



DALARNA
UNIVERSITY

Working papers in transport, tourism, information technology and microdata analysis

Speed prediction for Triggering Vehicle Activated Signs

Diala Jomaa
Siril Yella
Mark Dougherty
Editor: Hasan Fleyeh

Nr: 2016:01

Working papers in transport, tourism, information technology and microdata analysis

ISSN: 1650-5581

© Authors

Speed prediction for Triggering Vehicle Activated Signs

Diala Jomaa, Siril Yella and Mark Dougherty

Department of Computer Engineering, Dalarna University, Borlänge, 78170 Sweden

Email: djo@du.se, sye@du.se, mdu@du.se

Abstract-

Accurate speed prediction is a crucial step in the development of a dynamic vehicle activated sign (VAS). A previous study showed that the optimal trigger speed of such signs will need to be pre-determined according to the nature of the site and to the traffic conditions. The objective of this paper is to find an accurate predictive model based on historical traffic speed data to derive the optimal trigger speed for such signs. Adaptive neuro fuzzy (ANFIS), classification and regression tree (CART) and random forest (RF) were developed to predict one step ahead speed during all times of the day. The developed models were evaluated and compared to the results obtained from artificial neural network (ANN), multiple linear regression (MLR) and naïve prediction using traffic speed data collected at four sites located in Sweden. The data were aggregated into two periods, a short term period (5-min) and a long term period (1-hour). The results of this study showed that using RF is a promising method for predicting mean speed in the two proposed periods.. It is concluded that in terms of performance and computational complexity, a simplistic input features to the predictive model gave a marked increase in the response time of the model while still delivering a low prediction error.

Keywords- *vehicle activated signs; trigger speed; adaptive neuro-fuzzy inference systems; classification and regression tree; Random forest; multiple linear regression; mean speed; traffic flow.*

1. Introduction

Vehicle activated signs (VAS) are road warning signs that measure the speed of passing vehicles and when a driver exceeds a particular threshold, display a warning message typically a ‘slow down’ in combination with the current speed limit (Walter and Knowles 2008). The threshold which triggers the message to the driver is commonly based on a vehicle’s speed and, is called a trigger speed. Previous studies showed the optimal trigger speed i.e. the trigger speed that has the best effect on driver behaviour needs to be pre-determined according to the nature of the site and to the traffic conditions (Jomaa et al. 2014). If there are no traffic data available, it is hard to identify which trigger speed should be applied to the site. Hence dynamic signs can be used to find the optimal trigger speed responding to real traffic and road conditions. At the same time developing the VAS to be a dynamic sign requires fast processing, analysing and storing a large amount of traffic data making the system expensive and challenging to be implemented. To cope up with real time traffic management and time lags, an accurate predictive model based on historical traffic speed data is needed. Further A key requirement of such a model is to employ the appropriate trigger speed for traffic conditions featuring inappropriate speeds of occasional vehicles travelling very fast or slowly; because such occasional vehicles can distort speed prediction.

A large number of studies have attempted to predict short term traffic variables; where most of

them have been reported within the traffic flow prediction area. Such studies have mainly reported the usage of either statistical methods or artificial intelligence (AI) based techniques. Examples of statistical methods are Kalman filtering (KF), the auto-regressive integrated moving average (ARIMA) (Vlahogianni et al. 2005, Gazis et al. 2003, Van Der Voort et al. 1996) and several non-parametric regression models (Hamed et al. 1995, Smith and Demetsky 1997). Most of these statistical methods are mainly dependent on statistical distributions of traffic variables which from a practical point of view are irrelevant. AI techniques such as fuzzy logic, neural networks (ANN), genetic algorithms and intelligent multi-agents have also been reported; with ANN being reported as the most popular technique. Application of ANNs was investigated in various studies (Dougherty 1995, Dougherty M. and Cobbett 1997, Hooshdar and Adeli 2004, Park 2011). Wen et al. (2001) considered probabilistic neural networks (PNN) in incident detection using different patterns under a variety of flow conditions and traffic periods generated by a traffic simulation model. Fuzzy systems and fuzzy neuro systems have been employed in adaptive decision making of traffic control systems (Zhang and Ye 2008, Park 2002, Park et al. al. 2011, Li et al. 2008, Yin et al. 2002). Case base reasoning has been employed for real-time freeway traffic routing (Sadek et al. 2001). Adaptive hybrid systems; such as fuzzy rule based systems (Dimitriou et al.2008, Kwon and Stephanedes 1994, Lin and Lin 2007) and Genetic Algorithm (GA) combined with ANN models (Vlahogianni et al. 2005, Chiou et al. 2014) have also been reported in short term traffic flow prediction.

Although a considerable amount of literature focused on traffic flow, lack of studies attempting to deal with traffic speed prediction has been identified. More particularly the objective of this article is to investigate an appropriate model to predict speed one step ahead during all times of the day to be able to trigger and operate the VAS in good time i.e. before the drivers pass the sign.

In the current study adaptive neuro fuzzy inference system (ANFIS), classification and regression trees (CART) and random forests (RF) were employed to predict speeds. Results achieved by the aforementioned models were evaluated and compared to more robust techniques such as artificial neural network (ANN), multiple linear regression (MLR) and a naïve prediction model using traffic speed data collected at four sites located in Sweden for the sake of validation. Note that the naïve prediction model has simply assumed the current value as the predicted speed.

