JOURNAL OF ADVANCED TRANSPORTATION
J. Adv. Transp. 2016; **50**:647–666
Published online 14 March 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/atr.1368

Short-term prediction of traffic parameters—performance comparison of a data-driven and less-data-required approaches

Mohamed Badhrudeen, Jithin Raj and Lelitha Devi Vanajakshi*

Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India

SUMMARY

The travel decisions made by road users are more affected by the traffic conditions when they travel than the current conditions. Thus, accurate prediction of traffic parameters for giving reliable information about the future state of traffic conditions is very important. Mainly, this is an essential component of many advanced traveller information systems coming under the intelligent transportation systems umbrella. In India, the automated traffic data collection is in the beginning stage, with many of the cities still struggling with database generation and processing, and hence, a less-data-demanding approach will be attractive for such applications, if it is not going to reduce the prediction accuracy to a great extent. The present study explores this area and tries to answer this question using automated data collected from field. A data-driven technique, namely, artificial neural networks (ANN), which is shown to be a good tool for prediction problems, is taken as an example for data-driven approach. Grey model, GM(1,1), which is also reported as a good prediction tool, is selected as the less-data-demanding approach. Volume, classified volume, average speed and classified speed at a particular location were selected for the prediction. The results showed comparable performance by both the methods. However, ANN required around seven times data compared with GM for comparable performance. Thus, considering the comparatively lesser input requirement of GM, it can be considered over ANN in situations where the historic database is limited. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS: flow; highway and traffic engineering; intelligent transport systems; speed

1. INTRODUCTION

In order to make road travel convenient to users, it is essential to accurately predict necessary traffic variables that can help real-time applications such as advanced traveller information system, one of the functional areas of intelligent transportation systems. This information can help travellers take more informed decisions pre-trip or en-route. These decisions can affect the time of travel, route to be taken, mode of travel or a combination of the three, which may result in relieving congestion in the roadways and make road travel pleasant to the road user.

Various prediction methods that were reported as promising for the prediction of traffic parameters include historic averaging, regression analysis, Kalman filtering technique, time series analysis and machine learning techniques. However, in countries like India, where automated sensors and associated database are being implemented, such short-term prediction methods have not been developed or tested rigorously. Choosing the best prediction strategy, with several constraints in terms of database generation and storage, is one challenging task. Choice between data-driven and less-data-demanding approaches is one such question. With such limited data availability, a less-data-demanding approach would be more attractive, provided such an approach would not show a negative effect on prediction accuracy.

*Correspondence to: Lelitha Devi Vanajakshi, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India. E-mail: lelitha@iitm.ac.in

The present study explores this area and tries to answer this question using automated data collected from field. The main focus of this paper is to compare the short-term prediction performance of two different prediction tools that are extremely different in terms of data requirement. Artificial neural network (ANN), which is shown to be a good tool for prediction problems, is chosen for the data-driven approach. Grey model (GM), which is also reported as a good prediction tool, is selected for the less-data-demanding approach. The specific tools, namely, ANN and GM, were chosen based on various studies that reported a wide range of successful application of these approaches in various fields. Volume, classified volume, average speed and classified speed at a particular location were selected for prediction.

Artificial neural network is a machine learning technique inspired by biological nervous systems and is shown to be a good tool for prediction. It is composed of several processing units operating in parallel, making it a good prediction tool when the system under consideration is highly nonlinear. Each of these processing units is known as neurons. Neural networks are trained to adjust the weights of these neurons so that a particular input leads to a specific target. The Grey system theory, initiated by Julong Deng in the 80s [1], is a new methodology that focuses on the study of problems involving small samples and poor information [2]. It deals with uncertain systems with partially known information. In this study, short-term prediction of parameters using these two techniques is carried out and the performance is compared.

2. LITERATURE REVIEW

Several research studies have been reported on short-term traffic parameter prediction. In general, studies in this area can be broadly classified into two main classes, namely, (i) data-driven and (ii) less-data-demanding approaches. The data-driven approaches include methods like regression, time series and ANN, which try to capture the relationship between the available data and the corresponding output [3–5]. Some of the important studies coming under this approach are discussed first.

Regression analysis was attempted by many researchers to predict traffic parameters in the near future [6–10]. The time series analysis involves the examination of historical data, extracting essential data characteristics and effectively projecting these characteristics into the future [11–13]. ANN is one of the machine learning techniques that has been extensively used for the prediction of traffic parameters. Kisgyorgi and Rilett [14] used advanced neural network methods called modular neural networks to predict travel time from loop detector and probe vehicle data. Van Lint *et al.* [3] developed a State space neural network model for freeway travel time prediction under missing data conditions. Liu *et al.* [4] proposed an approach in which the optimal weight parameters of ANN were found using extended Kalman filter technique for urban arterial travel time prediction. Lam and Xu [15] developed an ANN-based traffic flow model for the estimation of average annual daily traffic, from short-period counts. A dynamic neural network model was proposed by Shen [16] for freeway travel time estimation. Yu *et al.* [17] proposed a variation-based online travel time prediction approach using clustered neural networks. Zou *et al.* [18] developed a multi-topology ANN model with a supplemental component of an enhanced k-nearest neighbour (k-NN) model for travel time prediction when the detectors were widely spaced. The purpose of the supplemental k-NN model was to detect potential errors and missing data. Other studies that used ANN for travel time prediction include Liu *et al.* [19], Van Lint [20], Zhu and Wang [21], and Lee [22].

Ozkurt *et al.* [23] presented a traffic density estimation and vehicle classification method using neural network. Cetiner *et al.* [24] reported prediction of traffic flow from historical data using ANN and showed a correlation coefficient in the range of 0.85 to 0.95. Mazloumi *et al.* [25] and Chen *et al.* [26] proposed an integrated framework to predict bus travel time using ANN. Yaghini *et al.* [27] used ANN to predict passenger train delay. The study compared the results with prediction methods like decision tree and multinomial logistic regression. Balogee *et al.* [28] studied the driver's response to en-route traffic information provided through radio. They used conventional discrete models for driver's response and evaluated the accuracy of this model using neural network model.

The less-data-demanding approaches include the use of filtering techniques and GM. A review of literature using these techniques is provided as follows. Kalman filter method was employed by Chen

DOI: 10.1002/atr

and Chein [29] to predict freeway travel time using data obtained from probe vehicles. Chein and Kuchipudi [30] used Kalman filter to predict travel time using real-time data and historic data. Gong and Zhang [31] proposed a combination of Kalman filter and Grey relational entropy methods to forecast traffic volume using the historic database. Okutani and Stephanedes [32] and Wang and Papageorgiou [33] used Kalman filter and extended Kalman filter techniques respectively to predict traffic volume. Under heterogeneous conditions, Vanajakshi *et al.* [34] employed Kalman filter to predict travel time. Zhang *et al.* [35] used hierarchical fuzzy rule system optimized by Genetic Algorithms (GAs) to develop an accurate and robust traffic congestion prediction system. Soriguera and Robuste [36] combined data from different sources to arrive at a reliable short-term prediction of highway travel time based on fuzzy logic and a probabilistic approach. Yin *et al.* [37] used fuzzy logic to classify input data and neural network to specify input–output relationships, to predict urban traffic flow. Sharma *et al.* [38] used fuzzy logic model under heterogeneous conditions to predict traffic volume and suggested to have extensive data to build a high-quality model. Yu *et al.* [39] proposed a hybrid model, based on support vector machine (SVM) and Kalman filter technique, to predict bus arrival times using historical data and the latest bus arrival information. Yu and Lam [40] proposed a model to predict traffic flow using multiple kernel learning and SVM methods. Comparison between SVM and ANN was carried out by Vanajakshi and Rilett [41], and concluded that ANN can be replaced by SVM if the training data available are less. Online-SVM regression was used by Neto *et al.* [42] for predicting traffic flow in both typical or normal and atypical conditions. Zhang and Liu [43] used least squares-SVM to predict travel time index and compared the results with other methods like Kalman filter and Radial Basis Function (RBF) neural network.

Grey model is reported as a promising prediction tool in several other research areas like agriculture, economy, hydrology and finance [44–52]. However, reported applications of GM in the field of traffic engineering are limited and are reviewed here. Sun *et al.* [53] proposed a new self-adapting GM to predict traffic flow at non-detector sections. GM combined with three-point average technique to predict vehicle fatality risk was reported to be highly efficient by Mao and Chirwa [54]. Xu *et al.* [55] carried out high-road traffic safety evaluations based on neural network and Grey systems theory. GM was used to predict traffic volume in [56, 57] and road traffic accident in [58]. He *et al.* [59] combined Grey method and regression method to forecast regional logistics demand and used Genetic algorithm to improve the accuracy of the model. Zhao *et al.* [60] improved GM forecasting by synthesizing GM with Back Propagation (BP) neural network.

