Negative Binomial Additive Models for Short-Term Traffic Flow Forecasting in Urban Areas

Yousef-Awwad Daraghmi, Chih-Wei Yi, and Tsun-Chieh Chiang

***Abstract*—Parallel, coordinated, and network-wide traffic management requires accurate and efficient traffic forecasting models to support online, real-time, and proactive dynamic control. Forecast accuracy is impacted by a critical characteristic of traffic flow, i.e., overdispersion. Efficiency depends on the time complexity of forecasting algorithms. Therefore, this paper proposes a novel spatiotemporal multivariate forecasting model that is based on the negative binomial additive models (NBAMs). Negative binomial is utilized to handle overdispersion, and additive models are used to efficiently smooth nonlinear spatial and temporal variables. To evaluate the model, it is applied to real-world data collected from Taipei City and compared with other forecasting models. The results indicate that the proposed model is an accurate and efficient approach in forecasting traffic flow in urban context where flow is overdispersed, autocorrelated, and influenced by upstream and downstream roads as well as the daily seasonal patterns, namely, low-, moderate-, and high-traffic seasons.**

***Index Terms*—Additive models, autocorrelation, multivariate, negative binomial (NB), overdispersion, seasonal patterns, short-term forecast, spatial correlation.**

I. Introduction

THE TRANSPORTATION network is a complex system where the behavior of the entire system cannot be determined by examining one part of the network [1]. In the new era of intelligent transportation systems (ITS), research has focused on managing the transportation network entirely through parallel or coordinated manners. In parallel or coordinated traffic management, continuous forecasting of traffic conditions for short time ahead is necessary to enable proactive dynamic traffic control [1], [2]. Forecasting not only provides information about future traffic conditions but also allows for prelaunch control decision evaluation by determining the future impact of these decisions on traffic. Forecasting models should have high accuracy and low time complexity to be applied practically and efficiently.

Focusing first on accuracy, the literature of ITS reveals that accuracy is improved by addressing multiple spatial and temporal characteristics of traffic flow; hence, multivariate forecasting models (e.g., [3]–[13]) are more accurate than univariate ones (e.g., [14]–[26]) since multivariate models account for the spatial correlations of flows on upstream and downstream roads [3]–[7]. They simultaneously address traffic flow temporal characteristics, including autocorrelation and seasonal patterns such as the weekly and daily patterns [3]–[5]. The most accurate and widely used multivariate traffic forecasting models were derived from various statistical or data mining methods. Examples of these models are the neural-network-based models (e.g., [3], [27], and [28]) and the time-series models such as smoothing- and autoregression (AR)-based models (e.g., [4], [5], and [9]).

Although previous studies provide the ITS with fundamental traffic theories and forecasting models, overdispersion has not been meticulously addressed. Some models did not consider the problem of overdispersion at all. Overdispersion occurs when the variance is greater than the mean, and it is a critical characteristic because it reduces accuracy and causes false judgment about the significance of correlated variables if it is not appropriately handled [29], [30]. Traffic flow exhibits overdispersion since vehicles in urban and signalized area have high fluctuating arriving and leaving rates during each traffic season [11], [20]. Therefore, overdispersion should be handled in traffic flow forecasting models. However, the used multivariate models such as smoothing- or AR-based models assume equidispersion or use distributions that do not assist in capturing rapid variations of flow in urban areas [3], [31]. Further, neural-network-based models require much time and extensive training of homogeneous traffic conditions to handle high variations [3].

Second, the time complexity concept quantifies the amount of time required to run a specific algorithm. Determining the time complexity of a forecasting model assists in judging whether this model can be applied to real-time traffic applications or not. Although this concept is crucial to the success of future ITS practical solutions, few studies [2], [5] have quantified forecasting computational time. Other studies (e.g., [4], [6], [14], and [32]) have implicitly referred to this concept as computational demand, efficiency, or time demand and specified that traffic systems require fast forecasting algorithms that process data quick enough to enable real-time decisions. To reduce computation cost, studies have used data from limited

Manuscript received January 31, 2013; revised May 10, 2013, July 10, 2013, and September 4, 2013; accepted October 11, 2013. Date of publication November 20, 2013; date of current version March 28, 2014. The work of C.-W. Yi was supported in part by the Ministry of Transportation and Communication under Grant MOTC-STAO-101-02; by Sinica under Grant AS-102-TP-A06; by the National Science Council under Grant NSC101-2628-E-009-014-MY3 and Grant NSC101-2219-E-009-028; by the Industrial Technology Research Institute under Grant 101-EC-17-A-05-01-1111, Grant 101-EC-17-A-03-01-0619, Grant B352BW1100, and Grant B301EA3300; by the Ministry of Education Aim for the Top University Plan; by D-Link, Inc.; and by HTC, Inc. The Associate Editor for this paper was Q. Kong. *(Corresponding author: C.-W. Yi.)*

Y.-A. Daraghmi and C.-W. Yi are with the Department of Computer Science, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: yousef@cs.nctu.edu.tw; yi@cs.nctu.edu.tw).

T.-C. Chiang is with the Division for Telematics & Vehicular Control System, ICL/ITRI, Hsinchu 31040, Taiwan (e-mail: tcchiang@itri.org.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2013.2287512

correlated roads instead of examining all roads. Further, studies have tried to adopt efficient forecasting methods. However, most widely used models such as AR- and neural-network-based models require much computational time [4], [5], [33].

In this paper, we adopt negative binomial additive models (NBAMs) derived from generalized additive models (GAMs) and negative binomial (NB) distribution, and we propose a novel spatiotemporal multivariate NBAM. The GAMs require low computational complexity to capture data nonlinearity and nonstationarity by nonparametric smoothing of all covariates [34], and the NB is the best overdispersion handling method [29], [30]. Therefore, the proposed model efficiently handles overdispersion and smooths multiple nonlinear spatial and temporal variables [30]. For evaluation, the proposed model is applied to data collected from Taipei City and compared with other classic forecasting models. The results show that the accuracy and the efficiency of the proposed spatiotemporal NBAM are very satisfactory. In summary, this paper first contributes accurate and efficient spatiotemporal multivariate NBAMs for forecasting traffic flow in low-, moderate-, and high-traffic conditions. Second, it outlines that overdispersion is an important characteristic that should be treated by forecasting models. Third, it reduces computational cost by using an efficient model and by employing statistical measures to identify all significant spatially correlated roads instead of assuming correlation from specific roads.

II. Theoretical Background

A. GAMs

The GAMs are based on the additive models, which assume that the mean of a response variable depends on predictors through a nonlinear function. To illustrate additive models, let $\mathbb{Y}$ be a random variable (RV) that is correlated with a set of other RVs (predictors) $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_r$. The dependence of $\mathbb{Y}$ on the predictors is defined as follows:

$$y_i = \sum_{j=1}^{r} s_j(x_{ij}) + \varepsilon_i \tag{1}$$

where y_i and x_i are the ith observations in data of size n, $s_j(x_{ij})$ is a smooth function of the ith observation of covariate j, and ε_i is the residual that is independent of the covariates where $\epsilon_i \sim NID(0, \sigma^2)$ [34]–[36]. If the dependence of $\mathbb{Y}$ on $\mathbb{X}$ is not linear, summarizing it with a straight line, as in linear regressions, would be inappropriate. Therefore, the additive models replace the usual linear function with an unspecified smooth function [34], [36].

