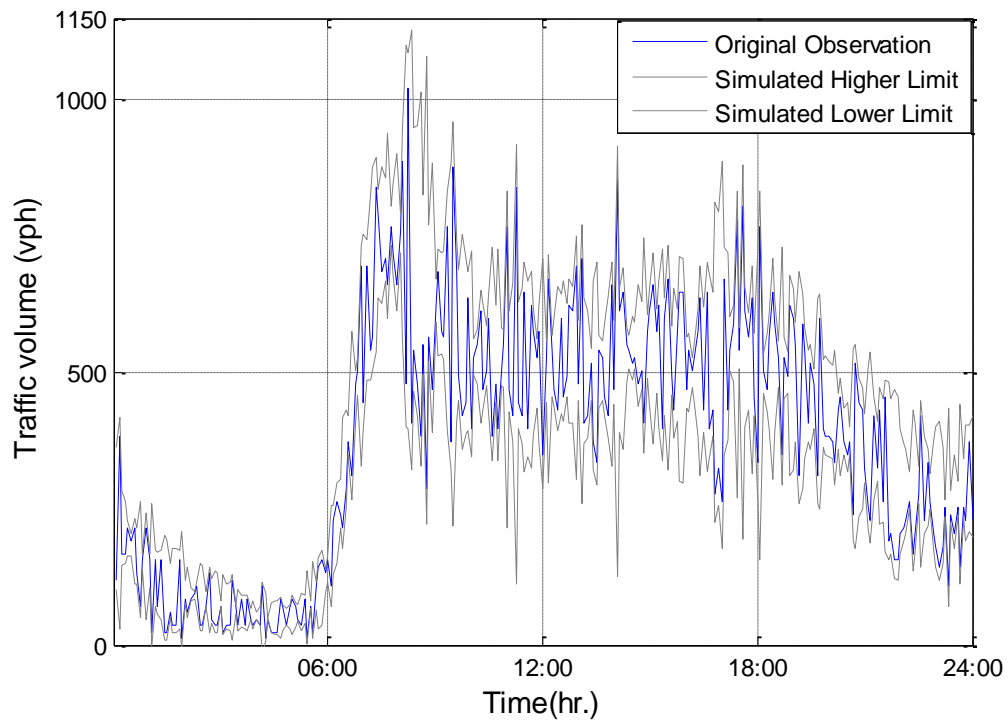




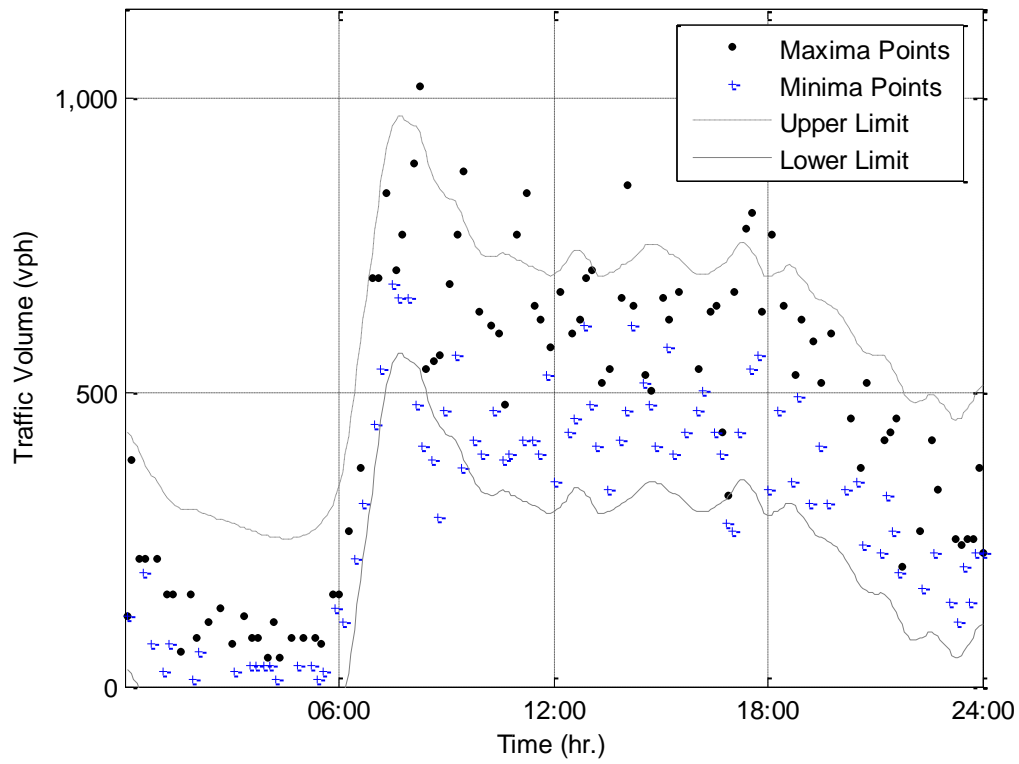
**Figure 8(B) Simulations of Values of  $\sigma$  for TCS 439.**



