Journal of Applied Sciences 8 (1): 146-151, 2008
ISSN 1812-5654

Neural Networks Based Modelling of Traffic Accidents in Interurban Rural Highways, Düzce Sampling

[1]Ercan Özgan and [2]Recep Demirci
[1]Department of Construction,
[2]Department of Electric, Faculty of Technical Education,
Düzce Üniversity, Konuralp, Düzce, Turkey

Abstract: In this study, alternatively, Artificial Neural Network (ANN) based modelling of traffic accidents on two line interurban rural highways in terms of number of accidents; injuries and dead have been presented. This study was conducted for D100/11 state highway section in Düzce. In this section of the highway, totally 783 traffic accidents occurred and 1396 vehicles involved in these accidents between 2002 and 2006 years. Using traffic accident reports data, ANN was applied for modelling of traffic accidents with respect to distance and months. As a result, it was observed that there was a perfect fit between the simulation results and actual data of accidents and the created neural network model of accidents resembles the actual data. Therefore, the developed model could be an alternative method for predictions of traffic accidents on interurban rural highways.

Key words: ANN, traffic accidents, identification, data mining, analysis, simulation

INTRODUCTION

The cause that affects traffic accidents can be a factor or combination of many factors. The basic factors, which cause or increase the severity of probable accidents, are drivers behaviour, vehicle features, highway characteristics, environmental effects and traffic characteristics (Ozgan, 2003). The acknowledgments of dangerous or risky section on interurban rural highway route for traffic and road safety are important. Because it will be possible to reduce, traffic accidents and made recovery studies by determining dangerous and risky section. The highway recovery studies can be generally determined by examining dangerous and risky kilometers on highways.

Statistical or crash prediction models have frequently been used in highway safety studies. They can be used to identify major contributing factors or establish relationships between crashes and explanatory variables, such as traffic flows, types of traffic control and highway geometric variables, with the aim that effective countermeasures could be implemented to reduce the number and severity of motor vehicle collisions occurring on different types of highway entities. The models can also be utilized to predict crash frequencies on sites that have not been used for estimating the original models or with different traffic flow and highway geometric conditions. The predicted results could be used in costs/benefit analyses and, if the predicted values are reliably estimated, could greatly help allocate the limited funds available to improve highway safety via the proper identification of hazardous sites (Hauer, 1996; Hauer *et al*., 2004; Miaou and Song, 2005). Previous studies that documented the development and application of crash prediction models have usually focused on statistical regression techniques. Most of these techniques are based on the Generalized Linear Modeling (GLM) framework. Hierarchical Bayes models have also been proposed for modeling motor vehicle collisions (Schluter *et al*., 1997; Tunaru, 2002; Miaou and Lord, 2003; Qin *et al*., 2005; Miaou and Song, 2005). These models have been found to offer superior statistical properties compared to GLMs when crash data are subjected to low sample mean values and small sample size (Lord, 2006).

The ANN methodology has been used in various communications traffic and transportation engineering related problems in the last decade. There are many applications of the ANNs in travel behaviour, traffic flow and traffic management (Himanen *et al*., 1998). Artificial neural networks were employed for modeling the relationship between driver injury severity and crash factors related to driver, vehicle, roadway and environment characteristics. The use of artificial neural

Corresponding Author: Ercan Özgan, Department of Construction, Faculty of Technical Education, Düzce Üniversity, Konuralp, Düzce, Turkey Tel: +90 380 542 11 33 Fax: +90 380 542 11 34

networks can lead to greater understanding of the relationship between vehicle, roadway and environment characteristics and driver injury severity (Abdelwahab and Abdel-Aty, 2001). A neural network-based system identification approach is used to establish an auto-adaptive model for simulating traffic flow dispersion (Fengxiang *et al*., 2001). Hongbin *et al*. (2002) developed a fuzzy-neural model (FNM) to predict the traffic flows in an urban street network.