A large number of input factors which impact current traffic situations have been frequently considered in previous studies (Smulders 1990, Smulders and Helleman 1998). Factors such as day of the week, time of the day, flow of the traffic i.e. number of vehicles for a certain period of time; mean speeds and standard deviation are often examined. Selecting a suitable combination of these inputs is very challenging and is a key to finding an appropriate prediction model; and correlation analysis has been employed for the purpose. Another important issue in short term traffic forecasting concerns the level of data aggregation (Dougherty and Cobbett, 1997). In the current case preliminary experiments using aggregation of the data in 1 min periods have been found to be unsuccessful because of the large stochastic variation. Additionally such short periods are not recommended in practice by traffic engineers. This is because rapidly varying the trigger speed of the VAS may confuse drivers or create unstable traffic speeds. Conversely a higher aggregation level (prediction of traffic speed in 1-hour period), might be appropriate in steady traffic conditions; typically in a road segment with relatively low speed limits (30/40 km/hr) but the same might be unsuitable in a highly varying situation. After preliminary experimentation data were aggregated into 5-min and 1-hour periods. The rest of the paper is organised as follows. Section 2 and 3 present data acquisition and input selection. Section 4 presents prediction models. Section 5 presents results and discussion. The paper presents concluding

remarks in section 6.

2. Data

Traffic speed data were collected 24 hours a day onsite at four different locations. First site¹ is located on highway E16 between Borlänge and Djurås in central Sweden (site-1) and is restricted to 60km/hr. Second² and third sites³ (sites-2 and 3) are both restricted to 40km/hr whereas the fourth site⁴ is restricted to 60km/hr (site-4). One single VAS has been used to collect all the data from the different sites to keep costs low. Further details such as number of observations and dates can be found in table 1 (see table 1). At this stage it is worth mentioning that the VAS is equipped with radar and a data logger to record the speed of passing vehicles 100m before the location of the VAS. The rationale behind such a VAS was to build an adaptive VAS which detects and records vehicles speed and predicts a trigger speed respective to previous traffic conditions.

Past studies in the area investigated speed prediction but in our opinion most of them have been rather selective i.e. by considering speed data recorded in certain periods of time typically excluding data collected during rush hours and night time (Dia 2001, Chen and Grant-Muller 2001, Yin et al 2002 and Stathopoulos and Karlaftis 2003). In the current study no such exceptions have been made i.e. all the observations have been included for further analysis.

Bearing in mind that traffic speed prediction is much challenging when compared to traffic flow and occupancy due to the presence of vehicles exhibiting unusual behaviour i.e. vehicles travelling very fast or slow and could thereby distort the prediction performance.

Table 1. Data acquisition

Test site	Speed limit	Date	Total number of observations
Site-1	60km/hr	2015/7/1-2015/7/22	239 127
Site-2	40km/hr	2013/7/3-2013/8/14	63757
Site-3	40km/hr	2012/11/21-2012/12/18	58 559
Site-4	60km/hr	2015/4/16-2015/5/8	179711

3. Input selection

In order to develop any of the prediction models, selection of input data points need to be investigated (see section 4). The output of each model was the speed prediction in the future (one step ahead) whereas the input of the network was found to be by no means a trivial choice task. There were five inputs (mean speed, flow, standard deviation, time of the day and day of the week) available for the prediction model. At this point it is worth mentioning that including a temporal time factor within the traffic prediction is a common practice. This gives several possible data input points which make the computational effort of the model to rise. Therefore, to specify the inputs to our prediction model, we had to consider two major questions as follows

(1) At which level the aggregation of the data will be appropriate to the speed prediction?

¹ Site-1 (Latitude: 60.558988, Longitude: 15.137701)

² Site-2 (Latitude: 60.476904, Longitude: 15.464145)

³ Site-3 (Latitude: 60.462058, Longitude: 15.467076)

⁴ Site-4 (Latitude: 60.497165, Longitude: 15.452249)

(2) Which input features would give a promising prediction results?

In the current study the level of aggregation of the data were assessed under two scenarios to be able to answer the first question.

Scenario 1-Short term level of aggregation (5-min periods)

In this scenario, the objective is to predict traffic speed in short term period of time. Speed prediction in a short time of period was often used in previous studies to respond to traffic dynamics that may significantly change the traffic pattern over time. However the most common short term traffic prediction was used in 5-min periods into the future. Therefore in this scenario the aggregation level of the data is chosen to be 5-min. The output of the model will be the speed prediction for the next 5 min ($t+1$), with t representing the current time. The inputs variables are the historical inputs from previous time back into the past represented by now (t), 15min ($t-3$) and 30min ($t-7$) (see fig 2).

