

PREDICTION OF PEDAL CYCLISTS AND PEDESTRIAN FATALITIES FROM TOTAL MONTHLY ACCIDENTS AND REGISTERED PRIVATE CAR NUMBERS

Kiarash Ghasemlou¹, Metin Mutlu Aydin², Mehmet Sinan Yıldırım³

¹Istanbul Technical University, Department of Civil Engineering, İstanbul, Turkey
e-mail: kiarash.ghasemlou@gmail.com

²Gümüşhane University, Department of Civil Engineering, Gümüşhane, Turkey
e-mail: metinmutluaydin@gmail.com

³Celal Bayar University, Department of Civil Engineering, Manisa, Turkey
e-mail: msyildirim35@gmail.com

Abstract: Accident prevention is relatively a complex issue considering the effectiveness of the injury prevention technologies as well as more detailed assessment of the complex interactions between the road condition, vehicle and human factor. For many years, highway agencies and vehicle manufacturers showed great efforts to reduce the injuries resulting from the vehicle crashes. Many researchers used a broad range of methods to evaluate the impact of several factors on traffic accidents and injuries. Recent developments lead up to capable for determining the effects of these factors. According to World Health Organization (WHO), cyclists and pedestrians comprise respectively 1.6% and 16.3% in traffic crash fatalities in 2013. Also in Turkey crash fatalities for pedestrian and cyclists are respectively 20.6% and 3% according to Turkish Statistical Institute data in 2013. The relationship between cycling and pedestrian rates and injury rates over time is also unknown. This paper aims to predict the crash severity with the traffic injury data of the Konya City in Turkey by implementing the Artificial Neural Networks (ANN), Regression Trees (RT) and Multiple Linear Regression modelling (MLRM) method.

Key words: traffic accident, cyclist, pedestrians, artificial linear network, regression trees, multiple linear regression

1. Introduction

Crash intensity prediction models are significantly important for transportation planning studies, and they are recurrently implemented in transportation safety issues. In this point, to assure a safety, traffic authorities need to understand the causes of a particular accident to propose the appropriate solutions. Based on the statistics of Turkish Statistical Institute (TUIK, 2013), more than 337.351 accidents reported in Turkey in 2013. In these accidents, 226.677 person died or injured. In literature, there exist many studies which investigated the role of demographic, socio-economic, land use and network characteristics on accidents. Also many crash prediction models were justified such as poisson models, negative binomial models, linear regression models and empirical analysis techniques (Abdel-Aty et al., 2000; Ivan et al., 2000; Noland et al., 2004; Caliendo et al., 2007; Miaou and Lord, 2003; Mitra and Washington, 2007; Ma et al., 2008; Naderan et al., 2010; Siddiqui

et al., 2012; Pulugurtha and Duddu, 2012). However, conventional methods such as regression analysis may be used for the traffic accident problems (Ozgan, 2008; Türe, 2008).

One of the important roles of safety key performance indicators are to evaluate performance of action plans. Therefore, prediction business as usual has significant role in evaluation of system. For this purpose, monthly or annually data set investigation was seen in literatures. However, these is a gap in prediction cyclist and pedestrian accidents in literatures.

Beside many studies, predicting the driver and vehicle characteristics is important for a better understanding the causes of accidents. Many studies are available in literature with dealing this topic. In the recent years, neural networks played a significant role for designing sophisticated models for traffic management, crash prediction and travel demand. Application of the ANN is well common solution for many of engineering problems by

