

# Improved Grey System Models for Predicting Traffic Parameters

Gurcan Comert<sup>a,\*</sup>, Negash Begashaw<sup>a</sup>, Nathan Huynh<sup>b</sup>

<sup>a</sup>*Computer Science, Physics, and Engineering Department, Benedict College, 1600 Harden St., Columbia, USA  
29204*

<sup>b</sup>*Department of Civil and Environmental Engineering, University of South Carolina, Columbia, SC 29208*

---

## Abstract

In transportation applications such as real-time route guidance, ramp metering, congestion pricing and special events traffic management, accurate short-term traffic flow prediction is needed. For this purpose, this paper proposes several novel *online* Grey system models (GM): GM(1,1| $\cos(\omega t)$ ), GM(1,1| $\sin(\omega t)$ ,  $\cos(\omega t)$ ), and GM(1,1| $e^{-at}, \sin(\omega t), \cos(\omega t)$ ). To evaluate the performance of the proposed models, they are compared against a set of benchmark models: GM(1,1) model, Grey Verhulst models with and without Fourier error corrections, linear time series model, and nonlinear time series model. The evaluation is performed using loop detector and probe vehicle data from California, Virginia, and Oregon. Among the benchmark models, the error corrected Grey Verhulst model with Fourier outperformed the GM(1,1) model, linear time series, and non-linear time series models. In turn, the three proposed models, GM(1,1| $\cos(\omega t)$ ), GM(1,1| $\sin(\omega t), \cos(\omega t)$ ), and GM(1,1| $e^{-at}, \sin(\omega t), \cos(\omega t)$ ), outperformed the Grey Verhulst model in prediction by at least 65%, 16% and 11%, in terms of Root Mean Squared Error, and by 82%, 58% and 42%, in terms of Mean Absolute Percentage Error, respectively. It is observed that the proposed Grey system models are more adaptive to location (e.g., perform well for all roadway types) and traffic parameters (e.g., speed, travel time, occupancy, and volume), and they do not require as many data points for training (4 observations are found to be sufficient).

**Keywords:** Grey system models, travel time, traffic speed, traffic volume, occupancy, transferable models.

---

## 1. Introduction

Traffic congestion has become a way of life for the majority of U.S. commuters in urban areas. With nearly 90 percent of commuters wanting to get to their destinations at the same time, it is simply a matter of supply not being able to accommodate the demand during peak-hour loads. Many solutions have been investigated to mitigate traffic congestion, including real-time route guidance, ramp metering, and congestion pricing. The cornerstone of these solutions is a proactive traffic management system that relies on accurate short-term prediction of traffic parameters, such as speed, travel time, volume, and occupancy (density). By their nature, these traffic parameters are time series, and therefore, many researchers have attempted to study their behavior using time series models. However, a shortcoming of time series models is their inability to deal with non-recurring abrupt changes such as those caused by accidents (Hamilton (2010);

---

\*Corresponding author

Email addresses: [gurcan.comert@benedict.edu](mailto:gurcan.comert@benedict.edu) (Gurcan Comert), [negash.begashaw@benedict.edu](mailto:negash.begashaw@benedict.edu) (Negash Begashaw), [huynhn@cec.sc.edu](mailto:huynhn@cec.sc.edu) (Nathan Huynh)

Clements and Harvey (2010); Hamilton (2008)). This is because traffic flow is nonlinear, stochastic and highly non-stationary.

To overcome the aforementioned limitation with time series models, this paper explores the use of Grey system theory. As the name implies, Grey (i.e., hazy or fuzzy) systems are those systems without complete information. Examples of Grey systems include the human body, economy and weather system. Grey system theory provides methods for resolving latent processes. As such, it uses discrete or continuous Grey numbers encoded by symbol  $\otimes$  to represent the unknown parameters of the system. Grey models are popular due to their ability to estimate the behavior of systems using only a small amount of data. This characteristic makes Grey models suitable for short-term traffic parameter estimation.

The Grey model GM(1,1) (GM(1,1) stand for "Grey Model First Order One Variable") is the most widely used and documented in the literature (Kayacan et al. (2010)). In Bezuglov and Comert (2016), first author explored the applicability of GM(1,1), Grey Verhulst, and their error corrected versions and showed that Grey models were capable of producing accurate forecasts for un-seasonal or cyclic average traffic speed and travel time data. Authors also noticed slow and excessive forecasts at change points. To improve the performance of the GM(1,1) model for short-term estimation of various traffic parameters, this study proposes to augment the GM(1,1) model with trigonometric functions. Incorporating the oscillating behavior of the sine and cosine functions, or their combinations, these Grey models will be more adept at predicting abrupt changes in traffic flow. In this paper, the following models are proposed:  $GM(1,1|\cos(\omega t))$ ,  $GM(1,1|\sin(\omega t),\cos(\omega t))$ , and  $GM(1,1|e^{-at},\sin(\omega t),\cos(\omega t))$ . To the best of our knowledge, these models have not been studied for short term traffic parameter estimation. To evaluate the performance of these models, they are compared with a set of benchmark models: GM(1,1) model, Grey Verhulst models with and without Fourier error corrections, seasonal time series model, and nonlinear time series model. The evaluations are done using data from California, Virginia, and Oregon. The models proposed in this paper outperformed the Grey Verhulst model in prediction by 65%, 16% and 11%, in terms of Root Mean Squared Error, and by 82%, 58% and 42%, in terms of Mean Absolute Percentage Error respectively. It is observed that the proposed Grey system models are more adaptive to location (e.g., perform well for all roadway types) and traffic parameters (e.g., speed, travel time, occupancy, and volume), and they do not require many data points for training (4 observations are found to be sufficient).

## 2. Literature Review

Grey system theory has been widely applied in different areas of applications for more than 30 years. Thus, there is much information about its theory and applications in the literature. Readers are referred to the work of Liu et al. (2010) for the theory and the work by Deng (1989).

In transportation related studies, the GM(1,1) model has been used to predict annual average daily traffic (AADT), hourly and monthly volume, traffic accidents, and pavement deterioration. To predict traffic volume, specifically AADT, Zhang (2010) combined the GM(1,1) model with a Markov transition matrix. The proposed model performed better in predicting AADTs. Badhrudeen et al. (2016) compared the performance of the GM(1,1) model against neural network models (NNs). Using about 100 minutes of observational data to train the GM(1,1) model, the results indicated that the GM(1,1) model outperformed the NNs by 1 to 3% in mean absolute percent error (MAPE). Similarly, An et al. (2012) compared the performance of the GM(1,1) model with back propagating NN and radial basis NN to predict AADT values. Their results indicated that the GM(1,1) model outperformed both types of neural network models. In contrast, Pan and Shi (2016) showed that the GM(1,1) model performed poorly against the NN model in

predicting traffic speed (about 6% on average in MAPE). Yu et al. (2015) examined the GM(1,1) model, the ARIMA (Autoregressive Integrated Moving Average) model, generalized NN regression model and their combinations. The authors tested different combinations of models with fixed and Elman NN-based weights, a framework similar to forecast encompassing (Clements and Harvey (2010)). Elman NN-based weighted combined model (i.e., GM(1,1), ARIMA and NN regression combined) showed the highest accuracy in predicting monthly traffic volume. Liu et al. (2014) combined cubic exponential smoothing with GM(1,1) model. In their study, they showed that the proposed model provided lower error in predicting the monthly traffic volume.

