Unobserved Component Model for Predicting Monthly Traffic Volume

Zheyong Bian[1]; Zhipeng Zhang[2]; Xiang Liu, Ph.D.[3]; and Xiao Qin, Ph.D., P.E.[4]

Abstract: Traffic volume prediction plays a critical role in transportation system and infrastructure management. This paper develops the first application of an unobserved component model (UCM) for monthly traffic volume forecasting. We compare the UCM model with simple linear regression, autoregressive integrated moving average (ARIMA), support vector machine (SVM), and artificial neural network (ANN) models based on monthly traffic volume data from a key corridor in New Jersey. As a general econometric method, the UCM decomposes the time series into trend, seasonal, and irregular components, exhibiting superiority for statistically modeling traffic data with cyclic or seasonal fluctuations. The numerical analysis shows that the UCM outperforms all of the other four models and generates reasonably accurate prediction results. This research indicates that UCM can be considered as an alternative approach to modeling traffic volumes. **DOI: 10.1061/JTEPBS.0000281.** *© 2019 American Society of Civil Engineers.*

Author keywords: Traffic volume prediction; Unobserved component model; Time series analysis.

Introduction

Monthly traffic volume provides Departments of Transportation (DOT) with directions to carry out more powerful transportation system management and apply more effective control measurements to maintain the freeway capacity, to weaken the traffic congestion, and to promote efficiency of the freeway networks. Transportation agencies (e.g. DOTs) make plans for infrastructure management based on varying traffic volumes in different time periods. The current practice necessitates medium-time or long-term traffic volume forecasting. For instance, Ng et al. (2009) and Frangopol and Liu (2007) optimized long-term infrastructure maintenance plan based on long-term traffic volumes. Bai et al. (2015) made an optimal pavement design and rehabilitation plan based on long-term traffic demand (e.g. year, quarter, or month). Hajibabai et al. (2014) optimized the freight facility location and pavement infrastructure rehabilitation using the network-level annual traffic data.

The focus of this research is to use a time series model [unobserved component model (UCM)] to predict future traffic volume based on the historical trend. Longer-term prediction focuses on monthly or even yearly traffic information which can be used for capital planning and transportation management (Lu 2014). The transportation agency that provides the data typically wants to forecast monthly traffic volume in order to make a new-year plan for transportation infrastructure management and capital planning. Thus, the scope of this paper is using historical data for the prediction of monthly traffic volume in one year.

The forecasting models proposed in the literature can be classified into several categories, including but not limited to (1) machine learning methods, such as artificial neural network (ANN), support vector machine (SVM), and deep learning models; (2) the autoregressive integrated moving average model (ARIMA); (3) the nonparametric regression model (e.g. kth nearest neighbor model); (4) Bayesian networks; and (5) hybrid methods (Table 1).

Most of the prior research has focused on short-term (e.g., hourly and daily) prediction of traffic volume, which is very important for traffic operations and intelligent transportation systems (ITS). In addition, traffic volumes in a longer time period (month, season, or year) can also provide information to support transportation planning, network design, and infrastructure management. In this context, artificial neural network, support vector machine, autoregressive integrated moving average, and other models have been used (Zhang and Qi 2005; Ma et al. 2015; Cawley and Talbot 2010).

Monthly traffic volumes could demonstrate regular, seasonal, or cyclical patterns, which may require a special type of time series model to account for these patterns. Many prior studies employed time series modeling for traffic volume prediction (e.g., Williams and Lester 2003; Min et al. 2010; Zhang et al. 2011; Kumar and Vanajakshi 2015). Along the same lines, the goal of this research is to explore the feasibility of UCM as a new, alternative time series approach to predicting longer-term traffic volumes. The UCM model was originally developed by econometricians in the 1990s (Harvey and Peters 1990) and has emerged as a promising approach to analyzing time series data that exhibits regular, seasonal, or cyclical trends in various applications, such as prediction of economic indicators (Kim 1993; Cowan and Joutz 2006; Paradiso and Rao 2012) and infrastructure management in engineering (Marquez et al. 2007; Pedregal et al. 2004). The UCM can filter, smooth, and extract signals, which enables the observed variables to be decomposed into different perceived features, differentiated by their spectral properties (Young 2011). This approach decomposes time series into trend,

[1]Ph.D. Student, Dept. of Civil and Environmental Engineering, Rutgers, State Univ. of New Jersey, Piscataway, NJ 08854-8018. Email: zb91@scarletmail.rutgers.edu

[2]Ph.D. Student, Dept. of Civil and Environmental Engineering, Rutgers, State Univ. of New Jersey, Piscataway, NJ 08854-8018. Email: zz282@scarletmail.rutgers.edu

[3]Assistant Professor, Dept. of Civil and Environmental Engineering, Rutgers, State Univ. of New Jersey, Piscataway, NJ 08854-8018 (corresponding author). Email: xiang.liu@rutgers.edu

[4]Associate Professor, Civil and Environmental Engineering, Univ. of Wisconsin, Milwaukee, WI 53201-0784. Email: qinx@uwm.edu

Note. This manuscript was submitted on July 25, 2017; approved on April 17, 2019; published online on October 4, 2019. Discussion period open until March 4, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part A: Systems***, © ASCE, ISSN 2473-2907.

Table 1. Selected models for traffic volume prediction

Prediction models	References
Artificial neural network	Dia (2001), Ishak and Alecsandru (2004), Vlahogianni et al. (2005), Xie and Zhang (2006), Zheng et al. (2006), Sun et al. (2012), Zhu et al. (2014), Kumar et al. (2015), and Goves et al. (2016)
Support vector machine	Dibike et al. (2001), Luo et al. (2005), Xu and Yang (2005), Qing-Fang et al. (2009), Chen et al. (2012), Guo et al. (2012), and Ahn et al. (2015)
Deep learning	Lv et al. (2015)
Autoregressive integrated moving average	Williams and Lester (2003), Min et al. (2010), Zhang et al. (2011), and Kumar and Vanajakshi (2015)
Nonparametric regression model (*k*th nearest neighbor model)	Zhong and Ling (2015), Yuan and Wang (2012), Guo et al. (2012), Wu et al. (2014), Dell'Acqua et al. (2015), Yu et al. (2016), and Xia et al. (2016)
Bayesian networks	Sun et al. (2005), Pascale and Nicoli (2011), Zhu et al. (2016), and Ahn et al. (2015)
Hybrid methods	Abdulhai et al. (2002), Hu et al. (2008), Dimitriou et al. (2008), McCrea and Moutari (2010), Chan et al. (2012), Zhang et al. (2014), Zou et al. (2015), and Hu et al. (2016)

seasonal, cyclical, and irregular components (Harvey and Peters 1990).

To our best knowledge, this is the first study that applies UCM to monthly traffic volume prediction. Using empirical data from one key corridor in New Jersey, we compare traffic volume prediction results based on UCM versus the other selected alternative methods, including the simple linear regression, ARIMA, SVM, and ANN models. As the following sections will show, the UCM outperforms the other four models for the data set used in this study.

