Scientific Research and Essays Vol. 5(19), pp. 2837-2847, 4 October, 2010
Available online at http://www.academicjournals.org/SRE
ISSN 1992-2248 ©2010 Academic Journals

Full Length Research Paper

A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics[1]

Darçin Akin[1]* and Bülent Akbaş[2]

[1]Department of City and Regional Planning, School of Architecture, Gebze Institute of Technology, Kocaeli, Turkey.
[2]Department of Earthquake and Structural Engineering, School of Architecture, Gebze Institute of Technology, Kocaeli, Turkey.

Accepted 12 July, 2010

In this paper, a neural network (NN) model was developed to predict intersection crashes in Macomb County of the State of Michigan (MI), USA. The predictive capability of the NN model was determined by grouping the crashes into these types: Fatal, injury and property damage only (PDO) () accidents. The NN approach was used to develop and to test multi-layered feedforward NNs trained with the back-propagation algorithm in order to model the non-linear relationship between the crash types and crash properties such as time, weather, light and surface conditions, driver and vehicle characteristics, and so on. 16000 cases of the crash data were used to train the NN model and the model testing was done by another set of 3200 crashes. A sensitivity analysis was performed to define the effect of crash properties on the crash types. The approach adapted in this study was shown to be capable of providing a very accurate prediction (90.9%) of the crash types by using 48 design parameters (selected based on statistical significance among crash properties defined in the data file). The results were considered to be very promising and encouraging for further research by the expanded data sets in order to estimate future year dependent variables with the model built.

Key words: Motor vehicle crashes, neural network (NN), intersection accidents, crash properties, driver, vehicle, and roadway surface characteristics.

INTRODUCTION

Classification of motor vehicle crashes is very important for the safety management of a highway network since the type of motor vehicle crashes occurring on a segment of highway network is a very significant safety indicator. The ability to accurately predict the type of motor vehicle crashes with input variables, such as time, location, weather, road type, and driver and vehicle characteristics, could significantly reduce highway casualties by designing the on- and off-road features, by providing highway departments information for the analysis as to the likelihood of a specific kind of crash at specified time, and highway location, condition, and feature.

The occurrence of accidents can be attributed to driver, vehicle and roadway characteristics, of which they may be blamed in various degrees for different types and characteristics of users, vehicles types and technologies, road characteristics, and climate conditions. Researchers al., 2005; Lee and Abdel-Aty, 2005; Schneider et al., 2004; Al-Ghamdi, 2002; Summala, 1996; Shankar et al., 1995) have explored the frequency of occurrence of road

*Corresponding author. E-mail: dakin@gyte.edu.tr. Tel: 90-262-605-1624. Fax: 90-262-653-8495.

[1]This paper was presented at the two following conferences and included in their proceedings: the 10th International Conference on Application of Advanced Technologies in Transportation, School of Civil Engineering, National Technical University of Athens, Greece, 2008; and the 3rd International Traffic & Road Safety Congress, Gazi University, Engineering & Architecture Faculty, Ankara, Turkey, 2006. pp. 753-767

accidents on the basis of various roadway geometrics,weather, and driver, vehicle and pedestrian characteristics and built models to explain the relationship between the accident occurrences and their reasons. These studies aimed at uncovering important determinants of accident frequency. By studying the relationship of accident types (fatal, injury and PDO accidents) with weather, roadway, driver and vehicle characteristics, the research offers insight into potential measures to counter the adverse effects of road conditions, road users and vehicle types on highway sections (alignments, intersections and interchanges) with proper geometrics and roadway elements which can provide safer transportation.

Modern neural networks are non-linear statistical data modeling tools that are usually used to model complex relationships between inputs and outputs or to find patterns in data. Neural Network (NN) analysis is considered to be an alternative approach for the investigation of non-linear relationships in engineering problems. A more realistic and accurate predictions can be obtained through NN analyses. NN applications in transportation engineering date back to the beginning of 1990s (Faghri and Hua, 1991; Ripley, 1992, 1994; Belgarovi and Blosseville, 1993; Gilmore et al., 1993; Watson, 1993; Huang and Prahlad, 1994; Nakatsuji et al., 1994; Dougherty, 1995; Hua and Faghri, 1995; Ledoux et al., 1995; Nakatsuji and Kaku, 1995; Ledoux, 1996; Ledoux, 1997; Himanen et al., 1998). Some researchers tested the predictive capability of NNs with statistical or econometric models in traffic safety (Akgüngör and Doğan, 2008 and 2009a, b; Hashemi et al., 1995), traffic forecasting (Danech-Pajouh and Aron, 1991; Clark et al., 1993; Dougherty et al., 1994; Dochy et al., 1995; Dougherty and Cobbet, 1996; Smith and Demetsky, 1996; Kirby e al., 1997), transportation planning (Shmueli, 1998; Hensher and Ton, 2000), and traffic operations (Khan and Ritchie, 1998; Murat, 2006).