Studies reporting the use of ANN for heterogeneous and less-lane-disciplined traffic are limited. Drakopoulos and Abdulkader [61] studied the neural network training for heterogeneous data and proposed data pruning (removal of noisy data) and ordered training (partitioning of data) as effective methods to deal with heterogeneous data. Padiath *et al.* [62] compared the performance of a historic technique, an ANN-based technique and a model-based technique in predicting traffic density under Indian traffic conditions. However, the study used limited amount of manually collected data. Use of GM under heterogeneous traffic condition was not reported. In addition, under Indian conditions, automated sensors were not available, and hence, prediction problems using approaches such as ANN and GM were not attempted exhaustively.

A summary of the above literature is shown in Table I. From Table I, it can be observed that majority of the studies, especially the ones that used machine learning techniques such as ANN, are from homogeneous traffic conditions. In addition, it can be noted that none of those studies analysed the effect of varying data size on the prediction process. Use of GM for such applications are very limited, and it can be observed that all of them are reported from homogeneous traffic conditions. It can also be noted that none of the studies reported a comparison of the performance of a data-driven and less-data-demanding techniques for the traffic variable prediction problems. In addition, although it was indicated that SVM performed better, compared with ANN, with relatively less data, the data used were in the order of 5 days to several months. However, GMs requirement is found to be 4–10 data points and is considered in this study as a good representation of less-data-demanding technique. Thus, although various tools have been reported for prediction of traffic variables, a systematic comparison of using less-data-demanding and data-driven techniques is not clearly reported. The present study carries out such a comparison of performance of a data-driven technique, ANN, and a less-data-demanding technique, GM, using automated data collected from field. The performance of these two selected techniques under heterogeneous traffic conditions for the prediction of volume, classified volume, average speed and

DOI: 10.1002/atr

Table I. Summary of literature in the area of traffic variable prediction.

Methods	References	Traffic conditions	Data used
Regression	Kwon *et al*. (2000), Zhang and Rice (2003), Rice and van Zwet (2004), Lee *et al.* (2006), Chang *et al.* (2010)	Homogeneous	20–34 days
Time series	Guin (2006), Xia (2006), Liu *et al.* (2014)	Homogeneous	2–6 months
Artificial neural networks	Kisgyorgi and Rilett (2002), Van Lint *et al.* (2005), Lam and Xu (2000), Shen (2008),Yu *et al.* (2008), Liu *et al.* (2007), Van Lint (2008), Zhu and Wang (2009), Lee (2009),Cetiner *et al.* (2010), Mazloumi *et al.* (2011)	Homogeneous	10 days–12 months
	Padiath *et al.* (2009)	Heterogeneous	5 days (1 hour each day)
Artificial neural networks combined with other models	Liu *et al.* (2006), Zou *et al.* (2008)	Homogeneous	3–6 months
Kalman filter	Chein and Kuchipudi (2003), Yi-shan and Yi (2013), Wang and Papageorgiou (2005), Okutani and Stephanedes (1984)	Homogeneous	7 days
	Padiath *et al.* (2009), Vanajakshi *et al.*(2009)	Heterogeneous	5–30 days
Fuzzy logic	Soriguera and Robuste (2011), Yin *et al.* (2002), Zhang *et al.* (2014)	Homogeneous	1 day–1 month
	Sharma *et al.* (2014)	Heterogeneous	
Support vector machine	Yu *et al.* (2014), Yu and Lam (2014), Neto *et al.* (2009), Zhang and Liu (2009), Vanajakshi and Rilett (2007)	Homogeneous	5 days–6 months
Grey model	Zhang (2010), Zhao *et al*, (2007), Guo *et al.* (2013)	Homogeneous	4–10 inputs

classified speed at a particular location was studied. The present study also includes a detailed analysis to find out the least number of data points required in ANN without compromising on the accuracy, and to find the number of data points required for similar accuracy in both the methods. It also aims to find out how the performance of GM would vary at different locations or under different road traffic conditions.

Thus, the present study, using the automated data available from sensors, would explore the possibility of using both these techniques for traffic variable prediction for intelligent transportation systems applications under Indian conditions. Another objective of the study is to compare the performance of these data-driven and less-data-demanding techniques to help user agencies make a suitable choice of prediction tool, based on the constraints in their data. ANN was selected as a sample data-driven technique based on its satisfactory performance reported in literature. Grey theory was selected as a technique with minimal data requirement because of its limited implementations in the field of transportation so far. Preliminary results of applying these two approaches for such a prediction were reported in [63].

It can be noted that there are several other prediction techniques, such as *k-NN, SVM,* dynamical systems approach and time series techniques, that can be applied to the same problem. However, the scope of the present study was limited to the comparison of performance of two sample techniques, namely,

DOI: 10.1002/atr

ANN and GM as representatives of data-driven and less-data-demanding techniques, based on their good performance in earlier studies. The contributions of this study can be detailed as follows.

This study is motivated to find tools or methods that could be used for prediction of traffic parameters, especially when the data available are limited. Thus, the present study is exploring the tools that can be used during the initial stages of such an implementation, when there is not enough historic information to build on. When the data available are limited, this study shows GM to be one of the possible methods that can be used. With more and more data added, tools like ANN would be a better option, as highlighted by the sensitivity analysis results of this study.

This is one of the first studies that compare the prediction performance of ANN and GM in the area of transportation engineering, paying special attention to the amount of data required for prediction. This study systematically analysed the variation in prediction accuracy with respect to the amount of data in both the cases. In ANN, the training data were varied from 6 to 24 hours and for GM, from 20 to 200 minutes. A comparison was carried out to show the data requirement of both these methods, to obtain equal prediction accuracy. This analysis would thus help in making a choice of the prediction method depending on data availability and accuracy preference.

As it can be seen from the literature review, there have not been many reported studies, which developed field implementable real-time solutions using automated sensor data under mixed traffic conditions based on advanced techniques such as GM and ANN. This is one of the first studies in this direction under Indian traffic conditions. The reported studies on ANN and GM for traffic variable prediction used the data collected in homogeneous and lane-disciplined traffic conditions. The heterogeneity and lack of lane discipline under Indian traffic conditions would add more complexity to the pattern, and hence there is no guarantee that the performance of ANN and GM would be the same as the reported studies. Furthermore, in countries like India, lack of historic and automated sensor data make conventional prediction tools unsuitable to use; the present study contributes here by helping organizations to choose a suitable prediction tool based on data availability and performance requirement. Thus, this is one of the first studies reporting the performance of ANN and GM for these types of applications using automated data under such high-variability traffic conditions.

3. DATA COLLECTION

The data required for the present study were collected using a location-based sensor [64] that uses infrared technology for the detection of vehicles. The infrared-based sensor has two units, Tx and Rx, which are the transmitter and the receiver, respectively. These units were placed at the wheelbase level across the road width. Transmitter transmits four Infra Red (IR) beams, two parallel beams and two crisscross beams. Whenever a vehicle passes through the beam, the wheels of the vehicle (not the body of the vehicle) cut through each of the beam. This causes a disturbance in the flow of the beam, and a movement is identified by the sensor. ‘Break beam event’ occurs when the front wheel cuts all four beams, and the times of these occurrences are noted. Similarly, ‘make beam event’ occurs when the wheel leaves the beams and the beam becomes whole and time of their occurrences is also recorded. Knowing the time taken by the front wheel to traverse the distance between the two parallel beams, the speed can be calculated. Using these, the sensor calculates the axle length of that vehicle and compares it with pre-defined vehicle classification to identify the type of vehicle that passed through. Thus, if two vehicles travel in parallel and the wheels of the second vehicle get hidden from both parallel and diagonal beams, the count or classification may go wrong. However, it is a rare situation.

The sensor was installed near Perungudi toll plaza, Rajiv Gandhi Salai, Chennai, India, facing the southbound traffic. The units were placed permanently on either side of a three-lane one-directional road of total width 10.4 meters. This road is a representative Indian road with heterogeneous and lane-less traffic conditions, with different classes of vehicles like two wheelers, three wheelers, passenger cars, buses and trucks travelling with no lane discipline. The data were being received from the sensor continuously. The percentage composition of different classes of vehicles at the study location was also analysed. The traffic was grouped into four classes, namely, two wheelers, three wheelers, light motor vehicles (LMVs) and heavy motor vehicles (HMV), and the proportions were found to be 56%, 8%, 31% and 5%, respectively. Small commercial vehicles, light commercial vehicles and passenger cars were grouped under LMV category. HMV includes vehicles like trucks, buses and multi-axle vehicles.

