View metadata, citation and similar papers at core.ac.uk
brought to you by CORE

MUHAMMED YASIN ÇODUR, Ph.D.
Corresponding Author
E-mail: mycodur@erzurum.edu.tr
Erzurum Technical University,
Faculty of Engineering and Architecture,
Civil Engineering Department
25070 Erzurum, Turkey
AHMET TORTUM, Ph.D.
E-mail: atortum@atauni.edu.tr
Atatürk University, Engineering Faculty,
Civil Engineering Department
25240 Erzurum, Turkey

Safety and Security in Traffic
Preliminary Communication
Submitted: May 22, 2014
Approved: May 27, 2015

AN ARTIFICIAL NEURAL NETWORK MODEL FOR HIGHWAY ACCIDENT PREDICTION: A CASE STUDY OF ERZURUM, TURKEY

ABSTRACT

This study presents an accident prediction model of Erzurum's Highways in Turkey using artificial neural network (ANN) approaches. There are many ANN models for predicting the number of accidents on highways that were developed using 8 years with 7,780 complete accident reports of historical data (2005-2012). The best ANN model was chosen for this task and the model parameters included years, highway sections, section length (km), annual average daily traffic (AADT), the degree of horizontal curvature, the degree of vertical curvature, traffic accidents with heavy vehicles (percentage), and traffic accidents that occurred in summer (percentage). In the ANN model development, the sigmoid activation function was employed with Levenberg-Marquardt algorithm. The performance of the developed ANN model was evaluated by mean square error (MSE), the root mean square error (RMSE), and the coefficient of determination (R^2). The model results indicate that the degree of vertical curvature is the most important parameter that affects the number of accidents on highways.

KEY WORDS

traffic accident prediction model; artificial neural network; highways of Erzurum/Turkey;

1. INTRODUCTION

Worldwide, 1.3 million people die and 20-50 million are injured each year due to traffic accidents. According to the World Health Organization (WHO) data collected from 178 countries, the traffic accidents are the ninth most common cause of death among all age groups [1]. Due to the rapid development of economy and continuous improvement in Turkey, the number of vehicles, transportation demand, and traffic accidents have increased during the last decades. Injuries and deaths are not only the result of traffic accidents, but also lead to the waste of social wealth. It is very important to reduce the highway traffic accidents in the developing countries like Turkey, and therefore the traffic safety should be improved by the analysis of accident characteristics.

The number of accidents on a given highway section during a certain period of time is probabilistic in nature and is a non-negative integer. Despite the fact that accidents are random and unpredictable at micro-level, statistical models can predict reliable estimates of expected accidents by relating aggregates of accidents to various explanatory measures of flow, site characteristics, and road geometry at macro-level. Numerous empirical relationships between vehicle accidents and these explanatory variables have been established in several previous studies [2, 3, 4].The reliability and results of traffic accident prediction models have an important meaning for the improvement of traffic safety management.

Traffic accident prediction models have been developed to understand the factors affecting traffic accidents and eventually to reduce traffic accidents by controlling and/or improving factors [5]. As far as the authors are aware the first accident-prediction models for multi-lane roads were devised by Persaud

and Dzbik [6]. The relationships between crash data and traffic flow, expressed both as average daily traffic (ADT) and hourly volume (HV), were proposed. The analysis was based on generalized linear models. The results showed that crash rate increases with increasing traffic flow expressed both as ADT and HV [7].

Knuiman examined the effect of median width of four-lane roads on crash rate using a Negative Binomial distribution. The findings indicated that crash rate decreases with increasing median width [8].

Fridstrøm related road accidents to four variables, namely: traffic flow, speed limits, weather and lighting conditions. They considered Negative Binomial regression. On the basis of the four independent variables, they were able to explain between 85 and 95% of the systematic component of variation [9].

Hadi proposed several accident-prediction models with regard both to multi-lane roads and two-lane roads of rural or urban designation. The dependent variables were total crash rate or injury crash rate. The values of these accident indicators were estimated as a function of annual average daily traffic (AADT) and road environmental factors. Poisson and Negative Binomial regression models were considered. By examining the effect of traffic flow on the crash rate the conclusions reached were that crash rate increases with increasing AADT on roads having higher levels of traffic, while it decreases with AADT on roads with lower traffic volumes [10].

Persaud presented one of the earliest studies for carrying out separate analyses for curves and tangents, albeit limited to two-lane roads. The dependent variable was crash frequency, while the independent variables were traffic flow and road geometry. The regression models were calibrated using generalized linear modelling. A dummy variable for “flat” or “undulating” terrain was also used. For curves, crash frequency was found to increase with: AADT, section length (L) and curvature (1/R). For tangents, the number of accidents per year increases with AADT and L. A higher accident number on undulating terrain than on a flat one was also shown [11].