Neural networks offer an alternative to regression analysis in the solution of nonlinear engineering problems. In fact, neural network analyses are nothing more than nonlinear regression analyses. However, the advantages of neural networks over regression analyses are that in regression analysis, the analyst has to choose a model to fit the data, while neural networks are not required to pre-select a model, and that in neural networks, sufficient hidden nodes can provide accuracy required for many different response surfaces. Recent research demonstrated the potential use of this technique in transportation engineering (Zhang et al., 1998; Hensher and Ton, 2000; Messai et al., 2002; Kalyoncuoglu and Tigdemir, 2004; Abdel-Aty and Pande, 2005; Delen et al., 2006; Murat, 2006). As an example for the superiority of predictive capability of the technique, Hashemi et al. (1995) compared the capability of a neural network (NN) model with multiple discriminant analysis and logistic regression to predict vessel accidents on the lower Mississippi River. The NN model's prediction power was 1.6 times better than the other two. Akgüngör and Doğan (2008) introduced accident prediction models using artificial neural network (ANN) and nonlinear regression approaches to estimate the number of accidents, injuries and deaths. Their study, which compares ANN and nonlinear models in terms of various error expressions, showed that ANN model had better results against the nonlinear regression models. In another study by Akgüngör and Doğan (2009a), artificial intelligence (AI) models, namely, ANN and genetic algorithm (GA) to predict the number of accidents, injuries and fatalities showed that the ANN models had minimum errors for training and testing data. Murat (2006) modeled vehicle delays at signalized intersections using the Neuro Fuzzy Delay Estimation (NFDE) model and Artificial Neural Networks Delay Estimation (ANNDE) model. The results of the models developed were compared with the Highway Capacity Manual (HCM), Akçelik's methods and the actual delay data collected from the intersections. The results showed that delay estimations by the ANNDE and NFDE models were promising and that the NFDE model was best fitted for the study. The Average Relative Error (ARE) rates of the NFDE model were 7% for under-saturated and 5% for over-saturated conditions, respectively. It was concluded the neuro-Fuzzy approach might be used as a promising method in vehicle delay estimation at signalized intersections.

This study is aimed at developing and testing multi-layered feedforward NNs trained with the back-propagation algorithm to assess accidents that occurred at intersections with different underlying reasons attributed to time of occurrence, weather and surface conditions, and user and vehicle characteristics. The intersection crash data classified by the type of traffic control device (no control, yield or stop sign, signal control or other warning) gave us the chance to have a clear view of the kind and number of accidents at different types of intersection control. Thus, the purpose of this study is to identify the most significant parameters that determine the possibility of an accident occurring at an intersection. Attention has been paid to identify the possible causes of accidents: Roadway conditions, visibility, weather, and the characteristics of vehicles and drivers.

NEURAL NETWORK DESIGN

A typical NN consists of a group of processing elements (PEs) (called neurons) linked together to construct a relation in an input/output set of learning patterns (Figure 1). A PE may be defined by computing the sum of their weighted inputs, subtracting its threshold from the sum, and transferring these results as function as follows (Haykin, 1994):

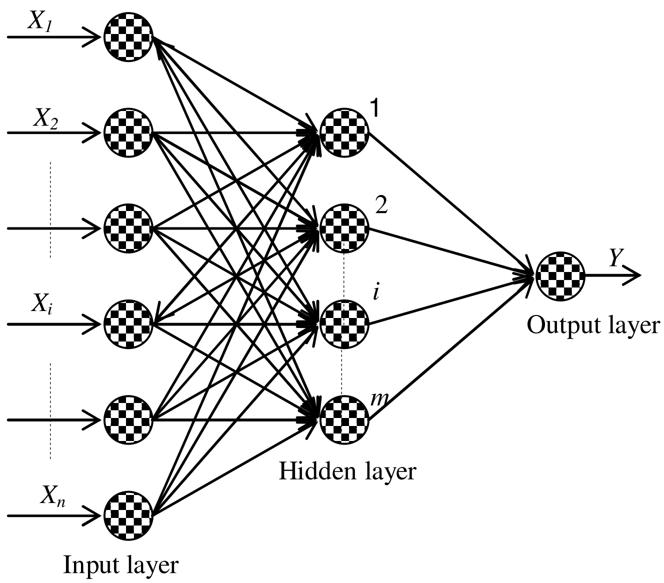

Figure 1. A typical neural network model structure.

$$u_i = \varphi(\sum_{j=1}^{n} w_{ij} x_j - \theta_i) \quad (1)$$

where u_i represents the output of a PE, w_{ij} represents the synaptic weights associated with PE *i*, x_j represents the input signal, θ_i represents the threshold value of the PE, and φ represents the transformation (or activation) function. The activation function is used to limit the amplitude of the output of a PE (Haykin, 1994).

The NN model in this study was developed in three phases: Modeling, training, and testing phases. The gathering of data, identification of input parameters, and the internal rules were considered in the modeling phase. The preparation of the data and the adaptation of the learning law for the training were performed during the training phase. And the prediction accuracy of the model was evaluated at the testing phase, that is, the comparison of the actual outputs and the estimated outputs.

ANNs typically start out with randomized weights for all their neurons. This means that weights are to be estimated for the solution of a particular problem. When a satisfactory level of performance is reached, the training ends and the network uses these weights to make a decision. Multi-layer perception (MLP) networks model is usually preferred in engineering applications because many learning algorithms can be used in MLP. One of the commonly used learning algorithms in ANN applications is the back propagation (BP) algorithm Akgüngür and Doğan, 2009; Tokar and Johnson, 1999), which is also used in this research.

In networks, the information about the error is provided backwards from the output to the input layer. The objective of a BP network is to find the weight that approximate target values of the output with a selected level of accuracy (Mohammadi et al., 2005).

The data

An official database of the accidents compiled by the State Police Department of the State of Michigan was used. First, a descriptive analysis of the data related to 26,939 crashes that occurred in Macomb County in the Southfield-Metro Region in 1995 revealed that 19,959 crashes (63%) occurred at intersections. The number of accidents that occurred at interchanges was 2,163 (8%). Non-motorized accidents were only 115 cases (0.4%). The highest number of accidents that occurred on one of the day during week (Friday) was by 17.8%. Then, Thursday, Wednesday, Tuesday, Saturday, Monday, and Sunday were as follows 16, 15, 14.2, 14, 13.6 and 9.3%, respectively. Accident occurrence during morning-peak hours (4.9% between 7 - 8 am and 4.2% between 8 - 9 am) was lower than that during afternoon-peak hours (8.6% between 4 - 5 pm and 9.2% between 5 - 6 pm). The percent from 11 pm to 7 am was the lowest ranging between 0.5 and 2.1%. Interestingly, the frequency of accidents that occurred in days of the month was evenly distributed with an average of about 3%. County roads and city streets took all the blame with respect to the highest representation of crash occurrences (61.3%) on different types of roads. The percent of crash occurrences on Interstate highways was 6% and that on MI routes (highways administrated by the State of Michigan) was 26.8%.

