Improved Dynamic Route Guidance based on Holt-Winters-Taylor Method for Traffic Flow Prediction

Yousef-Awwad Daraghmi[*], **Motaz Daadoo**

College of Engineering and Technology, Palestine Technical University Kadoorie (PTUK), Tulkarem City, Palestine

Abstract   Accurate short-term traffic flow prediction is necessary for the implementation of Dynamic Route Guidance as motorists need to know traffic conditions ahead. The accuracy of short-term traffic flow prediction depends on how prediction models handle traffic flow characteristics such as temporal correlation, overdispersion, and seasonal patterns. Several data mining methods have been proposed to model and forecast traffic flow for the support of congestion control strategies. However, these methods focus on some of the characteristics and ignore others. Some methods address the autocorrelation and ignore the overdispersion and vice versa. In this research, we propose a data mining method that can consider all characteristics by capturing the flow autocorrelation, trend, and seasonality and by handling the overdispersion. The proposed method adopts the Holt-Winters-Taylor (HWT) count data method. Data from Taipei city are used to evaluate the proposed method which outperforms other methods by achieving a lower root mean square error. Then the proposed method is used in a dynamic route guidance systems to enhance the efficiency of guidance.

Keywords   Autocorrelation, Holt-Winters, Negative Binomial, Overdispersion, Short-term prediction, Traffic flow

1. Introduction

Dynamic Route Guidance (DRG) as a part of Advance Traveller Information Systems (ATIS) aims to provide travelers with real-time information about traffic condition on routes. Due to the continuous change of traffic conditions in urban areas, travellers may select routes that will be congested during their trips. This problem can be overcome by predicting traffic flow for a short time ahead to determine the status of all roads on routes during trips. Prediction makes DRG adaptive to the traffic changes on roads and enables travellers to choose the best routes which has the shortest travel time.

Traffic flow is time series data consisting of sequences of values that are measured at equal or unequal time intervals. It has complex characteristics due to the complexity of trans-portation network particularly in urban cities which contain many intersections with traffic lights controlling traffic flow on upstream and downstream roads. These characteristics include autocorrelation as the flow time series is correlated with its values in the past and the future [1-3]. The flow may also have a trend which describes the general direction of the time series [3], [4]. Further, the flow has seasonal patterns that are repeated every week and every day [5]. Furthermore, the flow exhibits overdispersion which indicates that its variance ex-ceeds its mean [5]. Addressing these characteristics increases the accuracy of traffic flow modeling.

Several data mining methods have been proposed to model traffic flow. These methods include the exponential smoothing methods such as the Holt-Winters (HW) method [4], the Autoregressive Integrated Moving Average (ARIMA) [2], [4], [6], and the Multivariate Structural Time Series (MST) [3]. Although these methods can properly model autocorrelated data with trends and the seasonality, they are sensitive to the high variation. Therefore, the accuracy of these methods may deteriorate when data become overdispersed. Furthermore, a space-time Negative Binomial regression has been proposed to model overdispersed traffic flow [5]. However, this model does not account for trends and seasonal patterns.

The only forecasting method that handles time series trend and seasonality characteristics as well as overdispersion is the Holt-Winters method for count data proposed by Taylor and known as HWT count data method [7]. This method combines the double seasonal Holt-Winters with the Negative Binomial distribution. The double seasonal Holt-Winters method allows for the simultaneous capturing of small seasons that exist in large seasons such as the daily season within the weekly season. The Negative Binomial distribution allows the variance to exceed the mean during the modeling process by involving an overdispersion coefficient. The HWT count data method was shown more accurate than other smoothing methods which do not handle overdispersion [7].

* Corresponding author:
y.awwad@ptuk.edu.ps (Yousef-Awwad Daraghmi)

In this research, we propose a data mining method for modeling and forecasting traffic flow accurately. The proposed method utilizes the HWT count data method since the data used in this research are traffic flows lected from Taipei city and have all of the aforementioned characteristics. We model the weekly season as a large season and the daily season as a small season. The coefficients of overdispersion, trend, and seasonality are estimated during the training process using a likelihood function derived by Taylor [7].

The contribution of this research is firstly we propose a predictive DRG system by integrating an existing ATIS with prediction. We secondly propose a method that can be used to model and forecast the autocorrelated and overdispersed traffic flows. We thirdly compare the proposed method with the HW and the space-time NB regression methods, and we show that the proposed method outperforms the other two methods because it has smaller root mean square errors. This research also defines the temporal characteristics of traffic flow as seasonal patterns and categorizes these patterns into a weekly traffic season and a daily traffic season.