Abdel-Aty and Essam Radwan used Negative Binomial distribution to predict crash frequency as a function of: AADT, degree of horizontal curvature, section length, lane, shoulder and median widths, and urban/rural designation. The results showed that crash frequency increases with AADT, degree of horizontal curvature and section length. Accident frequency decreases, however, with lane, shoulder and median width [12].

Hauer developed statistical road safety modelling by using the Negative Binomial distribution. The dependent variable was the number of accidents per year, while the independent ones were geometric characteristics and traffic flow. The most innovative aspect of this study was the introduction of an alternative tool for measuring the goodness-of-fit of the predictive models, the so-called Cumulative Residuals (CURE) Method. This method consists of plotting the cumulative residuals as a function of the independent variable of interest, a good CURE plot being one oscillating around zero [13].

In a subsequent paper, Hauer applied the abovementioned statistical model to estimate crash frequency on undivided four-lane urban roads. The proposed models evaluated the number of accidents per year and carriageway as a function of the following independent variables: AADT, percentage of trucks, degree and length of horizontal curves, grade of tangents and length of vertical curves, lane width, shoulder width and type, roadside hazard rating, speed limit, access points (e.g. signalized intersections, stop-controlled intersections, commercial driveways and other driveways), the presence and nature both of parking and two-way-left-turn-lanes. The findings showed that significant variables were: AADT, the number of commercial driveways and speed limit [14].

Lately, ANN models have been developed based on traffic accident case studies. The national freeway of Taiwan was analyzed with the model of ANN by Chang [15]. The other study considered the relationship between the probabilities of involvement in a traffic accident and driver characteristics using ANN [16]. Akgüngör and Doğan developed traffic accident models using non-linear multiple regression and ANN in Turkey and in another prediction model for some big cities in Turkey using again ANN and genetic algorithm approaches [17, 18]. They also studied the accidents prediction model by using Modified Smeed, Adapted Andreassen, and ANN approaches [19]. Cansız estimated the number of fatalities in accidents using a non-linear accident model with Smeed equation and ANN model [20]. In these studies ANN models produced better results as compared with other models.

Statistical methods were commonly utilized in most previous studies, such as Linear Regression Model, Logistic Regression Model, Poisson Model, Negative Binomial Model, Zero-inflated Negative Binomial Model and Generalized Linear Regression Model methods. These methods are subject to strong assumptions and limitations in application. In contrast, ANN has been proven efficient and effective in many fields. In traffic safety, some studies have applied ANN to the forecasting of traffic accidents on highways but few have a large data set. Therefore, this paper establishes the traffic accident prediction model with the large data set and several parameters by the use of ANN. The ANN model was successfully expressed in the number of traffic accidents in comparison with the original dataset.

2. MATERIAL AND METHODS

2.1 Case study data

The collected data of the number of accidents covered a period of eight years from 2005 to 2012 and relate to the road network of the Province of Erzurum in Eastern Turkey. The traffic accident reports data used in this study are 7,780 complete accident reports which were collected from the Directory of Erzurum Traffic Region. Each accident report has various information such as the date, accident location, pavement type, vehicle type, driver's gender, driver's age, road surface condition, the day and time, weather condition, day or nighttime, the number of deaths, the number of injured persons, the number of involved vehicles, and the number of damaged vehicles. In addition to these data, geometric characteristics of the highway such as AADT, the degree of horizontal and vertical curvatures in each section, lane, median, and shoulder widths were collected from 12th Highway Regional Directorate of Erzurum. After eliminating the missing and erroneous data, 7,285 accident reports were utilized in this research. These data were categorized with 18 variables as shown in *Table 1* [21].

In order to analyze the traffic accidents on the highways, one needs to select highway(s) that possess a wide variety of geometric and traffic characteristics. The goal of this data collection exercise is to divide these highways into segments with homogenous characteristics. After reviewing several highways around Erzurum, it was decided that D950-03, D100-28, D052-03, and D100-29 four-lane median-divided highways were most appropriate for this task [21].

The highways include a total of 152 km of major principal arterials that connect the east and the west of Erzurum. These arterials are long enough to produce an adequate number of segments to develop the model. The information on highways includes geometric characteristics such as gradient of the highways, horizontal and vertical curves, shoulder widths, median widths, and traffic characteristics such as annual average daily traffic. D950-03, D100-28, D052-03, and D100-29 were divided into 5, 3, 4, and 4 highway segments, respectively and defined by any change in the geometric and/or highway variables (e.g. a new section would be identified when median changed from 2 to 4 m). Thus, each highway segment is uniform with respect to all the possible geometric and traffic features (*Figure 1*). The routes are as follows:

- D950-03, Erzurum-Tortum (52 km);
- D100-28, Erzurum South Ring Road (20 km);
- D052-03, Erzurum North Ring Road (30 km); and
- D100-29, Erzurum-Köprüköy (50 km).