The additive models are nonparametric since no parametric form is imposed on the smooth functions; however, they are iteratively estimated in the fitting phase by using scatterplot smoothers such as a regression spline or a cubic spline [34]–[36]. Smoothers fit a smooth function of one covariate with data by generating a continuous curve consisting of small sections joined together through knots [34], [36]. The total number of sections in one curve is z, and each small section has an equation such as linear regression or a cubic polynomial. Each equation is composed of a base function and a coefficient. The base functions of the curve form model matrix $\mathbf{X}_{n\times z}$, and the coefficients form parameter matrix $\beta_{z\times 1}$. Details on calculating the base functions and model matrix $\mathbf{X}$ are in [35] and [36]. The degree of smoothness is controlled by parameter λ, which is calculated in cross-validation procedures so that the mean square error is minimized [34], [36]. λ determines the number of sections (z) in a single continuous curve. When multiple covariates exist, their smooth functions are simultaneously fitted via a backfitting algorithm in an iterative manner where smoothers are called in each iteration [34], [36].

The idea behind GAMs is generalizing the additive models to allow the response variable to follow an exponential distribution or have a known mean variance relationship [34]–[36]. The GAMs extend the generalized linear models, which assume linear dependence, by allowing the dependence to be nonlinear [34], [36]. Consequently, they allow greater flexibility than the parametric models and enable the discovery of a hidden structure in the relationship between the response and the predictors. In GAMs, the mean μ of $\mathbb{Y}$, which is $\mu = E(\mathbb{Y}|\mathbb{X}_1 = x_1, \ldots, \mathbb{X}_r = x_r)$ is linked to the predictors by

$$G(\mu) = s_0 + \sum_{j=1}^{r} s_j(x_{ij}) + \varepsilon_i \tag{2}$$

where G is a link function, and s_0 is the intercept that is equal to the overall mean of the dependent variable $\mathbb{Y}$ [34], [36]. A likelihood function is usually formalized to enable better fitting with data. The smooth functions in GAMs are estimated using a local scoring algorithm that maximizes a likelihood function [34]–[36]. To avoid overfitting, the generalized cross-sectional validation is used [35].

B. NBAM

The NBAM is a special case of the GAMs. The NBAM was derived to overcome the main limitation of GAMs, which is the ability to model only data with an exponential distribution [30]. The response and the predictors may not only have nonlinear dependence but also follow other distributions mainly when overdispersion exists, i.e., the variance is greater than the mean. In this case, the best distribution is the NB [29]. The NBAMs extend the GAM framework to handle overdispersion by allowing $\mathbb{Y}$ to follow an NB distribution [30]. The NB distribution handles nonstationarity and high fluctuation of data by allowing the variance σ^2 of $\mathbb{Y}$ to exceed the mean as $\sigma^2 = \mu + \varphi\mu^2$, where φ is the overdispersion parameter [29]. To additively fit the NB model for $\mathbb{Y}$ given covariates $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_r$, a natural log link function is chosen. The NBAM can be written as

$$\log E(\mu) = s_0 + \sum_{j=1}^{r} s_j(x_{ij}) + \varepsilon_i \tag{3}$$

where the smooth functions are iteratively trained using a local scoring algorithm by estimating μ and φ that maximize a log-likelihood function given in [30].

Modeling a large data set from complex context such as urban traffic data, which contain many correlated variables,

requires much computational time. To minimize computational demand, it is necessary to consider only significant predictors for the forecast accuracy. This optimizes the NBAM to the best-fit model that contains only the significant covariates and fits the data with minimum error. Significant predictors are determined using evaluative measures, including the P-value, the Akaike information criterion (AIC) score, the generalized cross-validation (GCV) score, and the deviance explained (DE) score [29], [35]. A covariate is significant when its P-value is smaller than the significance level, and its significance increases when its P-value decreases [35]. The AIC, GCV, and DE are given by

$$AIC = 2(par) - 2\ln(\varrho(M_i)) \tag{4}$$

$$GCV = nD(n - Dof)^2 \tag{5}$$

$$DE = -2(\varrho(M_5 i - 1) - \varrho(M_i) \tag{6}$$

where M_i is the new model when a new covariate is added, $\varrho(M_i)$ is the value of the maximum-likelihood function of the model M_i, par is the number of estimated parameters, and Dof is the degree of freedom [29], [35].

The GAM forecasting models require a set of data, model matrix $\mathbf{X}$, and model parameter matrix β [35]. The set of data contains observations of predictors that are used to know the future values of the response. These data are used to calculate the model matrix, which is referred to here as prediction matrix $\mathbf{X}^p$. The prediction matrix maps the model parameters β to the forecast of the response ($\check{\mathbb{Y}}$) as

$$\check{\mathbb{Y}} = \mathbf{X}^p \beta \tag{7}$$

where the model parameters β are obtained in the training phase [35]. The total time complexity of the NBAM is $O(n)$, i.e., it depends mainly on the data size of one covariate [34], [36]. The GAM-based models are not affected by a small amount of missing data [29], [35]. Different examples about GAM-based models are shown in the literature as in [37]–[40].

III. Spatiotemporal NBAMs

Here, we accommodate the NBAM to vehicular traffic by proposing a spatiotemporal NBAM. To study the spatial correlation, all road segments are analyzed in turns. For a targeted road segment (dependent) R_i, we classify other segments into upstream segments, downstream segments, and adjacent road blocks. The upstream segments should be included in the model because they are the main source of traffic flow to the targeted segment. The downstream segments should be also included because the downstream traffic may affect the upstream during high traffic, mainly when road segments are short [4], [9], [41]. Finally, the adjacent road blocks containing merging/diverging maneuvers and parking lots should be included because they contribute to the traffic on the dependent road.

Let road i have sets of upstream roads, downstream roads, and adjacent road blocks U, D, and B, respectively. Let $v_{i,t}$ be the response variable corresponding to traffic flow on road i at time t and $v_{j\to i,t-1}$ be the predictors corresponding to traffic flow on any spatial segment j at time $t-1 (j \in U \cup D \cup B)$. The temporal autocorrelation can be controlled by including a one-step lag flow of the dependent road as a predictor, i.e., $v_{i,t-1}$. The proposed spatiotemporal NBAM can be written as

$$\log v_{i,t} = \mu_i + s_t^{Temp}(v_{i,t-1}) + \sum_{j \in U \cup D \cup B} s_{j\to i}^{Spat}(v_{j\to i,t-1}) + \varepsilon_t. \tag{8}$$

Lag values $v_{i,t-1}$ are included to account for the short-term persistence of flow and to continuously update the model with the data trend. Moreover, lag value $v_{j\to i,t-1}$ is included to account for travel time on road segments and the delay in data collection. Equation (8) is for fitting the smooth functions to data according to the NB distribution and calculating model $\mathbf{X}$ and parameter matrices β, which will be used for forecasting as in (7).

The step-forward approach, where covariates are added one at a time, is followed to find the best-fit model. The code in Algorithm 1 describes the procedure to compute the smooth functions and determine the significance of each covariate. Line 7 calls the local scoring and backfitting algorithms (proposed in [34]) to compute the smooth functions. Each time the smooth function of the covariate is estimated, the evaluative measures, including the P-value, AIC, GCV, and DE scores, are calculated to test the significance of the added covariate and the new model [29], [35]. This has two advantages: First, it reduces the computational demand of forecasting by incorporating only significant roads instead of dealing with a large number of variables. Second, it allows examining all correlated roads in one area instead of choosing some roads and ignoring others.