DOI: 10.1002/atr

However, the selected site was outside the urban centre and was downstream of a toll plaza. Thus, the location was not facing heavy congestion, and most of the time, the vehicles were moving without stop and go situation. A previous study reported the performance of this sensor at this location and summarized the findings on the accuracy of the sensor using different statistical measures [65]. In that study, it was shown that the sensor performs with maximum mean absolute percentage error (MAPE) of 2.68% for volume and speed. For the present study, five days' data were used initially for the analysis (named as Day 1 to 5). To test the performance of ANN over a longer period, another four days' data with a one month gap were also used (Day 6 to 9). Data given by the sensor include per vehicle details such as time stamp, classification, speed, and axle length. Basic data quality checks were carried out to remove any extreme outliers, like vehicles which were identified but not classified. The unclassified vehicles were then removed. It was found that 0.4% of total vehicles were unclassified vehicles. Missing data were then checked for. The missing data were imputed using spline interpolation technique. The percentage of missing data was found to be 0.8%. Only day time data were used for analysis.

The parameters considered for prediction in this study are volume, classified volume, average speed and classified speed. Five minutes was taken as the aggregation interval, and the number of vehicles identified was aggregated over every five minutes for getting the volume data. Similarly, for getting the speed data, speeds of individual vehicles were averaged over that time interval. Classified analysis was carried out only for two wheelers and LMVs, as the number of other classes was lesser as stated earlier. However, the analysis can be extended to all classes, if required.

4. PREDICTION TECHNIQUES

4.1. Artificial neural networks

Artificial neural network is a machine learning tool inspired by biological nervous systems and is composed of units operating in parallel. Neural network can be trained to perform a particular function by adjusting the weights of the connections between units. Each of these processing units is known as neurons. Neural networks are trained to adjust the weights of these neurons so that a particular input leads to a specific target. The basic architecture of neural networks is shown in Figure 1. In each neuron, scalar input p is transmitted through a connection that multiplies it by the scalar weight w and then

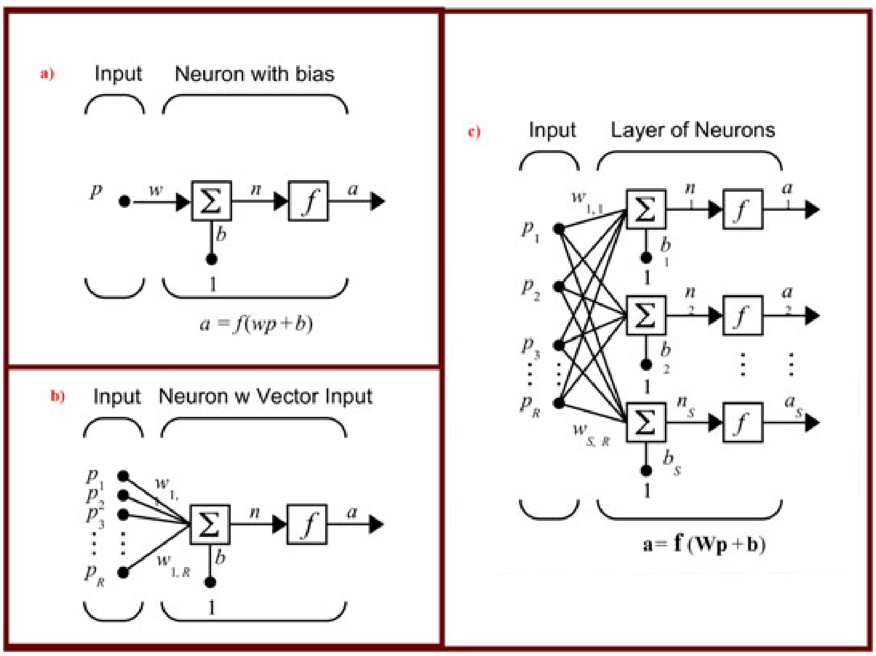

Figure 1. (a) Single neuron with single input; (b) single neuron with vector input; (c) layer of neurons with vector input (Source: [68]).

DOI: 10.1002/atr

added by a bias value b, to compute $wp + b$. This sum is the input for the transfer function f that gives an output $f(wp + b)$. The variables w and b can then be adjusted to obtain the desired output. The transfer function f can be hardlim, sigmoid or purelin based on the requirement. The output of this neuron will be the input of some other neurons based on connectivity and the process repeats. Thus, the variables w and b are continuously adjusted to obtain the desired output.

There are generally four steps in the ANN training; they are explained as follows in terms of the present study:

- Assembling the training data: Two sets of data are required for training—inputs and targets. Each input and the corresponding targets are used for identifying the weights and bias to produce the best result. In general, more training data lead to better performance. However, if the network is over trained, its generalizing capabilities will be reduced. Hence, a sensitivity analysis was carried out in the present study to identify the optimum size of training data and is explained in Section 5.2.
- Network architecture and initialization of weights: Feed forward network is one of the most commonly adopted architectures and hence was used for the present study. Feed forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. The number of hidden layers and neurons can be defined while creating the network. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. Weights are randomly initialized; as the training progresses, weights will be adjusted for best performance.
- Training the network: Training was carried out using back-propagation algorithm. This automatically recalculates and brings the error to the minimum in every iteration. The minimization is carried out by updating the network weights and biases in the direction in which the performance function decreases most rapidly.
- Testing: In this step, new data are given to the trained network for getting the corresponding output and to check the accuracy of prediction. Six days' data were used for this purpose in the present study.

Implementation of the previous steps was carried out using MATLAB.

4.2. Grey theory

The Grey system theory, established by Julong Deng in the early 80s [2], is a methodology that focuses on study of problems involving poor information and small samples. The basic assumptions of this approach are that the data used are positive and the time intervals are fixed [1]. In this theory, GM (n,h) represents the GM, where n is the order of the difference equation and h denotes the number of variables. Although various types of GMs are there, GM(1,1) model is widely used because of its computational efficiency [66].

GM(1,1), known as 'Grey model first order one variable', is the simplest model with first order differential equation and one variable. To use the GM(1,1) model, the minimum number of inputs required is four [1]. To reduce the randomness, a new set of data would be generated from the raw data, by applying an operator named accumulating generation operator (AGO). The differential equation is then solved to obtain the predicted value of the system. The inverse accumulating generation operator (IAGO) is applied finally to find the predicted values of the original data. These steps are explained in the succeeding text.

Consider a raw data series of equal time interval, $X^{(0)}$

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), \ldots, X^{(0)}(n)\} \quad n \geq 4$$

Applying the AGO to this raw series, the new set of inputs $X^{(1)}$ is obtained as

$$X^{(1)} = \text{AGO of } X^{(0)} = \{X^{(1)}(1), X^{(1)}(2), X^{(1)}(3), \ldots, X^{(1)}(n)\},$$

Here, AGO can be expressed as

DOI: 10.1002/atr

$$X^{(1)}(k) = \sum_{i=1}^{k} X^{(0)}(i), \qquad k = 1, 2, \cdots, n \tag{1}$$

Then the Grey difference equation of GM(1,1) is

$$X^{(0)}(k) + az^{(1)}(k) = b \qquad k = 1, 2, \cdots, n \tag{2}$$

where,

$X^{(0)}(k)$ is called the Grey derivative, a the development coefficient, b the Grey input. Here, $z^{(1)}(k)$ is the mean of $X^{(1)}(k)$ and $X^{(1)}(k\text{-}1)$. The coefficients a and b can be obtained using the least square method as

$$(a, b)^T = \left(B^T B\right)^{-1} B^T Y_n \tag{3}$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \tag{4}$$

$$Y_n = \left[X^{(0)}(2), X^{(0)}(3), \cdots, X^{(0)}(n)\right]^T \tag{5}$$

where, B matrix is called data matrix and Y_n is called constant term vector.

The whitening differential equation of GM(1,1) is

$$\frac{dX^{(1)}(k)}{dk} + aX^{(1)} = b \tag{6}$$

Equation (6) is solved to find $X^{(1)}(k)$ as

$$\widehat{X}^{(1)}(k+1) = \left[X^{(0)}(1) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a} \tag{7}$$

The IAGO is then carried out on (7) to obtain the restored series value $\widehat{X}^{(0)}(k)$ as

$$\widehat{X}^{(0)}(k+1) = \widehat{X}^{(1)}(k+1) - \widehat{X}^{(1)}(k) \tag{8}$$

where $\widehat{X}^{(0)}(1) = X^{(0)}(1)$.