2.2 Artificial neural network

Modelling of non-linear systems is far more difficult than linear systems. The disturbances influencing the system make the modelling task even more difficult.

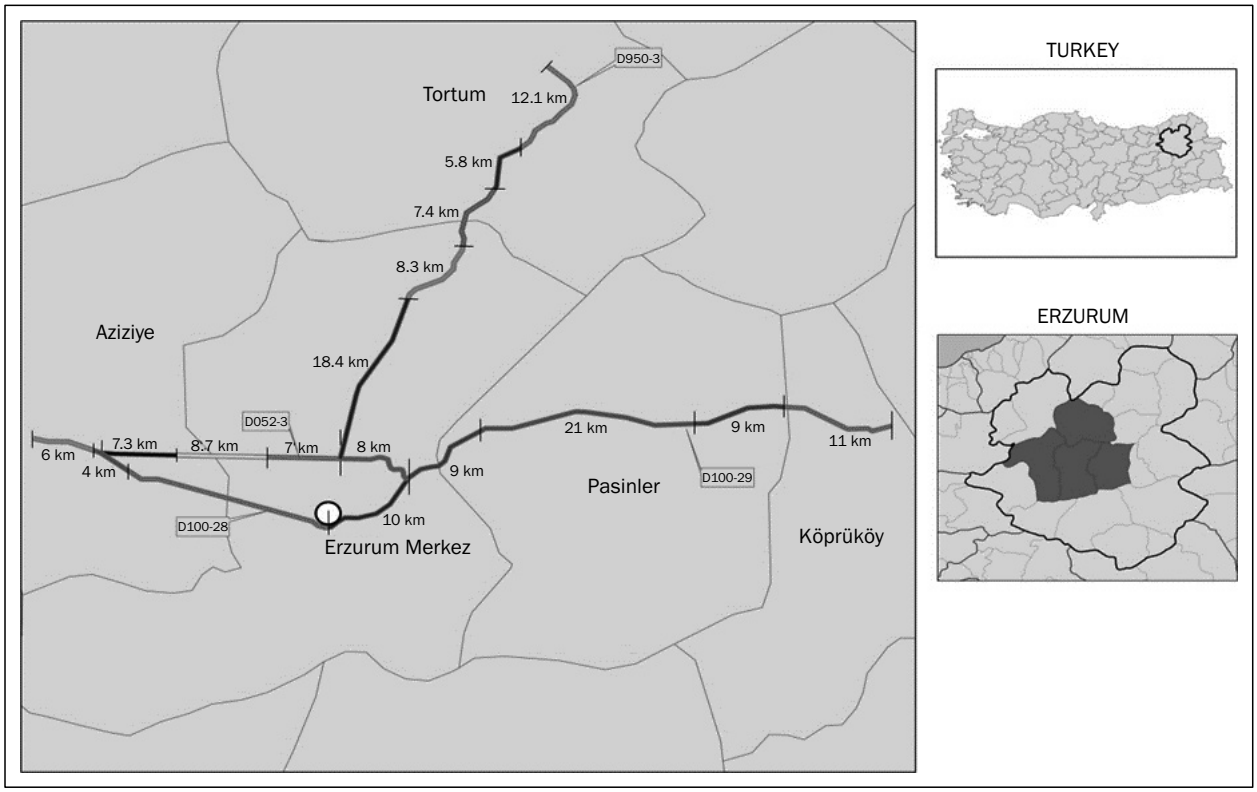

Figure 1 - Erzurum's highways and sections

Table 1 - A summary of input variables

Design parameters	Definition	Binary/ Numerical code
X1	Years	Categorical Value (2005-2006-2007-2008-2009-2010-2011-2012)
X2	Highway sections	Categorical Value (D 950-03, D100-28, D052-03, D100-29)
X3	Section length	Categorical Value (D 950-03(12.1-5.8-7.4-8.3-18.4), D100-28 (6-4-10), D052-03 (7.3-8.7-7-8.02), D100-29 (9-21-9-11))
X4	Driver's gender	Categorical Value (Male, Female)
X5	Day or night-time	Categorical Value (Daylight, night with illumination, night without illumination)
X6	Road surface condition	Categorical Value (Dry, wet, snow or ice)
X7	Shoulder width	Numerical Value
X8	Median width	Numerical Value
X9	Gradient of the highway	Numerical Value
X10	The degree of horizontal curvature	Numerical Value
X11	The degree of vertical curvature	Numerical Value
X12	Annual average daily traffic	Numerical Value
X13	Traffic accidents with heavy vehicles	Numerical Value
X14	Traffic accidents with light vehicles	Numerical Value
X15	Traffic accidents occurred in summer	Numerical Value
X16	Traffic accidents occurred in winter	Numerical Value
X17	Traffic accidents occurred in fall	Numerical Value
X18	Traffic accidents occurred in spring	Numerical Value
Y	Number of accidents	Numerical Value

Scientists have been studying non-linear system modelling for years and they have succeeded in teaching non-linear system dynamics to artificial neural networks without any mathematical modelling [22, 23, 24].