Algorithm 1 *Model Training* (dependent variable $\mathbb{Y}$, covariates $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_r$)

INPUT
1: y_i // flow on the dependent road i
2; $x_1, x_2, \ldots, x_r$ // flows on temporal and spatial covariates
BEGIN
3: $M = \varnothing$ // initializing the model
4: $s_0 = Mean(y_i)$
5: set j to covariate index
6: **for** $j = 1$ to r **do**
7: Compute $s_j(x_j)$, P-value of $s_j(x_j)$, AIC, GCV, DE
8: **if** P-value of $s_j(x_j) < 0.05$ **then**
9: **if** $|M| = 0$ **then**
10: $AICm = AIC; GCVm = GCV; DEm = DE$
11: $M = \{j\}$
12: **end if**
13: **if** $AICm < AIC$ and $GCVm < GCV$ and $DEm > DE$ **then**
14: $M = M \cup \{j\}$
15: **else**
16: $AICm = AIC; GCVm = GCV; DEm = DE$
17: **end if**
18: update model $y = s_0 + \sum_{j \in M} s_j(x_j)$
19: **end if**
20: **end for**
21: **return** model $y = s_0 + \sum_{j \in M} s_j(x_j)$

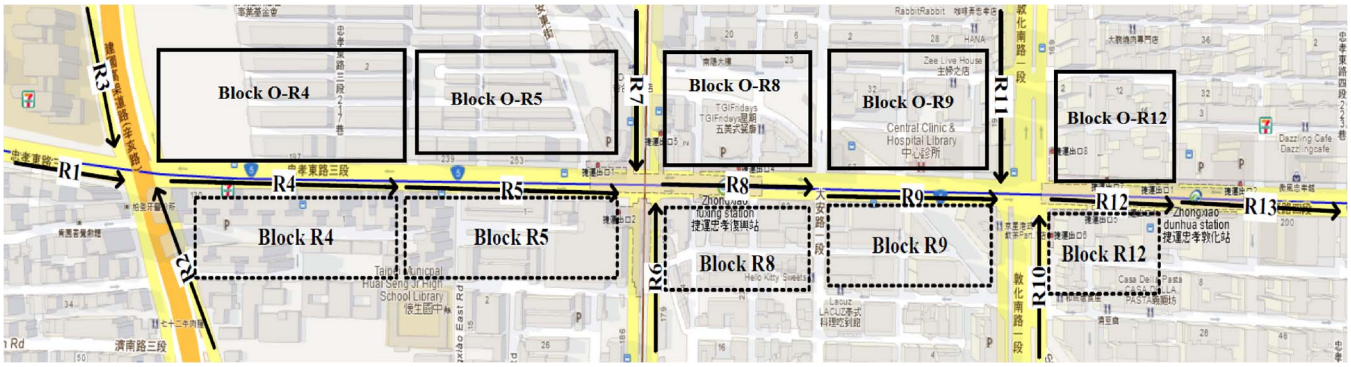

Fig. 1. Map of selected roads in Taipei City.

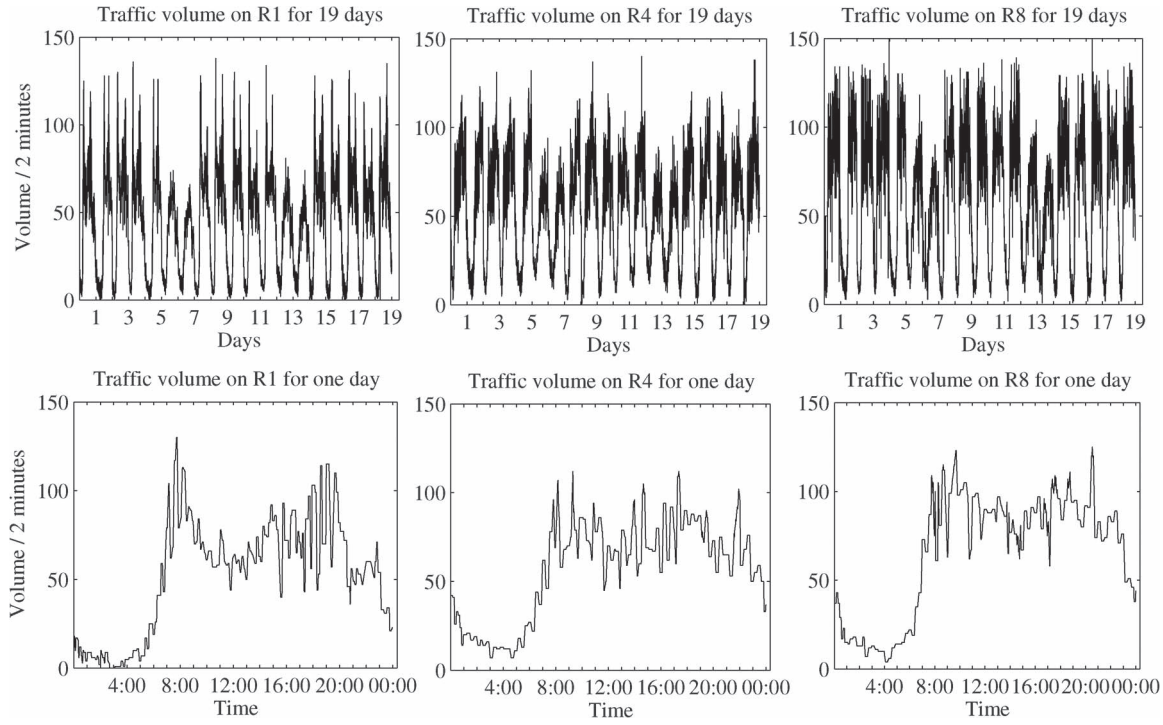

Fig. 2. Traffic flow on selected roads. (Upper part) Weekly pattern. (Lower part) Daily pattern.

Algorithm 1 is in the simplest form, whereas in reality, for obtaining improved results, a brute-force search is applied to select all possible significant covariates. The proposed model can be generalized to other networks as traffic flow on most urban networks shares the same characteristics: spatial correlation, autocorrelation, and M-shape traffic demand (as in [32]).

IV. Application of the Models

Here, we apply the proposed models and the other models for comparison to a data set, and we present the empirical findings.

A. Traffic Flow Data Set

The data consist of traffic volumes recorded every 2 min from 13 signalized arterial segments in Taipei City. Fig. 1 shows the traffic flow direction, the road segments, and the adjacent road blocks in the Zhongxiao Fuxing area. Microwave radar vehicle detectors were used to record the data in January, February, and April of 2008. The volumes on the road blocks are estimated based on the flow conservation principle by calculating the difference between the number of vehicles entering and leaving a road segment. The collection of dense data from three lanes per direction with a short distance between two measuring sites, i.e., from 200 to 400 m, and with short aggregation time, i.e., 2 min, is necessary to capture all flow patterns and variations in such a busy and crowded area.

Before starting the data analysis, the data were inspected, and invalid records were found, including missing values, negative-volume values, and zero-volume 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 (2%) per day, which is not large and does not affect GAM accuracy. To fix an invalid data record, we used an interpolation function that calculates the average of values measured at the same time of the day from the previous weekdays.

