



A COMPARISON OF THE PERFORMANCE OF ANN AND SVM FOR THE PREDICTION OF TRAFFIC ACCIDENT DURATION

B. Yu*, Y.T. Wang[†], J.B. Yao[‡], J.Y. Wang[§]

Abstract: The prediction of traffic accident duration is great significant for rapid disposal of traffic accidents, especially for fast rescue of traffic accidents and removing traffic safety hazards. In this paper, two methods, which are based on artificial neural network (ANN) and support vector machine (SVM), are adopted for the accident duration prediction. The proposed method is demonstrated by a case study using data on approximately 235 accidents that occurred on freeways located between Dalian and Shenyang, from 2012 to 2014. The mean absolute error (MAE), the root mean square error (RMSE) and the mean absolute percentage error (MAPE) are used to evaluate the performances of the two measures. The conclusions are as follows: Both ANN and SVM models had the ability to predict traffic accident duration within acceptable limits. The ANN model gets a better result for long duration incident cases. The comprehensive performance of the SVM model is better than the ANN model for the traffic accident duration prediction.

Key words: *prediction, freeway accidents duration, artificial neural networks, Support Vector Machines*

Received: July 16, 2015

DOI: 10.14311/NNW.2016.26.015

Revised and accepted: April 2, 2016

1. Introduction

1.1 Backgrounds

With the increasing of the highway mileage and vehicle quantity, the incidence of traffic accident also increases in our country these years. The traffic congestion caused by traffic accidents has brought a series of unfavorable results to people's lives. Moreover, this phenomenon is becoming more and more serious, and it has become a severity societal problem which has attached our attention. When an

*Yu B., School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, PR China

[†]Wang Y.T., Transportation Management College, Dalian Maritime University, Dalian 116026, PR China

[‡]Yao J.B. – Corresponding author, School of civil engineering, Beijing Jiaotong University, Beijing 100044, PR China, Email: jbyao@bjtu.edu.cn

[§]Wang J.Y., School of Management, Ocean University of China, Qingdao City 266100, China

accident happens, timely prediction of its duration is thought as a key role in the field of traffic controlling, vehicle guidance and casualties curing.

In addition, the incident duration is affected by many factors which are difficult to predict. Determining and understanding these factors can help the process of identifying and developing better strategies to reduce incident durations and alleviate traffic congestion [40]. Freeway incident duration is the time interval from the occurrence of the incident to the road clear [13]. Incident duration consists of three stages: reporting, response and clearance time (Fig. 1). It is not only a key index to estimate incident severity, but also an important factor to determine how to dispose traffic flow at the site of the incident [9]. In summary, predict incident duration timely can assist the manager to determine the best emergency rescue and traffic control strategies, so as to effectively reduce traffic delay and improve the level of rescue.

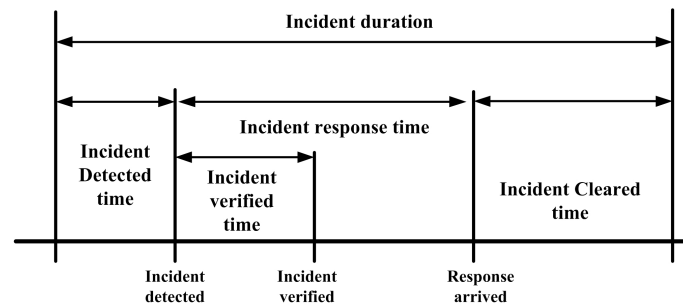


Fig. 1 Components of incident duration.

1.2 Literature review

A number of studies have been carried out in the area of accident duration prediction. Ozbay and Kachroo [17] focused on incident having major impact on traffic and proposed the use of decision tree method. The decision tree method is effective for an actual incident management process. However, in some cases, the decision trees can be unstable and insensitive to the stochastic nature of data.

Nam and Mannering [15] applied hazard-based duration models to statistically evaluate the time it takes detect/report, respond to, and clear incidents. The temporal characteristics, environmental characteristics, geographic information and incident characteristics were considered. The results showed that a wide variety of factors significantly affected incident duration, and different distributional assumptions for the hazard function were appropriate for the different incident times being considered.

Liu et al. [14] applied stepwise regression analysis method to establish a statistical model for estimating incident duration. The test result shows that over 85% of variations in incident duration can be forecasted by the eight factors involved in the regression model.

Ozbay and Noyan [18] applied Bayesian networks approach to study the estimation problem of incident clearance time. They try to develop a model which can

automatically learn emerging patterns in data to aid in the prediction of incident clearance times.

Chung [3] modeled the accident duration by hazard-based duration model, in particular the log-logistic accelerated failure time (AFT) metric. This study is based on a 2-year accident dataset from the Korean Freeway Systems. The estimated duration model based on the year 2006 was validated for the prediction accuracy with acceptable reasonableness.

As one of the most widely-used data mining technique for solving complex problems, artificial neural network (ANN) has being widely used in transportation studies. It is known as standard nonlinear estimators and its capability has been validated in different fields. The ANN processes information by means of interaction between many neurons and the different links between neurons have been associated with weights. Based on the highly interconnected neural computing elements, it has the ability to model complex relationships between inputs and outputs. Many studies demonstrate that ANN has the potential to accurately predict freeway traffic conditions [7, 31, 32, 30], traffic accidents prediction [36], and incident detection [10, 11].