Neural networks, with their remarkable ability to learn complicated relations from imprecise data can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an ''expert'' in the category of information that has been given to analysis. Artificial neural networks can be used in a variety of powerful ways: to learn and reproduce rules or operations from the given examples; to analyze and generalize from samples and to make predictions or to memorize features of given data and to match or make associations from new data to old data [25].

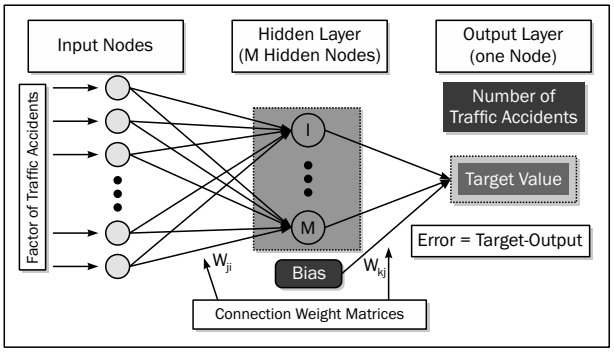

Figure 2 - Structure of artificial neural network model for traffic accidents

Figure 2 shows that there are nodes in the input layer and one node in the output layer which is called target value. To get strong reflection ability, a differential sigmoid function is used in the intermediate layer and a linear function in the output layer. There is no regulation or theory base to determine the number of neurons in the intermediate layer; it could be achieved by calculating iteratively.

Network training is a process by which the connection weights and biases of ANN are adapted through a continuous process of simulation by the environment in which the network is embedded. The primary goal of training is to minimize an error function by searching for a set of connection strengths and biases that causes ANN to produce outputs that are equal or close to targets. In other words, training aims at estimating the parameters (W_1, W_2, b_1, and b_2) by minimizing an error function, such as the mean square error (MSE) of the output values expressed as [26]:

$$MSE = \sum_{i=1}^{N} \frac{(t_i - a_i)}{N} \quad (1)$$

where N is the number of data.

2.3 Development and application of ANN models

Alyuda Neuro Intelligence is a neural network software for experts designed for intelligent support in applying neural networks to solve real-world forecasting, classification and function approximation problems; to use intelligent features to preprocess datasets; to find efficient architecture and to analyze performance and apply the neural network to new data. Experts can create and test their solutions much faster, increase their productivity and improve results [27].

To estimate the ANN model, there are a number of software packages ready to perform the Levenberg Marquardt algorithm, and Alyuda Neuro Intelligence was chosen for this study. In the ANN model, independent variables are named as the input, and dependent variables are named as the output. The input importance chart shows the relative importance of each input column. This chart can help the user decide what input columns can be removed without affecting the results. It also helps to understand the most important columns that had the biggest influence on the network. The input column importance is calculated as degradation in network performance after the input was removed and not used by the network [28]. Among the eighteen parameters applied in modelling, eight parameters were found to be significant based on those criteria. The significant parameters are years, highway sections, section length (km), annual average daily traffic (AADT), the degree of horizontal curvature, the degree of vertical curvature, traffic accidents with heavy vehicles (percentage), and traffic accidents occurred in summer (percentage). Here, there are 8 input variables containing 31 neurons which are the input variables representing the potential risk factors for accidents.

Before the training of the network, both input and output variables were normalized within the range 0-1 using a minimax algorithm. Categorical columns were automatically encoded during the data preprocessing using One-of-N method by Alyuda Neuro Intelligence program. The One-of-N encoding means that a column with N distinct categories (values) is encoded into a set of N numeric columns, with one column for each category. For example, for the Capacity column with values “Low”, “Medium” and “High”, “Low” will be represented as {1,0,0}, Medium as {0,1,0}, and High as {0,0,1}. The minimum and maximum of the dataset were found and scaling factors selected so that these were mapped to the desired minimum and maximum values. The mimimax algorithm is as follows:

$$H' = H\left\{\frac{1}{H_{max} - H_{min}}\right\} \cdot H\left\{\frac{H_{min}}{H_{max} - H_{min}}\right\} \quad (2)$$

where H_{max} and H_{min} indicate the largest and the smallest values of H, respectively, and H' the unified value of the corresponding H. Normalization of the data greatly improves learning speed and it is beneficial in reducing the error of the trained network.

The mathematical formulation of ANN is

$$Y = f_2\left\{[W_2]f_1\left\{[W_1]\begin{Bmatrix} Years \\ Highway_section \\ Section_Length \\ Annual_Average_Daily_Traffic \\ The_Number_of_Horizontal_Curvature \\ The_Number_of_Vertical_Curvature \\ Traffic_Accidents_With_Heavy_Vehicles \\ Traffic_Accidents_Occured_in_Summer \end{Bmatrix} + [b_1]\right\} + [b_2]\right\} \quad (3)$$

In Equation 3; Y is the number of accidents, W_1 and W_2 are weight matrices and b_1 and b_2 are bias vectors. The effectiveness of the back propagation training algorithm depends on the number of neurons in the hidden layer; various numbers of neurons (ranging from 1 to 29) in the hidden layer were tested. In this study, one output layer is called number of accidents for all highways.