2. Related Work

As shown earlier, prediction is necessary for DRG. This sections reviews the most relevant methods that have been used for traffic flow forecasting in urban areas.

The field of data mining consists of variety of methods that are used to model or forecast time series [8], [9]. In traffic context, the accuracy of data mining methods that are used to model and forecast traffic flow time series depends on how these methods deal with its various characteristics. Here, we give a brief review of these methods in terms of which method can properly traffic flow temporal autocorrelation, trends, seasonal patterns, and overdispersion.

Several ARIMA based methods were used in the traffic context as in [2], [4], [6]. These methods were used to model different traffic variables such as speed, flow, or travel time. The strength of these methods is their ability to model autocorrelated data and capture the trends in the data. Also, these methods can be formulated as multivariate models to address the traffic conditions in several locations. However, the ARIMA based methods are usually sensitive to high variations and they become less accurate when data is overdispersed [7]. In traffic context, these methods were shown less accurate and more complex than the HW method [3], [4].

The HW method was used to model autocorrelated traffic flows measured from a single site since it is a univariate method [4]. The HW can capture trends and seasonal patterns in a single time series. For more accurate results, trends and seasonal patterns from different sites were treated with a multivariate structural time series (MST) [3].

However, these methods did not address the high variation or the overdis-persion of traffic flow. Ignoring the overdispersion may cause inaccurate forecast results [10].

To overcome overdispersion, a negative binomial regression that allows the variance to be greater than the mean was used to model traffic flows in Taipei city [5]. This method is called the space-time multivariate NB regression and models traffic flows from a set of correlated roads N. The multivariate space-time NB regression model can be written compactly as

$$\ln y_{d,t} = \alpha + \sum_{i=1}^{N} \sum_{j=1}^{n} \beta_{i,j} y_{i,j} + \eta \qquad (1)$$

where $y_{d,t}$ is the traffic flow value of the dependent road at time t, $1 \leq d \leq N$, α is the intercept, $\beta_{i,j}$ is the regression coefficient corresponding to road i at time j, n is the size of data, and η is the regression error vector where the error is independent of all covariates and distributed with mean = 1 and a variance= $1/\varphi$. The results of this method were more accurate than the HW method and the MST method.

The main shortcoming of the space-time multivariate NB regression method is its ability to model limited autocorrelated data and consequently being incapable of addressing trends or seasonal patterns in the data. A better option is a method that can capture trends, seasonality, and overdispersion. To achieve this, we adopted the HWT count data method derived by Taylor [7] by combining the HW with the Negative Bino-mial distribution which is popular for modeling overdispersed counts.

3. The Proposed DRG System

We adopt an existing ATIS from Taipei city and add a prediction server so that the ATIS can provide predictive DRG services. The Traffic Control Center of Taipei city provides ATIS services as well as Intelligent Transportation Systems (ITS) services. The current ATIS consists of three blocks that are collecting data, processing data into information related to ITS and ATIS, and providing information to travellers [11]. The Taipei-ATIS relies on historical data and real time data of traffic flow and speed, which are stored in databases, to provide a web-based route planning service and information regarding traffic conditions. Traffic conditions of each road are displayed as coloured lines on the road network map.

The availability of future values of traffic flow enables the ATIS to provide better route choices and more realistic traffic condition. Such future values can be obtained by the prediction server which is added to the processing data block in the traffic control center servers. The ATIS can capture the high variations of traffic and adapt its output to the new traffic condition. Fig. 1 shows the overview architecture of the proposed ATIS.
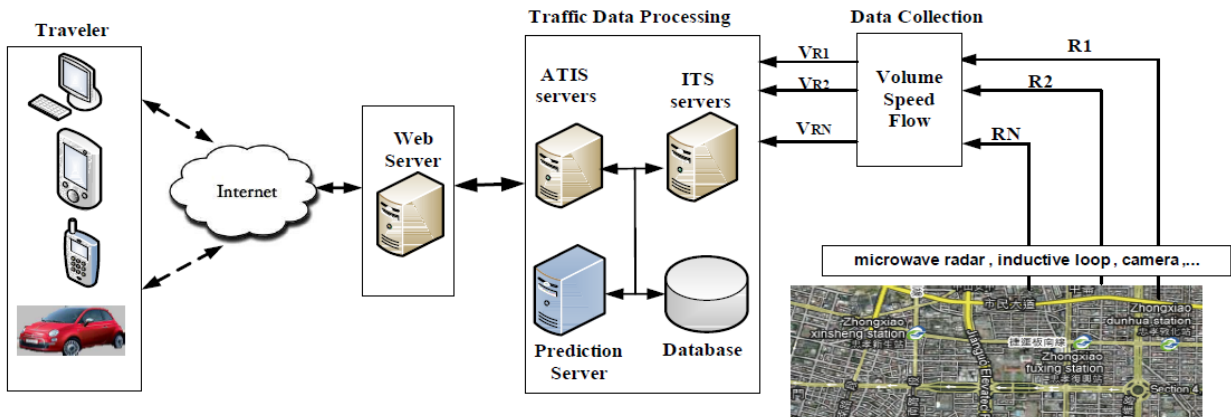
Figure 1. The overview architecture of the ATIS that includes the proposed prediction model within the prediction server