TABLE I
MEAN, VARIANCE, AND DISPERSION VALUES ON SELECTED ROADS DURING THE THREE TRAFFIC SEASONS

Road segment	R1	R4	R5	R8	R9	R12
daily mean μ	45.4	39.3	54.9	47.2	48.7	60.9
daily Var σ^2	865	830	853	1001	608	989
daily dispersion	1.8	2.4	2.1	1.6	2.9	3.1
low season μ	31	34	26	28	22	26
low season σ^2	520	645	414	491	526	539
low dispersion	2.9	2.7	2.5	2.7	3.6	3.7
moder season μ	57	61	62	64	59	56
moder season σ^2	407	412	488	372	428	451
moder dispersion	2.6	2.2	1.8	2.1	3.0	3.0
high season μ	88	84	79	87	91	85
high season σ^2	330	284	266	304	327	241
high dispersion	2.0	1.7	1.4	1.5	2.6	2.3

B. Traffic Flow Characteristics

The traffic volumes of roads are represented by time series that are plotted in Fig. 2 for selected roads during 19 days and a single day. The figure illustrates how the data have weekly and daily seasonal patterns, which are periodically repeated. The weekly pattern includes the workday pattern and the weekend pattern. The daily pattern includes different seasons, which we categorize as: 1) a low-traffic season when the volume is less than the mean and that often exists in the early morning from 00:00 to 06:00; 2) a high-traffic season when the volume is greater than the mean and that occupies time periods from 07:00 to 09:00 and from 17:00 to 19:00; 3) a moderate-traffic season when the volume is around the mean. These patterns are common in urban transportation and called the M-shape because the daily traffic demand is similar to the letter M [32].

Traffic flow in the data exhibits spatial correlation among different roads in the upstream and downstream directions. Further, as shown in Fig. 2, the flow is temporally autocorrelated and is not stationary as the daily seasons have different lengths, and within these seasons, the mean and variance are time dependent and have different values. These characteristics are presented in related studies which demonstrated that traffic flow is always autocorrelated and has nonstationary behavior.

Additionally, the traffic volume has overdispersion since the variance of the volume for each road is greater than the mean, as shown in Table I. The table presents the mean, variance, and dispersion for several roads during the three seasons. Overdispersion is identified when the dispersion value (the Pearson statistic (χ^2) divided by the degrees of freedom) is greater than 1 [29]. In the data set, all volumes are overdispersed since their dispersion values are greater than 1. In these data, fluctuation is also high during high-traffic seasons, which is similar to the results in [15].

The volume overdispersion is related to the volume variations within seasonal patterns and to volume fluctuation due to traffic lights. Red light increases the volume, and green light decreases it. The examined area has ten different timing plans in each weekday. This increases the variation mainly as the time for the red/green light in each plan depends on the flow patterns and the lane location. For example, in the middle lane, the green and red lights periodically switch every 60 s during low traffic (00:00–01:00), whereas the green light lasts 108 s and the red light lasts 92 s during high traffic (07:00–09:00). The left lane has different timing since vehicles must wait for the flow in the other direction to stop. In this paper, we consider only the flow in the middle lane for the analyses to control the value of overdispersion as the middle lane has the same light timing on roads R1, R4, R5, R8, R9, R12, and R13. The variations may be also caused by other effects such as changing weather conditions, road work, and driving behavior [15].

C. Training the Models

We perform analyses on roads R4, R5, R8, R9, and R12 because the data of their upstream roads, downstream roads, and road blocks are available. To train the models of these roads, we use the data recorded in January and February, excluding weekends because they have different patterns. The total number of days is 44 (23 Jan and 21 Feb). In all analyses, the significance level is set to 5% as the confidence interval is 95%.

1) Spatiotemporal NBAM: We model each traffic season separately using three two-month-long training data sets: a set for low-traffic season from 00:00 to 07:00, a set for moderate-traffic season from 09:00 to 17:00 and from 20:00 to 00:00, and a set for high-traffic season from 07:00 to 09:00 and from 17:00 to 20:00. Consequently, each season will have a forecast model. This is because each season has a different spatial correlation. For instance, the downstream segments have a significant impact on its upstream roads during high traffic but not during low or moderate traffic. In addition, a model for each season eliminates outliers and reduces the amount of variability and irregularity as the flow does not fluctuate between peak and off-peak.

The estimation starts by choosing a dependent road and estimating the smooth functions of the covariates. To simplify the estimation process, a new covariate is added in the following sequence: the temporal autocorrelation covariate, upstream roads, downstream roads, and road blocks. As we may have many correlated road segments in a complex traffic network, we need to select only the significant segments for the forecast accuracy using the evaluative measures, including the P-value, AIC, GCV, and DE scores [29], [35]. This optimizes the NBAM to the best-fit model that contains only the significant covariates and fits the data with minimum error. The analyses were conducted using the R $mgcv$ package introduced in [35]. This package contains all coding libraries for estimating smooth functions, evaluative measures, and forecast results.

The output of this stage is a smooth function for each covariate with an associated P-value, a model matrix $\mathbf{X}$, and a model parameter vector β as well as AIC, GCV, and DE for the final model. Due to space constraints, we only provide results for R8. We do not provide the model matrices and the coefficient vectors since they are very large. Each covariate has a λ value of around 0.002, making the number of sections in a curve very large. Table II presents the main outputs of the NBAM when R8 is the dependent variable. The values of φ in the table are greater than zero, which justifies the use of the NB to handle overdispersion.

The autocorrelation part of R8 is the most significant during the three seasons since the P-value is the smallest (see Table II).

TABLE II
MAIN OUTPUTS OF THE NBAMs WHEN R8 IS THE DEPENDENT ROAD

Traffic → / Covariates ↓	low P-value	Moderate P-value	high P-value
autocorrelation R_8	2×10^{-16}	2×10^{-16}	2×10^{-16}
upstream R_5	4×10^{-5}	7.2×10^{-5}	1.04×10^{-5}
upstream R_6	3.6×10^{-4}	2.5×10^{-4}	4.9×10^{-4}
upstream R_7	0.36	0.024	0.047
upstream R_4	6.4×10^{-3}	3.3×10^{-4}	2.5×10^{-4}
downstream R_9	0.091	0.064	0.014 *
downstream R_{12}	0.26	0.17	0.025 *
upstream R1	0.41	0.24	0.47
downstream R13	0.57	0.58	0.52
road block of R8	0.62	0.01 **	0.072
road block of O-R8	0.77	0.036 **	0.15
Intercept μ_{R8}	2.7	3.4	4.3
overdipsersion φ	0.41	0.43	0.39
n	9240	15840	7920
z	38	31	27
size of X	9240×38	15840×31	7920×27
The scores of the model using significant covariates only			
AIC	16750	15429	16.644
GCV	36.82	33.63	34.15
DE	94.2%	94.6%	94.3%

This indicates that the volume is mostly correlated with its previous values. The upstream segments (R4, R5, R6, and R7) are significant during all traffic seasons as their P-value is smaller than 5%. The significance of R5 and R6 is higher than the significance of R4 as their P-values are smaller than the R4 P-value. The significance of upstream roads decreases when the distance between the targeted road and the upstream road increases. We can also notice that although R7 is in the direct upstream, its significance is low since few cars turn left on that intersection. The downstream segments (R9 and R12) are significant during the high-traffic season only as the P-values ($*$ in Table II) become smaller than 5%. Segment R9 is more significant than R12. Road blocks R8 and O-R8 are only significant during the moderate-traffic season as indicated by the P-value ($**$ in Table II). The distant upstream and downstream road segments such as R1 and R13 are insignificant as their P-values are greater than 5%. Adding insignificant segments to the model increases the AIC and GCV scores and decreases the DE.

To evaluate the fitted models, we compare the fitted values of each covariate with its observed values as in Fig. 3, and we also perform residual analyses as in Fig. 4. The residual plots against the covariate show that the residual points are randomly dispersed around the x-axis. The last plot is the residual histogram, which illustrates that the residual has an approximately normal identical distribution. The residual analyses show that the models fit with data with minimum errors.