The fitness criterion specifies what network parameter should be used to distinguish the best network. It automatically finds the best architecture offering the user graphs of the search process and details for every tested neural network [28]. In the analysis, the sum of squares was adopted as the error function, activation function was taken to be logistic and the search method aroused to be exhaustive methods. The best five optimal architecture networks which were obtained from the software packages are shown in *Figure 3*. Considering *Table 2* and *Figure 3* the optimal network architecture was found to be 31x9x1.

Table 2 - The fitness criteria of the best five networks

Network	Fitness Criteria
31-2-1	0.726160
31-4-1	0.691016
31-7-1	0.701787
31-9-1	0.887323
31-10-1	0.800927

Data sets were divided into three sections: the training set, the verification set and the test set. Training algorithms do not use the verification or test sets to adjust network weights. The verification set may optionally be used to track the network’s error performance, to identify the best network and to stop training if over-learning occurs. The test set is not used in training at all, and it is designed to give an independent assessment of the network’s performance when an entire network design procedure is completed. Seventy percent of the data set was used to train the network, while the remaining thirty percent was employed for testing and verification. The assignment of cases to the training, verification and test subsets can sometimes affect the performance of the train-

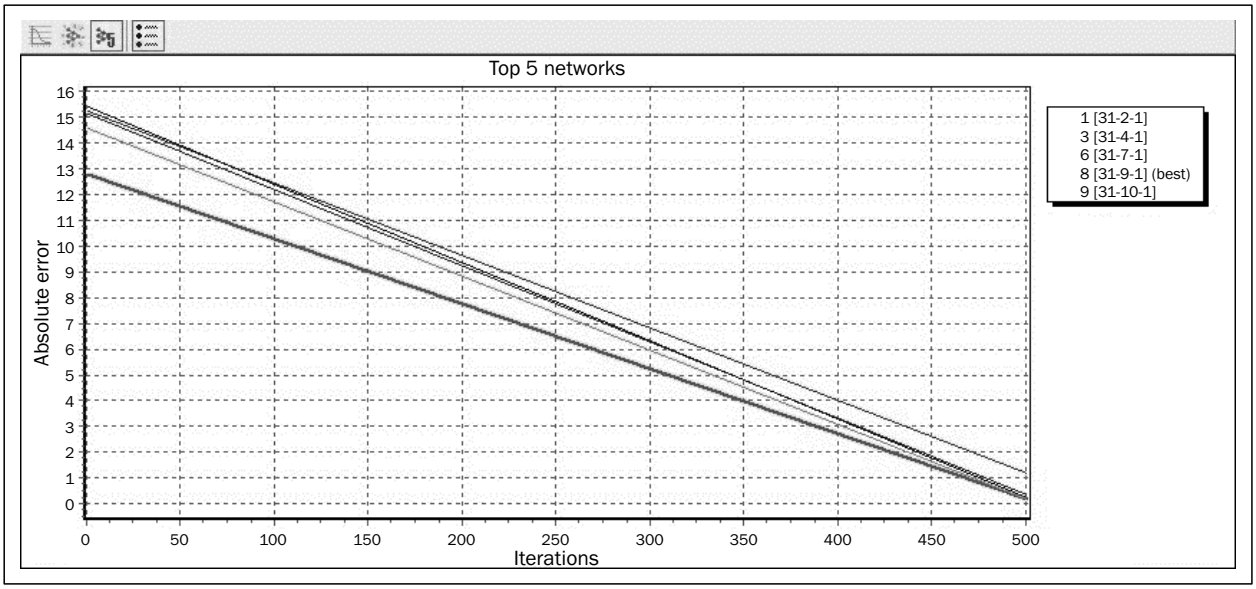

Figure 3 - The top five networks

ing algorithms. In order to eliminate this situation, the cases should be shuffled randomly between subsets. The cases can be left in their original order, or grouped together in the subsets. In this model, the cases were shuffled randomly between subsets (training, test and verification) [29].

2.4 Evaluation of ANN model

The coefficients of determination (R^2), mean square error (MSE), and the root mean square error (RMSE) are the main criteria that are used to evaluate the performance of ANN model. They are defined as follows:

$$R^2 = r^2 \tag{4}$$

$$Cor.Ratio(r) = = \frac{n(\sum actual \times predicted) - (\sum actual)(\sum predicted)}{\sqrt{\left[n\sum actual^2 - (\sum actual)^2\right] \times \left[n\sum predicted^2 - (\sum predicted)^2\right]}} \tag{5}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(t_{mi} - t_{gi})^2 \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N}MSE} \tag{7}$$

where t_{mi} is the *i*-th observation value; t_{gi} is the *i*-th model value; *N* is the number of trained data; *m* is the number of parameters in the model (the total number of weights and invariables in the net structure).