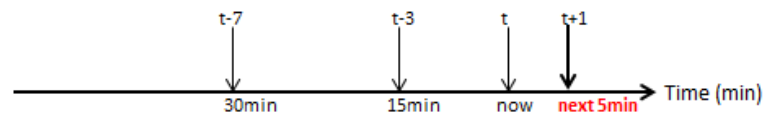


Fig.2 Short term speed prediction in 5-min period of time (Scenario 1)

Scenario 2- long term level of aggregation (1-hour period)

This scenario presents the long term speed prediction in 1 hour period of time. This prediction might be appropriate when traffic conditions have similar traffic characteristics where there is no need to change often the trigger speed. As per this scenario, the speed prediction is done for the next hour ($t+1$), with t representing the current hour by considering previous time step back into the past represented as now (t), 1day ($t-24$) and 1week ($t-168$) (see fig 3).

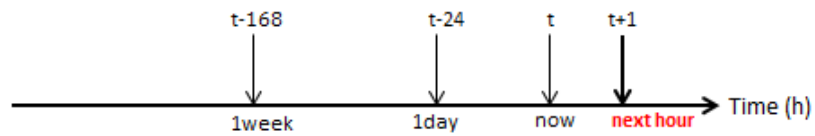


Fig.3 Long term speed prediction in 1-hour period of time (Scenario 2)

A practical answer to the second question (2) has been derived by computing the correlation matrix between the output (speed ($t+1$)) and the proposed inputs. When looking at the correlations matrix, presented in table 2 and 3, between the flow and the mean speed for the two time periods, it is clear that speed ($t+1$) is correlated to the previous speed and flow at time t . No clear correlation between the speed and standard deviation and the speed and time of the day. The correlation between the day of the week and speed is significantly low thus clearly shows unsuccessful input feature to the speed prediction. The correlations between the input features at time t and the output at time ($t+1$) are further computed in the tables below (table2 and table 3) for data collected at site 4.

Table 2: Correlation between inputs features at (t) and mean speed at (t+5) at site 1 for data aggregated in 5-min time period- Scenario 1

	Speed(t+1)	Flow (t)	Speed(t)	Std(t)	Hour(t)	Day(t)
Speed(t+1)	1.00	-0.64	0.60	-0.10	-0.24	-0.01
Flow (t)	-0.64	1.00	-0.65	0.10	0.41	-0.01
Speed(t)	0.60	-0.65	1.00	-0.20	-0.25	-0.01
Std(t)	-0.10	0.10	-0.20	1.00	0.07	0.03
Hour(t)	-0.24	0.41	-0.25	0.07	1.00	-0.01
Day(t)	-0.01	-0.01	-0.01	0.03	-0.01	1.00

Table 2: Correlation between inputs features at (t) and mean speed at (t+1) at site 1for data aggregated in 1-hour time period- Scenario2

	Speed(t+1)	Flow (t)	Speed(t)	Std(t)	Hour(t)	Day(t)
Speed(t+1)	1.00	-0.77	0.83	0.35	-0.35	-0.08
Flow (t)	-0.77	1.00	-0.81	-0.58	0.47	-0.01
Speed(t)	0.83	-0.81	1.00	0.32	-0.38	-0.08
Std(t)	-0.35	-0.58	0.32	1.00	-0.39	0.04
Hour(t)	-0.35	0.47	-0.38	-0.39	1.00	0.02
Day(t)	-0.08	-0.01	-0.08	0.04	0.02	1.00

As expected the absolute correlation generally shrinks as the time horizon goes back into the past. Therefore in this study the speed prediction for the next 5-min and for the next 1-hour are based on the present speed at time (t). The speed prediction based on time (t) was also compared to the speed prediction including mean speed for all previous steps (t, t-3 and t-7 for 5-min period and t, t-24 and t-168 for 1-hour period). Finally flow (t) and speed (t) have been chosen as the input variables for further analysis.

4. Prediction models

In the current study adaptive neuro fuzzy inference system (ANFIS), classification and regression trees (CART) and random forests (RF) were employed to predict speeds. Results achieved by the aforementioned models were evaluated and compared to more robust techniques such as artificial neural network (ANN), multiple linear regression (MLR) and a naïve prediction model using traffic speed data collected at four sites located in Sweden for the sake of validation. Data collected at each site data is split into training and testing sets. Following the practical rule of thumb 70% of the data used for training and 30% for testing; the training data set was used for determining the network parameters while the testing set was set for validating the performance of the trained models. Note that all the data sets were normalised for further analysis and evaluation. A brief description of the models has been provided for the benefit of the reader unfamiliar with the topic.

4.1 Artificial neural network-Multilayer perceptron

The multilayer perceptron is one of the most popular architecture of artificial neural network (ANN).

It consists of network of neurons called perceptron. The architecture of the perceptron is composed of three layers, input layer, hidden layer and an output layer. The perceptron computes a single output from real-valued inputs by forming combinations of linear relationships according to inputs weights using nonlinear transfer function. Hyperbolic tangent function and logistic sigmoid function are transfer functions that are often used in training the network. A logistic sigmoid function was chosen to this study. MLP are typically trained using the back propagation error algorithm in which the network's interconnecting weights are iteratively changed to minimize the predefined error (Detailed analysis of MLP and error back-propagation process can be found in Gardner and Dorling 1998, Dougherty 1995 and Rezaeianzadeh et al. 2013). In this study, a three-layer network with 50 neurons was trained with 60 epochs using the Levenberg–Marquardt algorithm. More description of Levenberg–Marquardt algorithm can also be found in (Nocedal and Wright 1999).