Unobserved Component Model

The unobserved component model is a type of the structural time series model which is a multiple regression model with time-varying parameters. The UCM decomposes time series into trend, seasonal, and irregular components (Harvey and Peters 1990). A general UCM model can be defined as follows:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{1}$$

where y_t = response variable (e.g., monthly traffic volume) at time t; μ_t = trend component at time t; γ_t = seasonal component at time t; and ε_t = irregular component, a Gaussian white noise process with variance σ_ε^2 at time t.

In general, the trend component can be viewed as a local approximation to a linear trend that has an upward or downward slope (Harvey and Peters 1990) as follows:

$$\text{Level: } \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \sim i \cdot i \cdot d \cdot N(0, \sigma_\eta^2) \tag{2}$$

$$\text{Slope: } \beta_t = \beta_{t-1} + \zeta_t, \quad \zeta_t \sim i \cdot i \cdot d \cdot N(0, \sigma_\zeta^2) \tag{3}$$

In this model, the trend is featured by *level* and *slope*. Eq. (2) represents the stochastic level of the trend (μ_t), and Eq. (3) represents the stochastic slope of the trend (β_t). η_t and ζ_t are distributed independently of each other and over time with mean zero and variances σ_η^2 and σ_ζ^2. The disturbance term ζ_t assigns to the slope a random parameter (Jalles 2009). The stochastic nature of the level at moment t derives from the presence of η_t.

The so-called seasonal effect at time t, represented by γ_t, is included in the UCM [Eq. (1)]. The seasonal component γ_t is associated with season $s = s(t)$, for $s = 1, 2, \ldots, S$, where S is the seasonal length. This paper uses the time-varying dummy seasonal pattern (Koopman and Ooms 2011) to formulate the seasonal effects that are allowed to change over time. We use a stochastic equation [Eq. (4)] to replace the summing-to-zero constraint ($\sum_j^{j+S} \gamma_j = 0$) in the fixed dummy seasonal pattern (Koopman and Ooms 2011). That is

$$\gamma_{t+1} = -\gamma_t - \gamma_{t-1} - \ldots - \gamma_{t-S+2} + \omega_t, \quad \omega_t \sim i \cdot i \cdot d \cdot N(0, \sigma_\omega^2) \tag{4}$$

where ω_t are identically and independently distributed variables, for $t \geq S - 1$. Note that the initial seasonal components $\gamma_1, \gamma_2, \ldots, \gamma_{S-1}$ are treated as unknown coefficients and will be estimated together with other coefficients in the model.

Model Calibration

Data set

Monthly traffic volumes of Class 1 vehicles (two-axle passenger cars) on a key corridor in New Jersey, from January 2006 to October 2016, are provided by a transportation agency in New Jersey. The data are used to develop and compare alternative traffic volume prediction models. In this research, the monthly traffic volume data in a whole year is predicted using the UCM, based on the practical use of the model by the data provider. Therefore, 1-year worth testing data is used. The remaining is the training data. The data from January 2006 to October 2015 is used as the training data, and the data from November 2015 to October 2016 is used as the testing data set for blind prediction. The empirical traffic volumes are shown in Fig. 1. The traffic volume distribution appears to have a temporally cyclic (i.e., seasonal) fluctuation. This empirical data pattern indicates that UCM, a structural time series model, might be a promising modeling technique.

Model Development and Diagnostics

For the model implementation, we use a linear Gaussian state space model. The model can be found in Koopman and Ooms (2011) to provide a unified representation of linear time series model consisting of a transition equation and a measurement equation. Then, Kalman filter is used to evaluate the Gaussian likelihood function via the prediction error decomposition (Harvey 1990). The parameters in the UCM are obtained by the maximum likelihood estimation. For the model diagnostics, we use the t-test to test the significance of variances of all components and use the Chi-square statistics test to determine the significances of all components. The variances of the disturbance terms in the evolution of μ_t, β_t, and γ_t and the variance of the irregular component ε_t are estimated in Table 2. These estimations and their corresponding t-values are

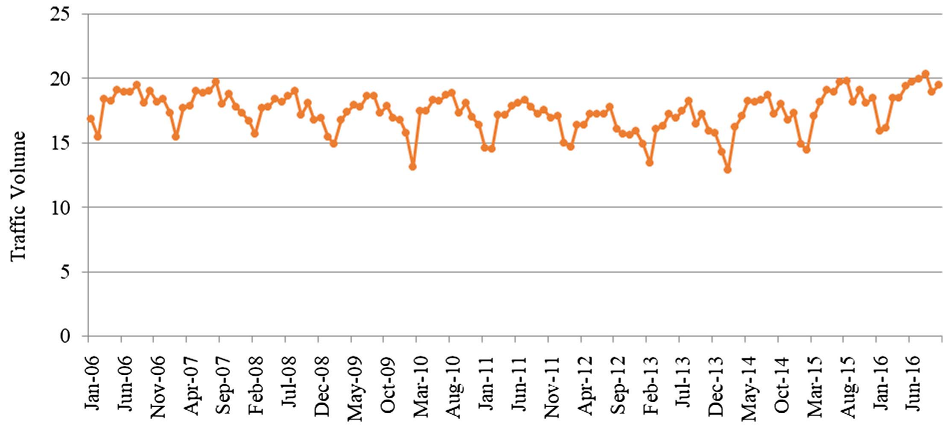

Fig. 1. Monthly passenger car traffic volume from January 2006 to October 2016 on a key transportation corridor in New Jersey.

Table 2. Statistical test of error variances

Components	Estimate	Approximate standard error	t-value	Approximate p-value
Irregular	0.09413	0.02288	4.11	<0.0001
Level	0.03621	0.01982	1.83	0.0677
Slope	0.00005	0.00009	0.58	0.5588
Season	6.943×10^{-9}	4.713×10^{-6}	0.00	0.9988

Table 3. Significance analysis of UCM component

Components	Chi-square values	P-values
Irregular	0.06	0.8126
Level	6871.90	<0.0001
Slope	1.48	0.2236
Season	930.81	<0.0001

used to infer whether the variances are significant (i.e., the corresponding components are stochastic or deterministic in the UCM model). Table 2 suggests that only the disturbance variance of the irregular component is significant, indicating that the irregular component ε_t is stochastic. The slope and season components appear to be deterministic (fixed) because their error variances are highly insignificant according to the theory developed by Harvey and Peters (1990).

Table 3 is another important output. The Chi-square statistics test the null hypothesis that the given component is not statistically significant. In other words, it tests whether a certain component should be included in the final state of the Kalman filter for UCM model calibration (Harvey 1990). If a component is deterministic, this analysis is equivalent to checking whether the regression effect is significant.

In our case, the *season* component and the *level* component are significant and should be retained in the model as deterministic components. The *slope* component is not statistically significant. Note that although the irregular term's contribution appears not statistically significant toward the end of the estimation span, we cannot remove it from the model because it is a stochastic component (Harvey 1990).

The goodness-of-fit statistics based on the raw residuals (residual = observed–predicted) is reported after a model is fitted. The adjusted $\mathrm{R}-\mathrm{square} = 0.90$, which suggests the model fits the empirical data reasonably well.