A. ATIS Components

The main components of the ATIS are:

- Traffic data collection: The traffic control center in Taipei uses different techniques such as microwave radar, inductive loops and video camera to collect traffic data mainly flow, flow and speed. Traffic flow measured on each road segment can be available to ATIS servers after a short time. These flows are considered as real time data because they describe the current traffic condition. The data are stored in a database and after sometime these data become historical data which can be represented by time series. Other information related to incidents can be also fed to the ATIS.

- Traffic data Processing: The historical and real time traffic data are processed by the ATIS servers to information that is useful to travellers and traffic control staff. Information includes traffic conditions on roads, travel time of routes, events on roads, traffic density, traffic speed and real time video of roads. This component also contains the prediction sever. It also contains the ITS servers which participate in data processing and provide information for congestion control, traffic light control and other ITS services.

- ATIS information delivery: Travellers can obtain traffic information such as route plan, travel time, and traffic conditions on demand via web. We initially tested the web-based ATIS although a large scale ATIS employs other means of communication to provide its services. The ATIS servers handle the information to web servers which is responsible for communicating with travellers. Travellers can access the ATIS website by desktops or ubiquitously by mobile devices such as smart phones and navigational devices.

B. The prediction method

The prediction method used in the proposed DRG shown in Fig. 1 is derived from the standard Holt-Winters (HW) which is also known as the triple exponential smoothing method. The HW method is the most commonly employed approach for modeling autocorrelated time series and seasonal patterns. It was introduced as a forecasting method in 1960 [12]. The HW method uses a time series model to make predictions assuming that the future will follow the same pattern as the past. Its mathematical forms include a model for additive seasonal patterns where the size of the seasonal variation is not affected by the value of the observation, and a model for multiplicative seasonal patterns where higher values of observations cause larger seasonal variations [13]. We focus on the multiplicative HW model only because the data in the research exhibit a multiplicative seasonal pattern.

In the HW method, a series value, Y_t, is decomposed into three components: the level (L_t), the trend (T_t) which de-scribes the long term direction of the series, and the seasonality (S_t) [13]. The multiplicative HW method is formulated as follows

$$\mu_t = L_{t-1}T_{t-1}S_{t-m} \qquad (2)$$

where μ_t is the mean value that is used to estimate the series, and the transition equations are

$$L_t = \alpha(y_t / S_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \qquad (3)$$
$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$
$$S_t = \gamma(y_t / L_{t-1}) + (1 - \gamma)S_{t-m}$$

where t is the time, α, β, γ are the HW smoothing coefficients and m is the number of periods in one seasonal cycle which is defined as any periodic pattern of fixed length [13]. The smoothing model error, ϵ_t, is assumed to be independent and identically distributed following a Gaussian distribution with mean = 0 and variance= σ^2, i.e., $\epsilon_t \sim NID\ (0, \sigma^2)$ [13]. The smoothing coefficients α, β and γ are usually restricted between zero and one, and they should be selected carefully to minimize residual errors. The equation to forecast future values, $\hat{Y}$, is given by

$$\hat{Y}_{t+q} = (L_t + qT_t)S_{t-m+q}. \qquad (4)$$

where q is the forecast horizon [13].

Usually data contain more than one cycle, hence Taylor developed a double seasonal Holt-Winters method that is called the HWT [14]. The developed method (HWT) improves the accuracy of the standard HW method [15]. The transition equations of the standard HW are modified to involve two seasonal indices, D_t and W_t instead of one seasonal index, S_t. The transition equations become

$$L_t = \alpha(y_t / D_{t-m1} \, W_{t-m2}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (5)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

$$D_t = \gamma(y_t / L_{t-1}W_{t-m2}) + (1 - \gamma)D_{t-m1}$$

$$W_t = \omega(y_t / L_{t-1}D_{t-m1}) + (1 - \omega)W_{t-m2}$$

where m_1 represent the number of periods in the daily season D_t and m_2 represent the number of periods in the weekly season W_t. Future values can be forecasted as follows

$$\hat{Y}_{t+q} = (L_t + qT_t) D_{t-m_1+q} W_{t-m_2+q} \quad (6)$$

where q is the forecast horizon.