Modeling the neural network

One hidden layer was used in the ANN model. The momentum coefficient of 0.7 for the hidden and output layers was performed very well. The step size in the learning rate was selected as 1.0 for the hidden and 0.2 for the output layers. The NN model's accuracy is significantly affected by the selection of the input variables. The variables related to the crashes analyzed; are time variables (hour, weekday, and month), roadway characteristics (route class, traffic control device), road conditions (alignment, light, surface condition, road/ pavement defect, visual obstruction), climate (weather, temperature), driver characteristics (age, sex, intent, violation, degree of injury, drinking or drug use), vehicle characteristics (vehicle type and make), and accidents characteristics (accident type, impact code, object hit, vehicle situation and condition, contributing circumstances, special tags).

To build the NN model, design parameters (statistically significant ones among crash properties) shown in Table 1 were fed into the input layer in order to evaluate output parameters (dependent variables): Y_1: Total number of type A injuries, Y_2: total number of type B injuries, Y_3: Total number of type C injuries, Y_4: Total number of occupants killed, Y_5: Total number of occupants injured, Y_6: Accident type (fatal=1, injury = 2, and PDO = 3). The 48 design parameters are: (1) X_1: Day of week, (2) X_2: Hour of occurrence, (3) X_3: Month of occurrence, (4) X_4: Weather condition, (5) X_5: Light condition, (6) X_6: Road surface condition, (7) X_8: Traffic control device, (8) X_{10}: Accident type, (9) X_{11}: Number of moving vehicles involved, (10) X_{13}: Type of vehicle 1, (11) X_{14}: Age of driver 1, (12) X_{141}: Age group of driver 1, (13) X_{15}: Sex of driver 1, (14) X_{17}: Intent of driver 1, (15) X_{18}: Violation of driver 1, (16) X_{19}: Contributing circumstances around vehicle 1, (17) X_{22}: Drinking or drug use in vehicle 1, (18) X_{23}: Object hit by vehicle 1, (19) X_{25}: Size of vehicle 1, (20) X_{26}: Impact code of vehicle 1, (21) X_{29}: Type of vehicle 2, (22) X_{30}: Age of driver 2, (23) X_{31}: Sex of driver 2, (24) X_{33}: Intent of driver 2, (25) X_{34}: Violation of driver 2, (26) X_{35}: Contributing circumstances around vehicle 2, (27) X_{36}: Visual obstructions to vehicle 2, (28) X_{38}: Drinking or drug use in vehicle 2, (29) X_{39}: Object hit by vehicle 2, (30) X_{40}: Situation of vehicle 2, (31) X_{41}: Size of vehicle 2, (32) X_{42}: Impact code of vehicle 2, (33) X_{43}: Condition of vehicle 2, (34) X_{44}: Type of trailer behind vehicle 2, (35) X_{45}: Type of vehicle 3, (36) X_{46}: Age of driver 3, (37) X_{47}: Sex of driver 3, (38) X_{49}: Intent of driver 3, (39) X_{50}: Violation of driver 3, (40) X_{51}: Contributing circumstances around vehicle 3, (41) X_{52}: Visual obstructions to vehicle 3, (42) X_{54}: Drinking or drug use in vehicle 3, (43) X_{55}: Object hit by vehicle 3, (44) X_{56}: Situation of vehicle 3, (45) X_{57}: Size of vehicle 3, (46) X_{58}: Impact code of vehicle 3, (47) X_{59}: Condition of vehicle 3, and (48) X_{60}: Type of trailer behind vehicle 3.

These forty-eight parameters are assumed to be the predominant independent variables of the NN model developed to classify the intersection crashes. The parameter of “road alignment (straight, curve, transition)” could not be included in the analysis because it was coded as “not known” for all cases in the database. The parameters of “where” and “how accidents occurred” were also not included in the analysis because they were coded as zero. Besides, the following parameters were excluded from the model by not finding them statistically significant to explain the dependent variables: (1) X_7: Road defect, (2) X_9: Special accident tags, (3) X_{12}: Number of persons uninjured, (4) X_{16}: Degree of injury to driver 1, (5) X_{20}: Visual obstructions to vehicle 1, (6) X_{21}: Direction of travel of vehicle 1, (7) X_{24}: Situation of vehicle 1, (8) X_{27}: Condition of vehicle 1, (9) X_{28}: Type of trailer behind vehicle 1, (10) X_{32}: Degree of injury to driver 2, (11) X_{37}: Direction of travel of vehicle 2, (12) X_{48}: Degree of injury to driver 3, and (13) X_{53}: direction of travel of vehicle 3.

The training phase

Standard back-propagation BP algorithm for the training of the NN network was employed in this study. Commercial NN software (Neuro Solutions, 2003) was used to implement this training method. The activation function used was the hyperbolic tangent function in the model. In this study, an epoch number of 5000 for all was found to be adequate for the final training process in a series of more than 50 runs.

Sample cases used in constructing the NN model are given in Table 2. A total of 16,384 cases which included six output and 48 input variables were divided into two sets. One set first 13108 cases in Table 2) was used for training of the model, and the other was for validating the performance of the trained network (testing). The data set used for training and testing was normalized.