Support vector machine (SVM), which were introduced by Vapnik [25, 26], is a relatively new structure in the data-driven prediction field. The SVM is a type of learning algorithms based on statistical learning theory, which can be adjusted to map the input–output relationship for the non-linear system. It has been applied to solve the forecasting problem with small samples and non-linear [1, 16, 20, 28]. In addition, the solution of SVM is always unique and globally optimal since training SVM is equivalent to solving a linearly constrained quadratic programming problem. Therefore, SVM shows the strong resistance to the over-fitting problem and the high generalization performance. Many scholars have done research on it, and have got successful applications in many aspects, such as bus arrival time prediction [34, 39], traffic speed prediction [27], and accurate internet traffic classification [8, 37]. However, there have been only a few studies related to the prediction of freeway accident duration.

These successful applications motivate us to apply ANN and SVM to the prediction of traffic accident duration, and the performances of the two models are compared to determine which model is better in terms of freeway accident duration forecasting.

Thus, this paper is organized as follows Section 2 provides the formulation of four proposed models for predicting freeway accident duration; Section 3 presents a case study together with results and analysis including performance evaluation of the two proposed models; and lastly, the conclusions and the suggestions for further study are given in Section 4.

2. Methodologies

2.1 Artificial neural network

The ANNs is a data-driven, self-adaptive, and nonlinear methodology [4]. Although there are numerous types of ANNs, the most commonly used type of ANN is the Multi-Layer Perceptron (MLP), which is made up of input, hidden, and output

layers with their nodes and activation functions [5]. It is generally easy to use and can approximate almost any input/output relationships. The mathematical expression of the MLPN is as follows:

$$y_j = f \left(\sum_{i=1}^N w_{ji}x_i + b_j \right), \quad (1)$$

where x_i is the i -th nodal value in the previous layer, y_j is the j -th nodal value in the present layer, b_j is the bias of the j -th node in the present layer, w_{ji} is a weight connecting x_i and y_j , N is the number of nodes in the previous layer, and f is the activation function in the present layer. Fig. 2 shows a schematic diagram of the MLP used in this study.

When inputting variables into the network, the weights from input layer to hidden layer are calculated. Through the transfer function in the hidden layer, the input data are rescaled as inputs to the output layer [6, 19]. Since a discrepancy might occur between the estimated output and the actual accident duration, the weights are adjusted repeatedly by a suitable training method until the resulting error is stabilized and negligible [2, 12, 23].

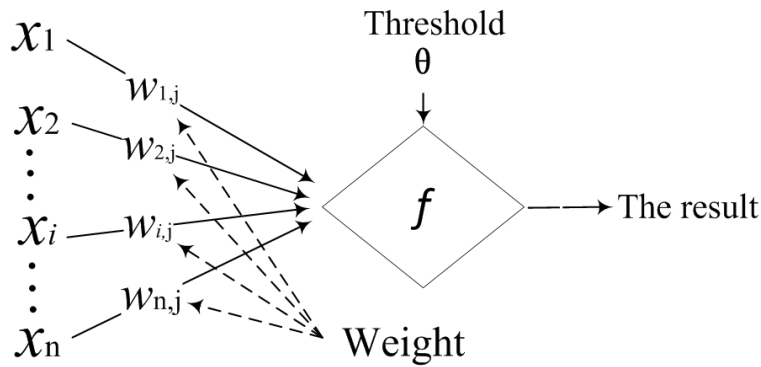


Fig. 2 Structure of ANN.

2.2 Support Vector Machine model

The SVM is a supervised learning machine in the framework of statistical learning theory, based on the Structural Risk Minimization Theory (SRM) developed by Vapnik [24]. The SVM model has the advantages of strong learning ability in small sample situation, fast learning speed and good generalization ability and so on.

Input vectors supporting the model structure are selected through a model training process described below. Given by a set of training samples of $\{x_i, y_i\}_{i=1}^n$, $x_i \in R^N, y_i \in R$, where x_i is an input vector of m components and y_i is a corresponding output value, an SVM estimator (f) on regression can be expressed as

$$f(x) = w \cdot \phi(x) + b, \quad (2)$$

where \emptyset denotes a nonlinear transfer function that maps the input vectors into a high-dimensional feature space in which the sample data are linearly separable. w is the weight and b is the offset, which can be estimated by minimizing Eq. (3):

$$\min R_{\text{SVM}}(C) = \frac{1}{2}ww^T + \frac{C}{n} \sum_{i=1}^n L(y_i, f(x_i)). \tag{3}$$

Define the ε -insensitive loss function:

$$L(y, f(x)) = |y - f(x)| = \begin{cases} 0 & \text{if } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{otherwise} \end{cases}. \tag{4}$$

$L(y, f(x))$ is the loss function. Loss function is used to measure the degree of prediction error. For a given input x_i , $f(x)$ gives the corresponding result which may be different from y . ε -insensitive can allow the presence of prediction error within a certain range and ensure that the model can find the optimal solution. Empirical risk is the difference between forecasted and real value. The first part of the empirical risk on the right in Eq. (3) can be estimated by the non-sensitive loss function given in Eq. (4), and the second part is the regularization confidence. Regularization confidence is a protection method which can avoid over fitting. Regularization confidence works by introducing certain restrictions which can reduce the complexity of the machine learning model. The structural risk minimization principle of the SVM is a compromise considering the empirical risk and confidence limit, minimizing the expected risk and preventing over-learning problems. The value of ε affects the support vector size, and C is the regularization parameter, which controls the degree of punishment beyond the error of the sample. By introducing a relaxation of nonnegative variables ξ_i and ξ_i^* , the objective function equation of the support vector regression machine (3) can be transferred into Eq. (5) as follows:

$$\begin{aligned} \min (w, b, \varepsilon, \varepsilon^*) \quad & \frac{1}{2}ww^T + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{subject to} \quad & y_i - w^T \emptyset(x_i) - b \leq \varepsilon + \xi_i, \\ & w^T \emptyset(x_i) + b - y_i \leq \varepsilon + \xi_i^*, \\ & \xi_i \geq 0, \xi_i^* \geq 0. \end{aligned} \tag{5}$$

