SHOR T-TERM TRAFFIC FLOW FORECASTING USING DYNAMIC LINEAR MODELS

Tiep Mai
PhD Student
Trinity College Dublin

Bidisha Ghosh
Lecturer
Trinity College Dublin

Simon Wilson
Professor
Trinity College Dublin

Abstract

Intelligent Transportation Systems (ITS) is an emerging concept which has been utilised to improve efficiency and sustainability of existing transportation systems. Short term traffic flow forecasting, the process of predicting future traffic conditions based on historical and real-time observations is an essential aspect of ITS. The existing well-known algorithms used for short-term traffic forecasting include time-series analysis based models. Among the time-series models, the Seasonal Autoregressive Integrated Moving Average (SARIMA) is one of the most precise statistical models in this field. In the existing literature SARIMA models are mostly used in its multiplicative form and the parameters of the model are mostly estimated using a frequentist approach. Estimation of the large scale multiplicative SARIMA model for traffic flow forecasting often proves to be complex and computationally expensive for researchers and end-users.

In this paper, an additive SARIMA model has been employed to predict traffic flow in short-term or near-term future. The Dynamic Linear Model (DLM) representation of the additive SARIMA model has been used here to reduce the number of latent variables. Traditionally in a frequentist approach, point estimations of the SARIMA model parameters are obtained by maximizing the likelihood, but in this paper the marginal posterior density of each of the parameters has been explored by applying a Bayesian inference framework. Markov Chain Monte Carlo (MCMC) sampling method has been used to develop the Bayesian inference framework. For such sampling method for SARIMA, a problem of serial correlation has proved to be quite serious; however in the additive form, with the help of a carefully designed Metropolis-Hastings algorithm (a type of MCMC algorithm) this problem has been mitigated. The efficiency of the proposed prediction algorithm has been evaluated by modelling real-time traffic flow observations available from a certain junction in the city-centre of Dublin.

I. Introduction

Intelligent Transportation Systems (ITS) is an emerging concept which has been utilized to improve efficiency and sustainability of existing transportation systems. Short-term traffic forecasting, the process of predicting future traffic conditions in short-term or near-term future, based on current and the past observations is an essential aspect of ITS. In the last decade, considerable research attention has been focused on developing precise, flexible, adaptable and universal short-term prediction algorithms for traffic variable observations. Several parametric and non-parametric techniques have been utilized to develop successful Short-Term Traffic Forecasting (STTF) algorithms [1,2].
The predominant parametric approach in STTF is time-series analysis techniques. Time-series analysis techniques which are popular in STTF are smoothing techniques [1], Autoregressive linear processes [3] and Kalman filtering [4]. Among these, the Autoregressive linear processes are the most developed and well-documented in this field. Ahmed and Cook [5] introduced the Auto-Regressive Moving Average (ARMA) class of models to the traffic flow forecasting literature. The next seminal step was extending the simple ARMA model to a seasonal format and accounting for the daily and/or weekly variability [6,7]. The aforementioned studies on applying ARMA model in developing STTF algorithms mainly focused on univariate structure; traffic data from any single station were modelled. In the last decade, the research attention has been shifted in developing more efficient prediction algorithms through the utilization of multivariate ARMA techniques which can model the spatial dependency and temporal evolution of traffic variables (such as, volume, speed and travel time) simultaneously.