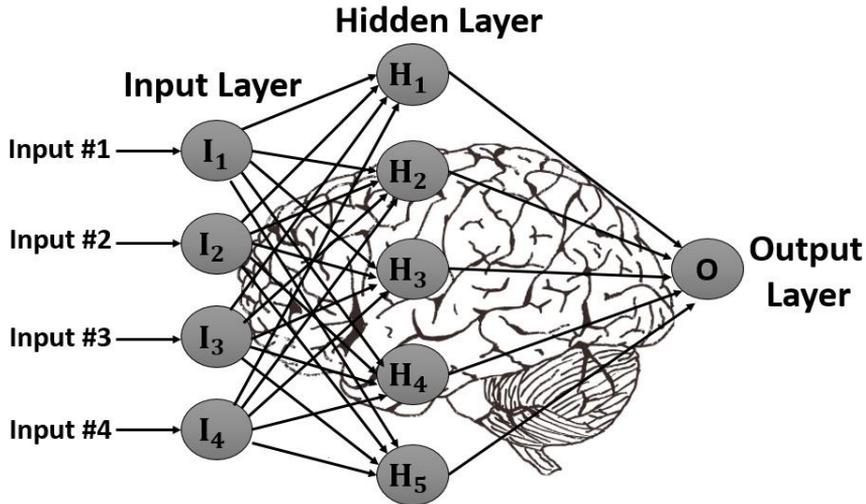


Fig. 1. A typical ANN structure with three layers (Consist of input, hidden and output).

modeling of nonlinear systems. In addition, the implementation of the ANN's in civil engineering problems in the field of transportation is quite common. Gilmore et al. (1993) applied the neural networks in traffic management studies. In another study, Nakatsuji et al. (1994) implemented the neural networks for signal timing optimization. Ledoux et al. (1995) and Yin et al. (2002) tried to model the traffic flow using neural and fuzzy-neural networks. Smith and Demetsky (1997), Wilde (1997), and Smith et al. (2002) studied on the short-term traffic flow forecasting by implementing neural networks. Ledoux (1996) integrated the neural networks on urban traffic systems. Srinivasan et al. (2007) studied on artificial intelligence based on congestion prediction. Delen et al. (2006) implemented an ANN application for predicting the contribution of the environmental factors, driver characteristics and road conditions to the road accidents. In a different study, Hashemi et al. (1995) compared the capability of a neural network (NN) model with multiple discriminant analysis and logistic regression to predict vessel accidents. Akgüngör and Doğan (2006) proposed an ANN model for prediction of the accidents (number of accidents, injuries and deaths). In their study, Altun et al. (2005) compared the ANN and nonlinear models according to various error expressions. The study results showed that the ANN model had better results according to nonlinear regression models.

Despite the aforementioned efforts, there exist no study in literature considered the application of neural networks in directly predicting crashes of pedestrians and cyclists. This paper aims to predict the crash severity with the traffic injury data of the Konya city in Turkey by implementing the artificial neural networks (ANN), Regression Trees (RT) and multilinear regression modelling (MLRM) method.

2. Methodology

Artificial Neural Network (ANN) is used for modeling the statistical data. This method is a reliable way of representing the non-linear relations between the inputs and outputs of a system. ANN can consider the complex relations inside the data, and it tries to generalize. A typical ANN model can be implemented for various problem types such as pattern recognition, classification, prediction and optimization considering a statistical dataset. One another advantageous of the ANN is that, there is no need of selecting a regression model to the data. In this way, ANN model generally fits well with the proposed data.

The ANN is basically a mathematical representation of the human brain that consist of numerous neurons connected to each other. Similar to the human brain, an ANN network is able to recognize the patterns in a data and suit well to the nonlinear characteristics of a system. An ANN consists of several data processing nodes called neurons. Basically, the

neurons or nodes are grouped in several layers called input, output and one or several hidden layers. A three-layer artificial neural network is shown in Figure 1.

In the ANN modelling, it is required to take attention of the large number of errors expected to be in the massive data set. The raw data used in the modelling progress in generally not clean. Errors are generally existed due to the equipment related problems and transcription errors while manually writing the accident records.

3. Data analyzing

In the scope of this study, the traffic accident data were gathered from local police department of Konya City in Turkey, and monthly registered passenger car's data set were obtained from Turkish Statistical Institute (TUIK) as shown respectively in Figure 2-3. The daily based data has interrupted properties. Therefore, this is the main reason for using monthly based data.

On the other hand, cyclist and pedestrian accidents were used as outputs of the models. These data were defined as dependent variables in two different cases. Total accident numbers and registered passenger car numbers were defined as input and independent variables. Hence, modelling was manipulated using ANN and MLR simultaneously.