To predict hourly volume, several researchers (e.g., Shuhua et al. (2012) and Shuhua and Xinping (2010)) have enhanced the GM(1,1) model by incorporating periodic sine and tangent trigonometric terms. According to Shuhua et al. (2012), the GM(1,1) model with the sine term included outperformed the GM(1,1) model by 7% in terms of MAPE in predicting hourly traffic volume. In the study by Shuhua and Xinping (2010), the authors optimized the trigonometric function periods for the GM(1,1) model using particle swarm optimization. The authors did not evaluate the performance of the proposed model.

Grey models have also been applied to predict traffic accidents and pavement deterioration. Studies in these areas are rather limited. Na et al. (2010) examined the contributing factors to accidents. The factors found to have an effect on accident probability, from lowest to highest are: traffic signal timing, driving behavior, weather condition, roadway surface, and visibility. Jie et al. (2015) applied a variation of the Grey Verhulst model to forecast annual high speed traffic accidents. They reported 9.5% accuracy in MAPE. Jiang and Li (2005) applied the GM(1,1) model to estimate pavement roughness index. The reported errors in MAPE ranged between 5 and 45%.

The above review highlights the effectiveness of the GM(1,1) model and its various enhancements. A particularly promising approach is the incorporation of trigonometric terms into the GM(1,1) model. This study further contributes to the literature by incorporating a cosine function, a combined sine and cosine function, and an exponential term into the GM(1,1) model for short-term prediction of different traffic parameters.

### 3. Methods

#### 3.1. GM(1,1) model

To model time series, Grey System theory provides a family of Grey models, where the first order Grey model is referred to as GM(1,1). The underlying theory of the GM(1,1) model is as follows. Notation was adopted from Kayacan et al. (2010); Bezuglov and Comert (2016) and presented here for completeness and initiation to the proposed new Grey models.

Let  $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$  denote a sequence of non-negative observations of a stochastic process and  $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$  be an accumulation sequence of  $X^{(0)}$  where

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad (1)$$

The original form of GM(1,1) is defined by the following equation.

$$x^{(0)}(k) + ax^{(1)}(k) = b \quad (2)$$

Let  $Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$  be a mean sequence of  $X^{(1)}$  where

$$z^{(1)}(k) = \frac{z^{(1)}(k-1) + z^{(1)}(k)}{2}, \forall k = 2, 3, \dots, n \quad (3)$$

The basic form of GM(1,1) is given by the following equation.

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (4)$$

If  $\hat{a} = (a, b)^T$  and

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}. \quad (5)$$

then, as in Liu and Lin (2006), the least squares estimate of the GM(1,1) model is  $\hat{a} = (B^T B)^{-1} B^T Y$  and the whitenization equation of the GM(1,1) model is given by,

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(k) = b \quad (6)$$

Suppose that  $\hat{x}^{(0)}(k)$  and  $\hat{x}^{(1)}(k)$  represent the time response sequence (the forecast) and the accumulated time response sequence of GM(1,1) at time  $k$  respectively. Then, the latter can be obtained by solving Eq. (6):

$$\hat{x}^{(1)}(k+1) = \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, k = 1, 2, \dots, n \quad (7)$$

From Eq. (1), the restored values of  $\hat{x}^{(0)}(k+1)$  are calculated as  $\hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ . Using Eq. 7, the following prediction equation is obtained.

$$\hat{x}^{(0)}(k+1) = (1 - e^a) \left( x^{(0)}(1) - \frac{b}{a} \right) e^{-ak}, k = 1, 2, \dots, n \quad (8)$$

Eq. (8) is the main forecasting equation that generates values  $\forall k = 2, 3, \dots, n$ . In the rolling window approach:  $x^{(0)}(k+1), x^{(0)}(k+2), \dots, x^{(0)}(k+w)$ , where  $w \geq 4$  is the window size (e.g.,  $w = 4$  found to produce very good results Bezuglov and Comert (2016)). Forecasts for next data points can be calculated as  $\hat{x}^{(0)}(k+w+1), \hat{x}^{(0)}(k+w+2)$ .

### 3.2. The Grey Verhulst model (GVM)

The Grey Verhulst model (GVM) is generally used for more nonlinearly behaving data (Liu et al. (2010)). Underlying structure of the GVM is given by the following equation Kayacan et al. (2010); Bezuglov and Comert (2016).

$$x^{(0)}(k) + az^{(1)}(k) = b (z^{(1)}(k))^2 \quad (9)$$

The whitenization equation of GVM is:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b (x^{(1)})^2 \quad (10)$$

Similar to the GM(1,1), the least squares estimate is applied to find  $\hat{a} = (B^T B)^{-1} B^T Y$ , where

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & z^{(1)}(2)^2 \\ -z^{(1)}(3) & z^{(1)}(3)^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & z^{(1)}(n)^2 \end{bmatrix}. \quad (11)$$

The forecasts  $\hat{x}^{(0)}(k+1)$  are calculated using Eq. (12).

$$\hat{x}^{(0)}(k+1) = \left[ \frac{ax^{(0)}(1)(a - bx^{(0)}(1))}{bx^{(0)}(1) + (a - bx^{(0)}(1))e^{a(k-1)}} \right] \left[ \frac{(1 - e^a)e^{a(k-2)}}{bx^{(0)}(1) + (a - bx^{(0)}(1))e^{a(k-2)}} \right] \quad (12)$$

### 3.3. $GM(1,1|sin(\omega t))$ model

The idea of this proposed model is to enhance prediction of the original  $GM(1,1)$  under oscillating behavior. This is accomplished by introducing trigonometric terms on the right hand side (RHS). In order to investigate possible improvements in the model performances, we used a single sine function, a single cosine function, a linear combination of sine and cosine functions, and a linear combination of sine and cosine function multiplied by an exponential term.  $GM(1,1|sin(\omega t))$  model is adopted from Shuhua et al. (2012) and derived via Eq. 13 using the same steps as  $GM(1,1)$  and GVM models.

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(k) = b_1 sin(\omega t) + b_2 \quad (13)$$

Model parameters  $(\hat{a}, \hat{b}_1, \hat{b}_2)^T = (a, b_1, b_2)^T$  are estimated by  $(\hat{a}, \hat{b}_1, \hat{b}_2) = (B^T B)^{-1} B^T Y$  where  $Y = [x^{(0)}(2), \dots, x^{(0)}(n)]^T$  is given above and  $B$  is:

$$B = \begin{bmatrix} -z^{(1)}(2) & sin(\omega 2) & 1 \\ -z^{(1)}(3) & sin(\omega 3) & 1 \\ \vdots & \vdots & \vdots \\ -z^{(1)}(n) & sin(\omega n) & 1 \end{bmatrix}. \quad (14)$$

Suppose that  $\hat{x}^{(0)}(k)$  and  $\hat{x}^{(1)}(k)$  represent the time response sequence (the forecast) and the accumulated time response sequence of  $GM(1,1|sin(\omega t))$  at time  $k$  respectively. Then, the latter can be obtained by solving the following equation.

$$\begin{aligned} x^{(1)}(t+1) = & (x^{(1)}(1) + \frac{b_1 \omega}{a^2 + \omega^2} - \frac{b}{a})e^{-at} + \\ & \frac{b_1 \omega}{a^2 + \omega^2} (a \sin(\omega t) - \omega \cos(\omega t)) + \frac{b}{a} \end{aligned} \quad (15)$$

Using the initial condition  $x^{(1)}(1) = x^{(0)}(1)$ ,  $\forall k = 2, 3, \dots, n$ , the restored values of  $\hat{x}^{(0)}(k+1)$  in Eq. (15) can be calculated as  $\hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ . Thus, using Eq. (15), the following prediction is obtained.

$$\begin{aligned} \hat{x}^{(0)}(k+1) = & (1 - e^a)(x^{(0)}(1) + \frac{b_1 \omega}{a^2 + \omega^2} - \frac{b}{a})e^{-ak} + \\ & \frac{b_1 \omega}{a^2 + \omega^2} (a(\sin(\omega k) - \sin(\omega(k-1))) - \\ & \omega(\cos(\omega(k-1)) - \cos(\omega k))) \end{aligned} \quad (16)$$

### 3.4. $GM(1,1|cos(\omega t))$ model

Similar to  $GM(1,1|sin(\omega t))$ , the model with cosine function is given by the following equation.