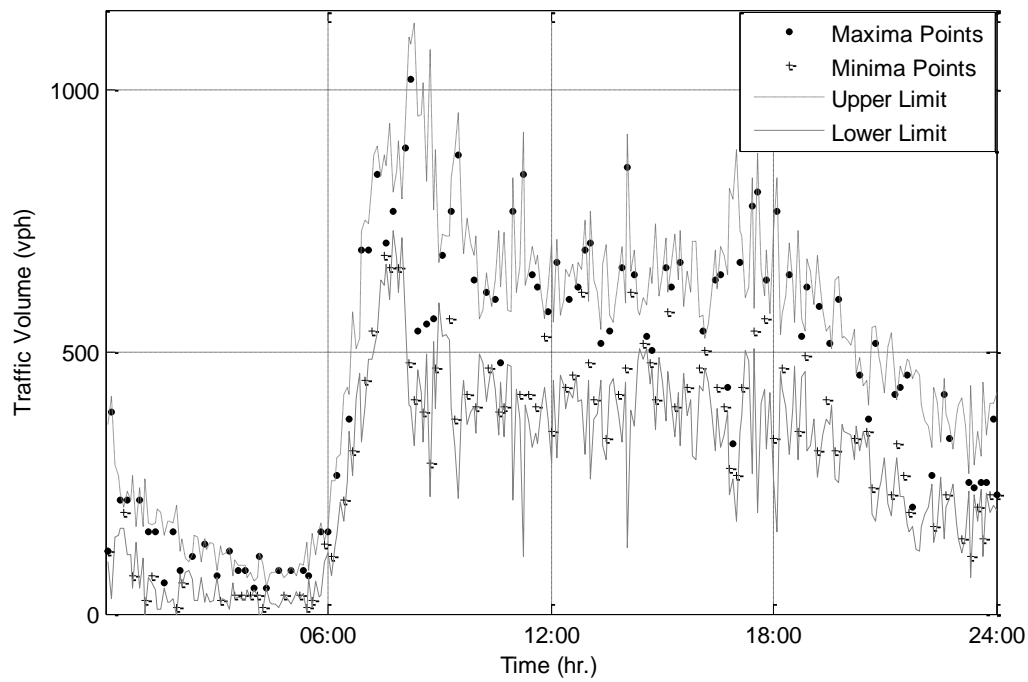
**Figure 9(A) Simulated and Original Traffic Volumes on 15-06-2005 at TCS 183.**



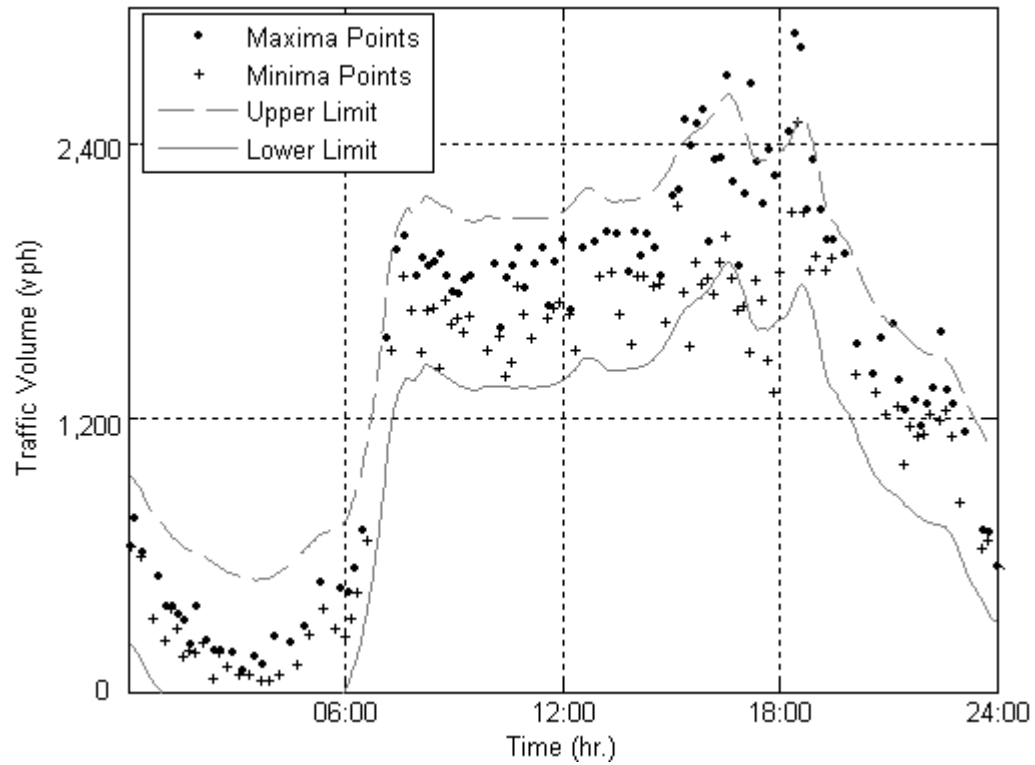
**Figure 9(B) Simulated and Original Traffic Volumes on 15-06-2005 at TCS 439.**