Modelling of traffic data based on driver characteristics in terms of ANN was proposed (Kalyoncuoglu and Tigdemir, 2004). Some researchers in various fields of research, including highway safety, have proposed the use of neural network models for modeling the phenomenon under study (Mussone *et al*., 1999; Abdelwahab and Abdel-Aty, 2002; Riviere *et al*., 2006). A series of artificial neural networks were used to model the potentially non-linear relationships between the injury severity levels and crash-related factors (Delen *et al*., 2006). ANNs were determined to be a viable alternative to regression for predict speeds on two-lane rural highways (McFadden *et al*., 2007). Artificial Intelligence (AI) approaches has showed a high level of performance in identifying different patterns of accidents in the training data and presented a better fit when compared to the regression model (Wael and Bruce, 2007). A series of models was estimated using data collected on rural frontage roads in Texas. The results of this study show that in general both types of neural network models perform better than the Negative Binomial (NB) regression model in terms of data prediction (Yuanchang *et al*., 2007).

In this study, D-100/11 State highway section between Bolu and Sakarya cities in Turkey has been focused on and the dangerous and risky sections were investigated detail by examining the reports that belong to accidents occurred on this highway in past years. For each kilometer, the number of accident, injury and dead were determined. In addition, Artificial Neural Network (ANN) was successfully performed based modelling of traffic characteristic of the highway in terms of number of accidents, injuries and dead.

MATERIALS AND METHODS

D100/11 state highway section: The accident data used in this study belongs to Turkey's most important arterial road, which includes Bolu mountain passage and D-100/11 State highway section as shown in (Fig. 1a, b). In this section of the highway, totally 783 traffic accidents has been occurred between 2003-2006. As a result of that, 1400 person were injured and 59 persons were died. Moreover, 1396 vehicles were involved in these accidents: 606 cars, 31 minibus, 134 small trucks, 80 buses, 294 trucks, 60 motorcycle, 36 bicycles, 13 tractors and 115 others. When the related data was studied, it was observed that the traffic accidents occurrences were focused in summer mounts (June, July and August). The highway investigated is 48 km long and joins Bolu and Sakarya cities. The Annual Average Daily Traffic (AADT) values in D-100/11 were given (Table 1).

Table 1: Values AADT of D100/11 State Highway Section

Years	Car	Bus	Truck	Trailer
2003	12055	1783	7331	1114
2004	9242	2186	4984	1579
2005	10978	1282	5842	1261
2006	11568	1348	6121	1304

Artificial Neural Networks (ANN): A general structure of a multi-layer NN was shown (Fig. 2). Such a neural network contains three layers: input layer, hidden layer(s) and output layer. Each layer is composed of several neurons. The number of neurons in the input and output layers depends on the system dynamics and the desired accuracy. All the neurons in adjacent layers are interconnected. The strength of the interconnection was determined by weighting vector of NN (Rumelhart and Mc Clelland, 1986).

Each neuron performs two functions as shown below Fig. 3. The first is to sum all the inputs from lower layer based on their weighting factors as given in Eq. 1. The second is to process this sum by a nonlinear sigmoid function as shown in Eq. 2. The input and output neurons may not contain nonlinear functions. The basic equations describing the dynamics of each neuron are;

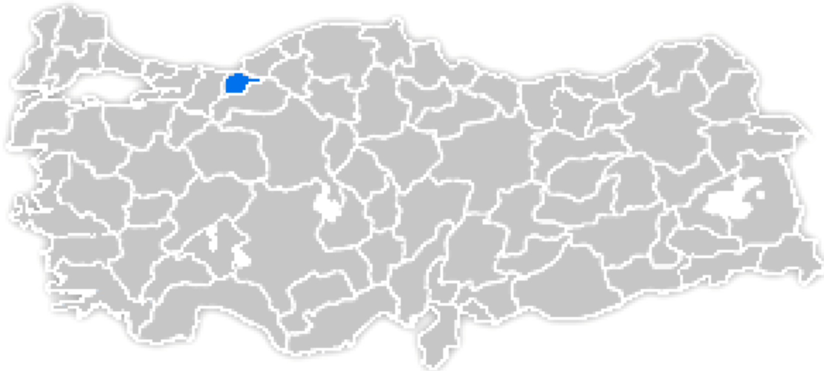
$$net_j = \Sigma_i W_{ij} O_i \tag{1}$$

$$O_j = f(net_j + \theta_j) \tag{2}$$

Where:

W_{ij} = Weight between the jth neuron and the ith neuron in two adjacent layers
θ_j = Threshold of the jth neuron
O_i = Output of the ith neuron
O_j = Output of the jth neuron
f(.) = Sigmoid function