The aforementioned results are also found for other roads (R4, R5, R9, and R12). The NBAMs should incorporate the temporal autocorrelation covariate and some spatial covariates. In agreement with traffic flow theories and previous studies [3], [4], [9], [27], the upstream roads are found to be always significant, and the downstream roads are only significant during high traffic due to the backward spread of traffic. Additionally, the results determine how far we can go in the upstream and downstream directions. This minimizes the computational complexity of the model by including only significant upstream and

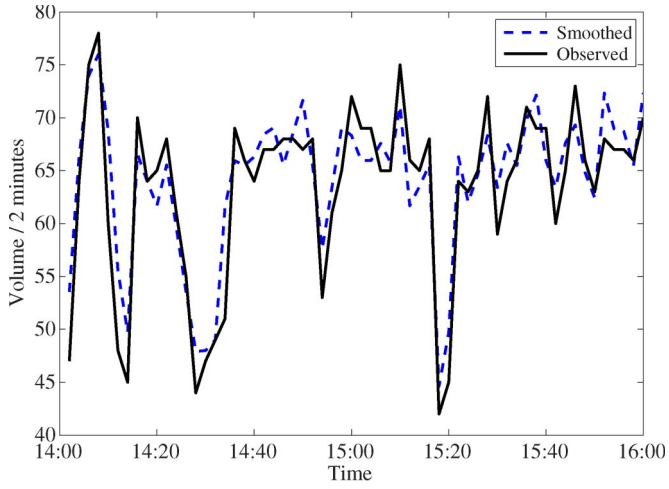

Fig. 3. Example of the smooth function of R5 during the moderate-traffic season.

downstream segments. The above analyses assist in reducing the three forecasting models to a single model by adding a categorical variable $k = 0$ or 1. The downstream segments are multiplied by $k = 1$ during high traffic and $k = 0$ for other seasons. The adjacent road block is multiplied by $k = 1$ during average traffic and $k = 0$ for other seasons.

2) Other Models: The HW, ARIMA, MSTARMA, and MST are trained, and their parameters are estimated using the assumptions in the original papers; however, they are accommodated to our training data set. For example, the HW coefficients for the level, trend, and seasonality are estimated as in [17] using one day cycle of length 720. The coefficients are presented in Table III.

As shown in Table III, the trend coefficient is zero, which means that the overall trend of the flow is zero, but there is only seasonal pattern expressed by the high seasonal coefficient.

The ARIMA model is estimated using one day cycle of length 720. Seasonal differencing is used to remove nonstationarity as in [17]. Different variations of ARIMA coefficients are used to find the best model that has the lowest AIC and MSE. The best variation is ARIMA$(2,1,2)(0,1,0)_{720}$. This means that the model contains two nonseasonal AR components, two nonseasonal moving-average components, and no seasonal lags. The ARIMA coefficients are shown in Table IV.

The MSTARMA is estimated using six templates corresponding to peak (daytime from 07:00 to 20:00) and off-peak (nighttime) for three sets of days (Mon and Fri; Tue, Wed, and Thu; and Sat and Sun) as in [5]. The time lag is set to two, which means that each template will have two spatial correlation matrices, e.g., $mat(a = 1, 1)$ and $mat(a = 1, 2)$. Because we have 13 road segments, the size of each matrix is 13 × 13, and an element (i, j) of $mat(a = 1, 1)$ is one if vehicles on road i reach road j within one time step (2 min in our data). The model in [5] is fitted to data using the training data set. Examples of the autocorrelation coefficients are shown in Table V for R8 in the "Tue, Wed, and Thu" peak template.

The table shows that not only the upstream roads affect R8 but also the downstream road R9 due to the backward spread of congestion.

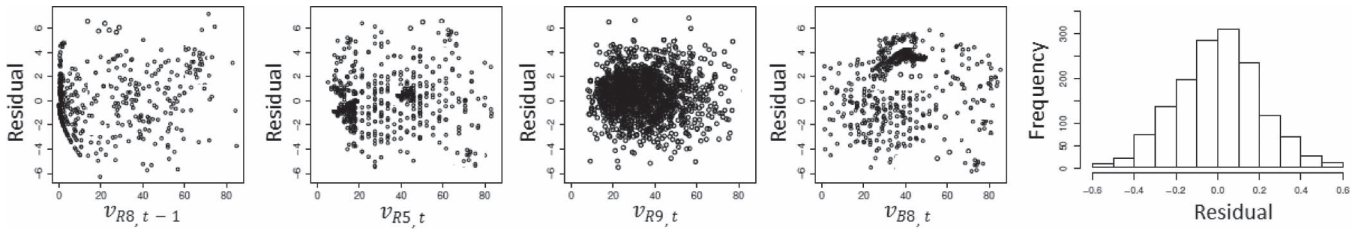

Fig. 4. Residual plots against the autocorrelation covariate (R8, t-1), the upstream road (R5), and the downstream road (R9) during high traffic and the road block (R8) during moderate traffic. The last plot is the residual histogram.

TABLE III
HW COEFFICIENTS

coef of	R4	R5	R8	R9	R12
Level	0.541	0.466	0.591	0.487	0.596
Trend	0	0	0	0	0
Season	0.811	0.827	0.831	0.844	0.796

TABLE IV
ARIMA COEFFICIENTS

coef	AR1	AR2	MA1	MA2	AIC
R4	0.506	0.045	-1.964	1.0	432.64
R5	0.593	0.181	-1.981	1.0	380.65
R8	1.136	0.144	-0.058	-0.94	355.85
R9	1.117	-0.420	-0.325	0.68	434.99
R12	0.611	-0.302	0.608	0.37	421.24

TABLE V
MSTARMA COEFFICIENTS FOR R8 IN THE "TUE, WED, AND THU" PEAK TEMPLATE

coef	R8	R5	R6	R7	R4	R9
lag-1	0.14	0.31	0.06	0.28	0.23	0.06
lag-2	0.08	0.17	0.15	-0.26	0.11	-0.14
lag-3	0.02	0.10	0.04		0.07	-0.04

TABLE VI
MST ESTIMATES FOR SELECTED ROADS

	R8	R5	R6	R9	R4
g_t	191.4	216.2	149.4	237.1	165.3
σ_ϵ	213.1	336.4	347.4	264.5	218.9
σ_η	37.6	81.1	45.7	53.5	62.8

TABLE VII
MST CORRELATION COEFFICIENTS FOR SELECTED ROADS

	R8	R5	R4	R9
R8	1	0.751	0.634	0.416
R5	0.751	1	0.846	0.356
R4	0.634	0.846	1	0.113
R9	0.416	0.356	0.113	1

Finally, the MST model is fitted to data as in [4] but using different lag values, i.e., $\tau = 1$ to 720. The relationships between variables are represented as any road is correlated with others at one lag, and therefore, a correlation coefficient exists for each road. The estimated trend g_t, standard deviation of error σ_ϵ, and standard deviation of stochastic trend σ_η are shown in Table VI, and the correlation coefficients for selected roads are shown in Table VII.

Table VI shows high standard deviation of the disturbances resulting from the high variation and the overdispersion of flow. Table VII shows that there is a high correlation between road segments and that the correlation decreases when the distance between the segments increases.