Known as a black box computational method, ANN is one of the most popular analyzing methods among high degree uncertainty modelling techniques. Therefore, it was used in analysis and predictions of the present research. Using different methods of radial based function (RBF) in comparison with generalized regression network (GRN), multi-linear regression (MLR) and regression trees (RT), four models were investigated in total. Then results were monitored by coefficient of determination (R^2) and mean squared percentage error (MSPE). In order to train the model, 80 percent of the data were used in an arbitrary manner (48 data for each inputs and output) and following to the procedure the rest of the data were used for prediction (20% or 12 data for each inputs and output). Using all this methods for model prediction, MATLAB© software was used for modelling ANN techniques, and SPSS software was used for MLR's. Models were developed to estimate the total number of cyclist and pedestrian crashes as well as crashes based on severity. A result of each case was presented in Figure 4(a-b).

The consequences of modelling and prediction due to calibration and prediction procedure manipulated carefully. Then results were extracted and compared with different aspects and ratios (See Table 1 and Figure 5).

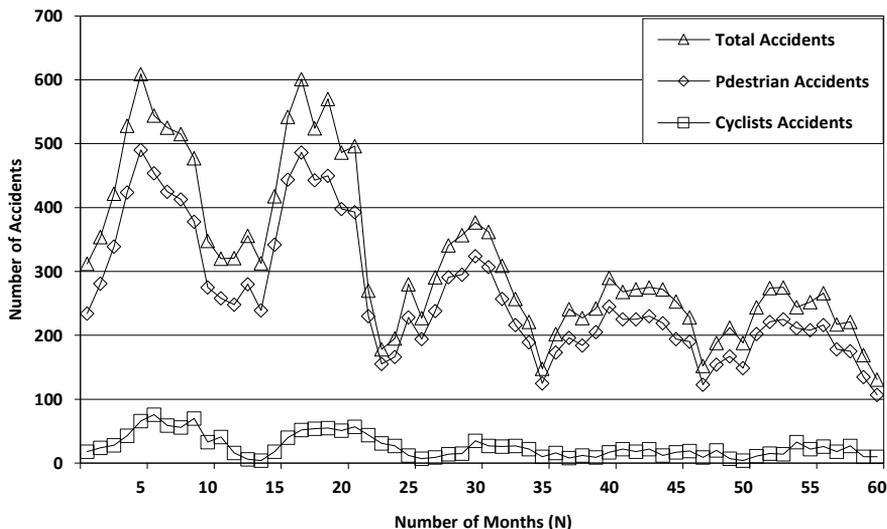


Fig. 2. Time series for pedestrians, cyclist and total accidents for 2009 to 2013