The HW and the HWT methods do not handle overdis-persion. An improvement was proposed in [7] by combining the HWT with the NB distribution to allow the HW to model overdispersed count data. Overdispersed count data means that the variance is greater than the mean as $\sigma^2 = \mu + \varphi\mu^2$ where μ is the mean, σ^2 is the variance, and φ is the overdispersion parameter [10]. In this case, Taylor showed that the smoothing coefficients α, β, γ and ω as well as the overdispersion parameters φ and ψ can be estimated by maximizing a likelihood function given by

$$f_{HWT} = \prod_{i=1}^{n} \frac{\Gamma(Y_i + \mu_i\psi / (1-\psi))}{Y_i! \, \Gamma(\mu_i\psi / (1-\psi))} \psi^{u_i\psi/(1-\psi)} (1-\psi)^{y_i} \quad (7)$$

where n is the size of data and Γ is the gamma function [7]. The parameter ψ controls the degree of overdispersion where $\psi = 1 / (1 + \mu_i\varphi)$ such that $0 < \psi < 1$.

The resulting method is referred to as HWT count data method, and it improves the accuracy of the HW method [7]. The HWT count data method also outperforms other smoothing methods which do not account for overdispersion [7].

4. Methodology

In this section, we show how we apply the HWT count data method to the data used in this research. We also present and discuss the results of the forecasting.

A. Data

1) *Data set*: The data consists of traffic flows of 13 signalized arterial road segments in Taipei city. Fig. 2 shows the traffic flow direction and the road segments. Microwave Radar Vehicle Detectors were used to record the data for 19 days from January 21, 2008 at 00:00 to February 8, 2008 at 11:58 including four weekend days. The collection of dense data with short distance between two measuring sites, 200m to 400m, and short time interval between two successive observations, two minutes, is necessary to capture all flow patterns and variations.

Initially, the data were inspected and invalid records were found including missing flow values, negative flow values and zero flow values when speed is greater than zero. The invalid data resulted from detector malfunction or failure. The number of the invalid records of each road is less than 20 records per day which is not large and does not affect the accuracy of the modeling. To replace an invalid data record, we used an interpolation function that calculates the average of the preceding value and the following value of that record.

2) *data characteristics*: The traffic flow of a selected road, e.g. R8, is represented by a time series that is plotted in Fig. 3 for 19 days and in Fig. ?? for a single day. Fig. 3 illustrates that the traffic flow has a weekly pattern which shows that the flow in workdays is different than the flow in weekends. Further, Fig. 3 illustrates that the traffic flow has daily seasonal patterns which includes intraday seasonal patterns, Fig. 4. The Intraday seasons are categorized into: a low-traffic season when the flow is less than the mean and often exists in the early morning from 00:00 to 6:00, a high-traffic season when the flow is greater than the mean and occupies time periods from 7:00 to 9:00 and 17:00 to 19:00, and an average traffic seas1on when the flow is around the mean. The flow in all roads follows the same patterns.

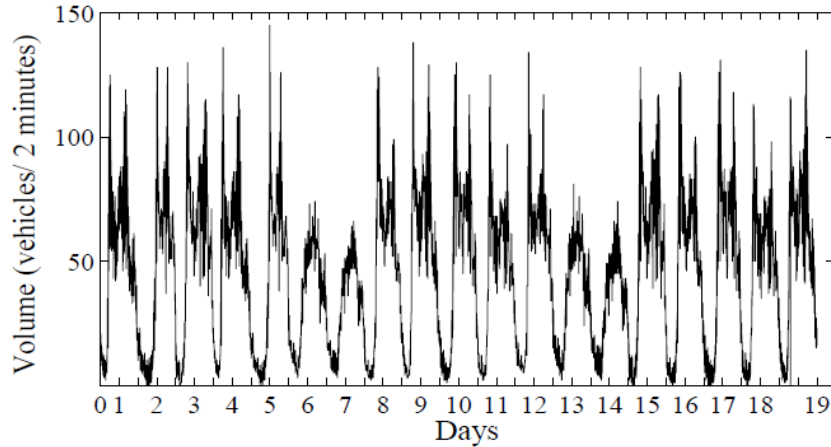**Figure 2.**   Map of selected roads in Taipei city