Fig. 2 presents the diagnostic plots based on residuals. The residual histogram [Fig. 2(a)] and the Q-Q (quantile-quantile) plot [Fig. 2(b)] indicate that the residuals approximately follow a normal distribution. The remaining plots check the whiteness of the residuals. The ideal scenario for an autocorrelation function (ACF) of residuals is that there are no significant correlations for any lag [Lag 0 is always 1, Fig. 2(c)]. If there's no correlation between lags, the bars lie within the 95% confidence range. The ideal scenario is the one in which the partial autocorrelation function (PACF) plot [Fig. 2(d)] is the same as the ACF plot [Fig. 2(c)]. In Fig. 2, the bars mostly lie within the 95% confidence range in the ACF and PACF plots. It indicates that there is no significant autocorrelation in the residuals. Therefore, the UCM model appears to well fit the empirical data.

Results

A developed UCM has the following structure:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \, \varepsilon_t \sim i \cdot i \cdot d \cdot N(0, \sigma_\varepsilon^2 = 94130) \quad (5)$$

where y_t represents the time series to be modeled (the responsive variable, which is monthly passenger vehicle traffic volume in this paper), μ_t is the trend component, γ_t is the seasonal component, and random error is represented by ε_t. ε_t are approximately identically, independently, and normally distributed variables for all times t with mean 0 and variance 94,130, using the training data set. The trend component μ_t only consists of the level term because of the insignificance of slope, and thus, μ_t is formulated as follows without a particular slope trend β_t:

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim i \cdot i \cdot d \cdot N(0, 36210) \quad (6)$$

where the level term η_t is composed of independent, identical, and normal distribution variables for all times t with mean 0 and variance 36,210. The seasonal fluctuations (γ_t) can be formulated as the following model:

$$\gamma_t = -\sum_{i=0}^{s-1} \gamma_{t-i} + \omega_t \quad \omega_t \sim i \cdot i \cdot d \cdot N(0, 6.943 \times 10^{-3}) \quad (7)$$

ω_t are identically and independently distributed normal distributions.

Fig. 3 shows differences between the predicted flow volumes and the actual observed volumes. There were 118 of 130 observations used to train the UCM model. Based on the training data, twelve months' traffic volumes are forecasted, marked as dots in

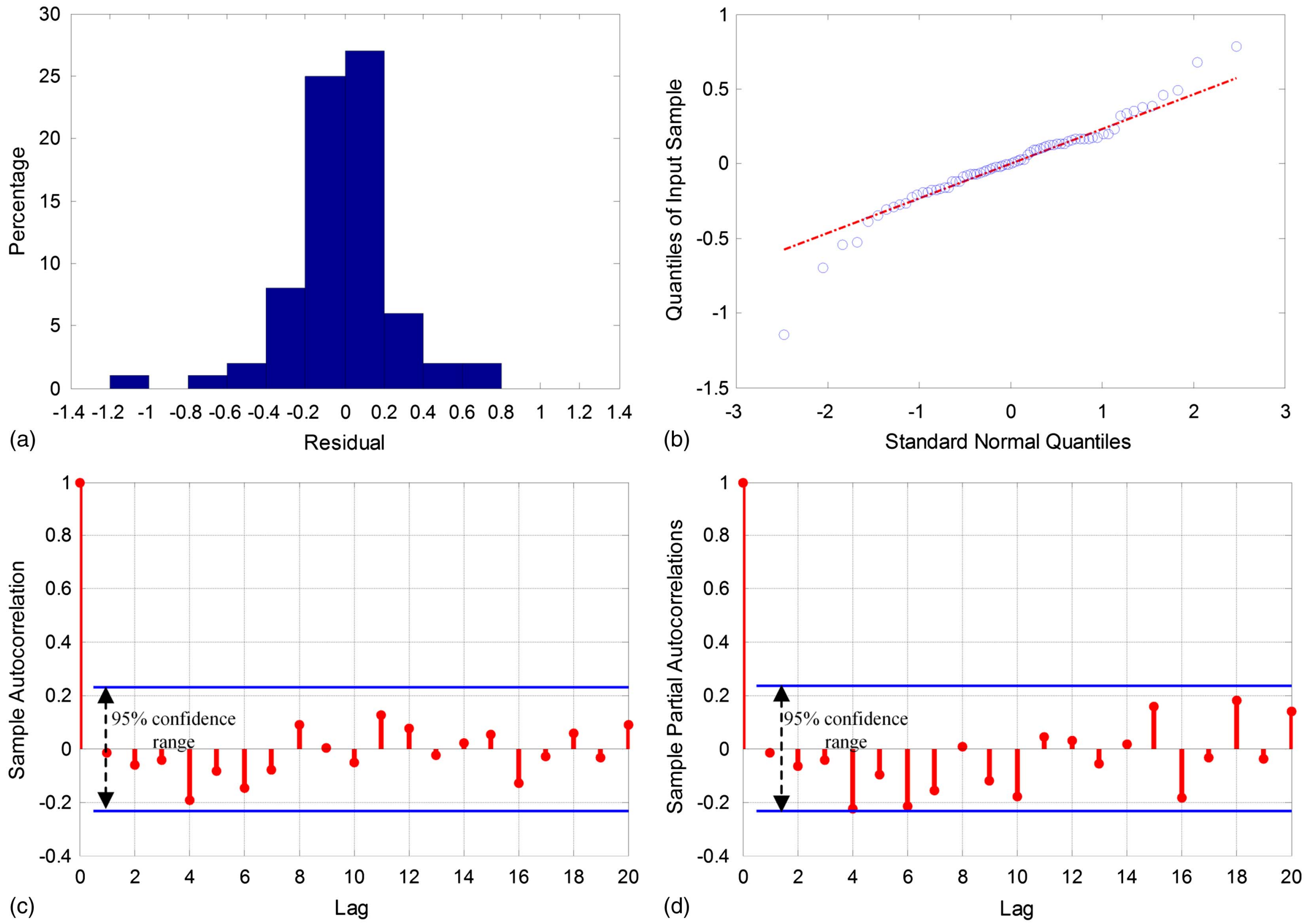

Fig. 2. Residual diagnosis: (a) residual histogram; (b) Q-Q plot of sample data versus standard normal; (c) sample autocorrelation function; and (d) sample partial autocorrelation function.

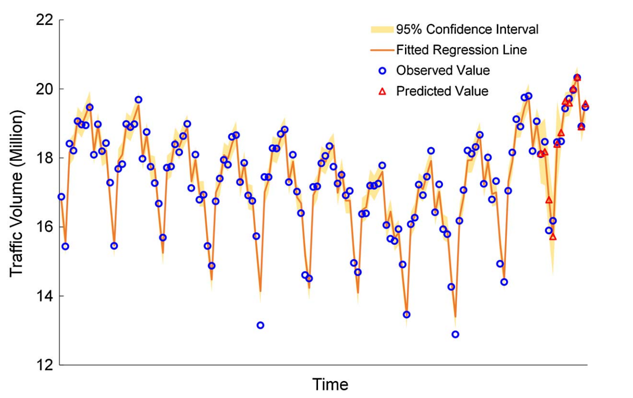

Fig. 3. Predicted flow volumes and the actual observed traffic volumes.

the figure. Although error between predicted values and observed values exists in each month, Table 4 shows that all the testing values are within the 95% confidence interval. In our prediction results, except January 2016, whose relative error is 5.597%, relative errors of all the other months are less than 2.80%. The average relative error is only 1.20%. The data provider indicates that they will be *satisfied* with the prediction accuracy of around 5% relative error when they are using monthly traffic volume data for capital planning and management purposes. Moreover, compared with the prediction accuracy in the literature [the average relative errors are 1.77% and 3.43% in Zhong and Sharma (2006) and Hou and Li (2016), respectively], the prediction accuracy of our model appears to be practically acceptable.