4.2 Adaptive neuro-fuzzy inference system

An adaptive neuro fuzzy system is a powerful system that combines the concepts of two approaches into one integrated system. In this integrated and fused system, ANN learning algorithms are used to determine the parameters of the Fuzzy inference system to share data structures and knowledge representations. A typical ANFIS structure, proposed first by Jang (Jang 1993), consists of 5 layers of nodes. The structure is used to map the input characteristics to the input memberships functions at the first layer to a set of rules at the second layer, rules to a set of output characteristics at the third layer, output characteristics to output memberships functions at layer 4, and in the end, the output memberships functions to a single output at the last layer (Khoshnevisan et al 2014, Jang and Sun 1995, Chang and Chang 2006, Ullah nd Choudhury 2013). In this paper, the fuzzy inference system is based on Takagi Sugeno methodology where the output membership functions are constant value. The network is trained using a hybrid learning algorithm based on two steps. In the first step (forward pass), the premise parameters, i.e. network parameters, are kept fixed and the information is propagated forward in the network using least square method to identify the consequent parameters for the current cycle through the training. Then, in the second step (backward pass), the error propagate backward while the premise parameters are modified using the gradient descended method and by keeping the consequent parameters fixed. The rule extraction method first uses the Fuzzy c-means (FCM) clustering function, known as genfis3, to determine the number of rules and membership functions for the antecedents and consequents. The Fuzzy c-means (FCM) clustering techniques (genfis3) was also used to optimise the result by extracting a set of rules that models the data and generate an initial FIS for ANFIS training.

4.3 Classification and Regression tree

Classification and regression tree (CART) is a nonparametric statistical methodology developed for analysing classification issues either from categorical or continuous dependent variables. If the dependent variable is categorical, CART produces a classification tree. When the dependent variable is continuous, it produces a regression tree. For the categorical and continuous variables, the CART generates binary decision trees which technically known as binary recursive partitioning. The development of these tree models is binary because each parent node is always split into exactly two child nodes and recursive because it is repetitive by treating each child node as a parent. Recursion for splitting each node in a tree will be based on a set of rules that decide when a tree is complete and decide what predicted value for each terminal node will be assigned (Breiman et al. 1985, Coppersmith et al. 1999, Mahjoobiab and Etemad Shahidi 2008, Almejalli 2010). Several aspects have been used for tuning CART performance. In this study, two parameters were used here to control these aspects. The first parameter is called *minsplit* which is the minimum number of observations that must exist in a node in order for a split to be attempted. The second one is the *complexity control*; the main role of this parameter is to save computing time by pruning off splits that are obviously not worthwhile.

4.4 Random Forest

A random forest is an ensemble machine learning proposed by Leo Breiman for building tree predictors and letting them vote for the most popular class. The algorithm of inducing random forest is based on bootstrap aggregation or so called bagging. For bagging, given a training set $X = x_1, \dots, x_n$ with responses variables $Y = y_1, \dots, y_n$, B times selects a random sample with replacement of the training set and fits tree regressions to these samples. After training, predictions can be made by averaging the predictions from all individual regression trees. The latter improve the stability and accuracy of machine learning algorithms in particular used in statistical classification and regression (James et al. 2013). It also reduces variance and helps to avoid overfitting. Random forests differ in only one way from the general bagging process. The algorithm use a modified tree learning algorithm that selects, at each candidate split in the learning process, a random subset of the features. This process is also called feature bagging. The reason for this is to overcome the correlation of the trees from an ordinary bootstrap sample. A detailed description can be found in (Breiman, 2001). For the implementation of the RF in this paper, the number of trees and the number of selected features were experimentally tuned. As the number of trees increases, the error converges to a limit where there are no presence of overfitting as in the case of MLP and other learning algorithms. The most important parameter is how many features to test for each split. The more useless features there are, the more features should be tried. Therefore this need to be tuned carefully by starting with a large number and then increasing and decreasing the number of features until the minimum error for the prediction is obtained.

5. Results and analysis

A key challenge in implementing the prediction models lies in identifying the optimal setting for each of the models. Following this, further search over the parameters which needs to be tuned was carried out on the different training and test data. Tuning parameters in such models are mainly to increase the predictive power of the model. In other words, through a specific range of parameters, the models were trained with the training data set and iterated. The trained models were then checked through

the test data in order evaluate the performance of the models. To evaluate the prediction performance of each of the models, root mean square error (RMSE) was employed (see equation 1):

$$RMSE = \sqrt{\frac{1}{n} \sum (P_i - A_i)^2} \quad (1)$$

Where P_i and A_i are the predicted and actual yield for the i th speed prediction and n is the total number of speed in the data set.

Table 2 and 3 summarises the RMSE performance for the next 5-min speeds and for the next 1-hour predicted by the models on the previous mean speed and flows on training and test data for the four sites. All models are also compared to a naïve prediction and to a multiple linear regression (MLR). As mentioned earlier the naïve prediction model in the current case has assumed current speed value as the predicted speed (see section 1.). When looking to the results on both tables, the speed prediction provided by RF appear the most successful model compared to other models. Unsurprisingly all the models reported better performance than the naïve model. When comparing the results between the two tables, the RMSE reveals some other interesting issues. First the prediction for the next 5min (in particular for the testing dataset) again performed better than the prediction for the next hour, thus confirming the speed prediction in long term appear to be inappropriate to the traffic dynamics that may vary over time of the day. Another interesting remark is that the results obtained at site 1 in table2 performed best of all sites, adapting well to the data patterns. One explanation for this may be that the location of this site on highway differs completely to the other sites located in town.