Substituting (7) in (8),

$$\widehat{X}^{(0)}(k+1) = \left[X^{(0)}(1) - \frac{b}{a}\right] e^{-a(k)} (1 - e^a) \tag{9}$$

Where $k = 1, 2, \ldots, n$.

$\widehat{X}^{(1)}$ denotes the predicted values using the smoothened data using AGO function, and $\widehat{X}^{(0)}$ denotes the predicted value for the actual raw data. This GM was implemented in the present study using MATLAB programming. Overall, the steps involved are given in the following text:

Step 1: Put the raw data as a sequence.
Step 2: Make the new sequence using AGO (Accumulating Generating Operation).

Step 3: Find the mean values of the adjacent values in the sequence.
Step 4: Establish the Grey Model equation.
Step 5: Obtain the data matrix B.
Step 6: Solve for the coefficients a and b.
Step 7: Predict the value for the interval ($k+1$).

To quantify the error in prediction, the statistical measure MAPE is used. MAPE is calculated as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Actual - Predicted|}{Actual}\times 100\ \% \tag{10}$$

where n is the number of time intervals used.

The steps mentioned earlier are illustrated using the first four volume data from one sample day as follows.

Step 1: Sequencing the raw volume data

$$X^{(0)}(k) = \{47\ 73\ 84\ 85\}$$

Step 2: Applying Accumulating Generating Operation (AGO) to the volume data (cumulative values)

$$X^{(1)}(k) = \{47\ 120\ 204\ 289\}$$

Step 3: Finding mean values of two adjacent volume data in $X^{(1)}$

$$Z^{(1)}(k+1) = \{83.5\ 162\ 246.5\}$$

Step 4: Establishing Grey model equation

$$X^{(0)}(k) + az^{(1)}(k) = b$$

$$73 + a\,(83.5) = b$$
$$84 + a\,(162) = b$$
$$85 + a\,(246.5) = b$$

Step 5: Obtaining the data matrix, B

$$\begin{bmatrix}73\\84\\85\end{bmatrix} = \begin{bmatrix}-83.5 & 1\\-162 & 1\\-246.5 & 1\end{bmatrix}\begin{bmatrix}a\\\mathrm{b}\end{bmatrix}$$

$$B = \begin{bmatrix}-83.5 & 1\\-162 & 1\\-246.5 & 1\end{bmatrix}\ Y_n = \begin{bmatrix}73\\84\\85\end{bmatrix}$$

Step 6: Solving for the coefficients a and b

$$\begin{bmatrix}a\\b\end{bmatrix} = \left(B^TB\right)^{-1}B^TY_n$$

$$\begin{bmatrix}a\\b\end{bmatrix} = \begin{bmatrix}-0.0728\\68.7219\end{bmatrix}$$

Step 7: Predicting the volume for desired interval, 5th interval

DOI: 10.1002/atr

$$\widehat{X}^{(0)}(k+1)=\left[X^{(0)}(1)-\frac{b}{a}\right]e^{-a(k)}(1-e^{a})$$

$$\widehat{X}^{(0)}(5)=\left[47-\frac{68.7219}{(-0.0728)}\right]e^{-a(4)}\left(1-e^{(-0.0728)}\right)$$

$$\widehat{X}^{(0)}(5)=[93]$$

Step 8: Calculating absolute percentage error

$$APE=\frac{|Actual-Predicted|}{Actual}\times 100\ \ \%$$

$$APE=\frac{|96-93|}{96}\times 100\ \ =1.60\ \%$$

As explained earlier, GM is designed to handle ‘small sample and incomplete information’ [67]. GM takes the original data and makes a new sequence of data using accumulating generating operation, a proxy for the original data. This generated data will be a monotonically increasing sequence, the rate of increase of which is referred to as the growth rate of the sequence. The GM makes the prediction based on this growth rate. Thus, if the number of data points is less, the change in the growth rate will be minimal, leading to the best performance. However, how much sample is actually ‘small sample’ is not well defined. Hence, to find the optimum point, a sensitivity analysis was carried out as explained in the next section. Figure 2 shows the framework of implementation of the methodologies adopted in this study. Broadly, this study has two parts, one is ANN-based prediction

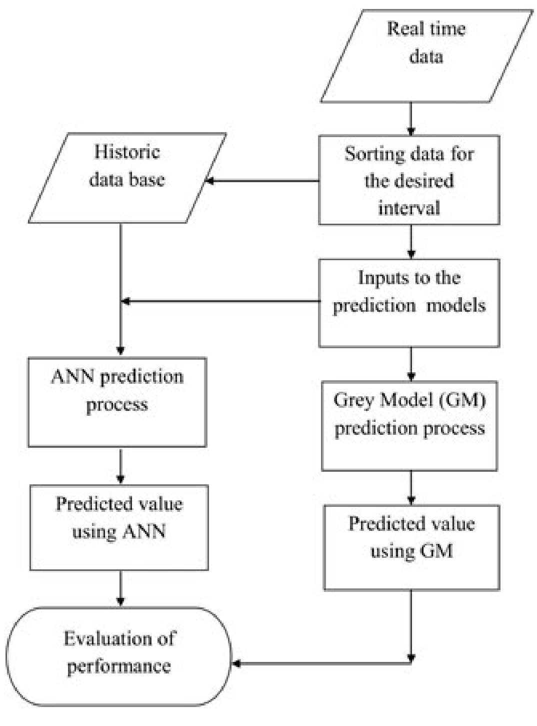

Figure 2. Implementation flow chart. ANN, artificial neural networks; GM, Grey model.

DOI: 10.1002/atr

and the other is GM-based prediction. In ANN-based prediction, a database is established to train the neural network with the desired number of inputs and time interval. Once the network is trained and ready, real-time data from the automated sensor are given as inputs, and the prediction is obtained. In the GM part, because it does not require a data base, the desired number of previous time interval data from the sensor is given directly to the model as inputs and prediction is carried out. Finally, the predicted values from the two models are evaluated using the actual value and errors are compared.

5. SENSITIVITY ANALYSIS OF GREY MODEL AND ARTIFICIAL NEURAL NETWORK

5.1. Grey model

As explained earlier, the minimum input required for prediction using GM is four. However, the optimum input size depends on the data used, and hence, a sensitivity analysis was carried out. The input was increased from 4 (20 minutes) to 40, at steps of 5 minutes. The MAPE obtained for each of this input size is shown in Figure 3. GM is designed to handle small sample and incomplete information. After an optimum number of inputs, the MAPE shows that more information causes the model to perform poorly. In this case, it can be seen that the minimum error of 13.67% corresponds to the input size of 19 (95 minutes). Hence, in this study, 19 inputs were used for prediction. Thus, while using GM (GM(1,1)), first 19 time steps are used for predicting the value at the 20th time step. Similarly, for the 21st time step, immediate previous 19 values are taken as inputs. It has to be noted here that unlike ANN, there is no training in this, and hence, a database is not required. With just 19 previous data points, the next interval value can be predicted.

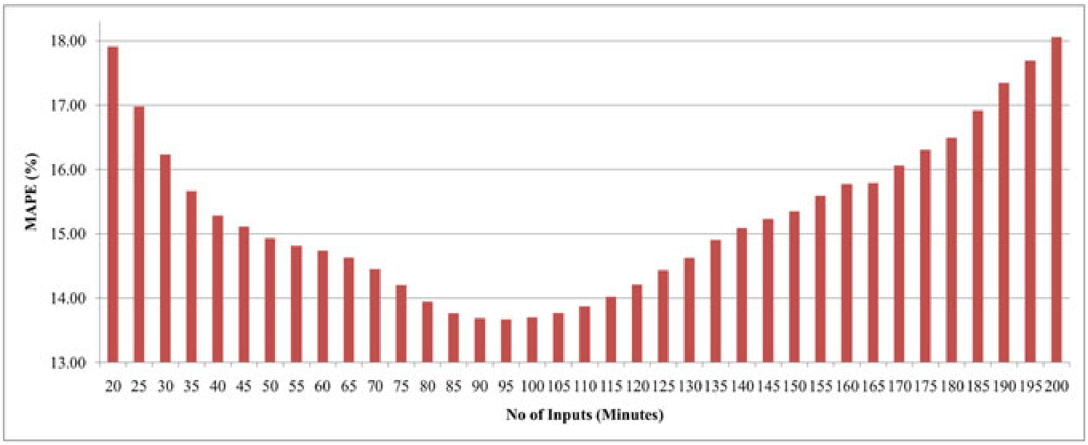

Figure 3. Change in mean absolute percentage error (MAPE) with increasing inputs in Grey model.