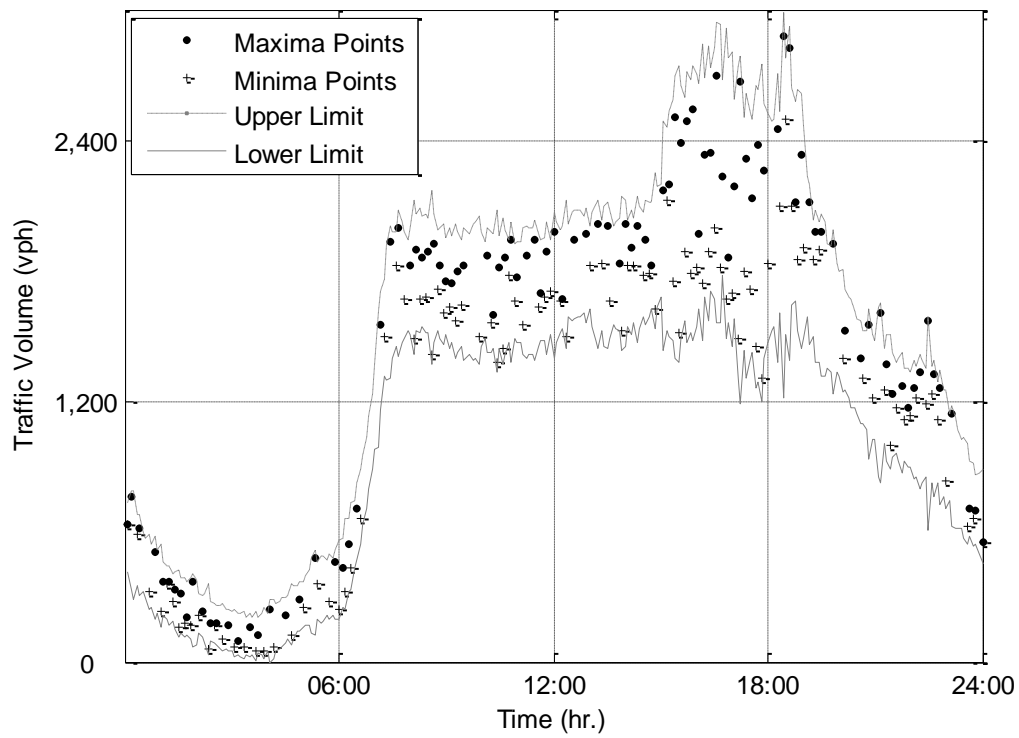
**Figure 10(a) Maxima and Minima points within simulated envelope from Trend + Gaussian Noise model for junction TCS 439.**



**Figure 10(b) Maxima and Minima points within simulated envelope from Trend + BHM model for junction TCS 439.**



**Figure 11(a) Maxima and Minima points within simulated envelope from Trend + Gaussian Noise model for junction TCS 183.**



**Figure 11(b) Maxima and Minima points within simulated envelope from Trend + BHM model for junction TCS 183.**

	<b>Mean of Original Observations (vph)</b>	<b>Standard Deviation of Residuals (vph)</b>
<b>TCS 183</b>	1301.75	188.28
<b>TCS 439</b>	379.12	103

**Table 1          Statistical Analysis of Residuals from the Chosen Intersections**



	<b>TCS 183</b>	<b>TCS439</b>
<b>Maxima Points Error</b> (Over a day)	16.12%	16.9%
<b>Minima Points Error</b> (Over a day)	19.5%	18.5%
<b>Maxima Points Error</b> (AM peak 7:30-8:30 am)	8%	8.5%
<b>Minima Points Error</b> (AM peak 7:30-8:30 am)	11%	11.9%
<b>Maxima Points Error</b> (PM peak 3:30-4:30 pm)	12.6%	11%
<b>Minima Points Error</b> (PM peak 3:30-4:30 pm)	15.8%	11%
<b>Points outside</b> (Trend + BHM model)	0	0
<b>Points outside</b> (Trend + Gaussian model)	19	16

**Table 2 Error of Extremes**