Finally, by introducing the Lagrange multiplier, the optimization problem is converted into a dual problem:

$$\begin{aligned} R(a_i, a_i^*) = \quad & \sum_{i=1}^n d_i(a_i - a_i^*) - \varepsilon \sum_{i=1}^n d_i(a_i + a_i^*) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i - a_i^*)(a_j - a_j^*)K(x_i x_j) \\ \text{subject to} \quad & \sum_{i=1}^n (a_i - a_i^*) = 0, 0 \leq a_i, a_i^* \leq C. \end{aligned} \tag{6}$$

The decision function shown in Eq. (2) is correspondingly converted into:

$$R(a_i, a_i^*) = \sum_{i=1}^n (a_i - a_i^*) K(x, x_i) + b. \tag{7}$$

In Eq. (7), $K(x, x_i)$ is the kernel function. Essentially, the kernel function is a mapping function. In order to reduce the algorithm complexity by using the kernel function, kernel functions can convert a nonlinear learning problem into a linear learning problem. a_i and a_i^* are the corresponding Lagrange multipliers, where $a_i, a_i^* = 0$ only if the corresponding data sample point of $a_i \neq a_i^*$ is defined as the SVM. SVM shows the strong resistance to the over-fitting problem and the high generalization performance [22]. It is mainly because SVM can construct a mapping from one-dimensional input vector into high-dimensional space by the use of reproducing kernels. The architecture of SVM is shown in Fig. 3.

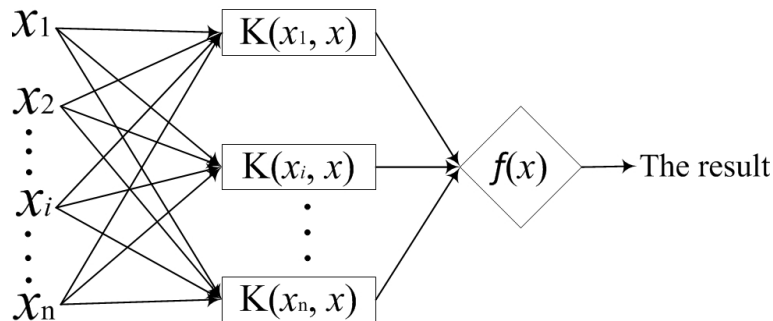


Fig. 3 Structure of SVM.

2.3 Traffic accident duration prediction based on SVM/ANN

When choosing the impact factors of freeway accident duration, the factors must be able to fully reflect the impact of freeway accident duration. In order to build a successful model, it is necessary to choose the appropriate variables.

When selecting the independent variables, in this paper, we pay attention to the two principles:

- 1) The independent variables should be closely related to the forecast object.
- 2) There is no strong linear relationship between the independent variables.

Using a large quantity of data features as the model inputs without careful processes may lead to the model significant noise. Therefore, data feature reduction with cluster method aims to decrease the number of model inputs and to preserve the relevant traffic characteristics with fewer inputs. Cluster analysis is a set of methodologies for automatic classification of samples into a number of groups using a measure of association, so that the samples in one group are similar and samples

belonging to different groups are not similar [21]. The distance between the two instances is used as a measure of similarity.

K-mean clustering method is adopted in this study. The procedure of the cluster analysis to reduce the data feature is as follows:

-
- Step 1: Select '*k*' initial cluster centroids or seeds, where '*k*' is the number of clusters desired.
 - Step 2: Assign each observation to the cluster to which it is the closest.
 - Step 3: Reassign or reallocate each observation to one of the *k* clusters according to a predetermined stopping rule.
 - Step 4: Stop if there is no reallocation of data points or if the reassignment satisfies the criteria set by the stopping rule. Otherwise go to Step 2.
-

In order to reduce the model significant noise, the factors with strong correlation should be excluded. By cluster analysis, the result shows that the traffic jam and the peak hour have a strong correlation, the burning vehicles and the hazard material involved have a strong correlation, the truck involved and the rollover vehicle involved. After that, the three variables (traffic jam, burning vehicles and truck involved) are excluded. The best combination of variables is determined for forecasting the freeway incident duration. They are night, casualties, peak hour, bad weather, facility damage, disabled vehicle, heavy tow truck, lay-by occupied, hazard material involved and rollover vehicle involved. Fig. 4 shows the structure diagram of freeway accident duration forecasting. The contribution of each variable in predicting is weighted and weights number is set reasonably. Tab. I shows the value of each significant variable.