Table II. Mean absolute percentage error variation of Grey model with different intervals.

10 minutes interval			15 minutes interval.		
No. of inputs	Inputs (minutes)	Mean absolute percentage error (%)	No. of inputs	Inputs (minutes)	Mean absolute percentage error (%)
6	60	14.4227	4	60	16.1267
8	80	13.3772	5	75	14.6634
10	**100**	**13.0189**	6	90	13.151
12	120	13.2612	7	**105**	**12.4879**
14	140	14.1873	8	120	12.9303

DOI: 10.1002/atr

Further analysis was carried out to check the effect of input interval size on the optimum value identified earlier. For this, the input time interval was increased to 10 and 15 minutes, and the analysis was repeated. Table II shows the results obtained, and it can be seen that the optimum value remained the same for around 100 minutes in both these cases (highlighted using bold face). Thus, it can be concluded that the optimum input size is independent of the time interval adopted. Hence, the present study used 5-minute interval data, and the input size was selected as 19.

5.2. *Artificial neural network*

The sensitivity analysis for GM, explained in the earlier section, showed 19 previous time intervals (previous 95 minutes data) as the optimum input set. Similarly, sensitivity analysis for ANN was also carried out, where the number of inputs was varied and the MAPE was observed. Table III shows the MAPE for varying input sizes, from 4 to 24. It can be seen from the table that for varying input sizes, the difference in MAPE is very minimal, which is around 13%. Hence, it was decided to use 19 inputs to make the comparison with GM easy.

Thus, ANN develops a relation between the data point at a particular time step with previous 19 time steps. For example, the first set of input for ANN is the first 19 data points and the corresponding output is the 20th data point. The second set of input is 19 data points from 2nd to 20th, and the corresponding output considered is the 21^{st} data point.

However, ANN needs a training data set, the size of which is not well defined. Hence, analysis was carried out to identify the optimum training set required for ANN with 19 inputs. The training data were varied from 6 to 48 hours, and the results obtained are shown in Figure 4.

From the figure, it can be seen that with increasing training data, there is a reduction in MAPE. However, it can be observed that the optimum size is around 36 hours of data. Hence, the present study used 36 hours of data as the training data set in further analysis.

Table III. Mean absolute percentage error for varying inputs in artificial neural networks.

Number of Inputs	Mean absolute percentage error (%)
4	13.29
8	13.26
12	13.18
16	13.6
20	13.52
24	13.46

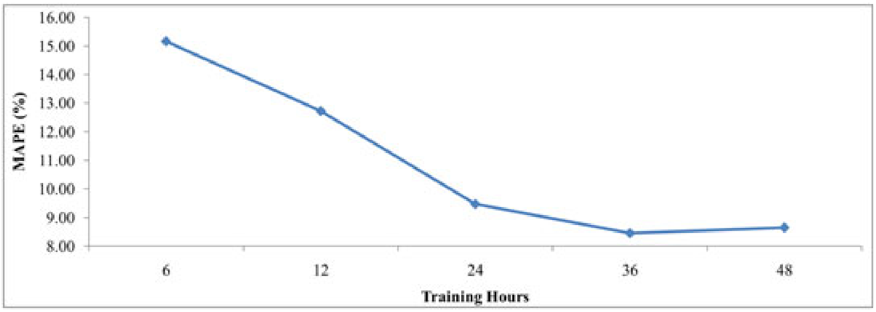

Figure 4. Variations of mean absolute percentage error (MAPE) for different amount of training hours in artificial neural networks.

DOI: 10.1002/atr

6. RESULTS

Using the identified inputs and training set from sensitivity analysis, the prediction was carried out, and results obtained are presented and discussed in this section.

6.1. Prediction of volume

The actual total volume varies from 50 to 450 vehicles per 5 minutes. Figure 5 shows a sample comparison of actual and predicted volume using both the techniques, GM and ANN (using 4 and 19 inputs), for one sample day, from 7:30 to 17:30 hours. From Figure 5, it can be seen that both the techniques are able to capture the variation in volume correctly. The corresponding MAPEs were 9.53% and 8.24% and 7.68% for ANN using 19 and 4 inputs, and GM, respectively. It can be seen that, overall, the predicted values obtained using both the methods closely match the actual values with GM's performance slightly better than ANN's.

Classified volumes for two wheelers and LMVs were also predicted. The volume per 5 minutes of two wheelers varied from 0 to 300 vehicles and that of LMV varied from 0 to 125 vehicles. Figure 6 shows an overall comparison of error results of the prediction of total and classified volume. It can be seen that the MAPE values of classified volume prediction also show GM to be slightly better than ANN.

Predictions were carried out for the remaining days also, and the results obtained for all 6 days are shown in Table IV. It can be seen that overall prediction performance of ANN and GM is comparable, with GM showing slightly better performance than ANN. However, it has to be noticed that ANN, even when tested with 1 month later data (day 6 to 9) was performing consistently. GM on the other hand does not need any training and hence uses only that day's data.

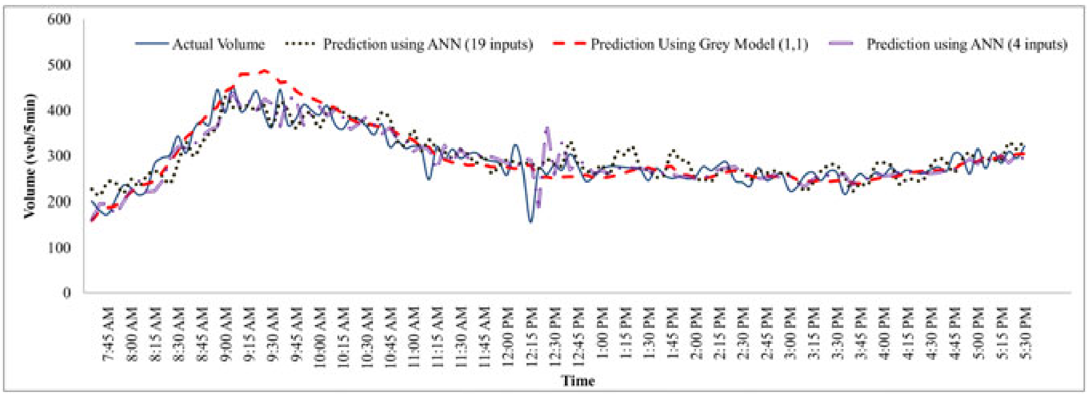

Figure 5. Comparison of actual and predicted volume. ANN, artificial neural networks.

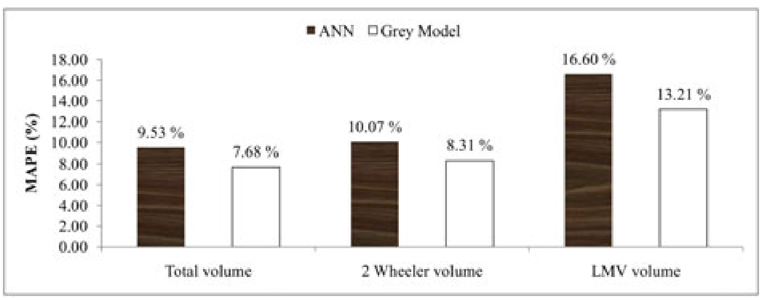

Figure 6. Comparison of mean absolute percentage error (MAPE) in prediction of volume. ANN, artificial neural networks.

DOI: 10.1002/atr

Table IV. MAPE obtained for volume prediction.

Days	Volume		Two-wheeler volume		Light motor vehicles volume	
	ANN MAPE (%)	GM MAPE (%)	ANN MAPE (%)	GM MAPE (%)	ANN MAPE (%)	GM MAPE (%)
Day 4	7.64	6.75	9.53	8.97	13.52	12.57
Day 5	9.53	7.68	10.07	8.31	16.6	13.21
Day 6	9.16	7.68	9.88	9.44	12.53	12.67
Day 7	8.79	7.31	8.92	9.41	15.67	12.94
Day 8	9.53	7.69	11.3	9.9	14.67	11.88
Day 9	9.68	8.47	9.57	9.85	14.28	13.93

ANN, artificial neural networks; MAPE, mean absolute percentage error; GM, Grey model.

6.2. Prediction of speed

The average speed in the study location varied from 25 to 50 kmph in the 5 days that were selected. Figure 7 shows the comparison of actual and predicted average speed for Day 5, from 7:30 to 17:30 hours. Similar to volume prediction, both the techniques were able to predict speed values closer to actual values.