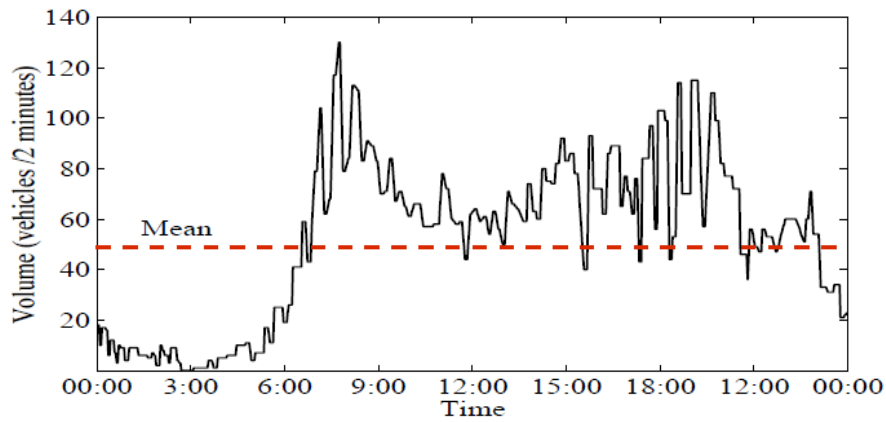**Figure 3.** The daily patterns of the traffic flow on R8 for 19 days

Figure 4. The intraday seasonal patterns of the traffic flow on R8

Additionally, the traffic flow has overdispersion since the variance of the flow for each road is greater than the mean as shown in Table 1. The flow overdispersion is tested using the dispersion value (the Pearson statistic (χ^2) divided by the degrees of freedom (*df*)), and data is overdispersed when the quotient is greater than one [10]. We find that all flows are overdispersed since the quotients for all roads are greater than one. The flow overdispersion is related to the flow variations within seasonal patterns and to flow fluctuation due to traffic lights. The green light phase decreases flow and the red light phase increases it, and the fluctuation is significant during rush hours, which agrees with the results in [16]. The variations may be also caused by other effects such as changing weather conditions, road work, and driving behavior.

Table 1. The mean, variance, overdispersion values of flows on selected roads

Roads	R1	R4	R5	R8	R9	R12	R13
Mean	45.4	39.3	54.9	51.2	48.7	60.9	62.9
Variance	865	830	853	1001	608	989	1270
(χ^2)/*df*	1.8	2.4	2.1	1.6	2.9	3.1	1.7

B. Modelling Process

The traffic flow contains weekly and daily seasons. Since the adopted method can model two seasons, we let D_t index represent the weekly season and W_t index represent the daily season. The data were recorded every two minutes, so the number of daily periods, m_1, is 720, and the number of weekly periods, m_2, is 5040.

The application of the adopted method includes training the models to estimate all required coefficients and then using the trained models for forecasting. We used 15-day data to train the models and four-day data to evaluate the forecast. The training process involves initializing the level and the trend indices as in [13], where the initial value of level L_0, which the overall mean, is given by

$$L_0 = (y_1 + ... + y_m)/m_2, \qquad (8)$$

and the initial value of trend, T_0, is given by

$$T_0 = \left[\left(y_{m2+1} + ... + y_{m2+m2} \right) - \left(y_1 + ... + y_{m2} \right) \right] / m_2^2. \quad (9)$$

The index of the large season (weekly season) is initialized by the ration of the observed data to the mean as

$$W_i = y_i / L_0 \text{ for } i = 1 \text{ to } m_2 \qquad (10)$$

and the index of the small season (daily season) is initialized as

$$D_i = y_i / A_0 \text{ for } i = 1 \text{ to } m_1 \qquad (11)$$

where A_0 is the daily mean of y [7].

The training process also involves estimating the smoothing and the overdispersion coefficients which can be estimated using the likelihood function, equation 7. The lower and upper boundaries of all coefficients were set to zero and one, respectively. A total of 10800 data records, lower and upper boundaries and initial values were used in a quasi- Newton algorithm to optimize the coefficients. The estimated coefficients are presented in Table 2. After the model was trained, we used it to forecast future values as in equation 6.

Table 2.   The smoothing coefficients estimated by the standard HW and the HWT-count data methods

Method	coef	Roads				
		R4	R5	R8	R9	R12
HW	α	0.541	0.466	0.591	0.487	0.596
	γ	0.811	0.827	0.831	0.844	0.796
HWT	α	0.657	0.553	0.657	0.566	0.745
	γ	1	1	1	1	1
	ω	0.898	0.918	0.892	0.917	0.874
	ϕ	0.291	0.381	0.367	0.381	0.299