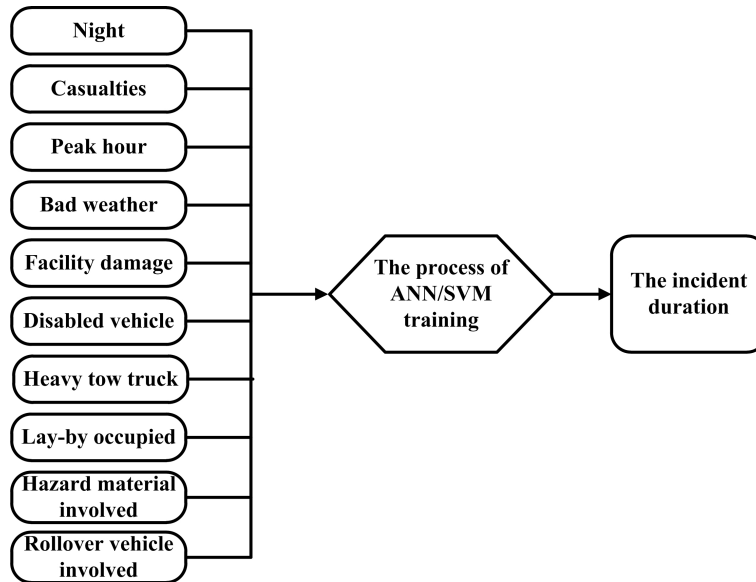


Fig. 4 The structure diagram of freeway accident duration forecasting.

Variables	Value
Night	1 = Yes; 0 = No
Casualties	1 = Yes; 0 = No
Peak hour	1 = Yes; 0 = No
Bad weather	1 = Yes; 0 = No
Facility damage	1 = Yes; 0 = No
Disabled vehicle	1 = Yes; 0 = No
Heavy tow truck	1 = Yes; 0 = No
Lay-by occupied	1 = Yes; 0 = No
Hazard material involved	1 = Yes; 0 = No
Rollover vehicle involved	1 = Yes; 0 = No

Tab. I The value of each significant variable.

Some variables that resulted significant for many models in other literature are found not determinant in this study. The reason for these apparently dissimilarity probably lies in the limited number of incident events.

3. Case study

The data was obtained from the database of one freeway management center in Liaoning. In this database, the incidents occurred on this freeway from 2012 to 2014 were recorded. The freeway with eight lanes in two directions is located between Dalian [35, 38, 33] and Shenyang (Fig. 5). The length of the freeway is 349 kilometers and its speed limit is 120km/h.

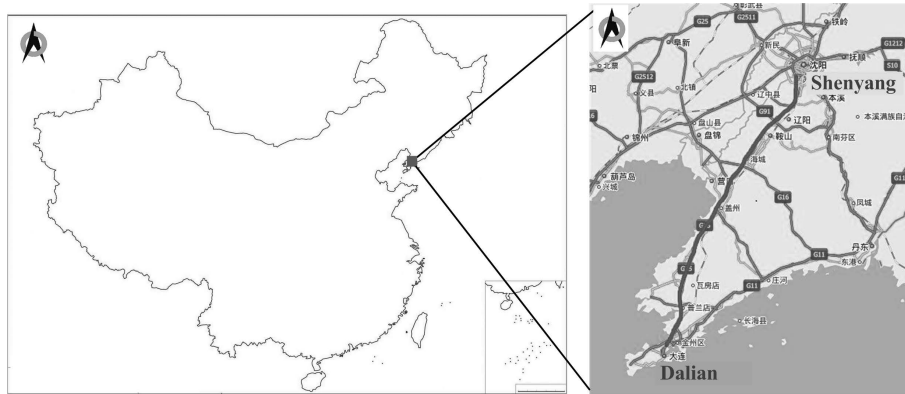


Fig. 5 The freeway from Dalian to Shenyang.

These data are normally used for monitoring incident management operations. All the records of the database contain the time of accident occurrence, the type of the accident, the location of the accident and so on.

As reported in previous studies, the models' prediction ability is heavily affected by the quality of the input data [29]. In order to improve the accuracy of prediction, in this study, the data items with serious errors or some important parameters missing are eliminated. After the previous processing, a data set consisting of 235 incidents is obtained. The frequency distribution of incident duration is shown in Fig. 6. The traffic accident statistics are shown in Tab. II.

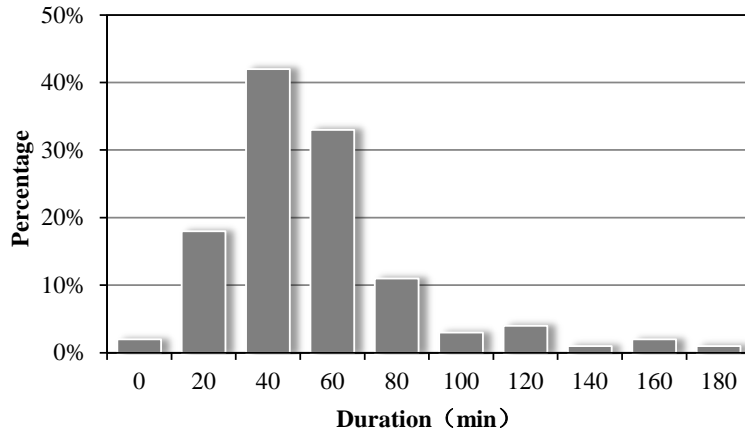


Fig. 6 Distribution of freeway incident duration.

About 30% of the incidents duration is 30 min or less, whereas 75% of the incidents duration is less than 60 min. Only 10% of the incidents duration is longer than 100 min.

3.1 Performance measures

The mean absolute error (MAE), the root mean square error (RMSE) and the mean absolute percentage error (MAPE) are used to evaluate the performances of the two measures as follows:

$$MAE = \frac{\sum |t - \bar{t}|}{N}, \tag{8}$$

$$RMSE = \sqrt{\frac{\sum (t - \bar{t})^2}{N - 1}}, \tag{9}$$

$$MAPE = \frac{1}{N} \sum \frac{|t - \bar{t}|}{t} \times 100\%, \tag{10}$$

where t is the actual time of the incident duration. \bar{t} is the predicted time of the incident duration. N is the number of the incident.