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(k) = b_1 \cos(\omega t) + b_2 \quad (17)$$

If  $(\hat{a}, \hat{b}_1, \hat{b}_2)^T = (a, b_1, b_2)^T$  and

$$B = \begin{bmatrix} -z^{(1)}(2) & \cos(\omega 2) & 1 \\ -z^{(1)}(3) & \cos(\omega 3) & 1 \\ \vdots & \vdots & \vdots \\ -z^{(1)}(n) & \cos(\omega n) & 1 \end{bmatrix}. \quad (18)$$

Solution of the differential equation in Eq. 17 is given by Eq. 19.

$$x^{(1)}(t+1) = Ke^{-t} + \frac{(a^2b_2 + b_2\omega^2 + a^2b_1\cos(\omega t) + ab_1\omega\sin(\omega t))/(a(a^2 + \omega^2))}{(19)}$$

where  $K$  is obtained from the initial condition  $x^{(1)}(1)=x^{(0)}(1)$  and is given by

$$K = e^a[x^{(0)}(1) - (a^2b_2 + b_2\omega^2 + a^2b_1\cos(\omega) + ab_1\omega\sin(\omega))/(a(a^2 + \omega^2))] \quad (20)$$

Following the same procedure to derive Eq. (16) above, using Eq. (19), the following prediction equation is obtained.

$$\begin{aligned} \hat{x}^{(0)}(k+1) = & Ce^{-ak}(1 - e^a) + \\ & \frac{1}{a^3 + a\omega^2}[-ab_1\omega(\sin(\omega k) - \sin(\omega(k-1))) + \\ & a^2b_1(\cos(\omega(k-1)) - \cos(\omega k))] \end{aligned} \quad (21)$$

### 3.5. $GM(1,1|\sin(\omega t), \cos(\omega t))$ model

The model with a linear combination of Sine and Cosine functions is given by Eq. 22.

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(k) = b_1\sin(\omega t) + b_2\cos(\omega t) + b_3 \quad (22)$$

If  $(\hat{a}, \hat{b}_1, \hat{b}_2, \hat{b}_3)^T = (a, b_1, b_2, b_3)^T$  and

$$B = \begin{bmatrix} -z^{(1)}(2) & \sin(\omega 2) & \cos(\omega 2) & 1 \\ -z^{(1)}(3) & \sin(\omega 3) & \cos(\omega 3) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ -z^{(1)}(n) & \sin(\omega n) & \cos(\omega n) & 1 \end{bmatrix}. \quad (23)$$

Solution of the differential equation in Eq. 22 is given by,

$$x^{(1)}(t+1) = Ke^{-t} + \frac{\left(\frac{b_3}{a} + \cos(\omega t)(ab_2 - b_1\omega) + (ab_1 + b_2\omega)\sin(\omega t)\right)}{a^2 + \omega^2} \quad (24)$$

where  $K$  is obtained from the initial condition  $x^{(1)}(1)=x^{(0)}(1)$  and is given by

$$\begin{aligned} K = & e^a[x^{(0)}(1) - \\ & \left(\frac{b_3}{a} + \cos(\omega)(ab_2 - b_1\omega) + (ab_1 + b_2\omega)\sin(\omega)\right)] + \end{aligned} \quad (25)$$

Similar to Eq. (16) above, Eq. (24) can reduce to Eq. (26).

$$\begin{aligned}\hat{x}^{(0)}(k+1) &= Ce^{-ak}(1-e^a) + \\ &\frac{1}{a^2 + \omega^2}(\cos(\omega k) - \cos(\omega(k-1)))(ab_2 - b_1\omega) + \\ &(ab_1 + b_2\omega)(\sin(\omega k) - \sin(\omega(k-1)))\end{aligned}\quad (26)$$

### 3.6. $GM(1,1|e^{-t}, \sin(\omega t), \cos(\omega t))$ model

The model with the product of a linear combination of Sine and Cosine functions with an exponential term is given by the following equation.

$$\frac{dx^{(1)}}{dt} + ax^{(1)}(k) = e^{-at}(b_1\sin(\omega t) + b_2\cos(\omega t)) + b_3 \quad (27)$$

If  $(\hat{a}, \hat{b}_3)^T = (a, b_3)^T$  and

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}. \quad (28)$$

Then, if  $(\hat{b}_1, \hat{b}_2)^T = (b_1, b_2)^T$

$$B = \begin{bmatrix} e^{-2\hat{a}}\sin(\omega 2) & e^{-2\hat{a}}\cos(\omega 2) \\ e^{-3\hat{a}}\sin(\omega 3) & e^{-3\hat{a}}\cos(\omega 3) \\ \vdots & \vdots \\ e^{-n\hat{a}}\sin(\omega n) & e^{-n\hat{a}}\cos(\omega n) \end{bmatrix}. \quad (29)$$

The solution of the differential equation in Eq. 27 is given by,

$$x^{(1)}(t+1) = Ke^{-at} - e^{-at}((b_1\cos(\omega t) - b_2\sin(\omega t))/\omega - (b_3e^{at})/a) \quad (30)$$

where  $K$  is obtained from the initial condition  $x^{(1)}(1)=x^{(0)}(1)$  and is given by

$$K = e^a[x^{(0)}(1) + e^{-a}((b_1\cos(\omega) - b_2\sin(\omega))/\omega - (b_3e^{-a})/a)] \quad (31)$$

Following a similar explanation as in Eq. (16) above, Eq. (30) can be reduced to Eq. (32).

$$\begin{aligned}\hat{x}^{(0)}(k+1) &= \frac{e^{-ak}}{(1-e^a)}[C + b_1(\cos(\omega(k-1)) - \cos(\omega k)) + \\ &b_2(\sin(\omega k) - \sin(\omega(k-1)))]\end{aligned}\quad (32)$$

### 3.7. Error corrections to Grey models

Suppose that  $\epsilon^{(0)}=\epsilon^{(0)}(1), \dots, \epsilon^{(0)}(n)$  is the error sequence of  $X^{(0)}$ , where  $\epsilon^{(0)}(k)=x^{(0)}(k) - \hat{x}^{(0)}(k)$ . These errors can be expressed using Fourier series (Tan and Chang (1996)) as in Eq. (??). And, expected errors  $\hat{\epsilon}^{(0)}(k)$  can be added to  $\hat{x}^{(0)}(k)$  (i.e., Grey system model predictions) for all  $k$  in order to error correct the predicted values of the signal. For the details of the error correction, readers are referred to Liu et al. (2010); Kayacan et al. (2010); Bezuglov and Comert (2016).

$$\hat{x}_f^{(0)}(k) = \hat{x}^{(0)}(k) + \epsilon^{(0)}(k), k = 2, 3, \dots, n \quad (33)$$

For simplicity in numerical examples, short notation for Grey models are given in Table 1.