The testing phase

For the testing purpose, 20% of the data (3276 cases) were selected at random order for the testing set for each training cycle (cases below the dotted line in Table 2). The NNs performance was measured by using the vehicle crash (VC) percentage error (PE_{VC}) formula as follows:

$$PE_{VC} = \frac{x(i) - X(i)}{X(i)} x100\% \quad (2)$$

To evaluate the entire NNs overall performance, weighted error (*WE*) was defined as follows (Hegazy and Ayed, 1998):

WE (%) = 0.5 (Average PE_{VC} for Training Set) + 0.5 (Average PE_{VC} for Testing Set) (3)

Average PE_{VC} for the 3276 testing cases was calculated as 8.74% (Y_is vs. NN-Y_is for the 3276 testing samples are given in Figure 2), while it was 9.45% for the training set (13108 cases). Thus, the weighted error (*WE*) was found to be 9.10%. To learn the effect of each input parameter on the output variables, a sensitivity analysis is carried out, during which the network weights are not subject to change, because the network training is disabled. The inputs to the network are shifted slightly and the corresponding change in the output is reported as a percentage, summing to 100% in total (Neuro Solutions, 2003). Sensitivity results of the design parameters are given in Figure 3 and they are discussed in the following -section.

Comments on modeling results

Crash data of 3276 cases out of a total of 16,384 crashes were used for the testing purpose. The modeling results showed a very high accuracy of 90.90% (=100% - 9.10%) in average. The most sensitive parameters were found to

Table 1. Design parameters.

Design parameters	Definition	Range
X_1	Day of week	1: Sunday thru 7: Saturday
X_2	Hour of occurrence	Midnight thru 23:00
X_3	Month of occurrence	1: January thru 12: December
X_4	weather condition	1: clear or cloudy, 2: fog, 3: rain, 4: snow, 5: other-unknown
X_5	Light condition	1: daylight, 2: dawn or dusk, 3: dark with street lights, 4: dark without street lights, 5: unknown
X_6	Road surface condition	1: Dry, 2: wet, 3: snow or ice, 4: other
X_8	Traffic control device	1: None, 2: stop sign, 3: signal, 4: regulator, 5: flasher, 6: yield, 7: school-zone, 8: no-pass zone, 9: other warning, 10: other
X_{10}	Accident type	10: Overturned, 141: head-on, 545: head-on left-turn, 645: dual left-turn, 646: dual right-turn, 144: angle straight, 147: rear-end, 244: angle-turn, 342: side-swipe same, 543: side-swipe opposite, 345: rear-end left-turn, 346: rear-end right-turn, 447: rear-end drive, 440: other drive, 48: backing, 49: parking, 20: with train, 30: with parked vehicle, 50: pedestrian, 90: bicycle, 60: fixed object, 70: other object, 80: animal
X_{11}	Number of moving vehicles involved	From minimum 1 vehicle to depend upon the severity of the crash
X_{13}, X_{29}, X_{45}	Type of vehicles 1, 2, 3	1: Passenger car, 2: truck, 3: motorcycle-scooter- moped, etc., 4: school bus, 5: commercial bus, 6: farm equipment, 7: construction equipment, 8: ambulance, police, emergency vehicle, 9: pedestrian, 10: pedal-cycle, 11: other
X_{14}, X_{30}, X_{46}	Age of drivers 1, 2, 3	Age of driver 1 based on the record of his/her drivers license, 99: unknown
X_{141}	Age group of drivers 1	1: 15 - 25, 2: 26 - 35, 3:36 - 45, 4:46 - 55, 5: 56 - 65, 6:66 and over
X_{15}, X_{31}, X_{47}	sex of drivers 1, 2, 3	1: male, 2: female
X_{17}, X_{33}, X_{49}	Intent of drivers 1, 2, 3	1: Going straight, 2: passing, 3: changing lanes, 4: making right turn, 5: making left turn, 6: making U turn, 7: slowing/stopping, 8: starting up, 9: entering park space, 10: leaving park space, 11: backing, 12: stopped, 13: police chasing, 14: avoiding object, 15: avoiding animal, 16: avoiding pedestrian, 17: lost load, 18: avoiding vehicle, 19: avoiding vehicle at angle, 20: other/unknown
X_{18}, X_{34}, X_{50}	Violation of drivers 1, 2, 3	1: No hazard action, 2: speeding, 3: slow moving, 4: failing yield row, 5: wrong way, 6: improper lane use, 7: improper turn no signal, 8: improper backing, 9: following too close, 10: other/unknown
X_{19}, X_{35}, X_{51}	Contributing circumstances around vehicles 1, 2, 3	0: Other/unknown, 1: duil or drugs, 2: reckless, 3: ill/fatigue/ill attention, 4: license restricted, 5: obscured vision, 6: defective equipment, 7: load shift/wind, 8: none, 9: skidding
X_{22}, X_{38}, X_{54}	Drinking or drug use in vehicles 1, 2, 3	1: Had been drinking, 2: had not been drinking, 3: unkown
X_{23}, X_{39}, X_{55}	Object hit by vehicles 1, 2, 3	1: no object hit, 2: guardrail, 3: highway sign, 4: pole, 5: culvert, 6: ditch, 7: bridge pier, 8: bridge rail, 9: tree, 10: signal ,11: building, 12: mailbox, 13: fence, 14: traffic island, 15: concrete median barrier, 16: other on-road object, 17: other off-road object, 18: overhead object, 19: unknown or unmoving object

Table 1. Contd.

X_{25}, X_{41}, X_{57}	Size of vehicles 1, 2, 3	1: passenger (<1500lbs), 2: passenger (1500-2499lbs), 3: passenger (2500-3500lbs), 4: passenger (>3500lbs), 5: carry station wagon, etc, 6: jeep, 7: pickup or pan truck, 8: st,dmp,stp,mthm, 9: tru,semi,rd trac, 10: other/unknown
X_{26}, X_{42}, X_{58}	Impact code of vehicles 1, 2, 3	0: rollover, 1: center front, 2: right front, 3: right side, 4: right rear, 5: center rear, 6: left rear, 7: left side, 8: left front, 9: other impact/miscellaneous
X_{36}, X_{52}	Visual obstructions to vehicles 2, 3	1: none, 2: within or on vehicle, 3: physical, 4: weather, 5: glare, 6: other/unkown
X_{40}, X_{56}	Situation of vehicles 2, 3	1: rebut from guardrail, 2: thru guardrail, 3: into median, 4: thru median, 5: hit object after collision, 6: run thru T-intersection, 7: none above, 8: hit & run
X_{43}, X_{59}	Condition of vehicles 2, 3	1: disabled vehicle, 2: blowout, 3: defective equipment, 4: no defect, 5: unknown
X_{44}, X_{60}	Type of trailer behind vehicles 2, 3	1: none, 2: utility trailer, 3: single bottom trailer combo, 4: double bottom trailer combo, 5: house trailer, 6: other/unknown, 7: towed vehicle after 1976