Classified analysis was also carried out, and Figure 8 compares the quantified error in prediction of speed for all the cases. Similar predictions were carried out for the remaining days, and the results obtained are shown in Table V. From the results, it can be seen that both the methods had prediction error within 8%, showing good performance. Comparing the two, both the methods showed comparable

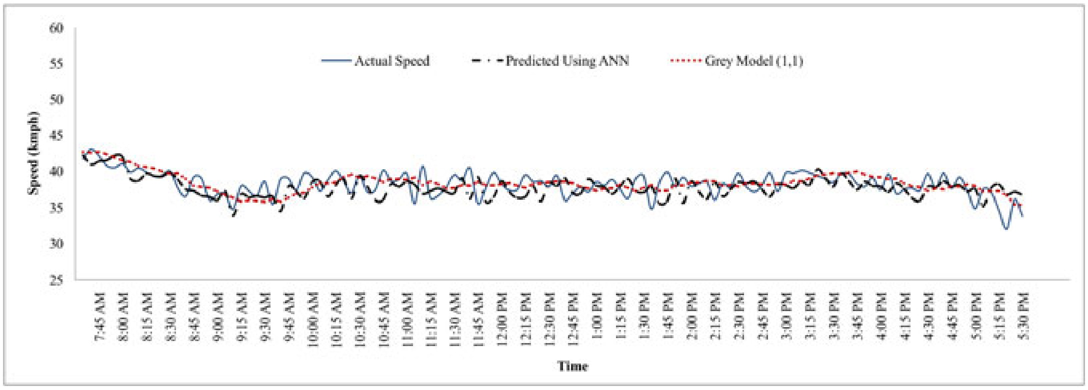

Figure 7. Comparison of actual and predicted average speed. ANN, artificial neural networks.

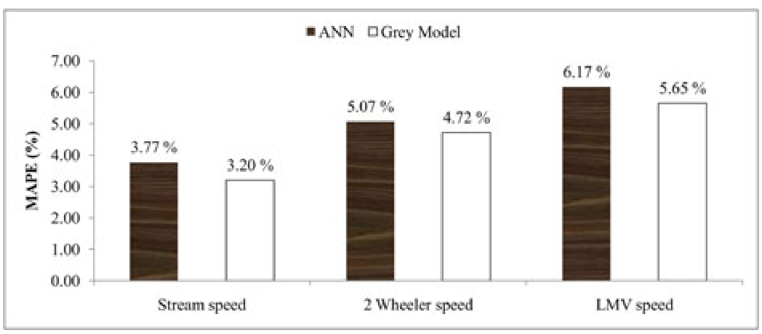

Figure 8. Comparison of mean absolute percentage error (MAPE) in prediction of average speed. ANN, artificial neural networks; LMV, light motor vehicles.

Table V. MAPE obtained for speed prediction.

Days	Speed ANN MAPE (%)	Speed GM MAPE (%)	Two-wheeler speed ANN MAPE (%)	Two-wheeler speed GM MAPE (%)	Light motor vehicles speed ANN MAPE (%)	Light motor vehicles speed GM MAPE (%)
Day 4	4.17	3.29	5.5	5.02	7.49	6.57
Day 5	3.77	3.2	5.07	4.72	6.17	5.65
Day 6	4.29	3.69	5.26	5.18	5.74	5.86
Day 7	4.15	3.53	5.46	5.28	6.64	6.23
Day 8	4.63	3.56	5.25	5.02	6.03	6.18
Day 9	4.32	3.21	5.42	5.17	5.64	5.36

ANN, artificial neural networks; MAPE, mean absolute percentage error; GM, Grey model.

performance in terms of MAPE. Thus, GM can be considered as a promising prediction tool, in the absence of a historic database, which is essential for ANN.

7. ARTIFICIAL NEURAL NETWORK AND GREY MODEL PREDICTION COMPARISON

As the main objective of the present study is to compare the performance of two different prediction tools, one which is data-driven and the other less-data-demanding, a comparison of the results of the sensitivity analysis was carried out to find the optimum data requirement for both. As discussed in the earlier section, GM was run with inputs varying from 4 to 40, and ANN was run for a training data of 6 to 48 hours. The MAPE obtained for each of this input size is tabulated and is shown in Table VI. It can be seen that for GM, the minimum error of 13.67% was obtained for an input size of 19 (95 minutes). Results obtained for ANN showed a comparable performance with a MAPE of 12.72% with a training data of 12 hours. Thus, it can be concluded that the amount of data required for both ANN and GM to perform equally is 95 to 100 minutes for GM and 12 hours for ANN (highlighted in Table VI using boldface). This result from ANN and GM clearly indicates that GM requires lesser data than ANN for comparable performance and hence can be considered under scenarios where availability of data is limited.

8. TRANSFERABILITY OF ARTIFICIAL NEURAL NETWORK AND GREY MODEL

To check the transferability of the two tools, two more days' data from two different locations in Chennai, India, namely, Gurunanak College and Tidel Park were considered. The volume data were compiled into every 5-minute interval. At Gurunanak College, the traffic flow is two-way, with no median and the volume count was taken in both directions. At Tidel Park, the traffic flow is one-way. As the data collected from these sites were limited, the already trained network using data from other sites was used for prediction of volume using ANN. Because of the same reason, GM was also run with the minimum required four inputs. Table VII tabulates the statistical parameters of all the sites, as well as the comparison of prediction performance at these locations. It can be observed that at Tidel Park location, the prediction errors are higher compared with other sites; this could be a result of the high coefficient of variation observed at that location. It can be observed that ANN is still able to perform better with a MAPE of 13.14%, even though the ANN network was trained using data from another

Table VI. MAPE for varying input size for both ANN and GM.

GM—no. of inputs (minutes)	GM MAPE	ANN—no. of inputs (hours)	ANN MAPE
50	14.93	6	15.15
100	**13.7**	**12**	**12.72**
150	15.35	24	9.47
200	18.06	36	8.45
		48	8.64

ANN, artificial neural networks; MAPE, mean absolute percentage error; GM, Grey model.

DOI: 10.1002/atr

Table VII. Descriptive statistics of volume data from different location.

Site details	Hourly volume (vehicles/hour)	Mean (vehicles/ minute)	Standard deviation (vehicles/minute)	Coefficient of variation (%)	GM MAPE (%)	ANN MAPE (%)
Tidel Park (Morning peak—8:00 to 8:45 hours)	2250	38	11	29.68	22.88	13.14
Perungudi (Morning peak—8:00 to 8:45 hours)	3550	60	13	22.43	10.43	9.41
Gurunanak College (Evening peak—17:30 to 18:15 hours)	2830	48	8	16.97	11.9	10.54
Perungudi (Evening peak—17:30 to 18:15 hours)	4070	68	11	16.06	9.2	4.7

ANN, artificial neural networks; MAPE, mean absolute percentage error; GM, Grey model.

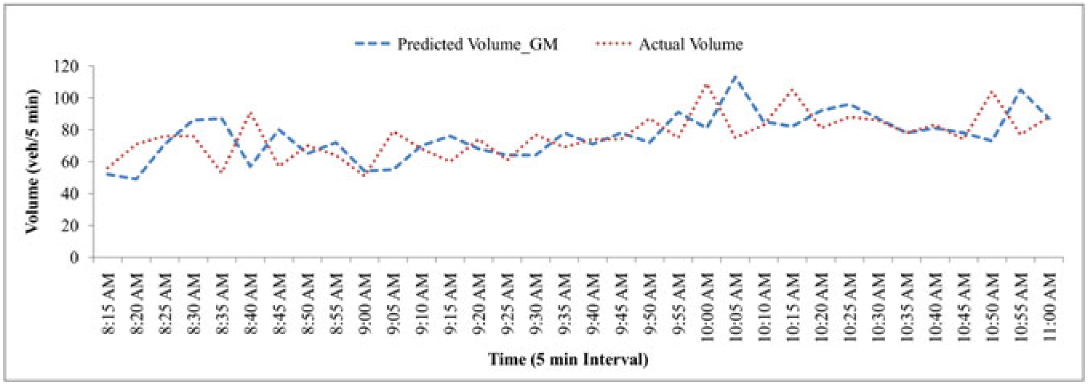

Figure 9. Comparison of predicted and actual volume for a highway location. GM, Grey model.

location. This may be due to the fact that all these roads are urban arterials of the same city and hence having comparable characteristics. The volumes were compared and were found to be in comparable ranges during corresponding periods.

The prediction performance of GM was tested for a highway traffic also. For this, data were collected in a highway near Achrapakkam, Chennai (National Highway 45). Because of data collection difficulty, data were collected only for a 3-hour period (from 7:45 to 11:00 hours). The traffic flow prediction was carried out using five inputs because of the limited data, and the results obtained are shown in Figure 9.