Table 1: Short notation for Grey models

Model	Short	Error Corrected Model
GM(1,1)	GM(1,1)	EFGM
GVM	GVM	EFGVM
GM(1,1 sin(ωt))	GM_S	EFGMLS
GM(1,1 cos(ωt))	GM_C	EFGM_C
GM(1,1 sin(ωt), cos(ωt))	GM_SC	EFGM_SC
GM(1,1 e <sup>-t</sup> , sin(ωt), cos(ωt))	GM_ESC	EFGM_ESC

### 3.8. Time series benchmark models

In addition to the basic Grey system models, linear and nonlinear time series models are used for comparison purposes. These include autoregressive (Linear) (Eq.(34)), ARIMA (Eq.(36)), SARIMA (Eq.(37)), self-exciting threshold autoregressive model (SETAR) (Eq.(38)) and additive nonlinear autoregressive model (AAR) (Eq.(39)) which is adopted from Di Narzo et al. (2015). These models were also compared in Bezuglov and Comert (2016) and Comert et al. (2016).

For speed forecasts using time series models, AR( $\infty$ ) representation in Eq. (35) is adopted which provides recursive calculation of coefficients that weighs past observations. For linear models, examples of fitted models for Eqs. (34)-(37) are given in Table 3.

$$Z_{t+1} = \phi + \phi_0 Z_t + \phi_1 Z_{t-\delta} + \dots + \phi_m Z_{t-(m-1)\delta} + \epsilon_{t+1} \quad (34)$$

$$\begin{aligned} \psi(B)(Z_t) &= a_t \\ \psi(B) &= \frac{\Phi(B)\phi(B)(1-B)^d(1-B)^D}{\theta(B)} \end{aligned} \quad (35)$$

Using Eq. (35), one can obtain the forecasting formula for ARIMA(1,1,2) as follows.

$$\begin{aligned} (1 - \phi B)(1 - B)(Z_t - \mu) &= (1 - \theta_1 B - \theta_2 B^2)a_t \\ \hat{Z}_{t+1} &= Z_t \sum_{i=1}^{\infty} B\psi_i \end{aligned} \quad (36)$$

where  $\psi_0 = 1$ ,  $\psi_1 = 1 + \phi - \theta_1$ ,  $\psi_2 = \psi_1\theta_1 - \phi - \theta_2$ , and  $\psi_i = \psi_{i-1}\theta_1 + \theta_2$  for  $i > 2$ . Similarly, SARIMA(1,0,3)(1,0,0) forecasts are calculated from Eq. (37). ARIMA and SARIMA model parameters (e.g.,  $\phi$  and  $\theta$ s) are trained per dataset and traffic parameter and kept constant otherwise.

$$\begin{aligned} (1 - \phi B)(1 - \Phi B^{18})(Z_t - \mu) &= (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)a_t \\ \hat{Z}_{t+1} &= (1 - \sum_{i=1}^{\infty} \psi_i \mu) + Z_t \sum_{i=1}^{\infty} B\psi_i \end{aligned} \quad (37)$$

where  $\psi_0 = 1$ ,  $\psi_1 = \phi - \theta_1$ ,  $\psi_2 = \psi_1\theta_1 - \theta_2$ ,  $\psi_3 = \psi_2\theta_1 + \psi_2\theta_2 - \theta_3$ , and  $\psi_i = \psi_{i-1}\theta_1 + \psi_{i-2}\theta_2 + \psi_{i-3}\theta_3$  for  $i > 3$ ,  $\psi_i = \psi_{i-1}\theta_1 + \psi_{i-2}\theta_2 + \psi_{i-3}\theta_3 + \Phi$ , for  $i = 18$ ,  $\psi_i = \psi_{i-1}\theta_1 + \psi_{i-2}\theta_2 + \psi_{i-3}\theta_3 - \phi\Phi$ , for  $i = 19$ .

$$Z_{t+1} = \begin{cases} \phi_1 + \phi_{10}Z_t + \phi_{11}Z_{t-\delta} + \dots + \phi_{1L}Z_{t-(L-1)\delta} + \epsilon_{t+1}, & X_t \leq th \\ \phi_2 + \phi_{20}Z_t + \phi_{21}Z_{t-\delta} + \dots + \phi_{2H}Z_{t-(H-1)\delta} + \epsilon_{t+1}, & X_t > th \end{cases} \quad (38)$$

where  $Z_t$  denotes observed traffic parameters at time  $t$ ,  $L = 1$  to 5 and  $H = 1$  to 5 low and high regimes,  $X_t$  is threshold function (the transition variable),  $\delta$  delay of the transition variable,  $th$  is the threshold value.

$$Z_{t+1} = \mu + \sum_{j=0}^{m-1} s_j(Z_{t-(j)\delta}) \quad (39)$$

In Eq. 39,  $s$  represents nonparametric univariate smoothing function that depends on  $Z_{ts}$ ,  $\delta$  is the delay parameter,  $m$  denotes embedding dimension,  $D$  is hidden layer of the neural network, and  $\beta_i, \gamma_{0j}, \gamma_{ij}$  represent the weights. Splines from Gaussian family are fitted in the form of  $Z_{t+1} \sim \sum_{i=0}^{m-1} s(Z_t, \dots, Z_{t-j})$ .

#### 4. Data Description

To assess the performance of the proposed Grey system models to predict various traffic parameters (i.e., speed, travel time, volume, and occupancy) relative to the benchmark models, four different datasets obtained from two different types of collection technologies are used: loop detectors and GPS-based probes. Datasets contain traffic parameters under both normal (without any incidents) and abnormal (with incidents) conditions. Summary of these dataset are given in Table 2.

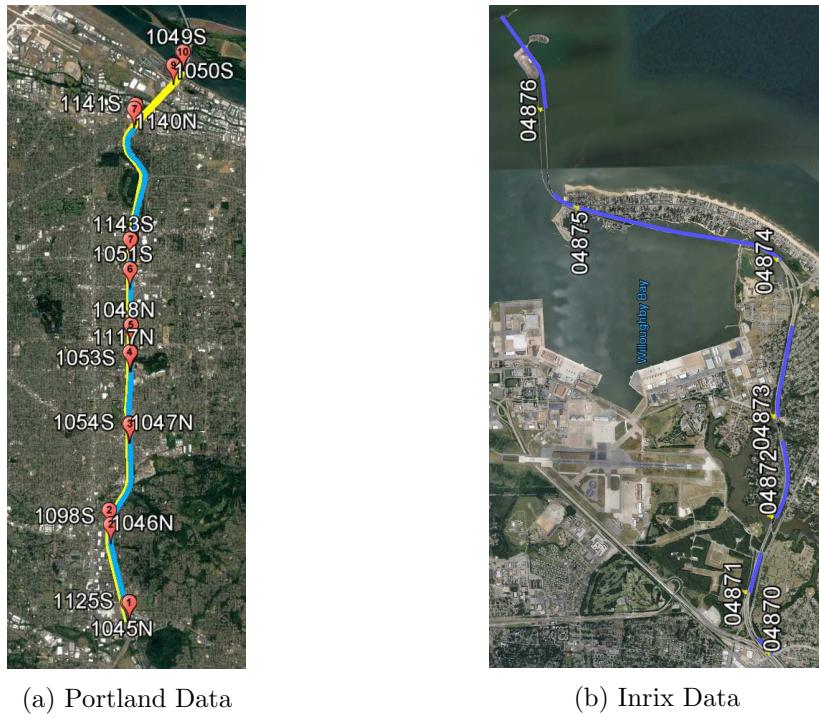


Fig. 1: Data Locations

The first dataset consists of 1-minute average traffic speeds (S), volume (V), and occupancy (O) collected from loop detectors by the California PATH (Partners for Advanced Transit and Highways) program on I-880 for the Freeway Service Patrol Project. The dataset is compiled from 12 locations for 24 days in February and March of 1993 between 5:01 AM to 10:00 AM and from 2:01 PM to 7:59 PM.