One of the initial multivariate models was developed by Stathopoulos and Karlaftis [8] using state-space methodology. The model provided superior forecasts to equivalent univariate ARMA models. A multivariate ARMA technique called space-time autoregressive integrated moving average (STARIMA) methodology was applied to develop a model to account for the spatial dependency of traffic data in an urban network [9]. The spatial dependency of the network were incorporated in the STARIMA model through the use of weighting matrices estimated based on the distances among the data collection points. A new class of time-series model called the Structural Time-Series Model was used to develop multivariate STTF algorithm; these models outperformed univariate seasonal ARMA models [10]. A direct multivariate extension of autoregressive linear processes has been first attempted by Chandra and Al-Deek [11]. This model utilizes a Vector Auto-Regressive (VAR) structure for STTF. Freeway traffic speed and volume had been predicted in this study. The model did not consider the correlation of the noise among multiple stations or data collection points as there does not exist a Moving Average (MA) part. Also, the seasonal nature of the traffic data has been modelled by eliminating seasonality through a seasonal difference and not by direct modelling of seasonality in a seasonal ARMA form. In this study the authors compared VAR with other univariate models and concluded that adding correlations among different locations improves the prediction result. Still, the results are restricted to VAR structure which is only a subclass of seasonal Vector ARMA (VARMA) model. In the same year, Multi-Regression Dynamic Model (MDM) was adopted to develop a multivariate algorithm for STTF by Queen and Albers [12]. The MDM consists of multiple independent regression equations often represented in Dynamic Linear Model (DLM) form. Similar to the previous study [11], MDM did not include the noise cross-correlation or the MA coefficients and moreover, MDM assumed spatially independent noise, which allows separate statistical inference for each station. Also, the spatial correlation among neighbouring stations evolved contemporaneously in the model and a temporal evolution of spatial cross-correlation was not modelled.

In this paper, a full multivariate extension of the most efficient univariate time-series model i.e. the seasonal ARMA has been proposed. Unlike the past studies this model involves a noise cross-correlation along with a seasonal form. In particular, an Additive Seasonal VARMA (A-SVARMA) model has been developed to predict traffic flow in short-term future in urban signalized arterial networks. A Bayesian framework has been proposed to estimate the parameters of the A-SVARMA model. The inference framework utilizes a Markov chain Monte Carlo (MCMC) sampling method. One serious problem of MCMC is the slow convergence caused by serial correlation. Hence, in order to have a better sampling of MA parameters, marginalization and adaptive MCMC are used. The proposed method has been applied to model traffic volume observations from multiple junctions situated at the city-centre of Dublin, Ireland. The results indicate that the proposed forecasting algorithm is an effective approach in predicting real-time traffic flow at multiple junctions within an urban transport network.

II. VARMA

Time series theory, including VARMA and DLM, is discussed in detail in [13-15]. A brief summary of VARMA is as follows. Denote the observations by . A
-variate VARMA is considered with mean and identical independent (iid) Multivariate Normal noise

\[ (1) \]

Where \( B \) is the backshift operator; \( \beta \); \( \phi \); is the lag index; Each element of \( \beta \) is a matrix. \( I \) is the identity matrix.

The spatial-temporal dependency of VARMA model is defined by the matrices and . However, using the full matrix is computationally costly. Hence, the dimension and computational cost have been reduced by adding neighbour information. Matrix denotes the neighbour dependency; \( \phi \). The matrix is for dependency.

In existing literature on STFF, the ARMA class of models in their seasonal form \([3]\) is always expressed in multiplicative form. However, the multiplicative form may not be the most appropriate for several reasons. Firstly, the physical interpretation of seasonal form is ambiguous due to non-commutative matrix multiplication \([16]\). Also, directly using additive form omits the complex and computationally expensive matrix multiplication which is calculated repeatedly in the inference process. Furthermore, the additive form specifies a seasonal VARMA with \( \beta \) components, rather than \( \phi \) components by multiplicative form. As a result, the serial correlation problem of MCMC sampling which will be described in details later is solved efficiently. So, in this paper, a sparse additive representation of VARMA with a seasonality effect is used.

For the additive form specification, vector is defined as: . Hence, Similarly, there is for the MA coefficients. So, for the VARMA model with seasonal period , can be set as . The Additive Seasonal VARMA is called by \( A \)-SVARMA from now on (The univariate version is called A-SARMA). An \( A \)-SVARMA is then specified with with mean , noise variance ; and are the ordered vectors of non-zero elements of and respectively.