TABLE VIII
RMSE VERSUS FORECAST HORIZON FOR R8 DURING THE THREE SEASONS

Traffic season	Method	Time horizon - minutes			
		10	20	30	40
High traffic	NBAMs	0.23	1.18	1.91	3.85
	ARIMA	3.27	4.2	7.04	9.34
	HW	1.86	2.16	5.17	8.61
	MST	1.06	2.74	3.47	5.36
	MSTARMA	1.38	3.17	4.12	5.62
Moderate traffic	NBAMs	1.07	2.12	2.74	4.30
	ARIMA	5.11	7.02	11.9	16.9
	HW	4.61	5.95	6.28	10.3
	MST	3.06	3.52	4.63	5.51
	MSTARMA	3.57	3.95	5.39	6.18
Low traffic	NBAMs	2.33	2.95	4.25	5.90
	ARIMA	8.64	11.0	14.2	17.9
	HW	8.22	9.42	12.6	15.1
	MST	4.48	5.87	7.15	9.54
	MSTARMA	5.16	6.69	7.83	10.8

D. Forecast Results

We evaluate the forecast during each traffic season separately using the models developed in the training phase. The NBAMs utilize the *predict* function in the R $mgcv$ package, which applies (7) to forecast future values. The forecast requires model coefficients β and data that are used to build prediction matrix $\mathbf{X}^p$. To forecast an hour ahead, traffic flows on the dependent road and the significant roads from the previous hour of the same season are used, i.e., the previous 30 records of a season are used to predict 30 records ahead during the same season. We choose only one hour ago since the flow does not have extreme fluctuations within one season. Moreover, considering an hour ago enables capturing the behavior of flow during the last hour and including real-time traffic conditions.

The models are evaluated in different days using the testing data set recorded in April 2008. The forecasted values are compared with real observations. The accuracy is evaluated using RMSE and the MAPE. We also compare the results of the spatiotemporal NBAM model with the results of other models during different traffic seasons. The HW uses the estimated coefficients and data from the current season $(m = 720)$ to forecast an hour ahead. The ARIMA and the MSTARMA consider the number of lags to forecast a single future value. To extend the forecast horizon, every time a $t + 1$ value is predicted, it is used to forecast the $t + 2$ value. The MST forecast uses the parameters in the current season $(m = 720)$ to forecast a single step or multiple steps ahead.

To test the forecast accuracy versus the forecast horizon, the RMSE values of R8 are calculated for different horizons, as shown in Table VIII. We notice that the RMSEs of the NBAMs

TABLE IX
COMPARISON OF FORECASTING MODELS BASED ON MAPE

Traffic season	Method	MAPE %			
		R4	R5	R8	R9
High traffic	NBAMs	2.11	2.62	1.74	2.44
	GAM ($\varphi = 0$)	2.47	3.02	1.92	2.63
	ARIMA	12.6	11.3	13.7	11.1
	HW	9.40	8.49	10.7	9.51
	MST	5.12	5.51	7.04	8.53
	MSTARMA	6.33	6.51	6.39	9.27
Moderate traffic	NBAMs	2.61	3.17	3.84	3.41
	GAM ($\varphi = 0$)	2.74	3.33	4.12	3.69
	ARIMA	17.2	16.5	14.3	14.9
	HW	10.2	10.8	12.7	10.5
	MST	8.11	7.36	8.75	8.89
	MSTARMA	8.61	8.34	9.11	8.51
Low traffic	NBAMs	5.3	7.95	9.24	8.13
	GAM ($\varphi = 0$)	5.46	8.66	10.5	10.3
	ARIMA	27.4	27.9	28.3	26.5
	HW	21.4	23.8	21.3	24.9
	MST	17.6	15.2	20.4	14.8
	MSTARMA	20.3	20.6	22.1	16.2

TABLE X
COMPARISON OF FORECASTING MODELS BASED ON RMSE

Traffic season	Method	RMSE			
		R4	R5	R8	R9
High traffic	NBAMs	1.16	1.26	1.18	1.04
	GAM ($\varphi = 0$)	2.13	1.16	1.37	1.22
	ARIMA	5.3	4.6	4.2	5.8
	HW	3.35	2.72	2.16	3.42
	MST	2.96	2.73	2.74	2.29
	MSTARMA	2.78	3.04	3.17	2.61
Moderate traffic	NBAMs	2.16	1.64	2.12	1.92
	GAM ($\varphi = 0$)	2.34	2.47	2.92	2.57
	ARIMA	6.13	7.26	7.02	7.91
	HW	5.86	6.62	5.95	7.61
	MST	3.76	3.10	3.52	3.64
	MSTARMA	4.25	3.72	3.95	4.33
Low traffic	NBAMs	2.45	2.09	2.95	2.37
	GAM ($\varphi = 0$)	2.95	2.68	3.44	2.77
	ARIMA	12.3	12.8	11.0	11.5
	HW	10.9	11.2	9.42	9.62
	MST	6.42	6.35	5.87	6.20
	MSTARMA	7.26	7.09	6.69	7.06

are smaller than those of the HW, ARIMA, MSTARMA, and MST models during the three traffic seasons for different horizons. The MAPE and RMSE values of the models applied to several roads for a 20-min forecast horizon are presented in Tables IX and X. For clearer comparisons, we let each model forecast flows for 20 min ahead. The results of the forecasts from the multivariate models during the three traffic seasons for R8 are shown in Figs. 5–7.

The tables and figures show that the spatiotemporal NBAMs perform better than the other models as its results are very close to the observed values, and its RMSE and MAPE values for all roads are less than the other models. The spatiotemporal NBAMs capture the flow trend and account for the flows on other roads. Modeling each traffic season separately assist in avoiding the flow fluctuation between low, moderate, and high traffic. The amount of variation left in each season is captured by the model through treating the overdispersion of data. The tables and figures also show that the multivariate models outperform the univariate models. All models perform better during the high- and moderate-traffic seasons, but the accuracy deteriorates during the low season.

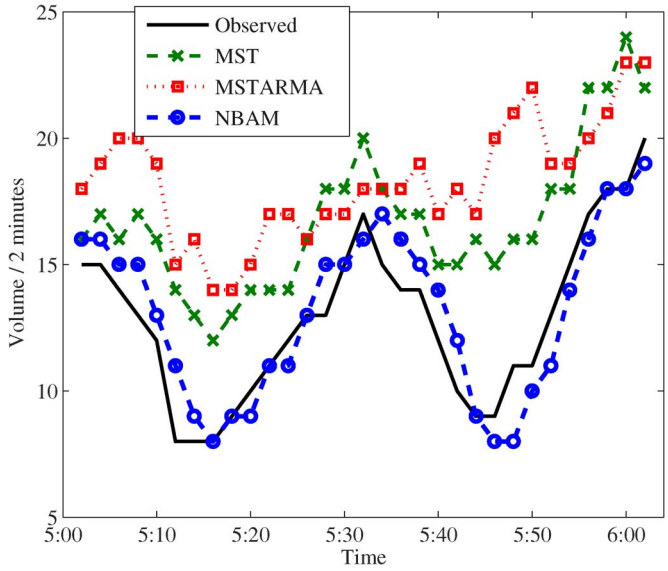

Fig. 5. Comparison between the NBAM, MST, and MSTARMA for a 20-min forecast during the low-traffic season.

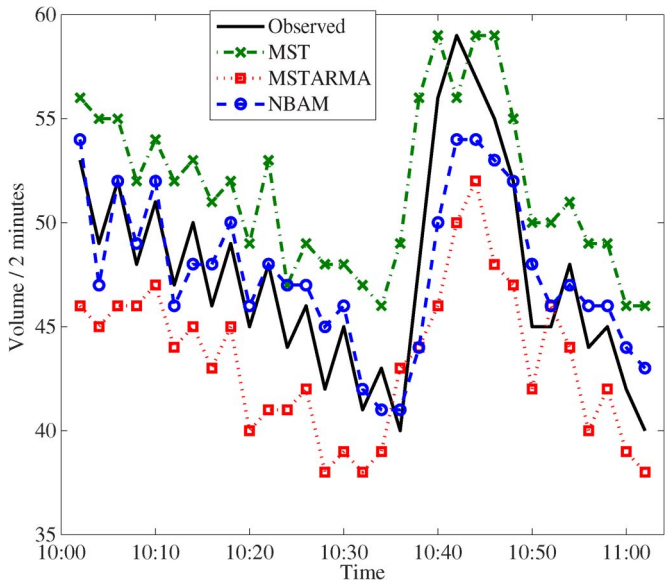

Fig. 6. Comparison between the NBAM, MST, and NBAM for a 20-min forecast during the moderate-traffic season.