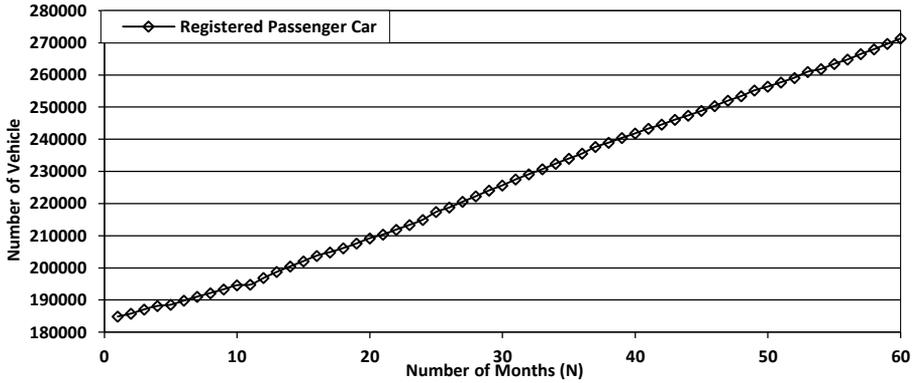


Fig. 3. Time series of monthly registered passenger cars for 2009 to 2013

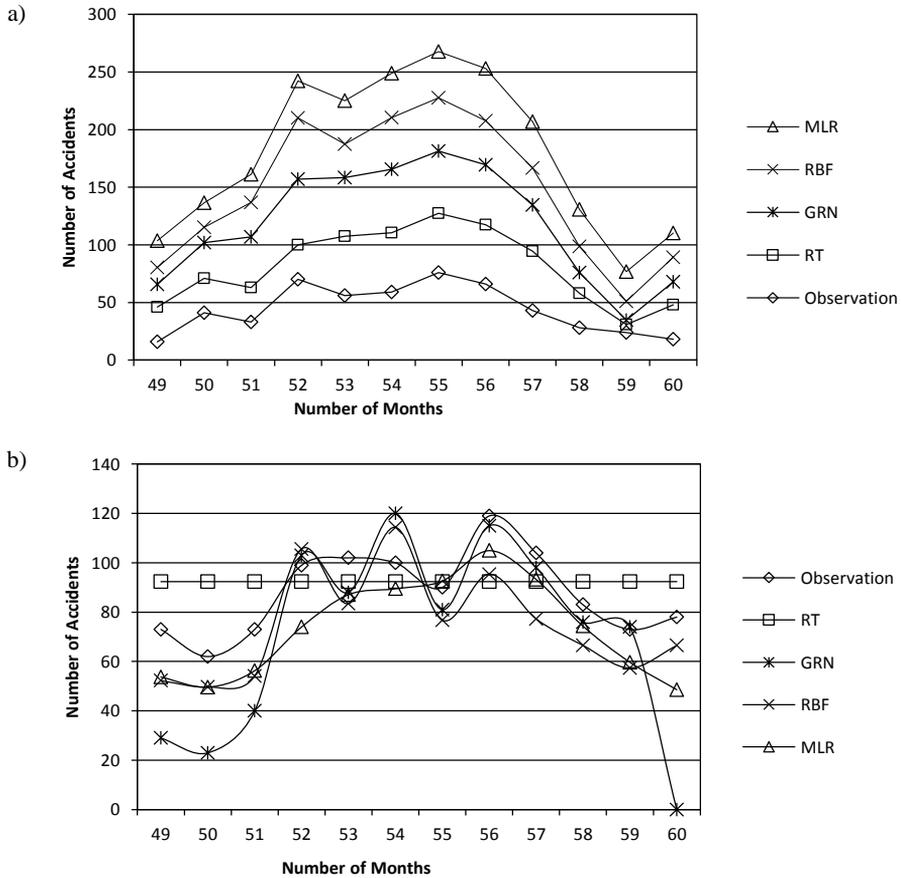


Fig. 4. (a) Model results for cyclist and (b) pedestrian accidents

Table 1. Methods and results

Pedestrians	Analysis Error	
	R ²	MSEP
RT	-	3.31%
GRN	0.68	19.11%
RBF	0.68	5.39%
MLR	0.83	4.93%
Cyclists	Analysis Error	
	R ²	MSEP
RT	0.77	7.12%
GRN	0.82	5.31%
RBF	0.66	12.20%
MLR	0.91	21.40%