Table 2: Traffic speed, travel time, and volume datasets

ID	Date	obs./day	# of days	Parameter (Agg)
I-880 Loop Data	Feb16-Mar19	560	23	S, V, O (1-min)
Inrix Probe Vehicle Data	Jan01-Jan31	1440	18	S, TT (1-min)
Portland Loop Data	Sep15-Nov15	288	31	S, TT, V (5-min)
Oregon Loop Data	Sep29-Nov04	288	6	S (5-min)

The second dataset consists of 20-second average traffic speeds and travel times (TT) collected from GPS-enabled probe vehicles by INRIX for various freeway segments in Norfolk and Hampton, Virginia, where the longest segment is 1.543 miles. This dataset is compiled from 18 weekdays from January of 2014. The data is aggregated to 1-min between 12:00 AM and 12:00 AM (24-hour). Long repeated sections of data are found in this dataset due to lack of probe vehicles especially overnight hours. In order to improve, they were added very low white noise (i.e.,  $N(0, 0.01^2)$ ) in order to differentiate the missing data.

The third dataset consists of two months of 20-second average data (September 15, 2011 through November 15, 2011) collected from dual-loop detectors on I-205, where the Northbound segment is 10.09 miles long and the Southbound segment is 12.01 miles long. The collected data include flow, occupancy, and speed as well as calculated travel times. This dataset is compiled from 11 different locations (see Fig. 1a)

The fourth dataset has the same traffic parameters as the Portland dataset collected using loop detectors for various locations in the state of Oregon (e.g., I-5, I-405, and I-84) between September 15, 2011 and November 15, 2011. From this dataset, a subset of 6 days between September 29, 2011 and November 4, 2011 was extracted. This particular dataset contains abnormal fluctuations that were intended to challenge the prediction models.

Table 3: Parameters for time series and Grey system models

Model	Parameters
SETAR(1,1,1)	$\phi_1 = 1.215 \phi_{10} = 0.302 \phi_{11} = 0.337 \phi_{12} = 0.221$
	$\phi_2 = 2.905 \phi_{20} = 0.748 \phi_{21} = -0.053 \phi_{22} = 0.196$
	$X_t = Z_t, \text{th}=12.29$
	low regime: 69.6 % % High regime: 30.4%
LINEAR(3)	$\phi = 0.346 \phi_1 = 0.637 \phi_2 = 0.146 \phi_3 = 0.193$
ARIMA(1,1,2)	$\phi_1 = -0.749 \theta_1 = -0.363 \theta_2 = 0.402$
SARIMA(1,0,3)(1,0,0)	$\phi_1 = 0.990 \theta_1 = 0.377 \theta_2 = 0.121 \theta_3 = -0.047$ $\Phi_1 = -0.071 \mu = 11.419$
GM_S	$\omega = 4.30$
EFGM_S	$\omega = 4.30$
GM_C	$\omega = 2.65$
EFGM_C	$\omega = 2.65$
GM_SC	$\omega = 9.30$
EFGM_SC	$\omega = 9.30$
GM_ESC	$\omega = 74.10$
EFGM_ESC	$\omega = 74.10$

The benchmark models were trained using the first series of every dataset (i.e., I-880, Inrix, Portland, and Oregon) and every traffic parameter (i.e., speed, travel time, volume, and occupancy). Models were chosen based on the detailed analysis in Bezuglov and Comert (2016); Comert et al. (2016). Examples of fitted models using February 16, 1993 I-880 loop occupancy data are presented in Table 3. For the Grey system models, the goal is to identify best  $\omega$  value for each GM model. The  $\omega$  values shown in Table 3 are optimized only once by grid search on February 16, 1993 I-880 loop speed data and used for other datasets and traffic parameters. The remaining series of each dataset were used to evaluate the models' performance.

## 5. Model Performance and Assessment

Figs. 2-3 show the prediction performance of the proposed Grey system models compared to actual data under an abnormal condition. To evaluate the performance of the proposed

models against the benchmark models, two types of metrics are used: root mean squared error (RMSE)  $RMSE_i = \sqrt{\frac{\sum_{k=1}^N (\hat{x}_i(k) - x_i(k))^2}{N}}$  and mean average percentage error (MAPE)  $MAPE_i = \frac{1}{N} \sum_{k=1}^N \left| \frac{\hat{x}_i(k) - x_i(k)}{x_i(k)} \right| 100\%$ . These metrics are calculated as follows. where  $\hat{x}_i(k)$  is the predicted and  $x_i(k)$  is the observed value at time  $k$  for time series  $i$ .