The accuracy of the other models is less than the spatiotemporal NBAMs. The Holt–Winters and the ARIMA methods do not handle overdispersion and do not address flow on other roads. Although the MST and the MSTARMA are multivariate and can include several spatially correlated road segments, their inability to handle overdispersion makes them less accurate than the spatiotemporal NBAMs. It is worth noticing that the NBAM outperforms GAM as indicated by the RMSE and MAPE values of the GAM ($\varphi = 0$, the variance and the mean are equal). This states that overdispersion is an important factor that should be handled in traffic flow modeling. Although the accuracy difference between GAM and NBAM is very small, this difference will become bigger if overdispersion is larger than the one in our data set.

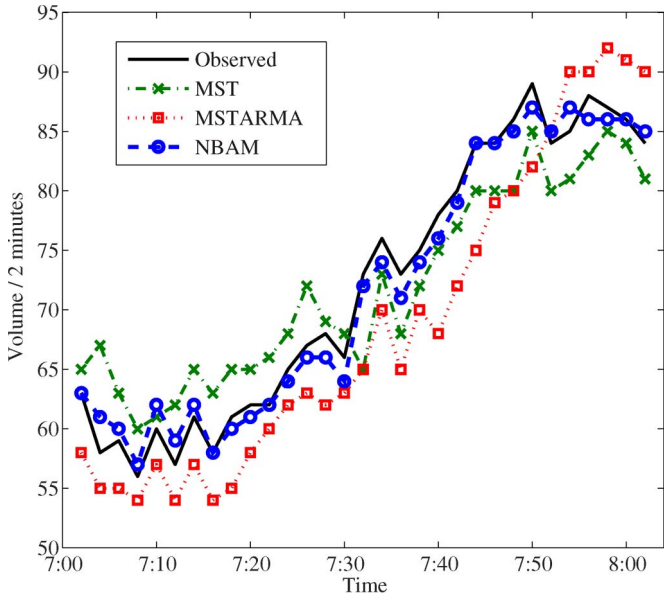

Fig. 7. Comparison between the NBAM, MST, and MSTARMA for a 20-min forecast during the high-traffic season.

TABLE XI
CPU TIME COMPARISON IN SECONDS

Model	Training	Forecasting
NBAM	6.3	0.4
GAM	4.7	0.3
ARIMA	47.1	6.3
HW	11.7	0.9
MST	14.2	1.4
MSTARMA	17.4	2.7

Our research results can be validated as the models used for comparison achieve similar results as in the published papers. We focus on the MAPE because it is not scale dependent as RMSE. The HW MAPE is 11.22% in rush hour [17]. The ARIMA MAPE values during high traffic in [14] and [17] are 11.28% and 8.82%, respectively. The MST MAPE value of different roads in [4] for a single step forecast (15 min) range from 3.84% to 12.81%. The HW and ARIMA MAPE values for a 15-min forecast during low traffic in [17] are 22.11% and 28.55%, respectively. In our results, the MAPE values of these models for a 20-min forecast during high and low traffic are very close to the published values and within the same range. The small difference is related to the data aggregation time, horizon difference (15 min versus 20 min), and overdispersion.

E. Time Complexity

In terms of the computational demand, the CPU time is measured on a 2.5-GHz Intel CPU, 46-bit operating system, and 4-GB RAM. The GAM-based model is the fastest compared with other models and costs little time during optimizing the smooth functions and forecasting future values, as shown in Table XI. GAM is slightly faster than NBAM, which makes GAM a better choice in practice. However, when overdispersion increases, little computation speed can be sacrificed for obtaining higher accuracy.

V. CONCLUSION AND FUTURE WORK

Traffic flow has spatial characteristics that include the spatial correlation with upstream/downstream roads and adjacent road blocks as well as temporal characteristics that include temporal autocorrelation and seasonal patterns. Additionally, flow exhibits overdispersion resulting from high and rapid variations in urban and signalized road network. Overdispersion influences forecast accuracy and requires special treatments. Therefore, we have proposed a novel spatiotemporal NBAM for short-term traffic flow forecasting in urban areas. The model addresses the spatial and temporal characteristics and handles overdispersion. The model is accurate and efficient during all traffic conditions.

We believe that the proposed forecasting model benefits the application of parallel and coordinated traffic management strategies. In these strategies, the traffic network is divided into subnetworks, where each subnetwork is controlled by a subcontroller. Subcontrollers communicate with each other to provide network-wide control decisions in a process that involves much data and computation. The latest information and communication technology that supports distributed large data processing and storage is cloud computing. Our final goal is to make use of cloud computing to provide parallel and coordinated congestion control. This is an ongoing project supported by the Ministry of Transportation and Communication of Taiwan. The proposed forecasting model will be utilized to provide adaptive and proactive dynamic traffic control. Its high accuracy enables estimating traffic conditions precisely and enables evaluating the impact of control decisions on future traffic. Its low computational demand guarantees the efficiency of computation and communication within the cloud system.

REFERENCES

[1] F.-Y. Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 630–638, Sep. 2010.
[2] S. Lin, B. D. Schutter, Y. Xi, and H. Hellendoorn, "Efficient network-wide model-based predictive control for urban traffic networks," *Transp. Res. Part C*, vol. 24, pp. 122–140, Oct. 2012.
[3] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks," *Comput.-Aided Civil Infrastructure Eng.*, vol. 22, no. 5, pp. 317–325, Jul. 2007.
[4] B. Ghosh, B. Basu, and M. OMahony, "Multivariate short-term traffic flow forecasting using time-series analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 2, pp. 246–254, Jun. 2009.
[5] W. Min and L. Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transp. Res. Part C*, vol. 19, no. 4, pp. 606–616, Aug. 2011.
[6] A. Stathopoulos and M. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transp. Res. Part C*, vol. 11, no. 2, pp. 121–135, Apr. 2003.
[7] Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches," in *Proc. 82nd Transp. Res. Board Annu. Meet.*, Washington D.C., USA, Jan. 2003, pp. 1–25.
[8] Q. Ye, W. Y. Szeto, and S. C. Wong, "Short-term traffic speed forecasting based on data recorded at irregular intervals," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1727–1737, Dec. 2012.
[9] Y. Kamarianakis and P. Prastacos, "Space–time modelling of traffic flow," *Comput. Geosci.*, vol. 31, no. 2, pp. 119–133, Mar. 2005.
[10] S. Dunne and B. Ghosh, "Weather adaptive traffic prediction using neurowavelet models," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 370–379, Mar. 2013.
[11] Y.-A. Daraghmi, C.-W. Yi, and T.-C. Chiang, "Space–time multivariate negative binomial regression for urban short-term traffic volume