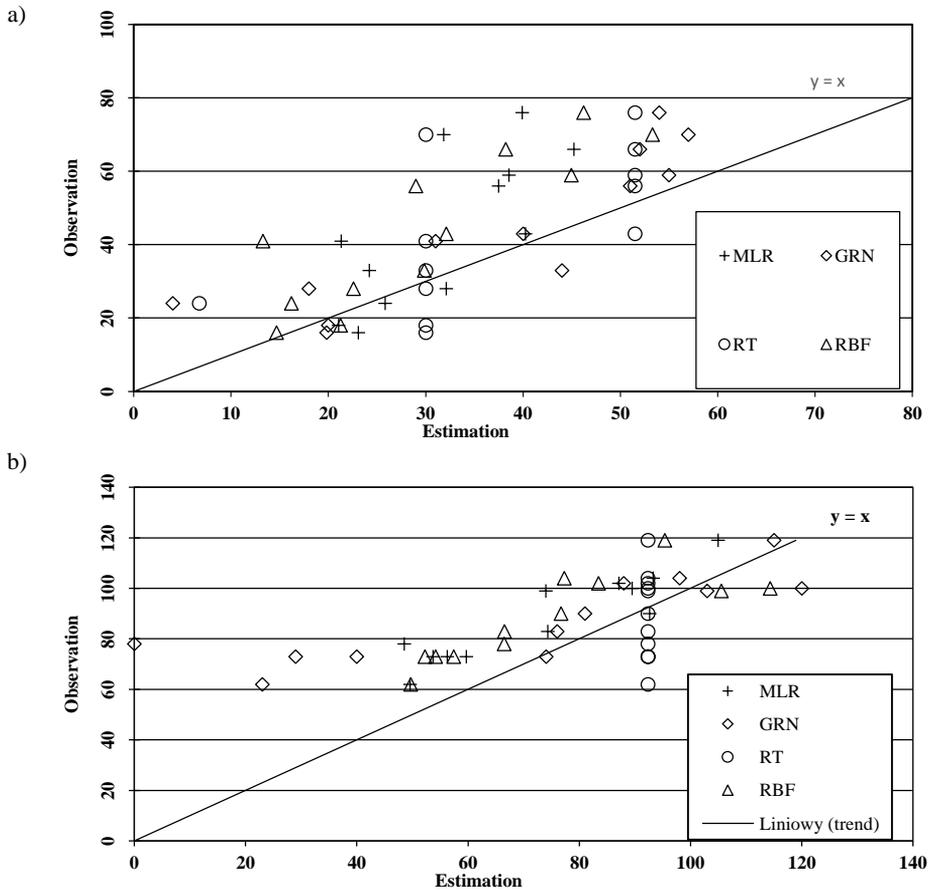


Fig. 5. Scatter diagrams of model outputs for (a) cyclist and (b) pedestrian accidents

As can be seen in Figure 5, all models presented unacceptable results except for GRN model for pedestrian accidents while the major predictions are over the trend line set. GRN model ensured significantly lower errors when compared to statistical and regression trees models for pedestrians. However, all models predicted lower values than observations especially for cyclist accidents. The main reason is the complex relationship between cyclist and vehicle accidents and also absence of value of Vehicle Travel Kilometer (VKT), Bike Travel Kilometer (BTK) and Pedestrian Travel Kilometer (PKT) in modelling.

4. Conclusions

In this study, the potential of an ANN model for simulating the accident properties was evaluated with successive comparison of the model performances with the LR model. The aim of this research was to investigate relation between injury and fatality accident with cyclist and pedestrian accidents. Also the focus of this research was to develop crash estimation models for cyclist and pedestrians as a function of total monthly accidents (death and injury) and registered passenger car. Also this study aimed to evaluate and analyze the application of neural network models for predicting the data came from traffic police department of Konya City that includes 14134 crashes between years of 2009 and 2013.

Suggested models were developed using the input parameters of monthly crashes and registered for passenger cars. The models were constructed based on cyclist and pedestrian crashes in total that occurred during 2009-2013 in the Konya City urban area. The ANN was applied during training, testing and validation of 14 input layer, 1 hidden layer and 1 output layer for each case (cyclists and pedestrians). Also dummy parameters are defined to include monthly trends in modelling. The best-fit model was chosen according to the R^2 value, mean square errors percentage (MSEP). The highest R^2 value was obtained for the GRN of ANN for pedestrian accidents, demonstrating that the ANN gave the best prediction results for pedestrian accidents.

The results from comparison of models for performance evaluation indicate that GRN model ensured significantly lower errors when compared to

statistical and regression trees models for pedestrians. However, all models predicted lower values than observations especially for cyclist accidents. The main reason is the complex relationship between cyclist and vehicle accidents and also absence of value of Vehicle Travel Kilometer (VKT), Bike Travel Kilometer (BTK) and Pedestrian Travel Kilometer (PKT) in modelling.