The speed prediction of each model is further depicted in Fig.2 and Fig.3. As it can be seen, the prediction accuracy of RF, ANFIS and ANN with respect to the actual speed is quite acceptable compared to MLR and CART. The only insignificant speed prediction refers to either rare high or low speed values when an extremely change in speed happens. Otherwise, it can be seen that those models tracks the actual speed profile smoothly, i.e. excluding the inappropriate speeds (too fast or too slow) from the dataset. By excluding such inappropriate speeds, the model will further cause a significant impact on the prediction of the trigger speed of VAS.

This study clearly demonstrates that simplistic selection of input quantity to the model yields a promising result. A good example of this can be seen with the comparative results between two different set of input data from site 1, i.e. the first input data is chosen at one time slice ($t-0$) and the second one is at three consecutive time slices. The model using all input data points, showed in table 4, did not exhibit better forecasting performance. However, in terms of computational complexity, it had a marked increase in the calculation time of the model.

Given these results, it worth to point out that RF are adequate model to predict trigger speed for the VAS in terms of computational performance (shorter calculation time) and in terms of efficiency (lower RMSE).

Table 2- Performance (RMSE) of the prediction of next 5 min speeds based on the previous 5 min at different sites respective to several methods- Scenario 1

Test site	Dataset	MLR	CART	RF	ANFIS	ANN	Naive
Site1	Training	0.05	0.04	0.03	0.05	0.05	0.07
	Testing	0.07	0.06	0.05	0.05	0.05	0.06
Site2	Training	0.07	0.07	0.06	0.07	0.07	0.10

Site3	Testing	0.08	0.07	0.06	0.07	0.07	0.09
	Training	0.06	0.06	0.06	0.07	0.07	0.09
Site4	Testing	0.08	0.09	0.06	0.08	0.08	0.11
	Training	0.06	0.07	0.06	0.08	0.08	0.09
	Testing	0.08	0.08	0.06	0.08	0.08	0.11

Table 3- Performance (RMSE) of the prediction of next hour speeds based on the previous hour at different sites respective to several methods- Scenario 2

Test site	Dataset	MLR	CART	RF	ANFIS	ANN	Naive
Site1	Training	0.06	0.02	0.02	0.06	0.06	0.09
	Testing	0.06	0.11	0.06	0.09	0.06	0.08
Site2	Training	0.09	0.04	0.05	0.07	0.09	0.13
	Testing	0.07	0.09	0.06	0.09	0.08	0.11
Site3	Training	0.04	0.02	0.02	0.03	0.03	0.04
	Testing	0.12	0.12	0.11	0.14	0.11	0.13
Site4	Training	0.09	0.06	0.05	0.09	0.09	0.11
	Testing	0.23	0.09	0.07	0.09	0.07	0.07

Table 4- Results of the performance measures of ANFIS, ANN and RF models respective to different input data sets at site 1 (testing data)- Scenario1

Data	ANFIS model		ANN model		RF model	
	RMSE	Time(s)	RMSE	Time(s)	RMSE	Time(s)
Data from (t-0)	0.05	1.96	0.05	102.51	0.05	2.12
Super Data from (t-0, t-3 and t-7)	0.06	7.59	0.05	175.57	0.05	4.90