Model Comparison

We select four models, simple linear regression, ARIMA, ANN, and SVM, which are commonly used for time series prediction to compare with UCM. The simple linear regression for forecasting the monthly traffic volume based on historical traffic volumes is used to understand whether a simple growth factor model may work. The ARIMA, ANN, and SVM models are derived from the literature. We use the ARIMA model proposed in Williams and Lester (2003). ANN is derived from Lisowski (2013) and Zhang and Qi (2005), and SVM is from Thissen et al. (2003) and Sapankevych and Sankar (2009). SVM and ANN demonstrate some advantages over other perdition models in some cases. More specifically, ANN does not need specific assumptions in the development of model. It can also be used in both linear and nonlinear models. Therefore, as a data-driven approach, ANN can solve some complex forecasting problems (Zhang and Qi 2005). For SVM, Sapankevych and Sankar (2009) pointed out that the major

Table 4. Summary of forecasted results for the testing data set (in millions)

Months	Observed value	Forecasted value	Gap between forecasted and actual values (%)	95% Confidence limits of prediction	
				Lower limit	Upper limit
November 2015	18.112	18.133	0.1159	17.261	19.006
December 2015	18.470	18.180	1.5701	17.210	19.151
January 2016	15.901	16.791	5.5971	15.729	17.853
February 2016	16.179	15.727	2.7937	14.575	16.879
March 2016	18.460	18.406	0.2925	17.167	19.645
April 2016	18.481	18.741	1.4041	18.277	18.990
May 2016	19.435	19.647	1.0927	19.191	19.905
June 2016	19.718	19.599	0.6064	19.217	19.931
July 2016	19.971	20.014	0.2145	19.609	20.323
August 2016	20.329	20.344	0.0770	19.933	20.646
September 2016	18.914	18.907	0.0362	18.473	19.187
October 2016	19.472	19.578	0.5459	19.079	19.794

advantage of SVM is the great improvement in computation time, compared with other prediction methods. In this paper, we compare these methods with UCM to examine whether the UCM is more appropriate given certain data patterns (e.g., seasonal, cyclic) as seen in our data set. In the development of models of ANN and SVM using training data, we select the models with *best* performance to be compared with the UCM. We use the linear regression analysis tool in Excel to implement the linear regression, and we use the packages of applied statistical time series analysis (ASTSA) (ARIMA), E1071 (SVM), and fit neural networks (NNET) (ANN) in R programming to implement the ARIMA, SVM, and ANN time series models, respectively.

Selected Compared Models

Simple Linear Regression

We use a simple method (linear regression) to predict the monthly traffic volume based on historical traffic volumes in the same month to understand whether a simple growth factor model may work. More specifically, we build 12 linear regression models to fit the traffic volumes in the 12 corresponding months from 2006 to 2016.

$$y_i = a_i x + b_i + \varepsilon_i \tag{8}$$

where y_i is the observed value of ith month's traffic volume, x is the independent value (i.e., year), a_i is the coefficient of the independent value for month i, b_i is the intercept for month i, and ε_i is the random error at month i. Similar to the UCM, the data from January 2006 to October 2015 is used as the training data, and the data from November 2015 to October 2016 is used as the testing data for blind prediction. For each month, we use the training data to fit a linear function of different years. Then, we use the fitted linear function to blindly predict the testing data. The predicted traffic volumes are compared with the observed testing data.

ARIMA

ARIMA is a regression analysis that generates short-term forecasts by examining the differences between time series values. It is a process that creates a transformed series that consists of differences between lagged series observations (Williams and Lester 2003). In the basic ARIMA model, the future value of a variable is a linear combination of past values and past errors, which can be formulated as

$$\begin{aligned} Y_t = \emptyset_0 + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} \\ + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \end{aligned} \tag{9}$$

where Y_t is the observed value, ε_t is the random error at time t, $\emptyset_i$ and θ_j are coefficients of the observed values and random errors, respectively, and integers p and q are used to represent different times (Ariyo et al. 2014). Lags of the differenced series are *autoregressive (AR)* and lags within forecasted data are the *moving average (MA)*. ARIMA tries to explain the movements in the form of a function combining autoregressive and moving average terms (Hyndman and Athanasopoulos 2014). ARIMA is *stationary*. For example, $\emptyset_1$ is fixed for any Lag 1 observations.

SVM

SVM is one of the most popular machine learning methods that can be used for both classification and regression analysis. SVM is based on the structural risk minimization criterion, and its goal is to find the optimal separating hyperplane where the separating margin should be maximized (Yao et al. 2013). In the field of data mining, the SVM represents one type of supervised learning models with associated learning algorithms that analyze data used for classification or regression. In this paper, we use SVM linear regression model to predict the monthly traffic volume data. SVM linear regression is to fit a linear function $y = \omega x + b$ by obtaining the parameters ω and b, where x is the input data and y is the response. The parameters ω and b are obtained by solving an optimization model

$$\begin{aligned} &\min \frac{1}{2}\|\omega\|^2 + C\sum_i(\xi_i + \xi_i^*) \\ &s.t. \\ &y_i - \omega \cdot x_i - b \le \varepsilon + \xi_i \\ &\omega \cdot x_i + b - y_i \le \varepsilon + \xi_i^* \\ &\xi_i, \xi_i^* \ge 0 \end{aligned} \tag{10}$$

where ξ_i and ξ_i^* are two groups of additional decision variables, and C is a weight factor. For the detailed explanation of the model,

Table 5. Eight SVM models

Models	Model description
1	Regression with lag-1 variable
2	Regression with lag-4 variable
3	Regression with lag-12 variable
4	Regression with season variable
5	Regression with lag-1 and season variable
6	Regression with lag-12 and season variable
7	Regression with lag-1 and lag-12 variable
8	Regression with lag-1, lag-12, and season variable

please refer to Thissen et al. (2003). In this paper, eight SVM models using different groups of variables for regression are developed. The model with the best performance is selected to compare with the UCM. The eight models are listed in Table 5.

ANN

ANN has useful pattern classification and pattern recognition capabilities for forecasting (Zhang et al. 1998). An ANN consists of a large collection of artificial neurons, each having multiple weighted inputs and an output. The output of the entire network, as a response to an input vector, is generated by applying certain arithmetic operations, determined by the neural networks, to that vector. Using a finite number of past observations, we can use the neural network to predict a reasonable future value. In this respect, the network can be treated as a function f that involves k inputs and one output

$$f{:}R^k \to R \tag{11}$$

ANN can also be used to find a proper network for forecasting in such a way that if an observed time series $Y = \{y_1, y_2, y_3, \ldots; y_i \in R\}$ consists of a certain number of observations, a group consisting of an input vector and a desired output value can be defined as (Lisowski 2013)

$$f{:}[y_{n-k}, \ldots, y_{n-3}, y_{n-2}, y_{n-1}] \to y_n \tag{12}$$

The weight values inside the network are adjusted by an algorithm that minimizes the root-mean square error between the expected output and current output. For the detailed ANN to model the time series, please refer to Lisowski (2013) and Zhang and Qi (2005).