3.2 Model identifications

In this paper, night, casualties, peak hour, bad weather, facility damage, disabled vehicle, heavy tow truck, lay-by occupied, hazard material involved, rollover vehicle

		Variables									
Number	Night	Casualties	Peak hour	Bad weather	Facility damage	Disabled vehicle	Heavy truck	Lay-by occupied	Hazard material involved	Rollover vehicle involved	Incident duration [min]
1	Yes	No	No	Yes	No	Yes	No	No	Yes	No	46
2	No	No	Yes	No	Yes	Yes	Yes	No	No	Yes	32
3	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No	99
4	Yes	No	No	No	No	No	No	Yes	Yes	Yes	75
5	No	Yes	No	Yes	Yes	No	Yes	No	No	No	67
6	No	Yes	No	No	No	Yes	No	No	Yes	No	25
7	No	Yes	No	No	Yes	No	Yes	No	No	Yes	110
8	Yes	No	No	Yes	No	Yes	No	Yes	No	No	37
...

Tab. II *The traffic accident statistics.*

involved are used as input data, and the incident duration is used as output data. As shown in Tab. I. To assess the predictive ability of the two models, the original dataset is separated into training and testing partitions. In this study, 165 incident cases are included in the training partition for the model construction process. 70 incident cases are used to evaluate the accuracy of the proposed models.

3.2.1 Artificial neural network

In this study, the number of neurons in the input layer is determined by the 10 most significant variables affecting incident duration. A single neuron in the output layer is the prediction of incident duration. ANN with two or three hidden layers was trained using the Levenberg-Marquardt back-propagation algorithm. After that, the better performing ANN architecture with three-layer network structure is adopted. When inputting variables into the network, the weights from input layer to hidden layer are calculated. Through the transfer function in the hidden layer, the input data are rescaled as inputs to the output layer. Since a discrepancy might occur between the estimated output value and the actual incident duration, the weights are adjusted repeatedly by a suitable training method until the resulting error is stabilized and negligible. The best performing ANN architecture is obtained with a single hidden layer of 15 neurons. Then a three-layer network structure is adopted. There are 10, 15 and 1 nodes in the input, hidden and output layers, individually. After model training, the mapping between variables and accident duration is formed. Then by inputting the test data into this trained model, the forecasted accident duration is produced.

The MAE value for this ANN model is equal to 17 min, and 32 out of the 70 incident cases have been predicted with an absolute error less than 25 min. The comparison of the ANN model results and the actual durations is presented in Fig. 7. The ANN model has a satisfactory accuracy for the incident whose duration is longer than 60 min. However the ANN model tends to overestimate the prediction values for the short duration incident cases (0–30 min). The MAE value is equal to 16 min.

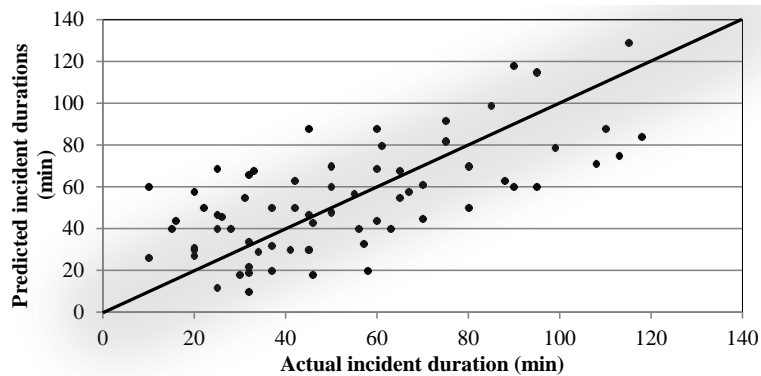


Fig. 7 ANN prediction results versus duration test data.

3.2.2 Support Vector Machine model

SVM is a type of learning algorithms based on statistical learning theory, which can be adjusted to map the input–output relationship for the non-linear system. In addition, the SVM can generalize well and can provide inferences at low computational cost. Before applying SVM, there are two parameters, C and ε , which are first determined. Parameter C is to determine the trade-off between the model complexity and the degree in the optimization equation. Parameter ε controls the width of the ε -insensitive zone which is used to fit the training data. ε -insensitive can allow the presence of prediction error within a certain range and ensure that the model can find the optimal solution. In order to identify the parameters in SVM, grid-search is used to pick up the optimal parameter values. Afterward, for the freeway accident duration forecasting, the two parameters (C, ε) are selected as (3, 0.2).

In this study, the SVM is trained by independent variables. With the use of Cauchy Function Kernel and the training dataset composed of the 10 significant explanatory variables, the best performing SVM is obtained with 40 support vectors. This model reaches the minimum MAE (13 min).

From the test results, in Fig. 8, we can find that SVM model performs better in the prediction of medium duration incident cases. The MAE is equal to 11 min. Only for the four longest duration incidents the prediction absolute errors are more than 25 min, while 38 out of the 70 incident cases have an absolute error less than 18 min.