Training of neuron network: The most common method of NN training is back error propagation algorithm. The algorithm is based on the gradient search technique that minimisation process is done by adjusting the weighting vector of the NN. Let the objective function (E) could be written as:



a

b

Fig. 1: (a) Map of Turkey and (b) location of D-100/11 State Highway

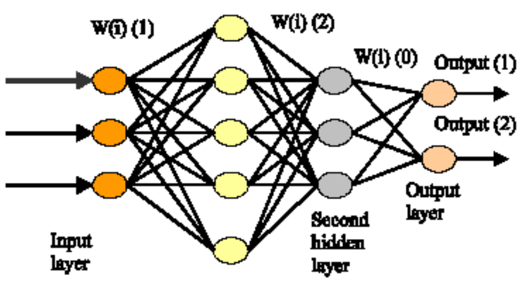


Fig. 2: Multi-layer neural networks

$$E = \frac{1}{2}\sum_{p}\sum_{j}\left(T_{pj} - O_{pj}\right)^2 \qquad (3)$$

Where:

T_{pj} = Target output of neuron j due to input pattern p

O_{pj} = NN output of the same neuron and for the same pattern

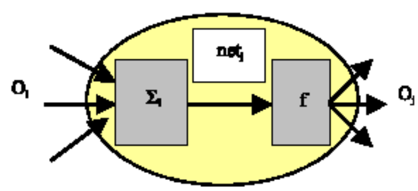


Fig. 3: A single neuron

Minimising Eq. 3 leads to a sequence of update of the weight vector. The weights of the interconnections between two adjacent layers could be updated based on the following formula:

$$W_{ij}(k+1) = W_{ij}(k) + \eta \frac{\delta E}{\delta W_{ij}(k)} + \alpha \Delta W_{ij}(k) \qquad (4)$$

Where:

k = Iteration number

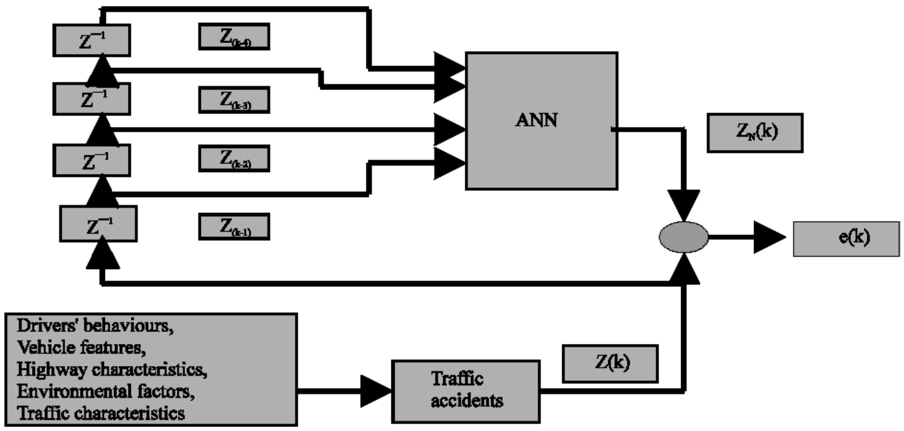


Fig. 4: Basic concept of ANN based modelling traffic accidents

η = Step size
α = Momentum gain
ΔW_{ij} (k) = Weight change based on the gradient of the cost function (Maren and Pap, 1990; Narendra and Parthasarathy, 1990; Yıldırım *et al.*, 1996).

Selection of NN structure: To model the database of accidents by using ANN, firstly the inputs and outputs of ANN have been determined in such a way that the number of accidents, injuries and dead at kth month or kilometre, $Z_{(k)}$ was taken as the output of ANN while the number of accidents, injuries and dead at $Z_{(k-1)}$, $Z_{(k-2)}$, $Z_{(k-3)}$ and $Z_{(k-4)}$ were taken as the inputs to ANN. The NN topology used in identification consists of four linear input neurons, 6 sigmoid hidden neurons and a linear output neuron. The bias inputs of neurons are not included into above numbers. The step size η and the momentum gain α were chosen to be 0.2 and 0.05, respectively (Fig. 4).