Model Fitting and Prediction Results

Linear Regression

Table 6 shows the outputs of the linear regression models for 12 months. We estimate monthly traffic volume based on historical traffic volumes in the same month. We find that, in the models for months April, May, June, July, August, September, October, and December, there is no evidence to demonstrate that the traffic volumes in these months have a statistically significant increasing or decreasing trend from the year of 2006 (p values of the t-test > 0.05, the bold numbers) in Table 6. Moreover, the R-squared values in most linear regression models are very small, indicating that the simple linear regression model does not fit the empirical data well. Thus, the simple *growth factor* model may not be suitable for the data set in this study. By comparison, more sophisticated statistical models such as ARIMA, SVM, ANN, and UCM perform better.

Table 7. MSE for various SVM models

Model description	Mean squared error (MSE)
Regression with lag-1 variable	1.05
Regression with lag-4 variable	1.97
Regression with lag-12 variable	0.48
Regression with season variable	0.95
Regression with lag-1 and season variable	0.44
Regression with lag-12 and season variable	0.42
Regression with lag-1 and lag-12 variable	0.37
Regression with lag-1, lag-12, and season variable	**0.27**

Note: Bold characters indicate that the model with lag-1, lag-2, and a season variable has the best statistical performance.

ARIMA

The method for looking for the best ARIMA is to maximize the R-squared values. First, we attempt to use the first-order model as a preliminary autoregressive model. However, the first-order model yields poor fit to empirical data since the R-squared values are small. This indicates that the assumption that the current month's traffic volume only depends on the traffic volume in the previous month is not satisfied in our data set. Therefore, we develop higher-order ARIMA models which yield a better fit. After several rounds of trial, the *final* ARIMA model is formulated below using the original data from January 2006 through October 2015. This model is then used to create predictions for the following months from November 2015 to October 2016

$$\text{ARIMA Model:}\ y_t = y_{t-1} + y_{t-12} - y_{t-13} + (e_t - 0.52e_{t-1} - 0.80e_{t-12} + 0.42e_{t-13}) \tag{13}$$

where y_{t-1}, y_{t-12}, y_{t-13} are actual values, and e_{t-1}, e_{t-12}, e_{t-13} are the random error values at time t for lag-1, lag-12, and lag-13, respectively.

SVM

Various SVM models are developed. The mean squared errors (MSE) of these models are shown in Table 7. Among the models, the model with lag-1, lag-12, and a season variable (bold characters in Table 7) has the *best* statistical performance because of its least

Table 6. Outputs of linear regression models

Months	Coefficients: Intercept	Coefficients: Independent variable (year)	P-value for parameter estimates: Intercept	P-value for parameter estimates: Independent variable (year)	R-squared values
January	625014171	−303130	3.64×10^{-11}	0.000739	0.78
February	472255792	−227704	1.50×10^{-9}	0.025753	0.48
March	411615023	−196234	1.51×10^{-11}	0.006710	0.62
April	217703387	−99643	1.19×10^{-10}	**0.187508**	0.21
May	181488120	−81203	1.91×10^{-10}	**0.319408**	0.12
June	190777567	−85872	1.76×10^{-10}	**0.287607**	0.14
July	112239491	−46625	3.62×10^{-10}	**0.590748**	0.04
August	162238428	−71353	2.53×10^{-10}	**0.406023**	0.09
September	156413933	−69194	2.91×10^{-10}	**0.391188**	0.09
October	211676783	−96369	4.12×10^{-9}	**0.411602**	0.09
November	466425910	−223654	6.43×10^{-10}	0.015390	0.59
December	370581860	−175972	3.39×10^{-9}	**0.085502**	0.36

Note: Bold numbers indicate that p-values of the t-test > 0.05.

Table 8. Prediction results obtained by simple linear regression, ARIMA, SVM, ANN, and UCM (in millions)

		Predicted values				
Months	Observed values	Simple linear regression	ARIMA	SVM	ANN	UCM
November 2015	18.112	15.762	18.003	17.331	17.553	**18.133**
December 2015	18.470	15.999	18.074	17.182	17.667	**18.180**
January 2016	15.901	13.904	**16.478**	16.650	17.481	16.791
February 2016	16.179	13.205	15.458	16.738	16.904	**15.727**
March 2016	18.460	16.007	18.137	16.650	16.526	**18.406**
April 2016	18.481	16.823	**18.582**	17.623	16.826	18.741
May 2016	19.435	17.782	**19.482**	18.519	17.414	19.647
June 2016	19.718	17.660	19.369	18.871	17.977	**19.599**
July 2016	19.971	18.244	19.792	19.053	17.148	**20.014**
August 2016	20.329	18.391	20.058	19.111	18.164	**20.344**
September 2016	18.914	16.920	18.563	18.125	17.594	**18.907**
October 2016	19.472	17.396	19.185	18.271	17.433	**19.578**

Note: Bold numbers are the closest predicted values to the observed values.

mean squared error. Thus, this model is chosen to predict the traffic volume.

ANN

Four ANN models are developed. They are regressions with a lag-1 variable; with lag-1 and lag-2 variables; with lag-1, lag-2, and lag-3 variables; and with lag-1, lag-2, lag-3, and lag-4 variables, respectively. After comparing these four models in terms of the residual sum of squares, we find that regression with the lag-1, lag-2, and lag-3 model has the *best* performance.

Table 8 exhibits the prediction results of the testing data obtained by linear regression, ARIMA, SVM, ANN, and UCM models. We compare the prediction accuracies of the five models based on the absolute deviation of the predicted values from observed values. The bold numbers in Table 8 are the closest predicted values to the observed values. Except for the months of January, April, and May in 2016, the UCM obtains the most accurate prediction results compared with the other four alternative models. We use the Mann-Whitney U test to test if the model prediction differences are statistically significant. The absolute deviations of the prediction values from observed values are used to conduct the Mann-Whitney U test (McKnight and Najab 2010). The p-values are 7.396×10^{-7}, 0.0885, 1.558×10^{-4}, and 7.658×10^{-5} for the tests of simple linear regression versus UCM, ARIMA versus UCM, SVM versus UCM, and ANN versus UCM, respectively. This indicates that the prediction difference between ARIMA and UCM is statistically significant given the Type I error over 8.85%. UCM outperforms linear regression, SVM, and ANN significantly. This paper focuses on modeling monthly traffic volume. In order to explore the potential use of UCM for shorter-term traffic prediction, hourly traffic data on another selected route is analyzed based on UCM, ARIMA, SVM, and ANN. The detailed comparison results are in the Appendix.

Model Performance Evaluation and Comparison

A number of criteria are used to measure the statistical performance of UCM versus the other four models, including mean squared deviation (MSD), mean absolute deviation (MAD), and mean absolute percentage error (MAPE) (Muttar 2008). Smaller values of MSD, MAD, and MAPE indicate more accurate predictions of traffic volumes. MSD, MAD, and MAPE are defined below in the following.

Mean Squared Deviation

MSD is a commonly-used measurement of the accuracy of fitted time series values. The equation is

$$\frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|^2}{n} \quad (14)$$

where y_t is the actual observation at time t, $\hat{y}_t$ is the forecast value of y_t based on a particular model, and n is the total number of observations.