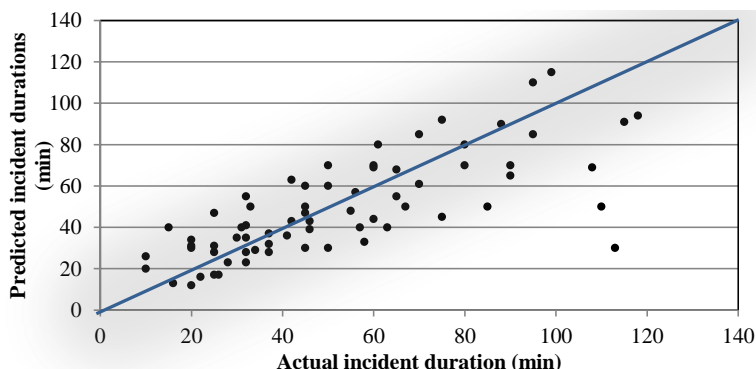


Fig. 8 SVM prediction results versus duration test data.

3.2.3 The comparison results

To evaluate the models performance, the difference between predicted value and true value of every incident is calculated. The prediction errors of the two models are shown in Tab. III. According to the MAE and RMSE values in Tab. III, the SVM is a relatively reliable model with smaller MAE and RMSE values. When it comes to the short duration cases, the MAPE of ANN becomes bigger, as shown in the previous section (Fig. 7). As listed in Tab. IV, 64% of ANN prediction errors

are less than 25 min, which is lesser than 73% of SVM. Moreover, 19% of ANN prediction errors are more than 30 min.

	ANN	SVM
MAE	17	11
RMSE	18	15
MAPE [%]	26	19

Tab. III MAE, RMSE and MAPE of the models.

Absolute error [min]	ANN [%]	SVM [%]
< 5	14	22
5 – 15	24	21
16 – 25	26	30
26 – 35	17	20
> 35	19	7

Tab. IV Distribution of absolute error for the two models.

In order to get more useful information from the data, the following work is conducted. Fig. 9 illustrates the MAE, RMSE and MAPE calculated for the five duration classes, and the following conclusions can be highlighted:

- 1) The two models tend to have a relatively low accuracy for incidents with long duration partly because the dataset has a relatively small number of fatal incidents.
- 2) The ANN model gives a better result for long duration incident cases. It is a relatively reliable model that can predict an incident longer than 80 min.
- 3) The SVM model is better than the ANN model in the prediction of medium/medium-long duration.
- 4) There exist some outliers with a large difference between recorded and predicted incident duration. The problem is largely due to the individual differences of incident management teams in responding to incidents.

4. Conclusions

For the sake of traffic accident duration prediction, this paper presents the performance of ANN and SVM. The ANN model is a multi-layer feed forward neural network and the SVM model is a support vector regression with radial basis kernel

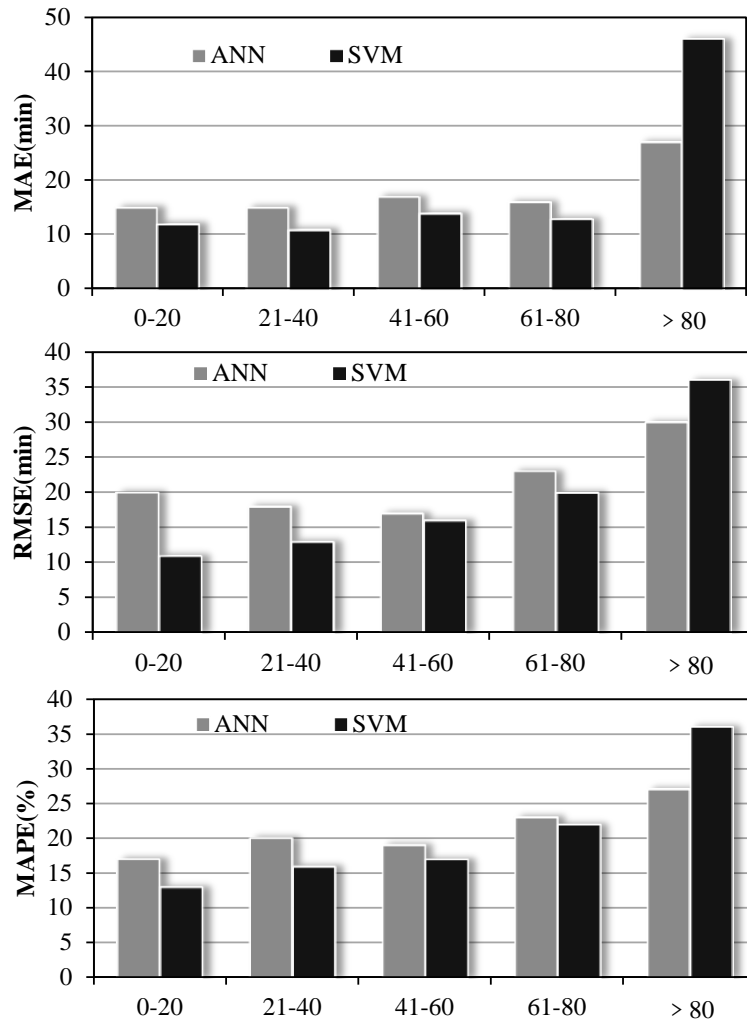


Fig. 9 The values achieved by the proposed model.

function. The two models have been developed and tested using a common incident data set. The predictive ability of the two models is compared. The TMAE, RMSE and MAPE are adopted to estimate the accuracy of the models.

In conclusion, although the ANN model gives a better result for long duration incident cases, the comprehensive performance of the SVM model is better than the ANN model for the traffic accident duration prediction. Nevertheless, we don't know whether it is appropriate in other geographical contexts. To ensure the proposed models can be applied to the other conditions, a deeper investigation needs to be done in the future.

Acknowledgement

This research was supported in National Natural Science Foundation of China 71571026, Liaoning Excellent Talents in University LR2015008, Higher Education Development Fund (for Collaborative Innovation Center) of Liaoning Province, China (20110116401).