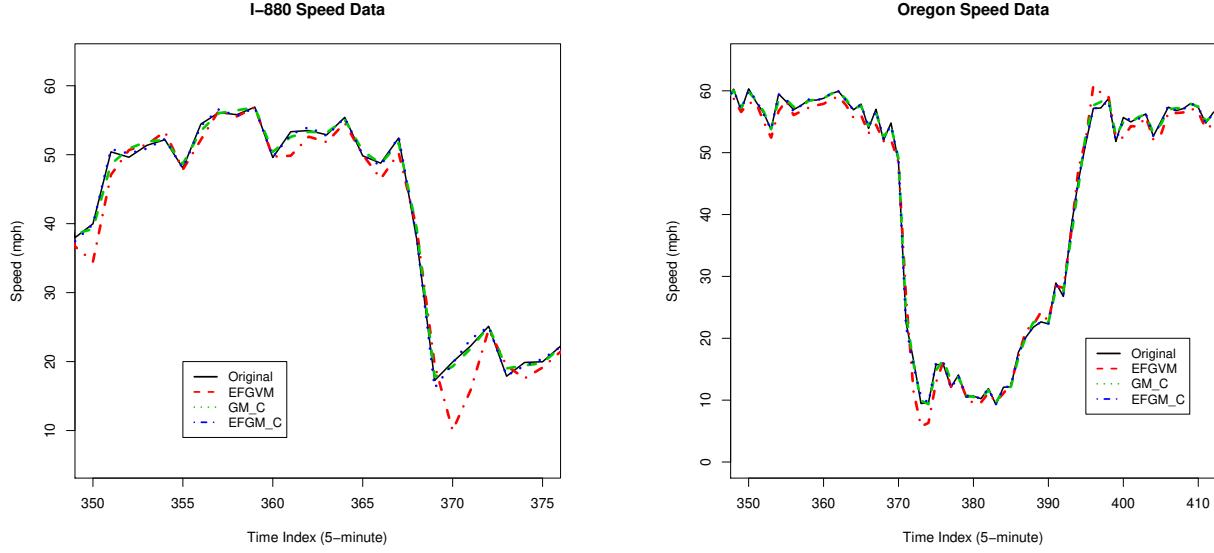


Fig. 2: 1-step speed predictions

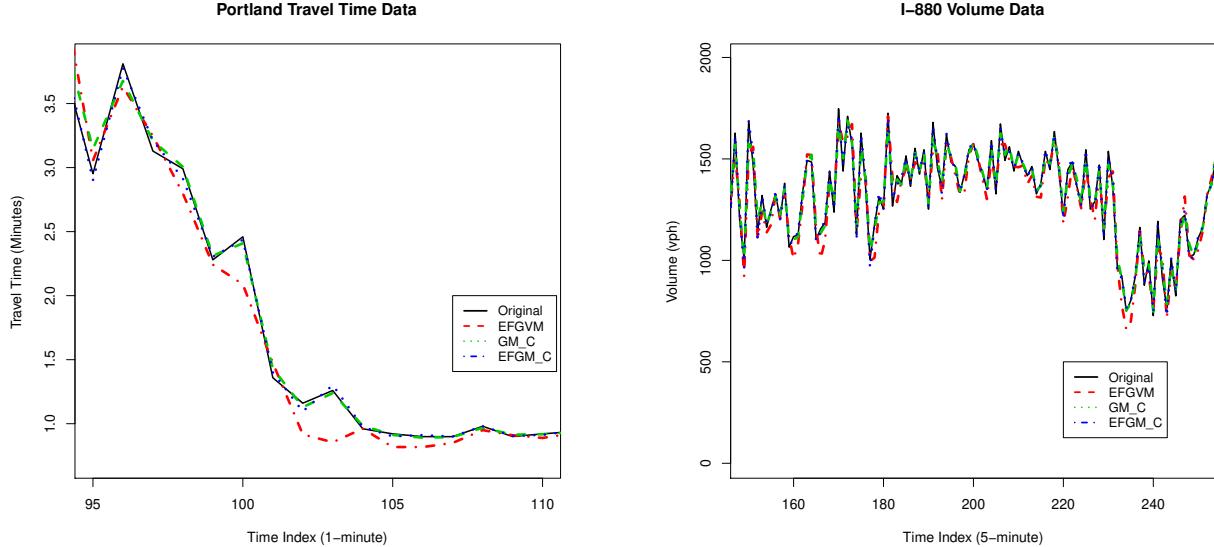


Fig. 3: 1-step travel time predictions

Figs. 2-3 illustrate 1-step predictions on loop speed and Inrix travel times. Fig. 2 demonstrates the effectiveness of the proposed GM models when the speed quickly increases (time step 220) and decreases (time step 230). The forecasts of best performing GM\_C, EFGM\_C, and EFGVM are shown along with the original values. At speed increases, EFGM overestimates the next value, whereas GM\_C and EFGM\_C provide more conservative estimates, and hence, more accurate forecasts. Since EFGM\_C model is specifically developed to deal with nonlinear sequences, it works better with the traffic speed data which is bound by the speed limits. Similarly, these models are effective in predicting abrupt changes in travel times as shown in Fig. 3.

Table 4: The accuracy of the models

Method	Portland S		Portland V		Portland TT		Inrix S		Inrix TT		Oregon S		I-880 S		I-880 V		I-880 O		CT
	RM	MA	RM	MA	RM	MA	RM	MA	RM	MA	RM	MA	RM	MA	RM	MA	RM	(s)	
GM(1,1)	2.01	2.39	4.30	8.05	0.11	8.41	1.98	1.74	0.10	1.68	1.87	2.99	2.14	3.15	112.80	6.93	1.06	8.10	0.25
EFGM	1.86	2.44	3.76	8.28	0.10	8.42	1.51	1.37	0.07	1.31	1.73	2.82	1.91	3.08	106.27	7.14	1.02	8.29	0.61
GVM	3.27	2.33	3.77	5.46	0.08	5.35	3.02	2.26	0.10	2.18	2.89	2.40	3.12	2.61	87.99	4.41	0.80	5.16	0.25
EFGVM	1.60	2.33	2.65	5.46	0.07	5.35	1.45	2.26	0.06	2.18	1.41	2.40	1.56	2.61	68.48	4.41	0.65	5.16	0.62
GM_S	1.53	1.71	3.39	5.75	0.08	5.70	1.62	1.50	0.08	1.41	1.45	2.20	1.61	2.24	72.52	4.60	0.77	5.53	0.27
EFGM_S	0.91	1.13	1.82	4.16	0.05	3.87	0.78	0.66	0.04	0.61	0.86	1.30	0.95	1.48	45.00	3.04	0.45	3.62	0.62
GM_C	0.65	0.71	1.35	2.47	0.04	2.59	0.57	0.41	0.03	0.40	0.49	0.79	0.58	0.86	33.81	2.18	0.34	2.54	0.28
EFGM_C	0.31	0.37	0.68	1.34	0.02	1.37	0.25	0.18	0.01	0.19	0.24	0.41	0.26	0.41	17.24	1.14	0.17	1.34	0.65
GM_SC	1.37	1.77	2.89	6.22	0.09	6.52	0.97	0.89	0.05	0.92	1.19	2.00	1.19	1.98	84.26	5.61	0.79	6.46	0.22
EFGM_SC	2.00	2.67	4.25	9.00	0.12	9.68	1.50	1.39	0.07	1.41	1.85	3.10	1.91	3.13	128.26	8.48	1.19	9.69	0.62
GM_ESC	1.36	1.84	2.77	5.93	0.08	6.25	1.13	1.32	0.05	1.17	1.26	2.08	1.39	2.29	94.91	5.50	0.76	6.18	0.21
EFGM_ESC	1.98	2.75	3.75	8.03	0.10	8.43	1.66	2.28	0.07	2.08	1.82	3.02	2.00	3.37	124.90	7.43	1.02	8.35	0.73
LINEAR	3.60	6.04	6.70	16.05	0.26	23.17	6.61	11.47	0.14	2.43	3.22	5.56	3.20	5.25	168.55	11.58	1.74	14.19	0.22
SETAR	3.40	5.20	6.81	15.90	0.28	24.04	7.39	12.97	0.14	2.49	3.26	5.68	3.20	5.22	170.82	11.65	1.77	14.22	0.23
AAR	4.20	6.48	6.88	16.07	0.27	24.16	3.89	5.58	0.16	2.63	3.39	6.03	3.43	5.35	169.06	11.57	1.73	13.76	2.56
SARIMA	3.52	5.62	6.89	15.01	0.26	23.12	6.74	11.72	0.14	2.44	3.21	5.56	3.19	5.24	160.29	10.76	1.72	13.90	0.56
ARIMA	3.56	5.43	7.00	14.89	0.19	14.61	2.98	3.35	0.13	2.18	3.20	5.15	3.20	5.20	160.51	10.85	1.73	13.98	0.01
% Imp	59%	69%	49%	55%	42%	52%	61%	82%	51%	82%	65%	67%	63%	67%	51%	51%	48%	51%	