Figure 9 shows the actual and predicted flows and the corresponding MAPE obtained was 16.9%, which is consistent with MAPE obtained from the urban locations analysed earlier. Thus, in the preliminary analysis, the performance seems to be comparable for the arterial and highway data. However, further analysis can be carried out with more data for arriving at the final conclusion. Because of data constraint, the highway data were not tested using ANN.

9. SUMMARY AND CONCLUSION

The aim of the present study was to compare the performance of data-driven and less-data-demanding tools for the traffic parameter prediction problem. ANN and GM were selected as sample tools. Short-term prediction of parameters such as volume and speed was carried out using ANN and GM, and the performances were compared. Classified parameters were also considered for prediction with two major classes—two wheelers and LMVs.

Sensitivity analyses were carried out first to identify the optimum inputs to be used for these methods. The results obtained from the sensitivity analyses showed that the optimum number of inputs for GM is 19 (95 minutes). For ANN, the optimum size of training data was found to be 36 hours. Using these inputs, predictions were carried out. The prediction results showed comparable

DOI: 10.1002/atr

performance from both the methods, with slight advantage to GM. Both the methods were able to follow the variations in the original data fairly well. The MAPE for volume prediction was found to be 7% to 17% for ANN and 6% to 14% for GM. For speed prediction, it varied from 3% to 8% for ANN and 3% to 7% for GM. The lesser MAPE for speed prediction is assumed to be a result of lesser variation in the data. Thus, the overall performance of both the methods is comparable, showing GM also to be a promising tool for traffic prediction problems along with ANN.

Comparison of input requirement of both the methods showed that, to obtain an identical performance by ANN and GM, ANN required 12 hours of data to train the network whereas GM required only around 95 to 100 minutes of data. Thus, considering the lesser input requirement of GM, it can be considered over ANN in situations where historic database is limited. Other prediction techniques such as support vector machines or dynamical systems approach can be applied to the same problem and performance can be compared and this will be considered as a possible future task.

10. LIST OF SYMBOLS AND ABBREVIATIONS

10.1. Symbols

p	Scalar input
w	Scalar weight
b (Figure 1)	Bias value
$f()$	Transfer function
n (Figure 1)	Net input
n	Order of the difference equation
h	Number of variables used in Grey Model
$X^{(0)}(k)$	Raw data series
$X^{(1)}(k)$	Accumulating generation operation data series
a	Development coefficient
b	Grey input
$z(k)$	Mean of two adjacent values in the data series
B	Data matrix
Y_n	Constant term vector
$[]^T$	Transpose of a matrix
$\widehat{X}^{(1)}$	Predicted data for $X^{(1)}$ series
$\widehat{X}^{(0)}$	Predicted data for $X^{(0)}$ series

10.2. Abbreviations

ANN	Artificial Neural Network
GM	Grey Model
k-NN	K Nearest Neighbour
SVM	Support Vector Machine
RBF	Radial Basis Function
BP	Back Propagation
Tx	Transmitter end of the sensor
Rx	Receiver end of the sensor
IR	Infra Red
LMV	Light Motor Vehicles
HMV	Heavy Motor Vehicles
MAPE	Mean Absolute Percentage Error
AGO	Accumulating Generation Operator
IAGO	Inverse Accumulating Generation Operator
kmph	Kilometer per hour
no.	Number

DOI: 10.1002/atr

ACKNOWLEDGGEMENTS

The authors acknowledge the support for this study as part of the sub-project CIE/10-11/169/IITM/LELI under the Centre of Excellence in Urban Transport project funded by the Ministry of Urban Development, Government of India, through letter no. N-11025/30/2008-UCD.

REFERENCES

1. Deng J. Introductions to Grey system theory. *The Journal of Grey System* 1989; **1**(1): 1–24.
2. Liu S, Forrest J, Yang Y. A brief introduction to Grey systems theory. *IEEE International Conference on Grey Systems and Intelligent Services (GSIS)*, Nanjing, China, September 2011; 1–9. DOI: 10.1109/GSIS.2011.6044018.
3. VanLint JWC, Hoogendoorn SP VanZuylen HJ. Accurate freeway travel time prediction with state-space neural networks under missing data. *Transportation Research Part C: Emerging Technologies* 2005; **13**(5–6): 347–369.
4. Liu H, VanLint H, VanZuylen H, Zhang K. Two distinct ways of using Kalman filters to predict urban arterial travel time. *Proceedings of the IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, September 2006; 845–850. DOI: 10.1109/ITSC.2006.1706849.
5. Zhang Y, Liu Y. Analysis of peak and off peak forecasts using combined models. *Journal of Advanced Transportation* 2011; **45**(1): 21–37.
6. Kwon J, Coifman B Bickel P. Day-to-day travel time trends and travel time prediction from loop detector data. *Transportation Research Record* 1717; **2000**: 120–129.
7. Zhang X, Rice JA. Short term travel time prediction. *Transportation Research Part C: Emerging Technologies* 2003; **11**(3–4): 187–210.
8. Rice J, van Zwet E. A simple and effective method for predicting travel times on freeways. *IEEE Transactions on Intelligent Transport Systems* 2004; **5**(3): 200–207.
9. Lee S, Lee YI Cho B. Short-term travel speed prediction models in car navigation systems. *Journal of Advanced Transportation* 2006; **40**(2): 122–139.
10. Chang H, Park D, Lee S, Lee H Baek S. Dynamic multi interval bus travel time prediction using bus transit data. *Transportmetrica* 2010; **6**(1): 19–38.
11. Guin A. Travel time prediction using a seasonal autoregressive integrated moving average time series model. *Proceedings of the IEEE Intelligent Transportation Systems Conference*, Toronto, Canada, September 2006; 493–498. DOI: 10.1109/ITSC.2006.1706789.
12. Xia J. Dynamic freeway travel time prediction using single loop detector and incident data. PhD Dissertation. University of Kentucky, Lexington, USA. 2006.
13. Liu X, Chien SI Chen M. An adaptive model for highway travel time prediction. *Journal of Advanced Transportation* 2014; **48**(6): 642–654.
14. Kisgyorgy L, Rilett LR. Travel time prediction by advanced neural network. *Periodica Polytechnica Ser. Civ. Eng* 2002; **46**(1): 15–32.
15. Lam WHK, Xu J. Estimation of AADT from short period counts in Hong Kong: a comparison between neural network method and regression analysis. *Journal of Advanced Transportation* 2000; **34**(2): 249–268.
16. Shen L. Freeway travel time estimation and prediction using dynamic neural networks. PhD Dissertation. Florida International University, Miami, Florida, USA. 2008.
17. Yu J, Chang GL, Ho HW, Liu Y. Variation based online travel time prediction using clustered neural networks. *11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China, October 2008; 85–90. DOI: 10.1109/ITSC.2008.4732594.
18. Zou N, Wang J, Chang GL. A reliable hybrid prediction model for real-time travel time prediction with widely spaced detectors. *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*; Beijing, China, October 2008; 91–96. DOI: 10.1109/ITSC.2008.4732664.
19. Liu H, VanLint HWC, VanZuylen HJ. Neural network based traffic flow model for urban arterial travel time prediction. *Proceedings of the 86th Annual Transportation Research Board Meeting,* Washington D.C., USA, January 2007.
20. VanLint JWC. Online learning solutions for freeway travel time prediction. *IEEE Transactions on Intelligent Transportation Systems* 2008; **9**(1): 38–47.
21. Zhu G, Wang X. Study on route travel time prediction based on RBF neural network. *First International Workshop on Education Technology and Computer Science*, Wuhan, Hubei, China, March 2009; 1118–1122. DOI: http://doi.ieeecomputersociety.org/ 10.1109/ETCS.2009.517.
22. Lee Y. Freeway travel time forecast using artificial neural networks with cluster method. *12th International Conference on Information Fusion,* Seattle, WA, USA, July 2009; 1331–1338.
23. Ozkurt C, Camci F. Automatic traffic density estimation and vehicle classification for traffic surveillance systems using neural networks. *Mathematical and Computational Application* 2009; **14**(3): 187–196.
24. Cetiner BG, Sari M Borat O. A neural network based traffic-flow prediction model. *Mathematical and Computational Applications* 2010; **15**(2): 269–278.
25. Mazloumi E, Rose G, Currie G Sarvi M. An integrated framework to predict bus travel time and its variability using traffic flow data. *Journal of Intelligent Transport Systems: Technology, Planning, and Operations* 2011; **15**(2): 75–90.