3. RESULTS

The conclusion summary has minimum, maximum, mean and standard deviation for target, output, absolute error (AE) and absolute relative error (ARE) along with R-squared and correlation parameters of the tested network which is given in *Table 3*. In order to determine the performance of the ANN model, the comparison of model prediction performance between the target and output is examined. *Table 3* shows the comparison results.

Table 3 - Conclusion summary

	Target	Output	AE	ARE
Mean	16.098039	16.143741	1.254079	0.139047
Std. Dev.	11.060892	10.275677	1.024443	0.215479
Min.	2	2.034983	0.000946	0.000376
Max.	53	48.679285	4.466516	1.337135
Correlation	0.991186			
R-Squared	0.982452			

The information criteria of the model are given in *Table 4*. R^2 is used to measure the closeness of fit. A perfect fit would result in $R^2 \approx 1$, a very good fit near 1, and a poor fit would be near 0. In the ANN model correlation ratio and R^2 are 0.991186 and 0.982452, respectively. When considering the low values of MSE, RMSE and high values of the correlation coefficients in ANN the superiority of the model will be understood. In the ANN model MSE and RMSE values are 4.110521 and 2.027442, respectively. This demonstrates that the ANN model is an appropriate methodology for analyzing traffic accidents.

Table 4 - Information criteria of the model

Model	Correlation	R-Squared	MSE	RMSE
ANN	0.991186	0.982452	4.110521	2.027442

Results showed that in *Table 5* the four important features, namely the degree of horizontal curvature, annual average daily traffic, the degree of vertical curvature, and section length were very effective in predicting the number of traffic accidents.

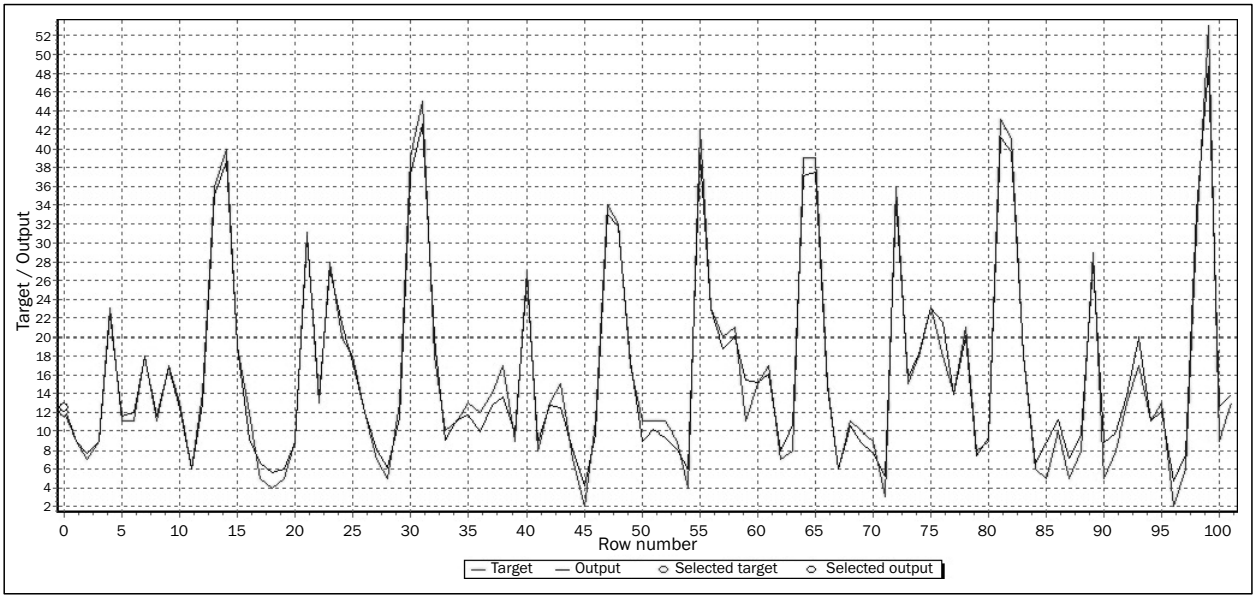

Figure 4 - Relationship between estimated and actual values

Table 5 - Importance of traffic accident parameters

Input Column Name	Importance %
The degree of vertical curvature	54.88
Annual average daily traffic	14.89
The degree of horizontal curvature	14.82
Section length	10.95
Highway location	2.48
Years	0.98
Traffic accidents occurred in summer	0.71
Traffic accidents with heavy vehicles	0.28

The output graph displays a line graph of the actual and network output values for records displayed in the output table. Horizontal axis displays the row number of the input dataset and vertical axis displays the range of the output values. In addition to this the ANN model trends were closer to actual values as shown in *Figure 4*.