prediction," in *Proc. 12th Int. Conf. ITS Telecommun.*, Taipei, Taiwan, Nov. 2012, pp. 35–40.
[12] S. Sun, C. Zhang, and G. Yu, "A Bayesian network approach to traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 124–132, Mar. 2006.
[13] A. Stathopoulos and L. Dimitriou, "Fuzzy modeling approach for combined forecasting of urban traffic flow," *Comput.-Aided Civil Infrastructure Eng.*, vol. 23, no. 7, pp. 521–535, Oct. 2008.
[14] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. Part C*, vol. 10, no. 4, pp. 302–321, Aug. 2002.
[15] T. Thomas, W. Weijermars, and E. van Berkum, "Predictions of urban volumes in single time series," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 1, pp. 71–80, Mar. 2010.
[16] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, Nov. 2003.
[17] B. Basu, B. Ghosh, and M. M. O'Mahony, "Time-series modelling for forecasting vehicular traffic flow in Dublin," in *Proc. 84th Transp. Res. Board Annu. Meet.*, Washington D.C., USA, Jan. 2005, pp. 1–22.
[18] P. Lingras and P. Mountford, "Time delay neural networks designed using genetic algorithms for short term inter-city traffic forecasting," in *Proc. 14th Int. Conf. Ind. Eng. Appl. Artif. Intell. Expert Syst., Eng. Intell. Syst.*, Budapest, Hungary, Jun. 4–7, 2001, pp. 290–299.
[19] T. T. Tchrakian, B. Basu, and M. OMahony, "Real-time traffic flow forecasting using spectral analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 519–526, Jun. 2012.
[20] Y.-A. Daraghmi, C.-W. Yi, and T.-C. Chiang, "Mining overdispersed and autocorrelated vehicular traffic volume," in *Proc. 5th Int. Conf. Comput. Sci. Inf. Technol.*, Amman, Jordan, Mar. 2013, pp. 194–200.
[21] J. Rice and E. van Zwet, "A simple and effective method for predicting travel times on freeways," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 3, pp. 200–207, Sep. 2004.
[22] 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 Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 60–69, Mar. 2009.
[23] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *J. Transp. Eng.*, vol. 117, no. 2, pp. 178–188, Mar. 1991.
[24] R. E. Turochy and B. D. Pierce, "Relating short-term traffic forecasting to current system state using nonparametric regression," in *Proc. 7th Int. IEEE Conf. Intell. Transp. Syst.*, Washington D.C., USA, Oct. 2004, pp. 239–244.
[25] B. L. Smith and M. J. Demetsky, "Traffic flow forecasting: Comparison of modeling approaches," *J. Transp. Eng.*, vol. 123, no. 4, pp. 261–266, Jul. 1997.
[26] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 644–654, Jun. 2012.
[27] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transp. Res. Part C*, vol. 13, no. 3, pp. 211–234, Jun. 2005.
[28] Y. Xu, Q.-J. Kong, and Y. Liu, "Comparison of urban traffic prediction methods between UTN-based spatial model and time series models," in *Proc. 15th Int. IEEE Conf. Intell. Transp. Syst.*, Anchorage, AK, USA, Sep. 2012, pp. 814–819.
[29] J. M. Hilbe, *Negative Binomial Regression,* 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[30] S. W. Thurston, M. P. Wand, and J. K. Wiencke, "Negative binomial additive models," *Biometrics*, vol. 56, no. 1, pp. 139–144, Mar. 2000.
[31] B. Abdulhai, H. Porwal, and W. Recker, "Short-term freeway traffic flow prediction using genetically optimized time-delay-based neural networks," California PATH Program, Institute of Transportation Studies, University of California, Berkeley, USA, 1999, pp. 1–50.
[32] C. Chen, Z. Liu, W.-H. Lin, S. Li, and K. Wang, "Distributed modeling in a MapReduce framework for data-driven traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 22–33, Mar. 2013.
[33] A. Bertoni and B. Palano, "Structural complexity and neural networks," in *Neural Nets*, M. Marinaro and R. Tagliaferri, Eds. Berlin, Germany: Springer, 2002, ser. Lecture Notes in Computer Science, pp. 190–216.
[34] T. Hastie and R. Tibshirani, "Generalized additive models (with discussion)," *Stat. Sci.*, vol. 1, no. 3, pp. 297–318, 1986.
[35] S. N. Wood, *Generalized Additive Models: An Introduction with R*. London, U.K.: Chapman & Hall, 2006.
[36] T. Hastie and R. Tibshirani, *Generalized Additive Models*. London, U.K.: Chapman & Hall, 1990.
[37] J. L. Pearce, J. Beringer, N. Nicholls, R. J. Hyndman, and N. J. Tapper, "Quantifying the influence of local meteorology on air quality using generalized additive models," *Atmos. Environ.*, vol. 45, no. 6, pp. 1328–1336, Feb. 2011.
[38] S.-R. Han, S. D. Guikema, and S. M. Quiring, "Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models," *Risk Anal.*, vol. 29, no. 10, pp. 1443–1453, Aug. 2009.
[39] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 134–141, Feb. 2012.
[40] X. Wang and D. E. Brown, "The spatia-temporal generalized additive model for criminal incidents," in *Proc. IEEE Int. Conf. ISI*, Beijing, China, Jul. 2011, pp. 42–47.
[41] J. Hu, I. Kaparias, and M. Bell, "Spatial econometrics models for congestion prediction with in-vehicle route guidance," *IET Intell. Transp. Syst.*, vol. 3, no. 2, pp. 159–167, Jun. 2009.

Yousef-Awwad Daraghmi received the B.E. degree in electrical engineering from An-Najah National University, Nablus, Palestine, in 2002 and the Master's degree in computer science and information engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2007. He is currently working toward the Ph.D. degree in the Department of Computer Science, National Chiao Tung University.

His research interests include intelligent transportation systems and vehicular ad hoc networks.

Dr. Daraghmi served as a Technical Program Committee Member for the International Conference on Connected Vehicles and Expo (ICCVE 2012 and 2013) and the International Conference on Intelligent Transportation Systems Telecommunications (ITST 2012). He received the Best Paper Award from the ITST in 2012.

Chih-Wei Yi received the B.S. degree in mathematics and the M.S. degree in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 1991 and 1993, respectively, and the Ph.D. degree in computer science from the Illinois Institute of Technology, Chicago, IL, USA, in 2005.

He is currently an Associate Professor with the Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan. His research interests include wireless ad hoc and sensor networks, vehicular ad hoc networks, intelligent transportation systems, network coding, and algorithm design and analysis.

Dr. Yi is a member of the Association for Computing Machinery and the Institute of Electronics, Information and Communication Engineers. He was a Senior Research Fellow with the Department of Computer Science, City University of Hong Kong, Hong Kong. He was awarded the Outstanding Young Engineer Award by the Chinese Institute of Engineers in 2009 and the Young Scholar Best Paper Award from the IEEE IT/COMSOC Taipei/Tainan Chapter in 2010.

Tsun-Chieh Chiang received the M.S. degree in electrical engineering and the Ph.D. degree in computer science from the Illinois Institute of Technology, Chicago, IL, USA.

He is the Division Director of the Telematics and Control System Division with the Information and Communications Research Laboratories, Industrial Technology Research Institute (ITRI), Taiwan. His division is responsible for the evolution of Telematics, EV control platform, and intelligent-transportation-system-related applications and services, focusing on communication and infotainment technology for Telematics service creation, implementation, delivery, operation, and maintenance. He has done the Next Generation Telematics project from 2008 to 2010 and, currently, is in charge of the Vehicle ICT Service and System Technology project from 2011 to 2014. Prior to his present job in ITRI, he was in an Architecture team for Lucent Bell Laboratories, Intelligent Network Unit (INU) and Global Professional Service organizations in Naperville, USA, participating in numerous forums, including industry conferences and standards meetings, leading the evolution planning for Lucent INU product into the packet and mobile technology areas. Since 2008, he has been an Adjunct Professor of vehicular networks and communications with National Chiao Tung University, Hsinchu, Taiwan.