RESULTS AND DISCUSSION

From data related to D100/11 state highway, accident numbers, injury and dead numbers for each kilometre were determined and distribution of accidents along the highway were shown (Fig. 5) while the allocation of injury and dead for each kilometre were shown (Fig. 6). It could be seen from Fig. 5 that the most dangerous and risky sections of the highway are 44th kilometre with 83 accidents, 21st kilometre with 54 accidents and 39th kilometre with 47 accidents. Moreover 185 dead and injuries were reported on 44th kilometre, 85 on 21st kilometre and 83 on 39th kilometre. Despite the accident characteristic with distance is quite nonlinear, the proposed ANN approach has ability to model it.

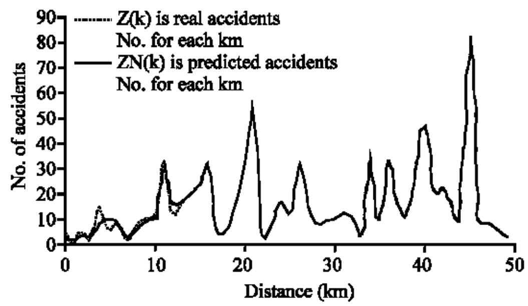


Fig. 5: ANN based modelling of traffic accidents with respect to distance

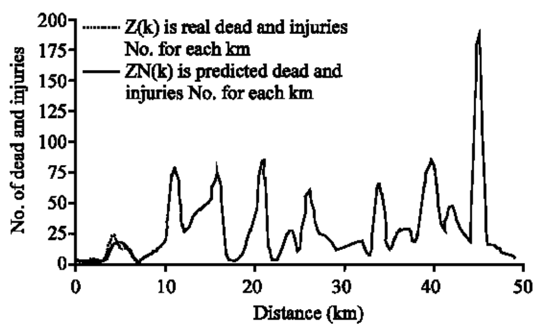


Fig. 6: ANN based modelling of dead and injuries with respect to distance

Accident numbers, injury and dead numbers for each month were also determined from relevant data. It was observed that the most of accidents were respectively occurred with 29 in July in 2006 (55th month), with 23 in July in 2002 (7th month) and with 20 in August in 2004 (32nd month). Contrary to, the least accidents were respectively occurred with 1 in April in 2006 (52nd month), with 3 in November in 2003 (23rd month) and with 5 in May in 2002 (5th month).

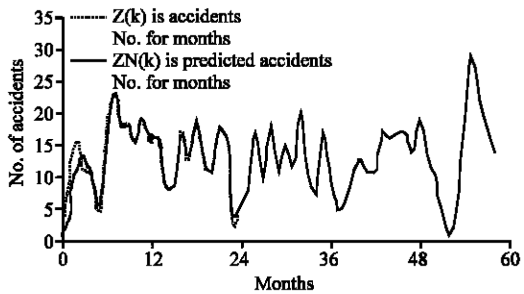


Fig. 7: ANN based modelling of accidents with respect to months

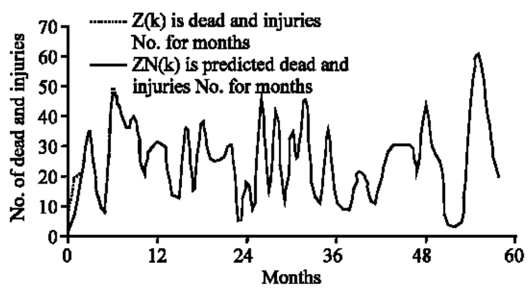


Fig. 8: ANN based modelling of dead and injuries with respect to months

It was observed that the most of dead and injured were respectively occurred with 60 in July in 2006 (55th month), with 50 in August in 2006 (56th month) and with 48 in June in 2002 (6th month). Contrary to, the least dead and injured were respectively occurred with 3 in April in 2006 (52nd month), with 5 in May in 2006 (53rd month) and with 5 in November in 2003 (23rd month).

The data obtained for accident numbers and dead and injuries with respect to months has been trained by ANN structure given Fig. 7 and 8 shows the actual data and ANN outputs for accident numbers. Consequently, although the accident characteristic with months is quite complex, the proposed ANN approach has been able to model them.