We compared the HWT count data method with the HW method since it also accounts for temporal autocorrelation and seasonal patterns and with the space-time Negative Binomial regression since it accounts for the overdispersion. The daily season only was used in the HW, and initial values were set as in the HWT but without using the W_t index, The HW smoothing coefficients were estimated by minimizing the mean square error, as in [13], instead of maximizing a likelihood function. The space-time NB regression coefficients were estimated by maximizing a log-likelihood function, derived in [10], using flows from correlated roads, as in [5]. The Root Mean Square Error (RMSE) was used to evaluate the forecast accuracy and compare the three methods. The RMSE is usually used to evaluate the forecast accuracy because it gives an idea about the forecast error. The RMSE is given by

$$RMSE = \sqrt{\frac{\sum (observed\text{-}predicted)^2}{n}} \qquad (12)$$

C. Results

In this subsection, we firstly show the results of estimation of the HW, the HWT count data, and the space-time Negative Binomial regression methods. The values of the smoothing coefficients of the HW and the HWT count data methods for some selected roads are presented in Table 2, and the NB regression coefficients for R8 as a dependent road are presented in Table 3.

We do not include β in Table 2 because it is zero that means the overall trend is zero. The values of φ in Table 2 can verify that the variance is greater than the mean ($\sigma^2 = \mu + \varphi\mu^2$), and these values are different from the overdispersion values in Table 1 as they are used to model the data while the values in Table 1 are used to test the existence of the overdispersion. The high values of γ and ω show that the variation is dominated by the seasonality while the overall trend is zero. The coefficients in the table are used to forecast future values for one cycle ahead which is one day.

The space-time Negative Binomial regression method assists in identifying the significant variables that affect the forecast results. In traffic context, several upstream and downstream roads exist, and it is important to identify which road is significant. A variable is considered significant if its P-value is smaller than the significance level and the significance increases when the P-value decreases. The spatial and temporal variables with significant coefficients are presented in Table 3, and the insignificant coefficients are represented by (-). It is shown that for a dependent road, R8, the significant spatially correlated variables (predictors) are R4, R5, R6, and R7. Also, the significant temporally correlated variables are $t - 1$, $t - 2$, and $t - 3$ only which indicates that the NB regression can capture limited temporal autocorrelation. The significant variables are the only ones used as predictors in a multistep forecast as in equation 1.

Secondly, we discuss the forecast results of each method. The forecast results show that the three methods can forecast for different horizons: one day for the HWT method and the HW method and 20 minutes for the space-time NB regression. As we are interested in short-term prediction only, which is less than 20 minutes ahead, because of the continuous changes in traffic, we compare the models for 20 minutes horizons. So, our goal is identifying which forecasting method is more accurate for 20 minutes ahead. We include comparison results during different traffic conditions: low, average and high traffics. The RMSE values resulting from the tests are averaged and presented in Table 4.

The HWT count data method reduces the residual error significantly during all traffic seasons and consequently it has the lowest RMSE values for 20-minute forecasts. The HWT count data method handles the overdispersion and can capture different seasonal patterns, so it outperforms the space- time NB regression because the latter does not account for autocorrelations and seasonal patterns. It is also better than the HW because the HW does not account for the overdispersion.

Table 3. The space-time NB regression coefficients when R8 is the dependent road

Variables	Coefficients β					
	low season		average season		high season	
	estimate	P-value	estimate	P-value	estimate	P-value
Intercept α	2.13	2×10^{-16}	2.33	3.5×10^{-16}	2.75	5.1×10^{-16}
autocorrelation R8, $t-1$	0.0830	0.62×10^{-16}	0.0852	1.4×10^{-16}	0.0892	4.1×10^{-16}
autocorrelation R8, $t-2$	0.0680	3.3×10^{-10}	0.00753	4.7×10^{-11}	0.0795	3.4×10^{-12}
autocorrelation R8, $t-3$	0.0461	1.2×10^{-5}	0.0512	3.8×10^{-6}	0.0563	9.1×10^{-7}
autocorrelation R8, $t-4$	-	0.063	-	0.059	-	0.068
direct upstream R5, $t-1$	0.0161	6.1×10^{-9}	0.0177	2.3×10^{-9}	0.0262	1.8×10^{-10}
direct upstream R5, $t-2$	0.0061	2.0×10^{-9}	0.0082	5.4×10^{-9}	0.0158	1.3×10^{-9}
direct upstream R5, $t-3$	0.0012	7.5×10^{-6}	0.0025	1.7×10^{-6}	0.0047	2.9×10^{-7}
direct upstream R6, $t-1$	0.0123	1.1×10^{-9}	0.0145	2.4×10^{-9}	0.0280	1.6×10^{-10}
direct upstream R6, $t-2$	0.0042	2.4×10^{-7}	0.0071	4.2×10^{-9}	0.0104	2.6×10^{-9}
direct upstream R6, $t-3$	0.0017	4.1×10^{-6}	0.0037	9.2×10^{-6}	0.0055	3.5×10^{-7}
direct upstream R7, $t-1$	0.0091	2.7×10^{-9}	0.0127	2.4×10^{-9}	0.0151	1.7×10^{-9}
direct upstream R7, $t-2$	0.0024	1.4×10^{-8}	0.0035	4.5×10^{-8}	0.0064	2.4×10^{-9}
direct upstream R7, $t-3$	0.0009	4.1×10^{-4}	0.0015	3.1×10^{-5}	0.0033	5.3×10^{-6}
indirect upstream R4, $t-1$	0.0037	2.6×10^{-3}	0.0045	5.1×10^{-4}	0.0062	4.8×10^{-4}
indirect upstream R4, $t-2$	0.0008	0.006	0.0014	0.003	0.0042	0.001
indirect upstream R4, $t-3$	-	0.61	-	0.3	-	0.21
direct downstream R9, $t-1$	-	0.49	0.0001	0.044	0.0013	0.008
direct downstream R9, $t-2$	-	0.33	-	0.72	0.0011	0.023
indirect downstream R12, $t-1$	-	0.12	-	0.24	0.0017	0.02
distant R1, $t-1$	-	0.23	-	0.24	-	0.37
distant R1, $t-2$	-	0.15	-	0.58	-	0.52
distant R1, $t-3$	-	0.45	-	0.71	-	0.82
# of observations	1800		600		600	