be impact code (rollover, center front, right front, right side, right rear, center rear, left rear, left side, left front, and so on), contributing circumstances (duil or drugs, reckless, ill/fatigue/ill attention, license restricted, obscured vision, defective equipment, load shift/wind, skidding), and object hit by vehicles (guardrail, highway sign, pole, culvert, ditch, bridge pier, bridge rail, tree, signal, building, mailbox, fence, traffic island, concrete median barrier, other on-road object, other off-road object, overhead object, and so on), which are followed by violations of driver (speeding, slow moving, failing yield row, wrong way, improper lane use, improper turn no signal, improper backing, following too close, and so on), visual obstructions (within or on vehicle, physical, weather, glare, and so on), number of moving vehicles involved, intent of driver (going straight, passing, changing lanes, making right and left turns, making U turn, slowing/stopping, starting up, entering and leaving park space, backing, stopped, police chasing, avoiding vehicle, object, animal and pedestrian, lost load, and so on), and situation of vehicles (rebut from guardrail, thru guardrail, into and thru median, hit object after collision, run thru T-intersection, hit and run). Type and size of vehicles and age of drivers were also important contributors to predict the type of crash as fatal, injury and PDO crashes. It should be interestingly noted that month of year was the least sensitive, the second least sensitive parameter was the sex of driver 1, and the third one was the type of traffic control devices. The parameter of road defect had no effect at all, so it was excluded from the analysis.

Conclusions

In this study, a neural network (NN) model was employed to predict intersections crashes by crash types as fatal, injury and property damage only (PDO) accidents by using the database, belonging to year 1995, compiled by the State Police of the State of Michigan. The data of 13108 cases were used to train the NN model. The testing of the NN was done by the data of 3276 cases. The approach adapted in this study was shown to be capable of providing a very accurate estimate (mean weighted error = 9.10%) of vehicle crashes by using 48 design parameters (selected based upon statistical significance among crash properties defined in the data file).

Main conclusions drawn from the descriptive analyses of the crash data and modeling of the design parameters (crash properties) are as follows:

1. In year 1995, out of 26,939; 63 and 8% of the crashes occurred at intersections and interchanges, respectively, in Macomb County in the Southfield-Metro Region, and thus, the likelihood of getting involved in an accident at an intersection seems to be higher, not in that degree though, than that in an interchange.
2. Non-motorized accidents were only 115 cases, which occur as the result of high motorization in the region.
3. The last working day of the week had the highest number of accidents by 17.8%, which is 1.3 times higher than that of the first working day probably due to the highest day of travel in average.
4. Accident occurrence (8.9%) during afternoon-peak hours was 1.9 times higher than that (4.6%) during morning-peak hours probably because of loss of attention during driving after a workday load.
5. The highest crash occurrences (61.3%) were observed on the county roads and city streets because of high travel demand on such roads as well as high number of such links in the highway network. (Michigan State)

Table 2. Sample cases used in constructing and testing the NN model *X* is: input variables, *Y* is: output variables.

Case no	X_1	X_2	X_3	X_4	X_5	X_6	X_8	$X_{10...}$	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
	For explanations see Table 1								Type A injuries	Type B injuries	Type C injuries	Occupants killed	Occupants injured	Accident type (fatal = 1, injury = 2, and PDO = 3)
1	7	2	1	1	3	1	2	60	0	0	0	0	0	3
2	6	3	1	3	4	2	10	0	0	0	0	0	0	3
3	1	21	3	3	4	2	10	60	0	0	0	0	0	3
4	6	16	3	1	1	1	3	545	0	0	0	0	0	3
5	7	23	6	1	4	1	10	49	0	0	0	0	0	3
6	2	21	6	1	1	1	2	144	0	0	0	0	0	3
7	1	3	6	1	4	1	10	60	0	0	0	0	0	3
8	3	17	7	1	1	1	3	147	0	0	0	0	0	3
9	3	18	7	1	1	1	3	147	0	0	0	0	0	3
10	7	16	8	1	1	1	3	147	0	0	0	0	0	3
11	4	11	10	1	1	1	3	545	0	0	0	0	0	3
12	5	9	10	1	1	1	3	48	0	0	0	0	0	3
13	5	11	10	1	1	1	3	543	0	0	0	0	0	3
14	4	13	10	1	1	1	10	444	0	0	0	0	0	3
15	4	16	10	1	1	1	3	147	0	0	0	0	0	3
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
13108	4	21	6	5	5	1	3	244	0	0	1	0	1	2
1	6	12	6	1	1	1	10	545	0	1	0	0	1	2
2	5	8	6	1	1	1	3	144	0	1	1	0	2	2
3	5	15	5	3	1	2	3	543	0	0	0	0	0	3
4	4	13	5	3	1	2	3	244	0	0	0	0	0	3
5	7	14	6	1	1	1	3	60	0	1	0	0	1	2
6	5	16	7	3	1	2	10	444	1	0	0	0	1	2
7	7	14	7	1	1	1	10	147	0	0	0	0	0	3
8	2	13	7	1	1	1	10	342	0	0	0	0	0	3
9	6	12	10	1	1	1	10	144	0	0	1	0	1	2
10	6	13	10	1	1	1	2	345	0	1	0	0	1	2
11	2	17	5	5	1	1	3	144	0	0	0	0	0	3
12	5	17	7	3	1	2	3	147	0	0	0	0	0	3
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
3276	2	14	8	1	1	1	3	144	0	0	0	2	0	2