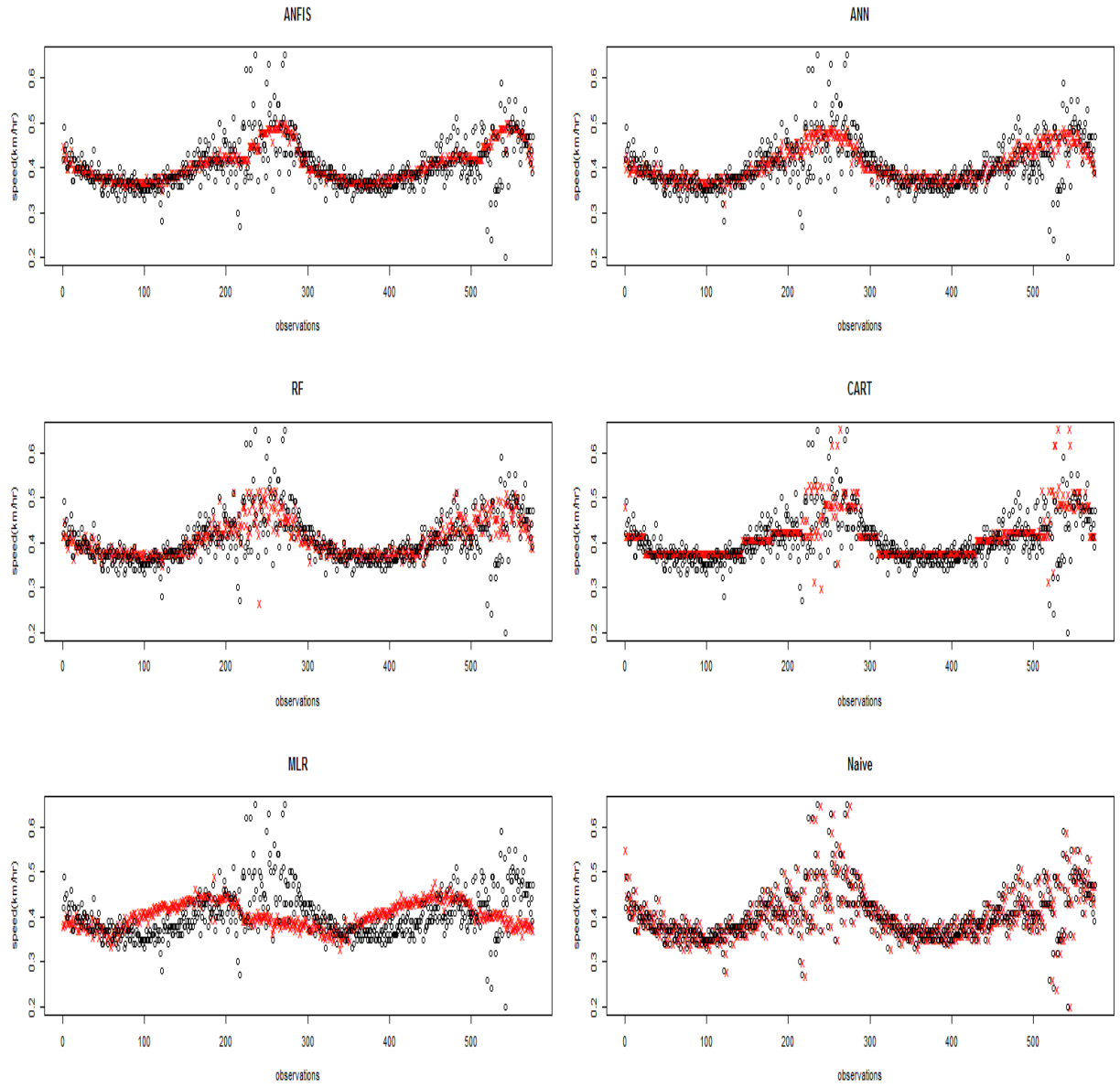


Fig.2. Speed prediction for next 5 min using ANFIS, ANN, RF, CART, MLR and Naïve models (Scenario 1)