### III. Bayesian Inference and MCMC sampling

In this paper, the \( A \)-SVARMA model has been estimated using DLM. By using the DLM representation, the number of initial random variables is reduced from to \([17]\). Define the initial state vector of DLM representation of \( A \)-SVARMA by . For the Bayesian inference, the prior of parameters is defined as follows:

\[ (2) \]

where denotes the inversed Wishart distribution with degree and inverse scale matrix ; denotes the Multivariate Normal distribution with mean and precision matrix . The conditional likelihood is:

\[ (3) \]

The closed-form conditional posterior of can be obtained by the conjugate prior and some transformations.

MCMC sampling is used to realize the Bayesian framework. Gibbs Sampling \([3]\) is used for each parameter block. Gibbs sampling of the conditional posterior of can be done by using standard distribution such as Normal and inversed Wishart. However, sampling suffers from the intractable posterior of and the serial correlation problem. The following method may be used. Firstly, is sampled analytically by exploiting the closed form standard distribution. Then, a Metropolis-Hastings \([3]\) is used with random walk for . The MCMC trace of such a method with simulated data is in Figure 1(a). It can be observed from the figures, this method suffers from the serial correlation. The serial correlation problem restricts the MCMC update and results in the slow MCMC convergence.
To solve the serial correlation problem, the following method is proposed: the posterior is analytically marginalized with respect to \( \theta \). Then, adaptive MCMC is used with the numerical evaluation of \( f(\mathbf{y}, \theta) \). Consequently, the MCMC convergence is significantly faster. Figure 1(b) shows the trace of this scheme which is much better than one of simple MCMC.

### IV. Data

The proposed Bayesian DLM methodology has been evaluated by modelling traffic volume observations from a busy thoroughfare in the city-centre of Dublin in Ireland. Three sites had been chosen for traffic data collection and the map of the chosen section of the transport network is provided in Figure 2. As seen in the figure, all the modelled sites are situated on/near Pearse Street which is one of the busiest roads in Dublin City.

The chosen sites will be referred as Stn 1, Stn 2 and Stn 3 from now on in this paper. The three sites are located in two signalized traffic intersections. To avoid unnecessary complexity in matrix and equation representation in an illustrative example, no other junctions were considered in this application of the DLM model. Both these junctions are four-legged cross-sections, with two-way traffic on all approaches barring the east-bound approach in Stn 3. Stn 1 and Stn 2 are two separate approaches on the second junction. These two approaches receive green time in separate phases. Both the junctions have three-phase signals with turning protection on Pearse Street. Both Stn 1 and Stn 3 have three lanes each and Stn 2 has two lanes. The data obtained collectively from all detectors on any approach, is used for the modelling. As the weekend traffic dynamics is very much unlike the traffic dynamics in the weekdays, the modelling is essentially carried out on the data observed during weekdays. Since, the used data set does not contain any missing data, no special treatment for missing data is required to be utilized here.

The data used for modelling was recorded from 22nd July 2010 to 3rd September 2010. 15-minute aggregate traffic volume observations were used in modelling purposes. Total 32 days of data from weekdays were used. The traffic observations from first 26 days were used for fitting and the rest of data was used for model evaluation.
In this section, the multivariate traffic time-series is denoted as, $\mathbf{X}$, where $n$ is the number of stations at which traffic data were measured. $n$ equals to 2 in this application and consequently, $x$ is the traffic volume time-series from Stn $i$. Stn 1 and Stn 2 are chosen in such a way that the observation from these two sites at do not have any obvious spatial correlation. The observations from both these stations are spatially correlated with observations at Stn 3. However, these visual observations were not utilized or implemented in the development of the matrices.