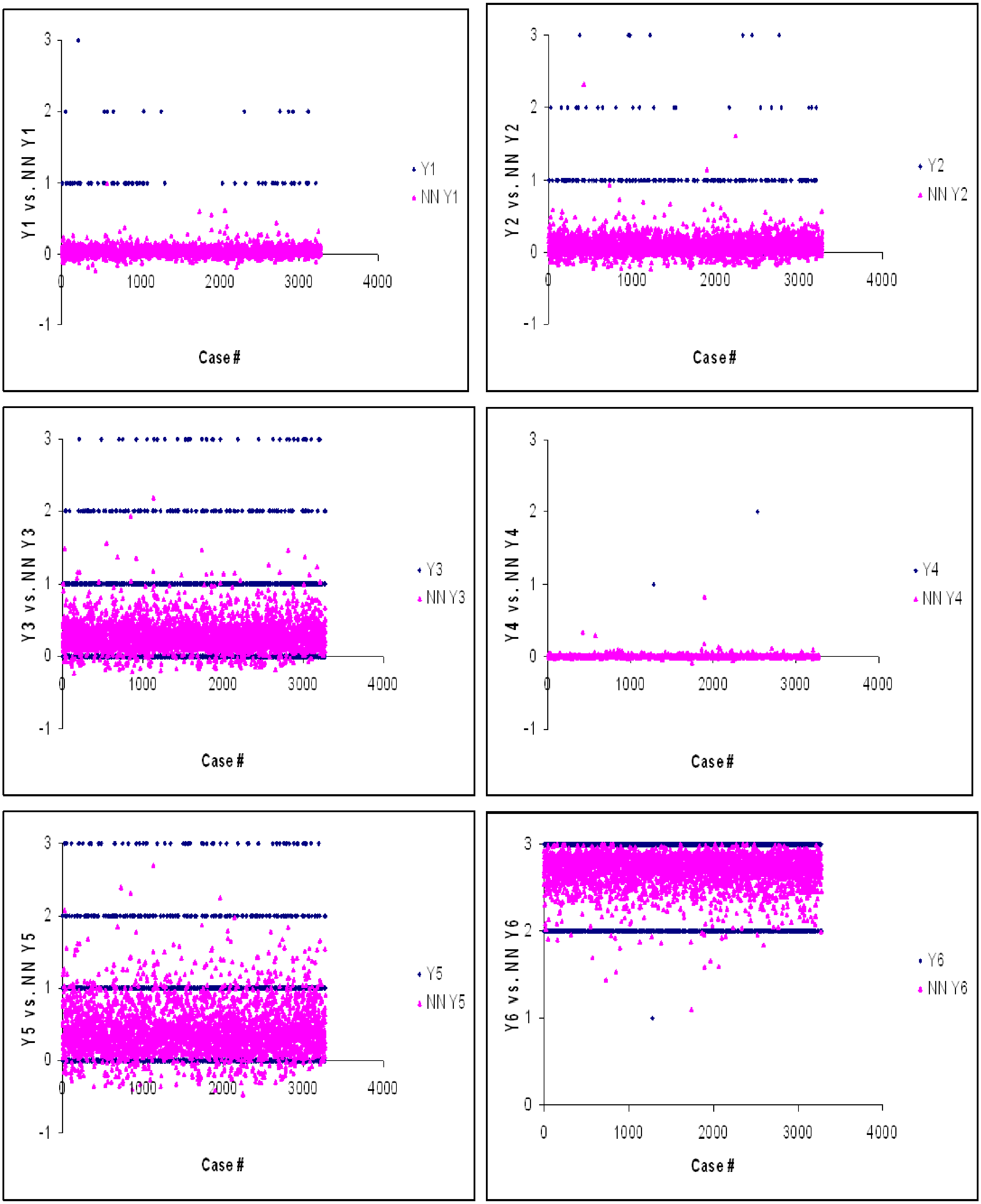

Figure 2. Y_is vs. NN-Y_is for the 3276 testing cases.