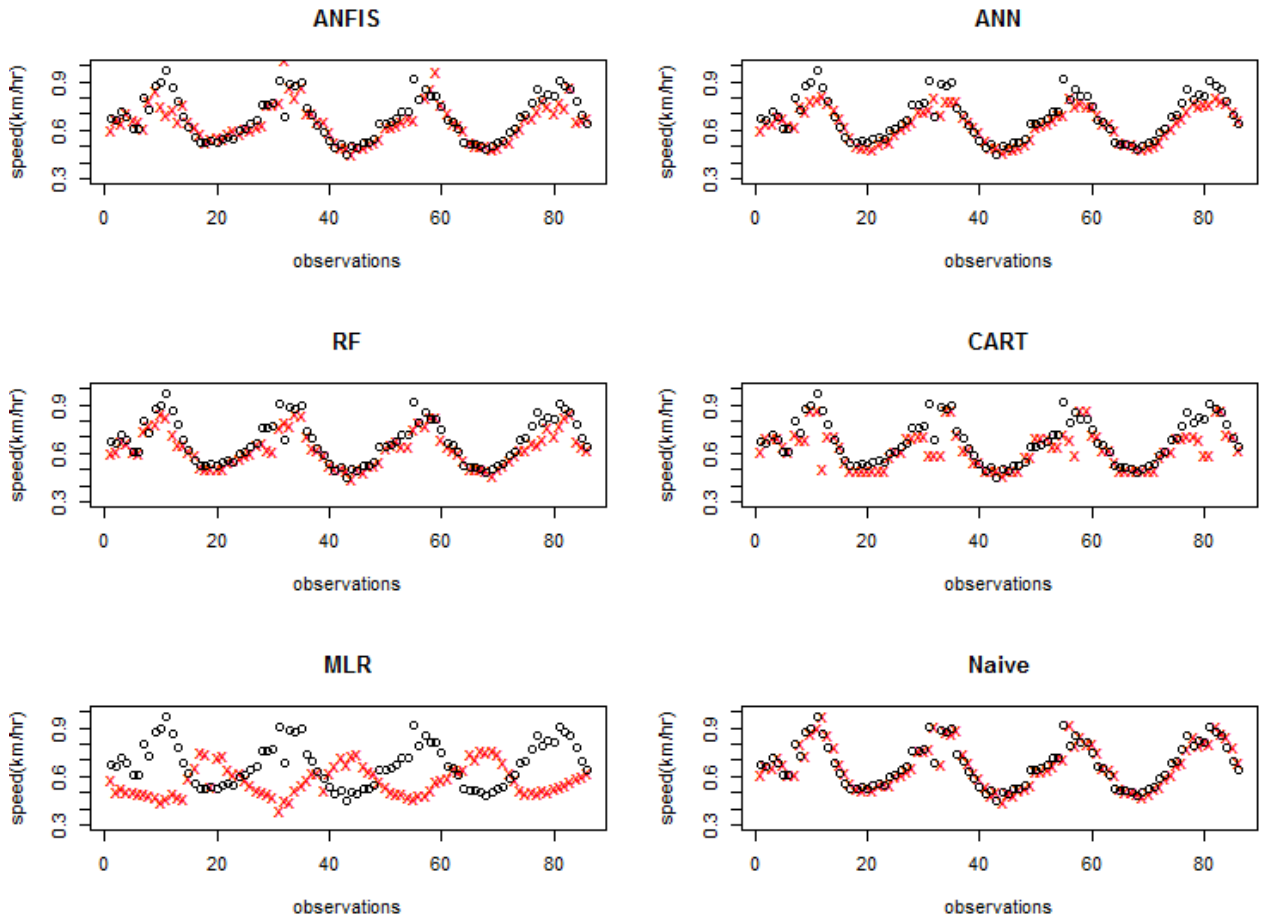


Fig.3. Speed prediction for next hour using ANFIS, ANN, RF, CART, MLR and Naïve models (Scenario 2)

6. Conclusion and future work

In this paper, accurate predictive models based on historical traffic speed data were investigated to derive the optimal trigger speed for VAS. ANFIS, CART and RF were employed to predict speeds and compared to the ANN, MLP and naïve base model. The inputs to the models were based on mean speed and traffic flow for one step data back into the past. The data were also aggregated in two time periods (5-min and 1-hour). The conclusions arising from this study are as follows:

- The results of using RF to predict mean speed show promising results in terms of performance. It is suggested that future work should investigate boosting random forest to maintain high performance particularly decreases the memory requirements of the system with a smaller number of decision trees.
- Analysis of the input selection is helpful in finding the simplistic inputs to the model. This simplistic input proved effective result in particular when shorter response time is required to the dynamic system. A filtering technique can be combined with the predictive model in order to automate the input selection.
- Regarding the incorporation of combination of the temporal characteristics, the use of one step back into the past was found to exhibit similar prediction results to the cases of including multiples steps backward. This result is promising when modelling adaptive system where time and memory requirements are necessary. While this results is more surprising compared to flow

predictions when approaches gave improved predictions including up at least to three steps back into the past.

- Selection of which level of aggregation of the data is strongly related to the purpose of the speed prediction. For the purpose of triggering the VAS, traffic engineering may attempt to use higher level of aggregation when there is no need to often change the trigger speed of such warning signs. Part of this can be attributed to the avoidance of causing ambiguity and confusion to drivers. In this context it is suggested that future work should investigate first a clustering algorithm such k-means or self-organising maps to group speed data with similar traffic characteristics and then prediction model should be sought.
- It was found that the correlation between the time of the day and the day of the week and the mean speed were weak. This was not surprising because the mean speeds are very similar in most traffic conditions. There are also similarities between the three geographical sites (2, 3 and 4). Therefore different sites need to be explored to generalise this conclusion.

References

- Almejalli, K.A. (2010). Intelligent real-time decision support systems for road traffic management, Doctor of Philosophy, Bradford, Great Britain.
- Breiman L. (2001). Random Forests, Machine Learning, 45, 5-32.
- Breiman L, Friedman JH, Olshen RA and Stone CJ. (1984). Classification and regression trees. Belmont: Wadsworth Statistical Press.
- Chang F.J and Chang Y-T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir, Advances in Water Resources, 29, 1–10.
- Chiou, Y.C., Lan, L.W., Tseng, C.M. (2014). A novel method to predict traffic features based on rolling self-structured traffic patterns. Journal of intelligent transportation systems. 18(4), 352-366.
- Chen H., Grant-Muller S. (2001). Use of sequential learning for short-term traffic flow forecasting. Transportation Research Part C, 9, 319–336
- Coppersmith, D., Hong S. J., and J. R. M. Hosking. (1999). Partitioning Nominal Attributes in Decision Trees." Data Mining and Knowledge Discovery, 3, 197–217.
- Dimitriou L., Tsekeris T. and Stathopoulos A. (2008). Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow Transportation Research Part C, 16, 554–573
- Dougherty, M. (1995). A review of neural networks applied to transport, Transportation Research Part C: Emerging Technologies, 3 (4), 247–260.
- Dougherty M. and Cobbett M. (1997). Short-term inter-urban traffic forecasts using neural networks, International journal of forecasting, 13(1), 21–31.