Mean Absolute Deviation

MAD expresses the accuracy in the same units, which helps measure the amount of error. The equation is

$$\frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n} \quad (15)$$

where y_t is the actual observation at time t, $\hat{y}_t$ is the forecast value of y_t based on a particular model, and n is the total number of observations.

Mean Absolute Percentage Error

MAPE measures the accuracy as a percentage of the error. The equation is

$$\frac{\sum_{t=1}^{n} |(y_t - \hat{y}_t)/y_t|}{n} \times 100(y_t \neq 0) \quad (16)$$

where y_t is the actual observation at time t, $\hat{y}_t$ is the forecasted value of y_t based on a particular model, and n is the total number of observations.

The three indexes, MSD, MAD, and MAPE are used to measure the performance of the five models in terms of training data and testing data (not used for model development). The comparison results are presented in Table 9. The bold numbers are the smallest values of three criteria. For the testing data, UCM has the smallest MSD, MAD, and MAPE values, indicating that the UCM better fits the empirical data than linear regression, ARIMA, ANN, and SVM.

UCM versus ARIMA

From the above comparison, UCM and ARIMA outperform ANN and SVM based on the traffic data in this study. ANN requires a large amount of data in practice. Furthermore, ANN could have difficulty effectively capturing seasonal or trend variations in untreated raw data (Zhang and Qi 2005). For SVM, the differences in MSD, MAD, and MAPE between testing data and training data are substantial, indicating that SVM might be overfitting the empirical data used. In this subsection, we focus on discussing the relationship and comparison between UCM and ARIMA.

Table 9. MSD, MAD, and MAPE for various models

Criterion	Linear regression	ARIMA	SVM	ANN	UCM
MSD					
Training	3.88×10^{11}	2.23×10^{11}	2.68×10^{11}	15.10×10^{11}	0.61×10^{11}
Testing	46.02×10^{11}	1.31×10^{11}	10.92×10^{11}	30.11×10^{11}	1.02×10^{11}
MAD					
Training	4.71×10^{5}	3.62×10^{5}	4.05×10^{5}	9.39×10^{5}	1.77×10^{5}
Testing	21.12×10^{5}	3.09×10^{5}	9.95×10^{5}	16.14×10^{5}	2.06×10^{5}
MAPE					
Training	2.77	2.18	2.41	5.68	1.07
Testing	11.49	1.72	5.32	8.56	1.20

First, ARIMA can be viewed as a *reduced form* of the generalized structural time series UCM model. In ARIMA, multiple disturbances can be treated as a single disturbance, whereas in the UCM, these disturbances are treated separately. One main advantage of an ARIMA approach is "applying convenient differences to the original series before performing the analysis, and thus eliminating the trend or seasonal components" (Jalles 2009). In principle, the ARIMA model can be deemed as a *black-box* approach in which the adopted model depends entirely on the data, without a prior analysis of the structure underlying the system. The UCM and other types of structural models are "more transparent because they allow to check if each predicted component corresponds to the expectation from the data" (Jalles 2009).

Second, UCM models are more flexible. The recursive nature of the model and the computation techniques used for its analysis allow the direct incorporation of known breaks in the system structure over time. On the contrary, ARIMA models are based on the assumption that different series are "stationary" (Harvey 1990; Jalles 2009).

Third, it is challenging to use an ARIMA approach to handle missing observations. By contrast, UCM is more robust against missing values (Jalles 2009). Jalles (2009) stated that the incorporation of explanatory variables, calendar effects, and structural breaks is not always immediate in ARIMA, in comparison with UCM. Although we do not experience missing data problems in this research, handling missing data might be another advantage of UCM in other transportation applications.

Concluding Remarks

This paper develops an unobserved component model to predict monthly traffic volume (for two-axle passenger cars) on a key corridor in New Jersey. The proposed UCM model outperforms linear regression, ARIMA, SVM, and ANN in terms of both the training data and testing data. UCM can be viewed as a more general form of ARIMA. UCM decomposes time series into trends, seasonal variations, and irregular components, accounts for component-specific disturbances, and is not restrained by stationary assumptions via direct incorporation of known breaks in the system structure. The preliminary results show that UCM could be a promising alternative statistical approach to the prediction of monthly traffic volume. Future study is needed to better understand the adaptation of UCM to a broader set of transportation engineering and prediction problems. Moreover, additional contributing factors (e.g., travel behavior and economic growth) across multiple routes and time periods will be used to more explicitly study the cause and effect when making traffic predictions. Finally, this paper focuses on monthly traffic volume prediction. The preliminary analysis indicates the promising application of this approach to shorter-term traffic modeling as well. In the future, more research will be conducted to explore the use of UCM for predicting traffic volume in an hourly or even shorter time interval.

Table 10. Model comparison for hourly traffic forecasting

Criterion	UCM	ARIMA	SVM	ANN
MSD	10,558.17	233,409.50	31,615.71	122,063.30
MAD	64.84	386.45	110.62	227.40
MAPE	21.22	87.22	31.99	45.76

Appendix. Hourly Traffic Volume Prediction

One-month (January 2016) hourly traffic volume on one route in the State of Wisconsin is analyzed to compare four prediction models, namely UCM, ARIMA, SVM, and ANN. In this data set with a total of 744 h, 672 h (28 days) are used as training data and the remaining 72 h (3 days) are used as the testing data. Four prediction models are developed with hourly traffic volumes. The goodness of fit of the model is measured by MSD, MAD, and MAPE (Table 10). According to the comparison results using this data set, UCM appears to outperform the selected alternative methods.

References

Abdulhai, B., H. Porwal, and W. Recker. 2002. "Short-term traffic flow prediction using neuro-genetic algorithms." *ITS J.-Intell. Transp. Syst. J.* 7 (1): 3–41. https://doi.org/10.1080/713930748.

Ahn, J. Y., E. Ko, and E. Kim. 2015. "Predicting spatiotemporal traffic flow based on support vector regression and bayesian classifier." In *Proc., 2015 IEEE 5th Int. Conf. on Big Data and Cloud Computing (BDCloud)*. New York: IEEE.

Ariyo, A. A., A. O. Adewumi, and C. K. Ayo. 2014. "Stock price prediction using the ARIMA model." In *Proc., 2014 UKSim-AMSSProc., 16th Int. Conf. on Computer Modelling and Simulation (UKSim)*. New York: . IEEE.

Bai, Y., O. E. Gungor, J. A. Hernandez-Urrea, Y. Ouyang, and I. L. Al-Qadi. 2015. "Optimal pavement design and rehabilitation planning using a mechanistic-empirical approach." *EURO J. Transp. Logist.* 4 (1): 57–73. https://doi.org/10.1007/s13676-014-0072-2.

Cawley, G. C., and N. L. Talbot. 2010. "On over-fitting in model selection and subsequent selection bias in performance evaluation." *J. Mach. Learn. Res.* 11 (Jul): 2079–2107.

Chan, K. Y., T. S. Dillon, J. Singh, and E. Chang. 2012. "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm." *IEEE Trans. Intell. Transp. Syst.* 13 (2): 644–654. https://doi.org/10.1109/TITS.2011.2174051.