4. CONCLUSION

In this study, the factors which cause accidents have been investigated, for providing road safety, and accident prediction models which include relations between these factors have been established. For the geometrical features of highways sections and traffic accident reports the years from 2005 to 2012 were used to form the database. The obtained data from the database in this study have been investigated with ANN as a tool of forecasting techniques. Since ANN method is a more flexible and assumption-free methodology and furthermore, capable of evaluating/comparing all of the traffic accident characteristics, it is selected for modelling the traffic accidents data. The low values of the MSE and RMSE and the high values of correlation coefficient and R^2 in ANN indicate the superiority of the model.

The model results indicate that the degree of vertical curvature with the high percentage (54.88%) is the most important parameter affecting the number of accidents on the highways. AADT and the degree of horizontal curvature have almost the same effect as the second important parameter. Section length is the third one and highway location, years, traffic accidents occurred in summer, and traffic accidents with heavy vehicles have small effect on the output parameter. Actually, traffic accidents occurred in summer are more important than in other seasons and in the same way, traffic accidents involving heavy vehicles are also more important than those involving light vehicles.

These results have implications for policy makers, transportation system designers, and researchers. Transportation safety designers cannot easily identify factors, make recommendations for incremental changes in the factor, and hope to achieve major differences in accident levels. The problems have to be analyzed and attacked from a multidimensional perspective: a wide variety of geometric and traffic characteristics. Researchers similarly may adopt techniques such as neural networks for analysis of such variables. This paper aims to offer and apply the great potential of neural network analysis through an example of the accident prediction model in the transportation engineering department.

Recommendations for future research

The modelling results obtained using the annual data (the years from 2005 to 2012) are encouraging for further research by the expanded data sets. By

setting up some random variables in the design parameters, it may be possible to predict the number of accidents for the years in the future. The future work might focus on how to improve the prediction performance of ANN models. It would also be interesting for future studies to predict the number of accidents on highways with using the technique of Genetic Algorithm and to see if the prediction performance could be improved.

ACKNOWLEDGEMENTS

The authors are grateful to the editor and reviewers for very helpful comments on the earlier version of the paper. This research was supported by the Ataturk University Grant No. 2011/410. We would also like to thank people working at the Directory of Erzurum Traffic Region and 12th Highway Regional Directorate of Erzurum.

MUHAMMED YASIN ÇODUR, Yrd. Doç. Dr.
Sorumlu Yazar
E-Mail: mycodur@erzurum.edu.tr
Erzurum Teknik Üniversitesi
Mühendislik ve Mimarlık Fakültesi
İnşaat Mühendisliği Bölümü
25070 Erzurum, Turkiye
AHMET TORTUM, Doç. Dr.
E-Mail: atortum@atauni.edu.tr
Atatürk Üniversitesi
Mühendislik Fakültesi
İnşaat Mühendisliği Bölümü
25240 Erzurum, Turkiye

ÖZET

YAPAY SINIR AĞI MODELI ILE KARAYOLLARINDA KAZA TAHMINI: TÜRKIYE/ERZURUM KARAYOLLARI ÖRNEK ÇALIŞMASI

Bu çalışmada Türkiye'de Erzurum karayollarında kaza tahmin modelleri Yapay Sinir Ağları (YSA) kullanılarak kurulmuştur. Karayollarında meydana gelen kaza sayılarını tahmin etmek için geçmiş yıllara ait verilerden (2005-2012) 8 yıllık 7780 adet kaza rapor tutanağı kullanılarak birçok YSA modeli geliştirilmiştir. Bu amaç için en iyi YSA modeli seçilmiş ve model için yıllar, kesim uzunluğu (km), yıllık ortalama günlük trafik (YOGT), yatay kurp sayısı, düşey kurp sayısı, ağır araçlardan kaynaklanan kaza sayısı (yüzde) ve yaz mevsiminde meydana gelen kaza sayıları (yüzde) girdi parametresi olarak alınmıştır. YSA modelinde sigmoid aktivasyon fonksiyonu Levenberg-Marquardt algoritması ile kullanılmıştır. Geliştirilen YSA modelinin performansı Ortalama Karesel Hata (OKH), Ortalama Karesel Hatanın Karekökü ve Korelasyon Katsayısı (R^2) kullanılarak değerlendirilmiştir. Model sonuçları, kaza sayılarını etkileyen en önemli parametrenin düşey kurp sayısı olduğunu göstermiştir.