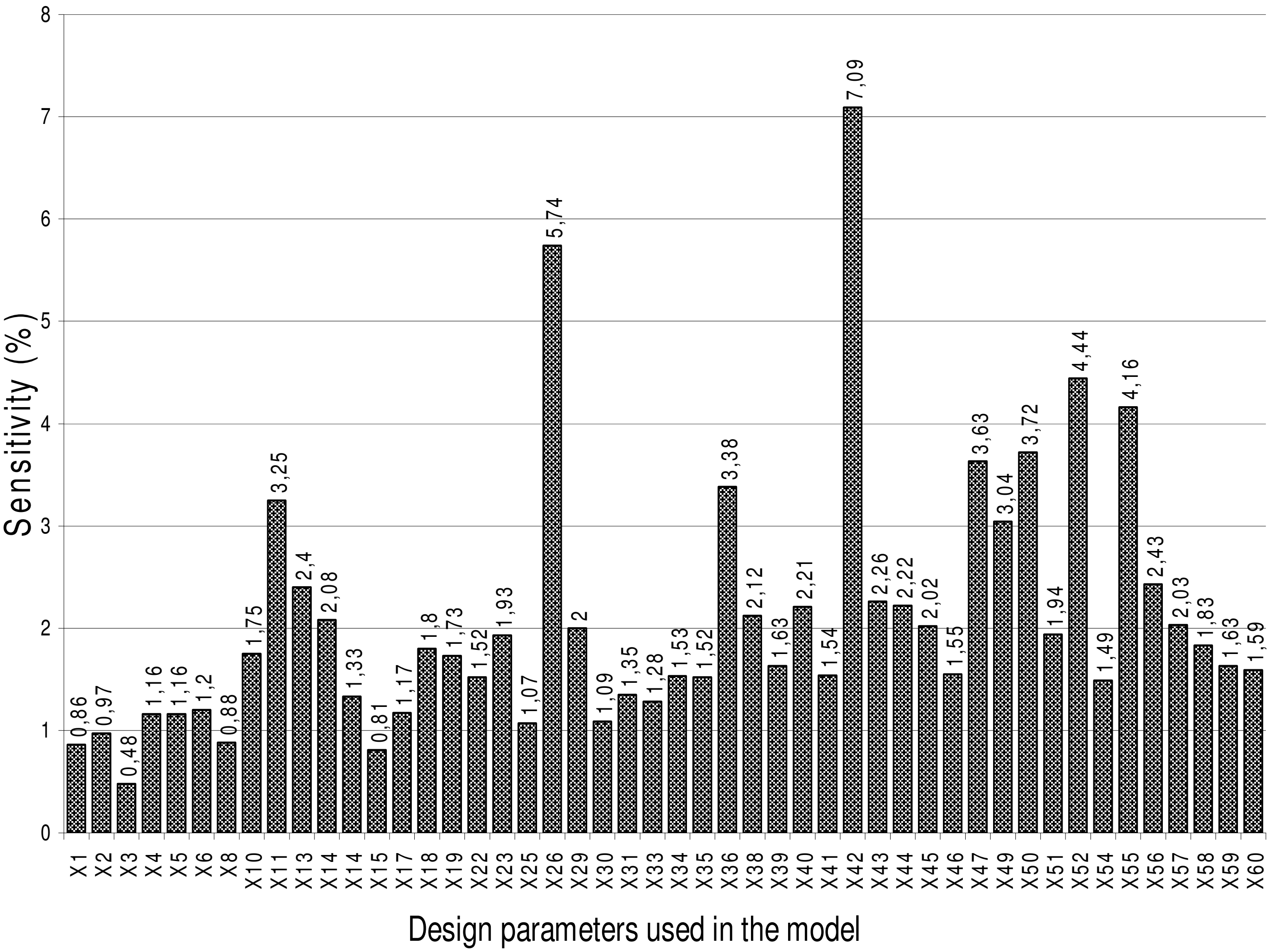

Figure 3. Sensitivity analysis of the design parameters.

(MI) highways had the second highest crash occurrences by 26.8%.

6. Modeling of the design parameters (48 crash properties) revealed that the most sensitive parameters in classifying the crashes into the types of fatal, injury and PDO accidents were impact code, contributing circumstances, and object hit by vehicles which are followed by violations of driver, visual obstructions, number of moving vehicles involved, intent of driver, and situation of vehicles.

7. Type and size of vehicles and age of drivers were also important contributors to predict the type of crash such as fatal, injury and property damage only (PDO) crashes. It should be interestingly noted that month of year is the least sensitive, the second least sensitive parameter was the sex of driver 1, and the third one was the type of traffic control devices.

8. The parameter of road defect hads no effect at all, so it was excluded from the analysis.

RECOMMENDATIONS FOR FUTURE RESEARCH

The modeling results obtained using a single year data (belonging to year 1995) 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 types of vehicle crashes based on various crash properties for a year in the future.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Dale Lighthizer of the Michigan Department of Transportation and Prof. Dr. Richard Lyles of Michigan State University for providing the crash data for the sole purpose of doing academic research and writing papers. Also the authors are grateful for the comments and discussions by two anonymous referees that helped to improve the previous versions of

this paper presented in the aforementioned conferences in the title page.

REFERENCES

Abdel-Aty M, Pande A (2005). Identifying crash propensity using specific traffic speed conditions. J. Safety Res., 36(19): 97-108.

Al-Ghamdi AS (2002). Pedestrian-vehicle crashes and analytical techniques for stratified contingency tables. Accid. Anal. Prev., 34: 205-214.

Akgüngör AP, Doğan E (2008). Estimating road accidents of Turkey based on regression analysis and artificial neural network approach. Adv. in Transportation Stud. An Int. J. Section, A 16: 11-22.

Akgüngör AP, Doğan E (2009a). An Application of Modified Smeed, Adapted Andreassen and Artificial Neural Network Accident Models to Three Metropolitan Cities of Turkey. Sci. Res. Essays, 4(9): 906-913.

Akgüngör AP, Doğan E (2009b). An artificial intelligent approach to traffic accident estimation: Model Dev. Applic. Transport, 24(2): 135-142.

Belgarovi B, Blosseville JM (1993). A road traffic application of neural techniques. Recherche Tranports Sec. Eng., 9: 53-65.

Clark SD, Dougherty MS, Kirby HR (1993). The use of neural networks and time series models for short term traffic forecasting: a comparative study. Presented at PTRC 21st Summer Annual Meeting, Manchester, London.

Danech-Pajouh M, Aron M (1991). ATHENA: a method for short-term inter-urban motorway traffic forecasting. Recherche Transports Sécurité. 6: 11-16 (English Issue).

Delen, D, Sharda R, Bessonov M (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. Accident Anal. Prev., 38(3): 434-444.

Dochy T, Danech-Pajouh M, Lechevallier Y (1995). Short term road traffic forecasting using neural networks. Recherche Transports Sécurité. 11:73–82 (English Issue).

Dougherty MS (1995). A review of neural networks applied to transport. Transportation Research Part C: Emerg. Technol. 3(4): 247-260.

Dougherty MS, Kirby HR, Boyle RD (1994). Using neural networks to recognize, predict and model traffic. Presented at Artificial Intelligence Applications to Traffic Engineering, VSP, Utrecht, The Netherlands, pp. 233-250.

Dougherty MS, Cobbet MR (1996). Short-term interurban traffic forecasts using neural networks. Int. J. Forecast, 13: 21-31.