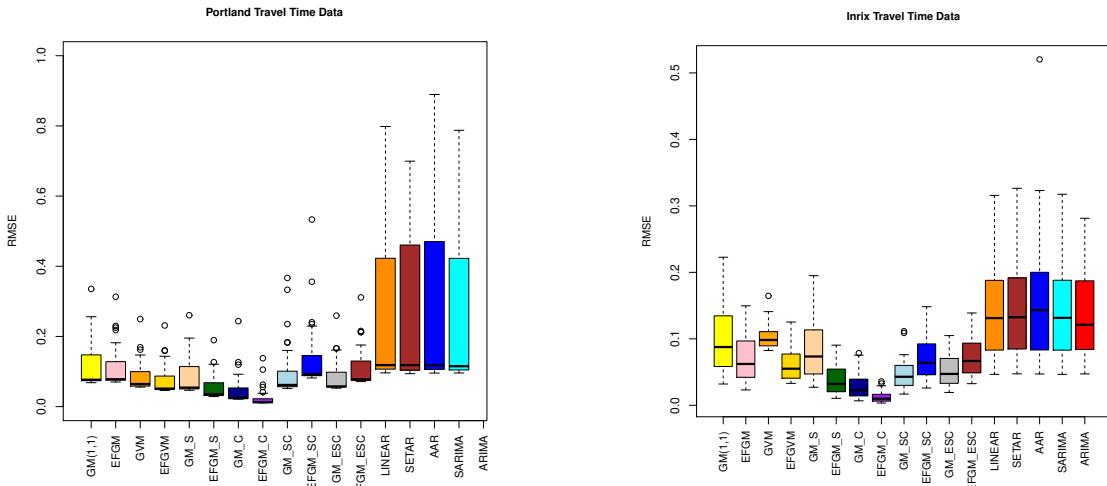


Fig. 4: RMSEs for 1-step travel time predictions

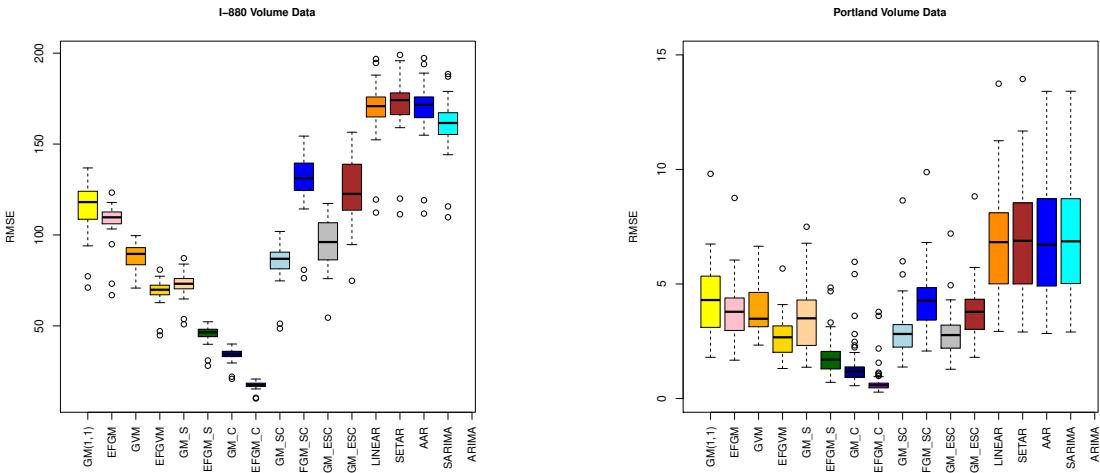


Fig. 5: RMSEs for 1-step volume predictions

The average RMSE and MAPE for the predicted speed (S), travel time (TT), volume (V),

and occupancy (O) for the different models are summarized in the Table 4. It can be seen that the Grey models outperformed the time series models using much less training data. Among the GM models, the error corrected by Fourier series GM( $1,1|\cos(\omega t)$ ) model (denoted by EFGM\_C) performed best. The results indicate that the error corrected Grey Verhulst model (EFGVM) is a better choice when data contains high noise and higher nonlinearity. However, when the noise level is very low and the system is not subject to drastic changes, the model without error correction GM( $1,1|\cos(\omega t)$ ) (GM\_C) can also be suitable. The last row of the table shows percent improvement of simpler trigonometric GM\_C model over best performing nontrigonometric EFGVM model. GM\_C model is able to surpass EFGVM by 59% in RMSE on Portland speed data and 82% in MAPE on Inrix speed and travel time data. The computation time needed to run the model on the INRIX dataset is provided in the last column (CT). These times are obtained on a PC with a Pentium i5 Quad-Core CPU and 8 GB of memory. The proposed GMs have higher computation times and they can be improved via metaheuristic.