26. Chen M, Yaw J, Chien SI Liu X. Using automatic passenger counter data in bus arrival time prediction. *Journal of Advanced Transportation* 2007; **41**(3): 267–283.
27. Yaghini M, Khoshraftar MM Seyedabadi M. Railway passenger train delay prediction via neural network model. *Journal of Advanced Transportation* 2013; **47**(3): 355–368.
28. Bagloee SA, Cedar A Bozic C. Effectiveness of en-route traffic information in developing countries using conventional discrete choice and neural-network models. *Journal of Advanced Transportation* 2014; **48**(6): 486–506.
29. Chen M, Chien S. Dynamic freeway travel time prediction using probe vehicle data: link based versus path based. *Transportation Research Board 80th Annual meeting,* Washington DC, USA, January 2001.
30. Chien SI, Kuchipudi CM. Dynamic travel time prediction with real time and historic data. *Journal of Transportation Engineering* 2003; **129**(6): 608–616.
31. Gong Y, Zhang Y. Research of short term traffic volume prediction based on Kalman filtering. *6th International Conference on Intelligent Networks and Intelligent Systems (ICINIS),* Shenyang, China, November 2013; 99–102.
32. Okutani I, Stephanedes YJ. Dynamic prediction of traffic volume through Kalman filtering theory. *Transportation Research Part B: Methodological* 1984; **18**(1): 1–11. DOI:10.1016/0191-2615(84)90002-X.
33. Wang Y, Papageorgiou M. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B: Methodological* 2005; **39**(2): 141–167. DOI:10.1016/j.trb.2004.03.003.
34. Vanajakshi L, Subramanian SC Sivanandan R. Travel time prediction under heterogeneous traffic conditions using global positioning system data from buses. *IET Intelligent Transport Systems* 2009; **3**(1): 1–9. DOI:10.1049/iet-its:20080013.
35. Zhang X, Onieva E, Perallos A, Osaba E Lee VCS. Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term congestion prediction. *Transportation Research Part C: Emerging Technologies* 2014; **43**(1): 127–142. DOI:10.1016/j.trc.2014.02.013.
36. Soriguera F, Robuste F. Highway travel time accurate measurement and short term prediction using multiple data sources. *Transportmetrica* 2011; **7**(1): 85–109.
37. Yin H, Wong SC, Xu J Wong CK. Urban traffic flow prediction using a fuzzy neural approach. *Transportation Research Part C: Emerging Technologies* 2002; **10**(2): 85–98. DOI:10.1016/S0968-090X(01)00004-3.
38. Sharma B, Katiyar VK Gupta AK. Fuzzy logic model for the prediction of traffic volume in weekdays. *International Journal of Computer Applications* 2014; **107**(17): 1–6. DOI:10.5120/18840-0026.
39. Yu B, Yang ZZ, Chen K Yu B. Hybrid model for prediction of bus arrival times at next station. *Journal of Advanced Transportation* 2014; **44**(6): 193–204.
40. Yu C, Lam KC. Applying multiple kernel learning and support vector machine for solving the multi-criteria and nonlinearity problems of traffic flow prediction. *Journal of Advanced Transportation* 2014; **48**(3): 250–271.
41. Vanajakshi L, Rilett LR. Support vector machine technique for the short term prediction of travel time. *IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, June 2007; 600–605. DOI: 10.1109/IVS.2007.4290181.
42. Neto MC, Jeong YS, Jeong MK Han LD. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications* 2009; **36**(3): 6164–6173. DOI:10.1016/j.eswa.2008.07.069.
43. Zhang Y, Liu Y. Traffic forecasting using least squares support vector machines. *Transportmetrica* 2009; **5**(3): 193–213. DOI:10.1080/18128600902823216.
44. Hsu LC. Applying the Grey prediction model to the global integrated circuit industry. *Technological Forecasting and Social Changes* 2003; **70**(6): 563–574.
45. Lin Y, Liu S. A historical introduction to Grey system theory. *IEEE International Conference on Systems Man and Cybernetics* 2004; **3**: 2403–2408.
46. Hui S, Yang F, Li Z, Liu Q Dong J. Application of Grey system theory to forecast the growth of larch. *International Journal of Information and System Sciences* 2009; **5**(3–4): 522–527.
47. Askari M, Askari H. Time series Grey system prediction based models: gold price forecasting. *Trends in Applied Sciences Research* 2011; **6**(11): 1287–1292.
48. Vishnu B, Syamala P. Grey model for stream flow prediction. *Aceh International Journal of Science and Technology* 2012; **1**(1): 14–19.
49. Feng SJ, Ma YD, Song ZL Ying J. Forecasting the energy consumption of China by the Grey prediction model. *Energy Sources, Part B: Economics, Planning and Policy* 2012; **7**(4): 376–389.
50. Mei HC, Hua SW Cong XX. Research on construction and application of the GM(1,1) forecast model of Olympics track and field achievements. *Grey Systems: Theory and Application* 2012; **2**(2): 178–196.
51. Delcea C, Bradea L, Maracine V, Scarlat E Cotfas LA. GM (1,1) in bankruptcy forecasting. *Grey Systems: Theory and Application* 2013; **3**(3): 250–265.
52. Zhang Q, Chen R. Application of metabolic GM(1,1) model in financial repression approach to the financing difficulty of the small and medium-sized enterprises. *Grey Systems: Theory and Application* 2014; **4**(2): 311–320.
53. Sun YC, Zhenguo SZ. Application of grey models to traffic flow prediction at non detector intersection. *Journal of Southeast University (Natural Science Edition)* 2002.
54. Mao M, Chirwa EC. Combination of Grey model GM (1,1) with three point moving average for accurate vehicle fatality risk prediction. *International Journal of Crashworthiness* 2005; **10**(6): 635–642.
55. Xu X, Chen B Gan F. Traffic safety evaluations based on Grey systems theory and neural network. *Proceedings of World Congress on Computer Science and Information Engineering* 2009; **5**: 603–607.
56. Zhang Y. Predicting model of traffic volume based on Grey–Markov. *CCSE- Modern Applied Science* 2010; **4**(3): 46–50.

57. Guo H, Xiao X Jeffrey F. Urban road short-term traffic flow forecasting based on the delay and nonlinear Grey model. *Journal of Transportation Systems Engineering and Information Technology* 2013; **13**(6): 60–66. DOI:10.1016/S1570-6672(13)60129-4.
58. Mohammadi A, Moradi L, Talebnejad A Nadaf A. The use of Grey system theory in predicting the road traffic accident in Fars province in Iran. *Australian Journal of Business and Management research* 2011; **1**(9): 18–23.
59. He FB, Chang J. Combined forecasting of regional logistics demand optimized by genetic algorithm. *Grey Systems: Theory and Application* 2014; **4**(2): 221–231.
60. Zhao Z, An S, Wang J. Prediction of traffic flow based on gray theory and BP neural network. *Proceedings of the First International Conference on Transportation Engineering*, Chengdu, China, July 2007; 2930–2935. DOI: 10.1061/40932(246)481.
61. Darkopoulos JA, Abdulkader A. Training neural networks with heterogeneous data. *Neural Networks* 2005; **18**(5–6): 595–601.
62. Padiath A, Vanajakshi L, Subramanian S, Manda H. Prediction of traffic density for congestion analysis under Indian traffic conditions. *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems,* St. Louis, MO, USA, October 2009; 1–6. DOI: 10.1109/ITSC.2009.5309716.
63. Badhrudeen M, Raj J, Vanajakshi LD. Short term prediction of traffic parameters—performance comparison of data driven and less data required approaches. *Transportation Research Board 93rd Annual Meeting,* Washington DC, USA, January 2014.
64. TIRTL Technical Overview (Version 2.1). TIRTL Manual (The Infra-red Traffic Logger). CEOS Industrial: Heidelberg. Available at: http://www.ceos.com.au/pdfs/TIRTLTechnicalOverview.pdf.
65. Raj J, Ramesh V, Varma SR, Vanajakshi LD. Evaluation and application of automated traffic sensor data under Indian conditions. *Transportation Research Board 93rd Annual Meeting,* Washington DC, USA, January 2014.
66. Demuth H, Beale M. *Neural Network Toolbox—For Use With MATLAB*. The MathWorks, Inc. 2000. http://www.image.ece.ntua.gr/courses_static/nn/matlab/nnet.pdf
67. Kayacan E, Ulutas B, Kaynak O. Grey system theory-based models in time series prediction. *Expert Systems with Applications* 2010; **37**(2): 1784–1789.
68. Yao T, Liu S, Xie N. On the properties of small sample of GM(1,1) model. *Applied Mathematical Modelling* 2009; **33**(4): 1894–1903.