Chen, L., Q. R. Li, X. Y. Tian, X. S. Chen, and R. X. Wang. 2012. "Paratactic spatial-temporal two dimension data fusion based on support vector machines for traffic flow prediction of abnormal state." In Vol. 532 of *Advanced materials research*. Zurich, Switzerland: Trans Tech Publications.

Cowan, A. M., and F. L. Joutz. 2006. "An unobserved component model of asset pricing across financial markets." *Int. Rev. Financial Anal.* 15 (1): 86–107. https://doi.org/10.1016/j.irfa.2004.10.004.

Dell'Acqua, P., F. Bellotti, R. Berta, and A. De Gloria. 2015. "Time-aware multivariate nearest neighbor regression methods for traffic flow prediction." *IEEE Trans. Intell. Transp. Syst.* 16 (6): 3393–3402. https://doi.org/10.1109/TITS.2015.2453116.

Dia, H. 2001. "An object-oriented neural network approach to short-term traffic forecasting." *Eur. J. Oper. Res.* 131 (2): 253–261. https://doi.org/10.1016/S0377-2217(00)00125-9.

Dibike, Y. B., S. Velickov, D. Solomatine, and M. B. Abbott. 2001. "Model induction with support vector machines: Introduction and applications." *J. Comput. Civ. Eng.* 15 (3): 208–216. https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208).

Dimitriou, L., T. Tsekeris, and A. Stathopoulos. 2008. "Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow." *Transp. Res. Part C: Emerging Technol.* 16 (5): 554–573. https://doi.org/10.1016/j.trc.2007.11.003.

Frangopol, D. M., and M. Liu. 2007. "Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost*." *Struct. Infrastruct. Eng.* 3 (1): 29–41. https://doi.org/10.1080/15732470500253164.

Goves, C., R. North, R. Johnston, and G. Fletcher. 2016. "Short term traffic prediction on the UK motorway network using neural networks." *Transp. Res. Procedia* 13: 184–195. https://doi.org/10.1016/j.trpro.2016.05.019.

Guo, F., R. Krishnan, and J. W. Polak. 2012. "Short-term traffic prediction under normal and abnormal traffic conditions on urban roads." In *Proc., Transportation Research Board 91st Annual Meeting*. Washington, DC: Transportation Research Board.

Hajibabai, L., Y. Bai, and Y. Ouyang. 2014. "Joint optimization of freight facility location and pavement infrastructure rehabilitation under network traffic equilibrium." *Transp. Res. Part B: Methodol.* 63: 38–52. https://doi.org/10.1016/j.trb.2014.02.003.

Harvey, A. C. 1990. *Forecasting, structural time series models and the Kalman filter.* Cambridge, MA: Cambridge University Press.

Harvey, A. C., and S. Peters. 1990. "Estimation procedures for structural time series models." *J. Forecasting* 9 (2): 89–108. https://doi.org/10.1002/for.3980090203.

Hou, Z., and X. Li. 2016. "Repeatability and similarity of freeway traffic flow and long-term prediction under big data." *IEEE Trans. Intell. Transp. Syst.* 17 (6): 1786–1796. https://doi.org/10.1109/TITS.2015.2511156.

Hu, C., K. Xie, G. Song, and T. Wu. 2008. "Hybrid process neural network based on spatio-temporal similarities for short-term traffic flow prediction." In *Proc., 2008 11th Int. IEEE Conf. on Intelligent Transportation Systems*. New York: IEEE.

Hu, W., L. Yan, K. Liu, and H. Wang. 2016. "A short-term traffic flow forecasting method based on the hybrid PSO-SVR." *Neural Process. Lett.* 43 (1): 155–172. https://doi.org/10.1007/s11063-015-9409-6.

Hyndman, R. J., and G. Athanasopoulos. 2014. *Forecasting: Principles and practice*. Melbourne, Australia: OTexts.

Ishak, S., and C. Alecsandru. 2004. "Optimizing traffic prediction performance of neural networks under various topological, input, and traffic condition settings." *J. Transp. Eng.* 130 (4): 452–465. https://doi.org/10.1061/(ASCE)0733-947X(2004)130:4(452).

Jalles, J. T. 2009. *Structural time series models and the Kalman Filter: A concise review.* FEUNL Working Paper No. 541. Amsterdam, Netherlands: Elsevier.

Kim, C.-J. 1993. "Unobserved-component time series models with Markov-switching heteroscedasticity: Changes in regime and the link between inflation rates and inflation uncertainty." *J. Bus. Econ. Statistics* 11 (3): 341–349. https://doi.org/10.1080/07350015.1993.10509962.

Koopman, S., and M. Ooms. 2011. *Forecasting economic time series using unobserved components time series models*, 129–162. Oxford, UK: Oxford University Press.

Kumar, K., M. Parida, and V. K. Katiyar. 2015. "Short term traffic flow prediction in heterogeneous condition using artificial neural network." *Transport* 30 (4): 397–405. https://doi.org/10.3846/16484142.2013.818057.

Kumar, S. V., and L. Vanajakshi. 2015. "Short-term traffic flow prediction using seasonal ARIMA model with limited input data." *Eur. Transp. Res. Rev.* 7 (3): 1–9. https://doi.org/10.1007/s12544-015-0170-8.

Lisowski, M. 2013. "An artificial neural network based on time series prediction system with distributed computing." *Zeszyty Naukowe Akademii Morskiej w Gdyni* 79: 32–40.

Lu, H. 2014. "Short-term traffic prediction using rainfall." *Int. J. Signal Process. Syst.* 2 (1): 70–73.

Luo, F., Y.-G. Xu, and J.-Z. Cao. 2005. "Elevator traffic flow prediction with least squares support vector machines." In Vol. 7 of *Proc., 2005 Int. Conf. on Machine Learning and Cybernetics*. New York: IEEE.

Lv, Y., Y. Duan, W. Kang, Z. Li, and F. Y. Wang. 2015. "Traffic flow prediction with big data: A deep learning approach." *IEEE Trans. Intell. Transp. Syst.* 16 (2): 865–873. https://doi.org/10.1109/TITS.2014.2345663.

Ma, T., Z. Zhou, and B. Abdulhai. 2015. "Nonlinear multivariate time–space threshold vector error correction model for short term traffic state prediction." *Transp. Res. Part B: Methodol.* 76 (Jun): 27–47. https://doi.org/10.1016/j.trb.2015.02.008.

Marquez, F. P. G., D. J. P. Tercero, and F. Schmid. 2007. "Unobserved component models applied to the assessment of wear in railway points: A case study." *Eur. J. Oper. Res.* 176 (3): 1703–1712. https://doi.org/10.1016/j.ejor.2005.10.037.

McCrea, J., and S. Moutari. 2010. "A hybrid macroscopic-based model for traffic flow in road networks." *Eur. J. Oper. Res.* 207 (2): 676–684. https://doi.org/10.1016/j.ejor.2010.05.018.

McKnight, P. E., and J. Najab. 2010. "Mann-whitney U test." In *Corsini encyclopedia of psychology.* New York: Wiley.

Min, X., J. Hu, and Z. Zhang. 2010. "Urban traffic network modeling and short-term traffic flow forecasting based on GSTARIMA model." In *Proc., 2010 13th Int. IEEE Conf. on Intelligent Transportation Systems (ITSC)*. New York: IEEE.