Faghri A, Hua J (1991). Evaluation of artificial neural network applications in transportation engineering. Transportation Research Record: Journal of the Transportation Research Board, TRB, National Research Council, Washington, D.C., 1358: 71-80.

Gilmore JF, Elibiary KJ, Abe M (1993). Traffic management applications of neural networks systems," Working notes, AAAI-93 Workshop on AI in Intelligent Highways Systems, pp. 85-95.

Hashemi RR, Le Blanc LA, Rucks CT, Shearry A (1995). A neural network for transportation safety modeling. Expert Syst. Appl., 9(3): 247-256.

Haykin S (1994). Neural Networks: A comprehensive Foundation, Macmillan College Publishing Company, Inc. NJ.

Hegazy T, Ayed A (1998). Neural Network Model for Parametric Cost Estimation of Highway Projects. J. Constr. Eng. Manage. 124(3): 210-218.

Hensher DA, Ton TT (2000). A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice. Transportation Research Part E: Logistics and Transportation Rev., 36(3): 155-172.

Himanen V, Nijkamp P, Reggiani A, Raitio J. (1998). Neural Networks in Transport Applications, Ashgate, USA.

Holdridge JM, Shankar VN, Ulfarsson GF (2005). The crash severity impacts of fixed roadside objects. J. Safety Res., 36(2): 139-147.

Hua J, Faghri A (1995). Development of neural signal control system—toward intelligent traffic signal control. Transportation Research Record: Journal of the Transportation Research Board. TRB.

National Research Council, Washington, D.C. 1497: 53-61.

Huang VP, Prahlad PD (1994). A simulation neural network model for evaluating dilemma zone problems. Presented at 73rd Annual Meeting of Transportation Research Board, Washington, DC, pp. 9-13.

Kalyoncuoglu SF, Tigdemir M (2004). An alternative approach for modelling and simulation of traffic data: artificial neural networks. Simul. Modelling Pract. Theory, 12(5): 351-362.

Khan SI, Ritchie SG (1998). Statistical and neural classifiers to detect traffic operational problems on urban arterials. Transportation Research Part C: Emerg. Technol., 6(5-6): 291-314.

Kirby HR, Watson SM, Dougherty MS (1997). Should we use neural networks or statistical models for short-term motorway traffic forecasting? Int. J. Forecast., 13(1): 43-50.

Ledoux C (1996). An urban traffic control system integrating neural networks. Presented at Eighth International Conference on Road Traffic Monitoring and Control, London, pp. 197-201.

Ledoux C (1997). An urban traffic flow model integrating neural networks. Transportation Research Part C: Emerg. Technol., 5(5): 287-300.

Ledoux C, Boillot F, Selam S, Gallinari P (1995). On the use of neural networks techniques for traffic flow modeling. Presented at Second World Congress on Applications of Transport Telematics and Intelligent Vehicle Highway Systems, Yokohama, Japan.

Lee C, Abdel-Aty M (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accid. Anal. Prev., 37(4): 775-786.

Messai N, Thomas P, Lefebvre D, El Moudni A (2002). Optimal neural networks architectures for the flow–density relationships of traffic models. Math. Comput. Simul., 60(3-5): 401-409.

Mohammadi K, Eslami HR, Dardashti D (2005). Comparison of Regression, ARIMA and ANN Models for Reservoir Inflow Forecasting using Snowmelt Equivalent (a case study of Karaj). J. Agric. Sci. Technol., 7: 17-30.

Murat YŞ (2006). Comparison of Fuzzy Logic and Artificial Neural Networks Approaches in Vehicle Delay Modeling. Transportation Research Part C- Emerging Technologies, 14(1): 316-334.

Nakatsuji T, Seki S, Shibuya S, Kaku T (1994). Artificial intelligence approach for optimizing traffic signal timings on urban road network. Presented at Vehicle Navigation and Information Systems Conference, Yokohama, Japan, pp. 199-202.

Nakatsuji T, Kaku T (1995). Development of a self-organizing traffic control system using neural network models. Transportation Research Record: Journal of the Transportation Research Board, TRB, National Research Council, Washington, D.C. 1324: 137-145.

Neuro Solutions (2003). Neurodimension, Inc., Version 4.24.

Neyens DM, Boyle LN (2008). The influence of driver distraction on the severity of injuries sustained by teenage drivers and their passengers. Accid. Anal. Prev., 40(1): 254-259.

Paleti R, Eluru N, Bhat CR (2010). Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accident Analysis and Prevention. Article in press.

Ripley BD (1992). Statistical aspects of neural networks. Presented at Semstat. Sandbjerg, Denmark, Chapman and Hall, London.

Ripley BD (1994). Neural networks and related methods for classification (with discussion). J. Royal Stat. Soc., 56(3): 409-456.

Schneider RJ, Ryznar RM, Khattak RM (2004). An accident waiting to happen: a spatial approach to proactive pedestrian planning. Accid. Anal. Prev., 36(2): 193-211.

Shankar V, Mannering F, Barfield W (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. Accid. Anal. Prev., 27(3): 371-389.

Shmueli D (1998). Applications of neural networks in transportation planning. Progr. Plann., 50(3): 141-204.

Smith BL, Demetsky MJ (1996). Short-term traffic flow prediction: neural network approach. Transportation Research Record: Journal of the Transportation Research Board, TRB, National Research Council, Washington, D.C., 1453: 98-104.

Summala H (1996). Accident risk and driver behavior. Safety Science, 22(1-3):103–117.

Tokar AS, Johnson PA (1999). Rainfall-Runoff Modeling Using Artificial Neural Network. J. Hydrol. Eng., 4(3): 232-239.

Watson, SM (1993). An introduction to adaptive forecasting techniques. ITS Technical Note 233. Institute for Transport Studies, University of

Leeds, Leeds, UK.
Zhang G, Patuwo BE, Hu MY (1998). Forecasting with artificial neural networks: The state of the art. Int. J. Forecast., 14: 35-62.