ANAHTAR KELIMELER

Trafik Kaza Tahmin Modeli; Yapay Sinir Ağları; Türkiye/Erzurum Karayolları;

REFERENCES

[1] World Health Organization. Global Plan for the Decade of Action for Road Safety 2011-2020.[Internet] [cited 2012 Dec 5].Available from: http://www.who.int/road-safety/decade_of_action/plan/plan_english.pdf.
[2] Maher MJ, Summersgill J.A comprehensive methodology for the fitting of predictive accident models. Accident Analysis and Prevention. 1996;28(3):281-296.
[3] Miaou SP. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention. 1994;26(4):471-482.
[4] Miaou SP, Lum H. Modelling vehicle accidents and highway geometric design relationships. Accident Analysis and Prevention. 1993;25(6):689-709.
[5] Lee JY, Chung JH, Son B. Analysis of traffic accident size for Korean highway using structural equation models. Accident Analysis and Prevention. 2008;40:1955-1963.
[6] Persaud B, Dzbik L. Accident prediction models for freeways. Transportation Research Record. 1993;1401:55-60.
[7] Caliendo C, Guida M, Parisi A. A crash-prediction model for multilane roads. Accident Analysis and Prevention. 2007;39:657-670.
[8] Knuiman MW, Council FM, Reinfurt DW. Association of median width and highway accident rates. Transportation Research Record. 1993;1401.
[9] Fridstrøm L, Ifver J, Ingebrigtsen S, Kumala R, Krogsgard Thomsen L. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Accident Analysis and Prevention. 1995;27:1-20.
[10] Hadi MA, Aruldhas J, Chow L, Wattleworth JA. Estimating safety effects of cross-section design for various highway types using Negative Binomial regression. Transportation Research Record. 1995;1500.
[11] Persaud B, Retting RA, Lyon C. Guidelines for the identification of Hazardous Highway Curves. Transportation Research Record 2000;1717:14-18.
[12] Abdel-Aty MA, Essam Radwan EA. Modelling traffic accident occurrence and involvement. Accident Analysis and Prevention. 2000;32:633-642.
[13] Hauer E. Statistical road safety modelling. Proceedings of the 83rd TRB Annual Meeting; 2004 Jan 11-15; Washington, D.C., USA.
[14] Hauer E. Safety models for urban four-lane undivided road segments. Proceedings of the 83rd TRB Annual Meeting; 2004 Jan 11-15; Washington, D.C., USA.
[15] Chang L. Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural. Safety Science. 43(8):541-557. doi:10.1016/j.ssci.2005.04.004
[16] Kalyoncuoglu SF, Tigdemir M. An alternative approach for modelling and simulation of traffic data: artificial neural networks. Simul Model Pract Theory. 12(5):351-362. doi:10.1016/ j.simpat.2004.04.002.
[17] Akgüngör AP, Doğan E. Estimating road accidents of Turkey based on regression analysis and artificial neural network approach. Advances in Transportation Studies International Journal. 2008 Nov;16:11-22.

[18] Akgüngör AP, Doğan E. An artificial intelligent approach to traffic accident estimation: model development and application. Transport.2009; 24(2):135-142. doi:10.3846/1648-4142. 2009

[19] Akgüngör AP, Doğan E. An application of modified Smeed, adapted Andreassen and artificial neural network accident models to three metropolitan cities of Turkey. Scientific Research and Essays. 2009 Oct;4(9):906-913.

[20] Cansız OF. Improvements in estimating a fatal accidents model formed by an artificial neural network. Simulation. 2011;87(6):512-522. doi:10.1177/0037549710370842

[21] Çodur MY. Traffic Accident Prediction Models: Applications for Surrounding Highways of Erzurum [PhD Thesis] [in Turkish]. Graduate School of Natural and Applied Sciences, Department of Civil Engineering, Ataturk University; 2012.

[22] Narendra KS, Parthasarathy K. Identification and control of dynamical systems using neural networks. IEEE Transaction Neural Networks. 1990;1(1):4-27.

[23] Narendra KS, Mukhopadhyay S. Adaptive control using neural networks and approximate models. IEEE Transactions on Neural Networks. 1997 May;8(3):475-485.

[24] Singh P, Deo MC. Suitability of different neural networks in daily flow forecasting. Applied Soft Computing. 2007;7(3):968–978.

[25] Ertuğrul S, Hizal NA. Neuro-fuzzy controller design via modelling human operator actions. Journal of Intelligent Fuzzy Systems. 2005;16:133-140.

[26] Hamed M, Khalafallah M, Hassanien E. Prediction of Wastewater Plant Performance Using Artificial Neural Networks. Environmental Modelling and Software. 2004;19(10):919-928.

[27] Svetoslavova Kovacheva Z. Application of Neural Networks to Data Mining. SQU Journal for Science. 2007;12(2):75-85.

[28] Alyuda Neuro Intelligence Manual; 2004.Available from: http://www.alyuda.com

[29] Bates JJ, Quarmby DA. An economic method for car ownership forecasting in discrete areas. Mathematical advisory unit note 219. London: Department of the Environment; 1971.