The traffic volume observations from the abovementioned junctions form non-stationary time-series datasets [3]. The initial step in analyzing this multivariate time-series dataset is to eliminate the seasonality and the trend of $x$ through filtering and/or transformations. Stationarity or weak stationarity was attempted through removal of trend and seasonal patterns of the traffic data. In this paper, multiple variations of DLM have been developed and fitted to identify the most suitable one. In the next section, the modelling variations are discussed in further detail.

V. Model Variations

Three variations of DLM model were fitted to the traffic volume observations using the aforementioned Bayesian using the aforementioned Bayesian inference framework. The previous studies [3, 7, 10] showed that there exists a daily seasonality in the weekday traffic volume datasets. Hence, the first variation of the DLM utilizes seasonal difference in an attempt to attain stationarity. This model is named the SD model. For a daily model with 15-minute aggregate data, the season period is 24. In the SD model:

$$\text{for } t = 1, \ldots, T$$

The second variation of the A-SVARMA model studied utilizes a first and a seasonal difference to attempt a stationary behavior. The model is named as DSD model and follows an equation:

$$\text{for } t = 2, \ldots, T$$

The third variation of the A-SVARMA model is slightly different from the previous models as it does not attempt to attain stationarity through differencing at the initial step. Rather, the model is developed considering the within day variability of the traffic data which is modelled as:

$$\text{for } t = 1, \ldots, T$$

is the empirical mean/average which varies with the time of the day and is obtained from fitting the past traffic observations. This model has been named as the MP model. $\mu_1$ and $\mu_2$ are further modelled using A-SVARMA...
The A-SVARMA structure for the three multivariate models (SD, DSD, MP) are the same:

\[
(7)
\]

where . Notice that for model M, with such and , the transformed (at Stn 3 is dependent on and (at Stn 1 and Stn 2). For comparison purposes, along with each variation of the A-SVARMA model, a univariate model of similar ARMA structure has been fitted to the traffic volume data . The univariate variation of SD model is named as SD-U which is essentially a univariate SARIMA model with seasonal differencing. The univariate variation of DSD model is named as DSD-U which is a univariate SARIMA model with both first and seasonal differencing. The MP-U is the univariate variation of MP model. The Equation (7) for univariate ARMA models changes slightly; and are still but . All six models are used for forecasting short-term traffic volume and the results are discussed in the following subsection.

VI. Result

All univariate and multivariate models are compared based on their predictive performances while forecasting traffic data at Stn 3. For any model M, at any time instant , the v-step ahead forecasts are denoted by on . can be generated by performing inverse operations on . The models have been used to generate -step ahead or one-day ahead forecasts. The prediction error is defined as the Mean Accumulated Error (MAE). MAE for model M for steps ahead forecasts is:

\[
(8)
\]

where, for this study as all the models are used for predictions over six days; can be estimated by MCMC samples. The prediction accuracy of the SD, DSD and MP are represented in Table 1. The prediction accuracy is expressed in MAE values in vehicles/15 minutes. In Table 2, the predictive performance of the univariate models while used for forecasting traffic volume at Stn 3 are tabulated in the form of MAE values. In Figure 3, plots of the MAE values in vehicles/15 minutes are plotted against the steps of forecast for Stn 3 for all six models. In the figure, the MAE values increase with steps of forecast. However the rate of this increase is different for different variations of DLM model.

In general, multivariate models give better prediction than univariate models and the MP model provides the best forecast among all. However, the univariate variation in this case, the MP-U model, provides only slightly inferior prediction results. This is due to the fact that 15-minute aggregate data were considered in developing the models. As a car can travel a considerable distance within a 15-minute time interval, the spatial correlation among the stations are not as high as can be seen from high resolution traffic observations. It is expected that for a dataset of higher resolution the impact of multivariate modelling will be more significant. Among all three variations the predictions from the DSD model and its univariate counterpart is the worst.
Table 1: Summary of all stations for multivariate model