References

- [1] ABDEL-ATY, M.A. and RADWAN, A.E., 2000. Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5), pp. 633-642.
- [2] AKGÜNGÖR, A.P. and DOĞAN, E., 2008. Estimating road accidents of Turkey based on regression analysis and artificial neural network approach. *Advances in transportation studies*, 2008(16), pp. 11-22.
- [3] ALTUN, H., BILGİL, A. and FIDAN, B., 2007. Treatment of multi-dimensional data to enhance neural network estimators in regression problems. *Expert Systems with Applications*, 32(2), pp. 599-605.
- [4] CALIENDO, C., GUIDA, M. and PARISI, A., 2007. A crash-prediction model for multilane roads. *Accident Analysis & Prevention*, 39(4), pp. 657-670.
- [5] DELEN, D., SHARDA, R. and BESSONOV, M., 2006. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, 38(3), pp. 434-444.
- [6] GILMORE, J.F., ELIBIARY, K.J. and ABE, N., 1993. Traffic Management Applications of Neural Networks, *Working Notes, AAAI-93 Workshop on AI in Intelligent Vehicle Highway Systems 1993*, pp. 85-95.
- [7] HASHEMI, R.R., LE BLANC, L.A., RUCKS, C.T. and SHEARRY, A., 1995. A neural network for transportation safety modeling. *Expert Systems with Applications*, 9(3), pp. 247-256.
- [8] IVAN, J.N., WANG, C. and BERNARDO, N.R., 2000. Explaining two-lane highway crash rates using land use and hourly exposure.

- Accident Analysis & Prevention*, 32(6), pp. 787-795.
- [9] LEDOUX, C., 1996. An urban traffic control system integrating neural networks, *Eighth International Conference on Road Traffic Monitoring and Control*, 23-25 April 1996 1996, IET, pp. 197-201.
- [10] MA, J., KOCKELMAN, K.M. and DAMIEN, P., 2008. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention*, 40(3), pp. 964-975.
- [11] MIAOU, S. and LORD, D., 2003. Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. *Transportation Research Record: Journal of the Transportation Research Board*, 1840, pp. 31-40.
- [12] MITRA, S. and WASHINGTON, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. *Accident Analysis & Prevention*, 39(3), pp. 459-468.
- [13] NADERAN, A. and SHAHI, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. *Accident Analysis & Prevention*, 42(1), pp. 339-346.
- [14] NAKATSUJI, T. and KAKU, T., 1991. Development of a self-organizing traffic control system using neural network models. *Transportation Research Record*, 1324, pp. 137-145.
- [15] NOLAND, R.B. and OH, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. *Accident Analysis & Prevention*, 36(4), pp. 525-532.
- [16] OZGAN, E., 2008. KARAYOLU ARAÇ TİPİ VE KAZA ŞEKLİ İLE KAZA SONUÇLARI ARASINDAKİ İLİŞKİLERİN ANALİZİ. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 23(3), pp. 97-104.
- [17] PULUGURTHA, S.S., DUDDU, V.R. and KOTAGIRI, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. *Accident Analysis & Prevention*, 50, pp. 678-687.
- [18] SIDDIQUI, C., ABDEL-ATY, M. and HUANG, H., 2012. Aggregate nonparametric safety analysis of traffic zones. *Accident Analysis & Prevention*, 45, pp. 317-325.
- [19] SMITH, B.L. and DEMETSKY, M.J., 1997. Traffic flow forecasting: comparison of modeling approaches. *Journal of Transportation Engineering*, 123(4), pp. 261-266.
- [20] SMITH, B.L., WILLIAMS, B.M. and OSWALD, R.K., 2002. Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 10(4), pp. 303-321.
- [21] TÜİK, 2013. *Traffic Accident Statistics*. Ankara, Turkey: Türkiye Statistical Institute.
- [22] TÜRE, K.F., 2008. *Trafik Kazaları ve Trabzon Bölünmüş Sahil Yolu Örneğinde Kaza Tahmin Modelinin Oluşturulması*, Fen Bilimleri Enstitüsü, Karadeniz Teknik Üniversitesi, Trabzon.
- [23] YIN, H., WONG, S., XU, J. and WONG, C., 2002. Urban traffic flow prediction using a fuzzy-neural approach. *Transportation Research Part C: Emerging Technologies*, 10(2), pp. 85-98.