Muttar, T. R. 2008. "A proposed technique for the problem of selecting the best forecasting model in time series: A case study." *Iraqi J. Stat. Sci.* 14: 1–20.

Ng, M., D. Y. Lin, and S. T. Waller. 2009. "Optimal long-term infrastructure maintenance planning accounting for traffic dynamics." *Comput. Aided Civil Infrastruct. Eng.* 24 (7): 459–469. https://doi.org/10.1111/j.1467-8667.2009.00606.x.

Paradiso, A., and B. B. Rao. 2012. "Flattening of the Phillips curve and the role of the oil price: An unobserved component model for the USA and Australia." *Econ. Lett.* 117 (1): 259–262. https://doi.org/10.1016/j.econlet.2012.05.022.

Pascale, A., and M. Nicoli. 2011. "Adaptive Bayesian network for traffic flow prediction." In *Proc., 2011 IEEE Statistical Signal Processing Workshop (SSP)*. New York: IEEE.

Pedregal, D. J., F. P. García, and F. Schmid. 2004. "RCM 2 predictive maintenance of railway systems based on unobserved components models." *Reliab. Eng. Syst. Saf.* 83 (1): 103–110. https://doi.org/10.1016/j.ress.2003.09.020.

Qing-Fang, M., C. Yue-Hui, and P. Yu-Hua. 2009. "Small-time scale network traffic prediction based on a local support vector machine regression model." *Chin. Phys. B* 18 (6): 2194. https://doi.org/10.1088/1674-1056/18/6/014.

Sapankevych, N. I., and R. Sankar. 2009. "Time series prediction using support vector machines: A survey." *IEEE Comput. Intell. Mag.* 4 (2): 24–38. https://doi.org/10.1109/MCI.2009.932254.

Sun, S., R. Huang, and Y. Gao. 2012. "Network-scale traffic modeling and forecasting with graphical lasso and neural networks." *J. Transp. Eng.* 138 (11): 1358–1367. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000435.

Sun, S., C. Zhang, and Y. Zhang. 2005. "Traffic flow forecasting using a spatio-temporal bayesian network predictor." In *Proc., Int. Conf. on Artificial Neural Networks*. Berlin: Springer.

Thissen, U., R. Van Brakel, A. P. De Weijer, W. J. Melssen, and L. M. C. Buydens. 2003. "Using support vector machines for time series prediction." *Chemom. Intell. Lab. Syst.* 69 (1): 35–49. https://doi.org/10.1016/S0169-7439(03)00111-4.

Vlahogianni, E. I., M. G. Karlaftis, and J. C. Golias. 2005. "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach." *Transp. Res. Part C: Emerging Technol.* 13 (3): 211–234. https://doi.org/10.1016/j.trc.2005.04.007.

Williams, B. M., and A. H. Lester. 2003. "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results." *J. Transp. Eng.* 129 (6): 664–672. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664).

Wu, S., Z. Yang, X. Zhu, and B. Yu. 2014. "Improved k-NN for short-term traffic forecasting using temporal and spatial information." *J. Transp. Eng.* 140 (7): 04014026. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000672.

Xia, D., B. Wang, H. Li, Y. Li, and Z. Zhang. 2016. "A distributed spatial–temporal weighted model on MapReduce for short-term traffic flow forecasting." *Neurocomputing* 179 (Feb): 246–263. https://doi.org/10.1016/j.neucom.2015.12.013.

Xie, Y., and Y. Zhang. 2006. "A wavelet network model for short-term traffic volume forecasting." *J. Intell. Transp. Syst.* 10 (3): 141–150. https://doi.org/10.1080/15472450600798551.

Xu, Q. H., and Rui Yang. 2005. "Traffic flow prediction using support vector machine based method." *J. Highway Transp. Res. Dev.* 22 (12): 131–134.

Yao, Y., Y. Liu, Y. Yu, H. Xu, W. Lv, Z. Li, and X. Chen. 2013. "K-SVM: An effective SVM algorithm based on K-means clustering." *JCP* 8 (10): 2632–2639.

Young, P. C. 2011. "Introduction." In *Recursive estimation and time-series analysis*, 1–9. Berlin: Springer.

Yu, B., X. Song, F. Guan, Z. Yang, and B. Yao. 2016. "k-nearest neighbor model for multiple-time-step prediction of short-term traffic condition." *J. Transp. Eng.* 142 (6): 04016018. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000816.

Yuan, Z.-W., and Y.-H. Wang. 2012. "Research on K nearest neighbor non-parametric regression algorithm based on KD-tree and clustering analysis." In *Proc., 2012 4th Int. Conf. on Computational and Information Sciences (ICCIS)*. New York: IEEE.

Zhang, G., B. E. Patuwo, and M. Y. Hu. 1998. "Forecasting with artificial neural networks: The state of the art." *Int. J. Forecasting* 14 (1): 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7.

Zhang, G. P., and M. Qi. 2005. "Neural network forecasting for seasonal and trend time series." *Eur. J. Oper. Res.* 160 (2): 501–514. https://doi.org/10.1016/j.ejor.2003.08.037.

Zhang, N., Y. Zhang, and H. Lu. 2011. "Seasonal autoregressive integrated moving average and support vector machine models: Prediction of short-term traffic flow on freeways." *Transp. Res. Rec.* 2215 (1): 85–92. https://doi.org/10.3141/2215-09.

Zhang, Y., Y. Zhang, and A. Haghani. 2014. "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model." *Transp. Res. Part C: Emerging Technol.* 43 (Part 1): 65–78. https://doi.org/10.1016/j.trc.2013.11.011.

Zheng, W., D.-H. Lee, and Q. Shi. 2006. "Short-term freeway traffic flow prediction: Bayesian combined neural network approach." *J. Transp. Eng.* 132 (2): 114–121. https://doi.org/10.1061/(ASCE)0733-947X(2006)132:2(114).

Zhong, J. T., and S. Ling. 2015. "Key factors of k-nearest neighbours non-parametric regression in short-time traffic flow forecasting." In *Proc., 21st Int. Conf. on Industrial Engineering and Engineering Management 2014*. Paris: Atlantis Press.

Zhong, M., and S. Sharma. 2006. "Matching hourly, daily, and monthly traffic patterns to estimate missing volume data." *Transp. Res. Rec.* 1957 (1): 32–42. https://doi.org/10.1177/0361198106195700106.

Zhu, J. Z., J. X. Cao, and Y. Zhu. 2014. "Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections." *Transp. Res. Part C: Emerging Technol.* 47 (Part 2): 139–154. https://doi.org/10.1016/j.trc.2014.06.011.

Zhu, Z., B. Peng, C. Xiong, and L. Zhang. 2016. "Short-term traffic flow prediction with linear conditional Gaussian Bayesian network." *J. Adv. Transp.* 50 (6): 1111–1123. https://doi.org/10.1002/atr.1392.

Zou, Y., X. Hua, Y. Zhang, and Y. Wang. 2015. "Hybrid short-term freeway speed prediction methods based on periodic analysis." *Can. J. Civ. Eng.* 42 (8): 570–582. https://doi.org/10.1139/cjce-2014-0447.