References

- [1] ALENEZI A., MOSES S.A., TRAFALIS T.B. Real-time prediction of order flow times using support vector regression. *Computers & Operations Research*. 2008, 35(11), pp. 3489–3503, doi: [10.1016/j.cor.2007.01.026](https://doi.org/10.1016/j.cor.2007.01.026).
- [2] BISHOP C. *Neural networks for pattern recognition*. Oxford University Press, 1995.
- [3] CHUNG Y.S. Development of an Accident Duration Prediction Model on the Korean Freeway Systems. *Accident Analysis and Prevention*. 2009, 42(1), pp. 282–289, doi: [10.1016/j.aap.2009.08.005](https://doi.org/10.1016/j.aap.2009.08.005).
- [4] CHERKASSKY V. The Nature of Statistical Learning Theory. *IEEE Transactions on Neural Networks*. 1997, 8(6), pp. 1564, doi: [10.1007/978-1-4757-2440-0](https://doi.org/10.1007/978-1-4757-2440-0).
- [5] DURSUN D., SHARDA R., BESSONOV M. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*. 2005, 38(3), pp. 434–44, doi: [10.1016/j.aap.2005.06.024](https://doi.org/10.1016/j.aap.2005.06.024).
- [6] DOUGHERTY M.S. Applications of neural networks in transportation. *Transportation Research Part C*. 1997, 5(5), pp. 255–257, doi: [10.1016/S0968-090X\(97\)00013-2](https://doi.org/10.1016/S0968-090X(97)00013-2).
- [7] DHARIA A., ADELI H. Neural network model for rapid forecasting of freeway link travel time. *Engineering Application of Artificial Intelligence*. 2003, 16(7), pp. 607–13, doi: [10.1016/j.engappai.2003.09.011](https://doi.org/10.1016/j.engappai.2003.09.011).
- [8] ESTE A., GRINGOLI F., SALGARELLI L. Support vector machines for TCP traffic classification. *Computer Networks*. 2009, 53(14), pp. 2476–2490, doi: [10.1016/j.comnet.2009.05.003](https://doi.org/10.1016/j.comnet.2009.05.003).
- [9] GARIB A., RADWAN A.E., AL-DEEK H. Estimating magnitude and duration of incident delays. *Journal of Transportation Engineering*. 1997, 123(6), pp. 459–466, doi: [10.1061/\(ASCE\)0733-947X\(1997\)123:6\(459\)](https://doi.org/10.1061/(ASCE)0733-947X(1997)123:6(459)).
- [10] GHOSH-DASTIDAR S., ADELI H. Wavelet-clustering neural network model for freeway incident detection. *Computer-Aided Civil and Infrastructure Engineering*. 2003, 18(5), pp. 325–38. doi: [10.1111/1467-8667.t01-1-00311](https://doi.org/10.1111/1467-8667.t01-1-00311).
- [11] GHOSH-DASTIDAR S., ADELI H. Neural network wavelet micro-simulation model for delay and queue length estimation at freeway work zones. *Journal of Transportation Engineering ASCE*. 2006, 132(4), pp. 331–41, doi: [10.1061/\(ASCE\)0733-947X\(2006\)132:4\(331\)](https://doi.org/10.1061/(ASCE)0733-947X(2006)132:4(331)).
- [12] HAYKIN S. *Neural Networks: a comprehensive foundation*. Prentice Hall, 1999.
- [13] JONES B., JANSSEN L., MANNERING F. Analysis of the frequency and duration of freeway accidents in Seattle. *Accident analysis and prevention*. 1991, 23(4), pp. 239–255, doi: [10.1016/0001-4575\(91\)90003-N](https://doi.org/10.1016/0001-4575(91)90003-N).
- [14] LIU W.M., GUAN L.P., YIN X.Y. Prediction of Incident Duration Based on Multiple Regression Analysis. *Journal of Highway and Transportation Research and Development*. 2005, 22(11), pp. 126–129, doi: [1002-0268\(2005\)11-0126-04](https://doi.org/10.1002-0268(2005)11-0126-04).
- [15] NAM D., MANNERING F. An exploratory hazard-based analysis of highway incident duration. *Transportation Research Part A*. 2000, 34(2), pp. 85–102, doi: [10.1016/S0965-8564\(98\)00065-2](https://doi.org/10.1016/S0965-8564(98)00065-2).
- [16] NOSOV A.V. *An introduction to support vector machines*. China: Machine Press, 2005.
- [17] OBZAY K., PUSHKIN K.P. *Incident management in intelligent transportation systems*. Information Retrieval. 1999, 5(4), pp. 63–67.