<table>
<thead>
<tr>
<th>Station</th>
<th>Model</th>
<th>DSD</th>
<th>SD</th>
<th>MP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>StdY (veh/15min)</td>
<td>11.94</td>
<td>13.20</td>
<td>9.26</td>
</tr>
<tr>
<td></td>
<td>E_ML1 (veh/15min)</td>
<td>5.78</td>
<td>5.41</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td>E_ML20 (veh/15min)</td>
<td>6.64</td>
<td>5.73</td>
<td>5.29</td>
</tr>
<tr>
<td></td>
<td>E_ML90 (veh/15min)</td>
<td>7.71</td>
<td>5.83</td>
<td>5.57</td>
</tr>
<tr>
<td>2</td>
<td>StdY (veh/15min)</td>
<td>20.42</td>
<td>18.94</td>
<td>13.86</td>
</tr>
<tr>
<td></td>
<td>E_ML1 (veh/15min)</td>
<td>10.02</td>
<td>9.05</td>
<td>8.42</td>
</tr>
<tr>
<td></td>
<td>E_ML20 (veh/15min)</td>
<td>11.58</td>
<td>9.12</td>
<td>8.67</td>
</tr>
<tr>
<td></td>
<td>E_ML90 (veh/15min)</td>
<td>12.38</td>
<td>9.17</td>
<td>8.76</td>
</tr>
<tr>
<td>3</td>
<td>StdY (veh/15min)</td>
<td>31.32</td>
<td>40.75</td>
<td>29.64</td>
</tr>
<tr>
<td></td>
<td>E_ML1 (veh/15min)</td>
<td>15.00</td>
<td>14.79</td>
<td>13.11</td>
</tr>
<tr>
<td></td>
<td>E_ML20 (veh/15min)</td>
<td>18.88</td>
<td>15.86</td>
<td>14.45</td>
</tr>
<tr>
<td></td>
<td>E_ML90 (veh/15min)</td>
<td>20.75</td>
<td>16.25</td>
<td>15.12</td>
</tr>
</tbody>
</table>

Table 2: Summary of all stations for multivariate model

<table>
<thead>
<tr>
<th>Station</th>
<th>Model</th>
<th>DSD-U</th>
<th>SD-U</th>
<th>MP-U</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>StdY (veh/15min)</td>
<td>31.32</td>
<td>40.75</td>
<td>29.64</td>
</tr>
<tr>
<td></td>
<td>E_ML1 (veh/15min)</td>
<td>15.15</td>
<td>15.40</td>
<td>13.23</td>
</tr>
<tr>
<td></td>
<td>E_ML20 (veh/15min)</td>
<td>19.07</td>
<td>16.06</td>
<td>14.48</td>
</tr>
<tr>
<td></td>
<td>E_ML90 (veh/15min)</td>
<td>20.90</td>
<td>16.71</td>
<td>15.22</td>
</tr>
</tbody>
</table>

Figure 3: Accumulated error

VII. Conclusion

In this paper, for the first time a seasonal ARMA model has been used to develop STTF algorithm in a multivariate paradigm. A Bayesian inference framework for estimating the parameters of A-SVARMA in DLM form has been developed for the STFF algorithm. The Bayesian estimation provides flexibility to introduce expert knowledge in the model with the use of prior densities. MCMC sampling is applied to realize the Bayesian estimations. In such sampling method, marginalization and adaptive MCMC are proposed to solve the problem of serial correlation. As a result, the MCMC sampling converges much faster.

The proposed model is also the first attempt in modelling correlation of spatial noise (MA components) among multiple traffic data collection points or junctions for STTF. The model proves that there exists such spatial correlation and it is beneficial to model such behavior to improve the applicability, efficiency and robustness of STFF algorithms.
The study compares three variations of DLM model. For illustrative purposes, real-time traffic data from a small network in the city Dublin, Ireland had been considered and the traffic volume observations from the network has been modelled successfully using the proposed variations of DLM model. A time-variant mean process model (MP) provides the best prediction accuracy. This model outperforms all other multivariate and univariate models.

References