CONCLUSIONS

The use of ANN to model the traffic accidents on interurban highway has been confirmed effective in terms of its nonlinear mapping capabilities. Simulation results were verified through relevant data and ANN model was proven to be reasonable accurate. The advantages of the model developed here are that no a priori knowledge is required (mathematical model or equations).

REFERENCES

Abdelwahab, H.T. and M.A. Abdel-Aty, 2001. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalizes intersection. TRR 1746, Paper No. 01-2234, pp: 6-13.

Abdelwahab, H.T. and M.A. Abdel-Aty, 2002. Artificial neural networks and logit models for traffic safety analysis of toll plazas. TRR., 1784: 115-125.

Delen *et al*., 2006. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. Accid. Anal. Prev., 38 (3): 434-444.

Fengxiang, Q. *et al*., 2001. Intelligent simulation and prediction of traffic flow dispersion. TR Part B: Methodological., 35 (9): 843-863.

Hauer, E., 1996. Identification of sites with promise. TRR., 1542: 54-60.

Hauer, E. *et al*., 2004. How to best rank site with promise. TRR., 1897: 48-54.

Himanen, V. *et al*., 1998. Neural Network in Transport Application. Ashgate Pub. Co., 184014808X., USA., pp: 311-340.

Hongbin, Y. *et al*., 2002. Urban traffic flow prediction using a fuzzy-neural approach. TR Part C: Emerging Technologies, 10 (2): 85-98.

Kalyoncuoglu, S.F. and M. Tigdemir, 2004. An alternative approach for modelling and simulation of traffic data: Artificial neural networks. Simulation Modelling Practice and Theory, 12: 351-362.

Lord, D., 2006. Modeling motor vehicle crashes using *Poisson gamma* models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accid. Anal. Prev., 38 (4): 751-766.

Maren, C.H. and R. Pap, 1990. Handbook of Neural Computing Applications. Academic Press, London.

McFadden, J. *et al*., 2007. Application of artificial neural networks to predict speeds on two-lane rural highways. TRR., 1751: 9-17.

Miaou, S.P. and D. Lord, 2003. Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form and *Bayes versus* empirical Bayes. TRR., 1840: 31-40.

Miaou, S.P. and J.J. Song, 2005. Bayesian ranking of sites for engineering safety improvements: Decision parameter, treatability concept, statistical criterion and spatial dependence. Accid. Anal. Prev., 37 (4): 699-720.

Mussone, A. *et al*., 1999. An analysis of urban collisions using an artificial intelligence model. Accid. Anal. Prev., 31 (6): 705-718.

Narendra, K.S. and K. Parthasarathy, 1990. Identification and control of dynamical systems using neural networks. IEEE Trans. Neural Networks, 1 (1): 4-27.

Ozgan, E., 2003. Multi-variate clinical aspects of traffic accidents at Sivas environmental state highway. Ph.D Thesis, Gazi University, Turkey.

Qin, X. *et al*., 2005. Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov Chain Monte Carlo modeling. J. Transport. Eng., 131 (5): 345-351.

Riviere *et al*., 2006. A Bayesian neural network approach to estimating the energy equivalent speed. Accid. Anal. Prev., 38 (2): 248-259.

Rumelhart, D.E. and J.L. McClelland, 1986. Parallel Distributed Processing. Vol. 1. The MIT Press, Cambridge.

Schluter, P.J. *et al*., 1997. Ranking and selecting motor vehicle accident sites by using a hierarchical Bayesian model. Statistician, 46 (3): 293-316.

Tunaru, R., 2002. Hierarchical Bayesian models for multiple count data. Aust. J. Stat., 31 (2/3): 221-229.

Wael, H.A. and N.J. Bruce, 2007. Prediction models for truck accidents at freeway ramps in Washington State using regression and artificial intelligence. TRR., 1635/1998: 30-36.

Yıldırım, S. *et al*., 1996. Design of adaptive NNs-Robust-PID controller for a robot control' proceedings of IEEE. International Symposium on Intelligent Control (ISIC). Dearborn, Michigan, USA., pp: 508-513.

Yuanchang, X. *et al*., 2007. Predicting motor vehicle collisions using Bayesian neural network models an empirical analysis. Accid. Anal. Prev., Article in Press.