- [18] OZBAY K., NOYAN N. Estimation of Incident Clearance Times Using Bayesian Networks Approach. *Accident Analysis and Prevention*. 2006, 38(3), pp. 542–555, doi: [10.1016/j.aap.2005.11.012](https://doi.org/10.1016/j.aap.2005.11.012).
- [19] PARK D., RILETT L.R. Forecasting freeway link travel times with a multi-layer feed forward neural network. *Computer Aided Civil and Infrastructure Engineering*. 1999, 14, pp. 358–367, doi: [10.1111/0885-9507.00154](https://doi.org/10.1111/0885-9507.00154).
- [20] RAJASEKARAN S., GAYATHRI S., LEE T.L. Support vector regression methodology for storm surge predictions. *Ocean Engineering*. 2008, 35(16), pp. 1578–1587, doi: [10.1016/j.oceaneng.2008.08.004](https://doi.org/10.1016/j.oceaneng.2008.08.004).
- [21] SHARMA S. *Applied Multivariate Techniques*. New York: Wiley Press, 1996.
- [22] Tipping M. Sparse Bayesian learning and the relevance vector machine. *The journal of machine learning research*. 2001, 1, pp. 211–244, doi: [10.1162/15324430152748236](https://doi.org/10.1162/15324430152748236).
- [23] VALENTI G., LELLI M., CUCINA D. A comparative study of models for the incident duration prediction. *European Transport Research Review*. 2010, 2(2), pp. 103–111, doi: [10.1007/s12544-010-0031-4](https://doi.org/10.1007/s12544-010-0031-4).
- [24] VAPNIK V.N. *The Nature of Statistical Learning Theory*. New York, USA: Springer-Verlag, 1995, doi: [10.1007/978-1-4757-2440-0](https://doi.org/10.1007/978-1-4757-2440-0).
- [25] VAPNIK V.N. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*. 1999, 10(5), pp. 988–999, doi: [10.1109/72.788640](https://doi.org/10.1109/72.788640).
- [26] VAPNIK V.N. The Nature of Statistical Learning Theory. The nature of statistical learning theory /, 2000, pp. 988–999, doi: [10.1007/978-1-4757-2440-0](https://doi.org/10.1007/978-1-4757-2440-0).
- [27] WANG J., SHI Q. Short-term traffic speed forecasting hybrid model based on chaos-wavelet analysis-support vector machine theory. *Transportation Research Part C*. 2013, 27, pp. 219–232, doi: [10.1016/j.trc.2012.08.004](https://doi.org/10.1016/j.trc.2012.08.004).
- [28] WANG P., ZHU X.Y. Model Selection of SVM with RBF Kernel and its Application. *Computer Engineering and Applications*. 2003, 39(24), pp. 72–73, doi: [10.1002/8331-\(2003\)24-0072-02](https://doi.org/10.1002/8331-(2003)24-0072-02).
- [29] WANG W., CHEN H., BELL M.C. Vehicle breakdown duration modelling. *Journal of Transportation and Statistics*. 2005, 8(1), pp. 75–84.
- [30] WEI C.H., CHEN Y.C. Review of artificial neural network research and applications in transportation. *Transportation Planning Journal Quarterly*. 2001, 30(2), pp. 324–48.
- [31] WEI C.H., LEE Y. Applying data fusion techniques to traveler information services in highway network. *Journal of the Eastern Asia Society for Transportation Studies*. 2005, 6, pp. 2457–75, doi: [10.11175/easts.6.2457](https://doi.org/10.11175/easts.6.2457).
- [32] WEI C.H., LEE Y. Sequential forecast of incident duration using artificial neural network models. *Accident Analysis and Prevention*. 2007a, 39(5), pp. 944–54, doi: [10.1016/j.aap.2006.12.017](https://doi.org/10.1016/j.aap.2006.12.017).
- [33] YAO B.Z., HU P., LU X.H., GAO J.J., ZHANG M.H. Transit network design based on travel time reliability. *Transportation Research Part C*. 2014, 43, pp. 233–248, doi: [10.1016/j.trc.2013.12.005](https://doi.org/10.1016/j.trc.2013.12.005).
- [34] YAO B.Z., WANG Z., ZHANG M.H., HU P., YAN X.X.. Hybrid Model for Prediction of Real-Time Traffic Flow. In: *Proceedings of the Institution of Civil Engineers –Transport*. 2016, 169(2), pp. 88–96, doi: [10.1680/jtran.14.00015](https://doi.org/10.1680/jtran.14.00015).
- [35] YAO B.Z., YU B., HU P., GAO J.J., ZHANG M.H. An improved particle swarm optimization for carton heterogeneous vehicle routing problem with a collection depot. *Annals of Operations Research*. *Annals of Operations Research*. 2015, pp. 1–18, doi: [10.1007/s10479-015-1792-x](https://doi.org/10.1007/s10479-015-1792-x).
- [36] YONG G.J., JONG H.L. Automobile Traffic Accidents Prediction Model Using by Artificial Neural Networks. In: *Proceedings of the 6th International Conference (ICHIT 2012)*, Daejeon, Korea. Communications in Computer and Information Science. Berlin-Heidelberg: Springer, 2012, 310, pp. 713–719, doi: [10.1007/978-3-642-32692-9_89](https://doi.org/10.1007/978-3-642-32692-9_89).

- [37] YUAN R., LI Z., GUAN X., XU L. An SVM-based machine learning method for accurate internet traffic classification. *Information Systems Frontiers*. 2010, 12(2), pp. 149–156, doi: [10.1007/s10796-008-9131-2](https://doi.org/10.1007/s10796-008-9131-2).
- [38] YU B., KONG L., SUN Y., YAO B.Z., GAO Z.Y. A bi-level programming for bus lane network design. *Transportation Research Part C*. 2015, 55, pp. 310–327, doi: [10.1016/j.trc.2015.02.014](https://doi.org/10.1016/j.trc.2015.02.014).
- [39] YU B., SONG X.L., YANG Z.M., YAO B.Z. k-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition. *Journal of Transportation Engineering*. 2016, 142(6), 04016018, doi: [10.1061/\(ASCE\)TE.1943-5436.0000816](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000816).
- [40] ZHAN C.J. *Development of Prediction Models for Freeway Incident Durations Using Data Mining Techniques*. Miami, USA: Florida International University Press, 2006.