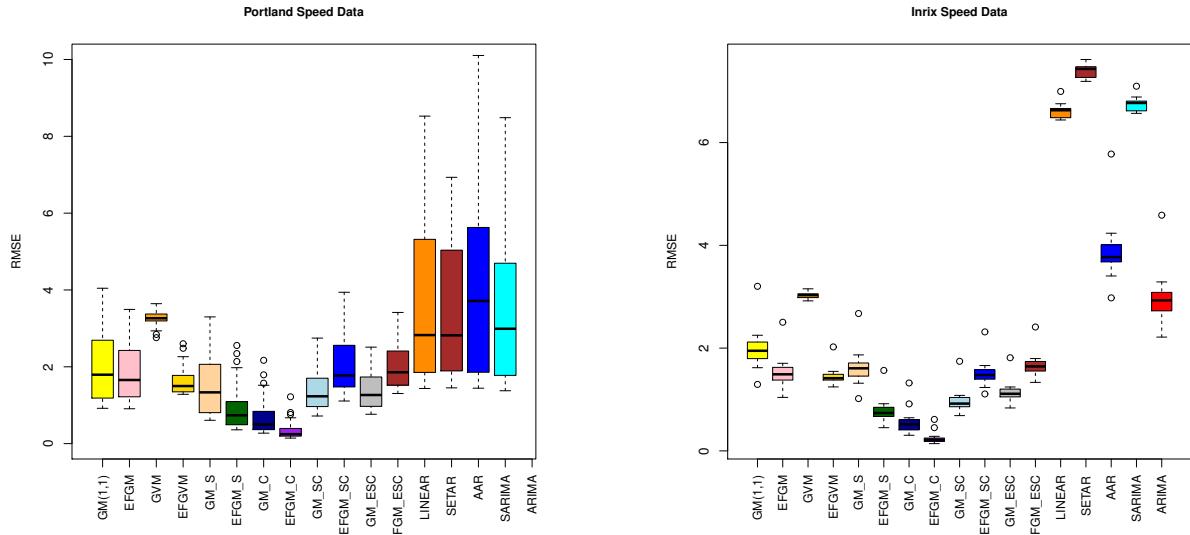


Fig. 6: RMSEs for 1-step speed predictions

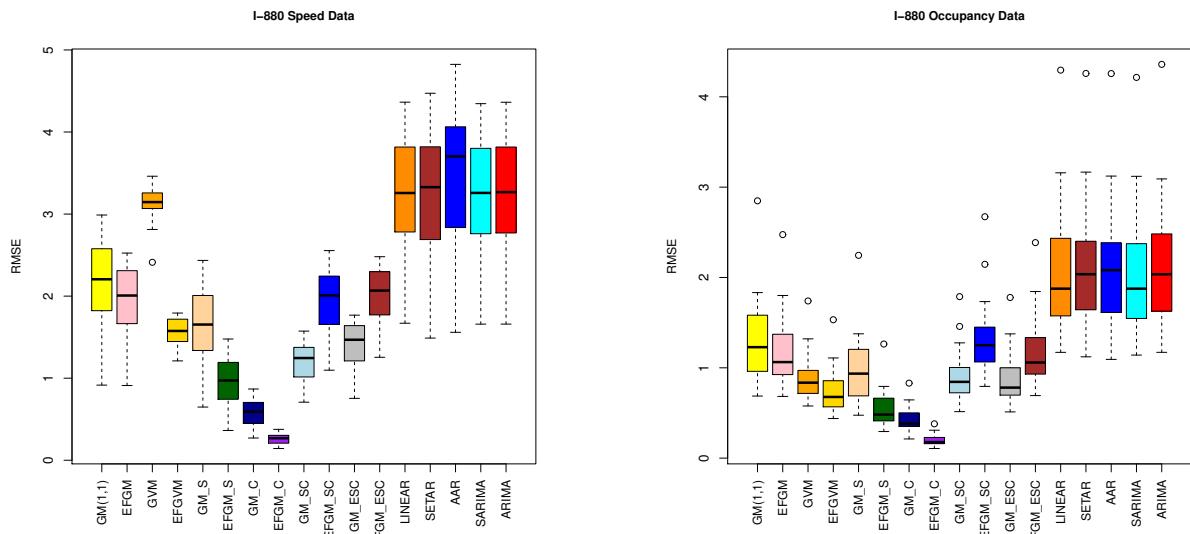


Fig. 7: RMSEs for 1-step speed and occupancy predictions

The proposed Grey models are compared to each other to determine which one would produce

better results more consistently. Figs. 4-7 show RMSEs for each time series and for each model. As shown in the figure, RMSE values are below 3 (mph) for the GM(1,1) model and error corrected GM(1,1) (EFGM); and below 2 (mph) for EFGVM. EFGVM model yielded more accurate predictions for all datasets among models that do not contain trigonometric terms. Further examination shows that EFGM\_C follows the true values closest (see Figs. 2-3), especially when there is significant sudden change. The model errors are shown graphically via Box plots in Figs. 4-7.

## 6. Conclusions

This paper introduced new Grey system models to forecast traffic parameters and compared them to other benchmark models. Overall, the proposed Grey models yielded more accurate forecasts. The GM(1,1| $\cos(\omega t)$ ) model outperformed other Grey models when the system has abrupt changes. The key advantages of using the proposed Grey system models are: (1) simple expert system with well-defined mathematics and straightforward adaptation to include multidimensional data such as weather and other location specific observations, and (2) simple implementation that would allow the model to be implemented and deployed on portable electronic devices such as smart phones and tablets (Vlahogianni et al. (2014)).

Future work in this area could include exploration of whitenization functions, long-term predictions, and the use of multivariable Grey models. These models could be further extended for use on interrupted facilities (i.e., highways instead of freeways).

## Acknowledgments

This study is supported by the Center for Connected Multimodal Mobility ( $C^2M^2$ ) (USDOT Tier 1 University Transportation Center) Grant headquartered at Clemson University, Clemson, South Carolina, USA. The authors would also like to acknowledge U.S. Department of Homeland Security (DHS) Summer Research Team Program Follow-On, and National Science Foundation (NSF, No. 1719501) grants. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of ( $C^2M^2$ ), USDOT, DHS, or NSF and the U.S. Government assumes no liability for the contents or use thereof.

## References

An, Y., Cui, H., Zhao, X., 2012. Exploring gray system in traffic volume prediction based on tall database. *International Journal of Digital Content Technology & its Applications* 6 (5).

Badhrudeen, M., Raj, J., Vanajakshi, L. D., 2016. Short-term prediction of traffic parameters performance comparison of a data-driven and less-data-required approaches. *Journal of Advanced Transportation* 50 (4), 647–666.

Bezuglov, A., Comert, G., 2016. Short-term freeway traffic parameter prediction: Application of grey system theory models. *Expert Systems with Applications* 62, 284–292.

Clements, M. P., Harvey, D. I., 2010. Forecast encompassing tests and probability forecasts. *Journal of Applied Econometrics* 25 (6), 1028–1062.

Comert, G., Bezuglov, A., Cetin, M., 2016. Adaptive traffic parameter prediction: Effect of number of states and transferability of models. *Transportation Research Part C: Emerging Technologies* 72, 202–224.

Deng, J. L., Nov. 1989. Introduction to grey system theory. *J. Grey Syst.* 1 (1), 1–24.  
 URL <http://dl.acm.org/citation.cfm?id=90757.90758>

Di Narzo, A. F., Aznarte, J. L., Stigler, M., 2015. R package:tsdyn manual.

Hamilton, J. D., 2008. Regime-Switching Models, 2nd Edition. Palgrave McMillan Ltd.

Hamilton, J. D., 2010. Calling recessions in real time. *International Journal of Forecasting* 27 (4), 1006–1026.

Jiang, Y., Li, S., 2005. Gray system model for estimating the pavement international roughness index. *Journal of performance of constructed facilities* 19 (1), 62–68.

Jie, C., Hongyan, M., Chaoqing, Y., Sifeng, L., Peng, L., Xiaodong, Z., 2015. Novel grey verhulst model and its prediction accuracy. *Journal of Grey System* 27 (2), 47 – 53.

Kayacan, E., Ulutas, B., Kaynak, O., 2010. Grey system theory-based models in time series prediction. *Expert systems with applications* 37 (2), 1784–1789.

Liu, S., Lin, Y., 2006. Grey information: theory and practical applications. Springer Science & Business Media.

Liu, S., Lin, Y., Forrest, J. Y. L., 2010. Grey systems: theory and applications. Vol. 68. Springer Science & Business Media.

Liu, W., Qin, Y., Dong, H., Yang, Y., Tian, Z., 2014. Highway passenger traffic volume prediction of cubic exponential smoothing model based on grey system theory. In: 2nd International Conference on Soft Computing in Information Communication Technology. Atlantis Press.

Na, Z., Shuangwei, W., Jianfeng, X., Chaoyang, Y., Xiaoyan, W., 2010. Grey synthetic degree of incidence entropy and its application in traffic accidents analysis. *Journal of Grey System* 22 (3), 257–264.

Pan, Y., Shi, Y., 2016. Short-term traffic forecasting based on grey neural network with particle swarm optimization. In: Proceedings of the world congress on engineering and computer science. Vol. 2.

Shuhua, M., Xinping, X., 2010. Short-term traffic flow grey forecasting model gm (1, 1 vertical bar tan (k-tau) p, sin (k-tau) p) of single-cross-section and its particle swarm optimization. *JOURNAL OF GREY SYSTEM* 22 (4), 383–394.

Shuhua, M., Ye, C., Xinping, X., 2012. City traffic flow prediction based on improved gm (1, 1) model. *Journal of Grey System* 24 (4).

Tan, C., Chang, S., 1996. Residual correction method of fourier series to gm (1, 1) model. In: Proceedings of the first national conference on grey theory and applications, Kauhsing, Taiwan. pp. 93–101.

Vlahogianni, E. I., Karlaftis, M. G., Golias, J. C., 2014. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies* 43, 3–19.

Yu, Z., Sun, T., Sun, H., Yang, F., 2015. Research on combinational forecast models for the traffic flow. *Mathematical Problems in Engineering* 2015.

Zhang, Y., 2010. Predicting model of traffic volume based on grey-markov. *Modern Applied Science* 4 (3), p46.