Table 4. Comparison between the forecast methods for different roads in different traffic seasons

Traffic	Method	RMSE of 20 minutes forecast			
seasons		R4	R5	R8	R9
Low	HWT	2.24	2.21	1.4	2.65
	HW	3.35	2.72	2.16	3.42
	NB	3.13	2.26	1.79	2.45
Average	HWT	4.02	3.23	2.85	4.12
	HW	5.86	6.62	5.95	7.61
	NB	3.53	4.54	4.59	4.18
High	HWT	4.37	4.67	3.62	4.06
	HW	10.9	11.2	9.42	9.62
	NB	6.5	5.73	5.47	4.97

To illustrate how the methods behave in different forecast horizons, we plot the RMSE values of the methods against different prediction horizons during the low, average and high traffic seasons for R8, Fig. 5. We notice that the accuracy of all methods decreases when the horizon increases. The HWT count data method has the lowest RMSE during all traffic seasons and for all time horizons. The slope of the HWT count data method curve is also the lowest which means that the RMSE increase little over time. We decided a RMSE value of four as a threshold value where smaller RMSE values can be acceptable. The accuracy of the HWT count data method starts to deteriorate (RMSE > 4) after a 36-minute horizon during the low and average traffic seasons and after 24-minute horizon during the high traffic season. Meanwhile, the other methods behave similarly but they have larger RMSE values and start to deteriorate earlier than the HWT count data method. Clearly, the HWT count data method is the most accurate one.
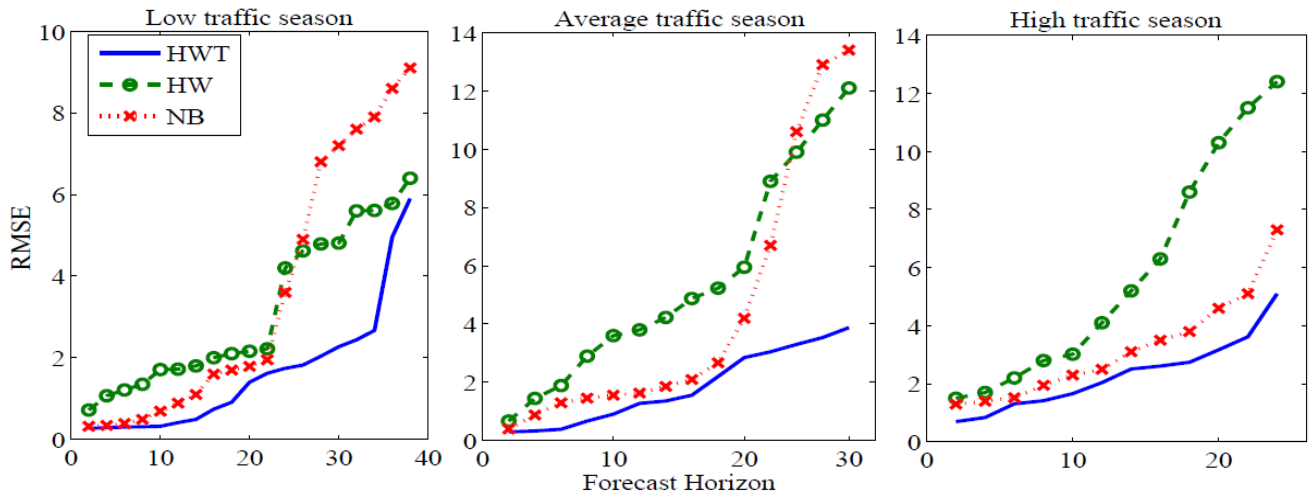
Figure 5.   RMSE against different forecast horizons during the three traffic seasons