- Gardner M.W., Dorling S.R. (1998). Artificial neural networks (the multi-layer perceptron) – a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32, 2627–2636
- Gazis, D. and Liu, C. (2003). Kalman filtering estimation of traffic counts for two network links in tandem. *Transportation Research Part B*. 37(8), 737–745.
- Hamed, M.M., Al-Masaeid, H.R. and Bani Said, Z.M. (1995). Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering*, 121 (3), 249–254.
- Hooshdar S., Adeli H. (2004). Toward intelligent variable message signs in freeway work zones: A neural network approach. *Journal of Transportation Engineering*, 130(1). 83–93.
- Dia H. (2001). An object-oriented neural network approach to short-term traffic forecasting. *European Journal of Operational Research* 131, 253–261.
- Jang JSR. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern*, 23(3), 665–685.
- Jang JSR. and Sun CT. (1995). Neuro-fuzzy modelling and control. *Proc IEEE* 83(3), 378–406.
- James G., Witten D., Hastie T. and Tibshirani R. (2013). *An Introduction to Statistical Learning*. Springer. pp. 316–321.
- Jomaa D., Dougherty, M, Yella S. and Edvardsson K. (2014). Effectiveness of Trigger Speed of Vehicle Activated Signs on Mean and Standard Deviation of Speed. *Journal of Transportation Safety & Security* (in press).
- Khoshnevisan, B., Rafiee S.H., M. Omid, and H. Mousazade. (2014). Development of intelligent based on ANFIS for predicting wheat grain yield on the basis of energy inputs. *Information processing in agriculture*. 1(1), 14-22.
- Kwon, E., Stephanedes, Y.J., (1994). Comparative evaluation of adaptive and neural-network exit demand prediction for freeway control. *Transportation Research Record*. 1446, 66–76.
- Lin, S. W., & Lin, Y. (2007). Urban expressway control based on fuzzy logic and variable speed limit. *Journal of Transportation Systems Engineering and Information Technology*, 2(1): 50 - 56.
- Li T, Zhao D, Yi J. (2008). Adaptive Dynamic Neuro-Fuzzy System for Traffic Signal Control. In: *IEEE International Joint Conference on Neural Networks, IJCNN08*, Hong Kong. 1840-1846.
- Mahjoobiab J., Etemad-Shahidi A. (2008). An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Applied Ocean Research* 30, 172-177.
- Nocedal J. and Wright S.J. (1999). *Numerical Optimization*. Springer, New York.
- Park B. (2002). Hybrid Neuro-Fuzzy Application in Short-Term Freeway Traffic Volume Forecasting. *Transportation Research Record: Journal of the Transportation Research Board*. 1802(361):190-196.

- Park J., Li D., Murphey Y., Kristinsson J., McGee R., Kuang M., and Phillips T. (2011). Real time vehicle speed prediction using a neural network traffic model, The 2011 International Joint Conference on Neural Networks (IJCNN). IEEE, San Jose, CA. 2991–2996.
- Rezaeianzadeh M., Tabari H., Yazdi A. A., Isik S. and Kalin L. (2013) Flood flow forecasting using ANN, ANFIS and regression models, Neural Comput & Applic, DOI 10.1007/s00521-013-1443-6.
- Sadek A, Smith B, and Demetsky M. (2001). A Prototype Case-Based Reasoning System for Real-Time Freeway Traffic Routing. Transportation Research Part C. Emerging Technologies. 9(5):353-380.
- Sisiopiku, V., Sullivan A., and Fadel G. (2009). Implementing active traffic management strategies in the U.S., Management and safety of transportation, UTCA, 08206, Birmingham.
- Smith, B.L. and Demetsky, M., (1997). Traffic flow forecasting: comparison of modelling approaches. Journal of Transportation Engineering. 123 (4), 261–266.
- Smulders, S. 1990. Control of Freeway Traffic Flow by Variable Speed Signs. Transportation Research Part B, 24(2), 111-132.
- Smulders, S. and Helleman, D. E. (1998). Variable speed control: State-of-the-art and synthesis. In proceedings on Road Transport Information and Control IEEE, 454,155-159.
- Stathopoulos A. and Karlaftis M. G. (2003). A multivariate state space approach for urban traffic flow modeling and prediction. Transportation Research Part C, 11(2), 121-135.
- Ullah N. and Choudhury P., (2013) Flood Flow Modeling in a River System Using Adaptive Neuro-Fuzzy Inference System, Environmental Management and Sustainable Development, 2(2), 2164-7682.
- Van Der Voort M., Dougherty M., Watson S. (1996). Combining Kohonen maps with ARIMA time series models to prediction traffic flow. Transportation Research Part C. 4 (5), 307–318.
- Vlahogianni, E.I., Karlaftis M.G. and Golias J.C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach, Transportation Research Part C: Emerging Technologies, 13 (2), 211–234
- Vlahogianni E.I., Karlaftis M.G., Golias J.C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach, Transportation Research Part C: Emerging Technologies. 13 (2), 211–234
- Walter, L.K. & Knowles, J. (2008). Effectiveness of Speed Indicator Devices on reducing vehicle speeds in London. TRL report TRL 314, Transport Research Laboratory.
- Wen H, Yang Z, Jiang G, Shao C. (2001). A New Algorithm of Incident Detection on Freeways. In: IEEE International Vehicle Electronics Conference, IVEC01, Tottori, Japan. 197-202.
- Yin, H.B., Wong, S.C., Xu, J.M., Wong, C.K. (2002). Urban traffic flow prediction using a fuzzy-neural approach. Transportation Research Part C. 10 (2), 85–98.

Zhang Y., Ye Z. (2008). Short-term traffic flow forecasting using fuzzy logic system methods. *Journal of Intelligent Transportation Systems*. 12 (3), 102–112