5. Conclusions

The success of DRG requires accurate short-term prediction with low computational demand [17]. Accurate short-term prediction models enables DRG to provide adaptive services in real-time. Travellers will be able to select the best route pre-trip or on-trip based on current and future traffic conditions. Therefore, this paper has proposed a method for short-term traffic flow prediction in urban areas where flows are autocorrelated and overdispersed. The method addresses the most important characteristics of traffic flow including autocorrelation, trend, seasonality and overdispersion by adopting the HWT count data method. The proposed method can capture the weekly season and the daily season simultaneously. The comparison between the proposed method with the HW and the space-time NB regression methods states that the proposed method outperforms the others. This paper also has presented an ATIS that incorporates the proposed prediction model into its architecture.

The limitation of the proposed method is that it only models the weekly season and the daily season. Future work will investigate the possibility of using multiple seasonal methods to model all of the intraday seasons simultaneously. Further, the proposed method needs to account for not only one road segments, but also other correlated segments. This can be achieved by deriving a multivariate HW method that can handle overdispersion and capture trends and seasonal patterns of multiple flows on different correlated roads.

REFERENCES

[1]   A. Stathopoulos and M. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transportation Research Part C*, vol. 11, no. 2, pp. 121–135, April 2003.

[2]   Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington D.C., USA, January 2003, pp. 1–25.

[3]   B. Ghosh, B. Basu, and M. OMahony, "Multivariate short-term traffic flow forecasting using time-series analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 2, pp. 246–254, June 2009.

[4]   B. Basu, B. Ghosh, and M. M. O'Mahony, "Time-series modelling for forecasting vehicular traffic flow in Dublin," *in Proceedings of the 84th Transportation Research Board Annual Meeting*, Washington D.C., USA, January 2005, pp. 1–22.

[5]   Y.-A. Daraghmi, C.-W. Yi, and T.-C. Chiang, "Space-time multivariate negative binomial regression for urban short-term traffic volume pre- diction," *in 12th International Conference on ITS Telecommunications*, Taipei, Taiwan, November 2012.

[6]   M.-C. Tan, S. C. Wong, J.-M. Xu, Z.-R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 60–69, March 2009.

[7]   J. W. Taylor, "Density forecasting of intraday call center arrivals using models based on exponential smoothing," *Management Science*, pp. 1– 16, October 2011.

[8]   J. Han and M. Kamber, *Data Mining: Concepts and Techniques. Morgan Kaufmann*, 2006.

[9]   C. Chatfield, *Time Series Forecasting*. Chapman and Hall/CRC, 2000.

[10] J. M. Hilbe, *Negative Binomial Regression*, 2nd ed. Cambridge: Cambridge University Press, 2011.

[11] *Taipei Traffic Control Center*. [Online]. Available: http://tms.bote.taipei. gov.tw/main.jsp?lang=en.

[12] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol. 6, no. 3, pp. 324–342, April 1960.

[13] R. Hyndman, A. B. Koehler, K. Ord, and R. D. Snyder, "Linear inno- vation state space models," *in Forecasting with Exponential Smoothing: The State Space Approach*, 1st ed., R. Hyndman, A. B. Koehler, K. Ord, and R. D. Snyder, Eds. Berlin: Springer, 2008, ch. 3, pp. 33–51.

[14] J. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of Operational Research Society*, vol. 54, no. 8, pp. 799–805, August 2003.

[15] J. W. Taylor, "Exponentially weighted methods for forecasting intraday time series with multiple seasonal cycles,"

International Journal of Forecasting, vol. 26, pp. 627–646, 2010.

[16] T. Thomas, W. Weijermars, and E. van Berkum, "Predictions of urban volumes in single time series," *IEEE Transactions on Intelligent Trans- portation Systems*, vol. 11, no. 1, pp. 71–80, March 2010.

[17] S. Lin, B. D. Schutter, Y. Xi, and H. Hellendoorn, "Efficient network- wide model-based predictive control for urban traffic networks," *Trans- portation Research Part C*, vol. 24, pp. 122–140, October 2012.