Mapping patterns of pedestrian fatal accidents in Israel

Carlo Giacomo PRATO [a *], Victoria GITELMAN [b], Shlomo BEKHOR [c]

[a] Department of Transport
Technical University of Denmark
Bygningstorvet 116 Vest
2800 Kgs. Lyngby, Denmark
Telephone: +45.45256595, Fax: +45.45936593
E-mail: cgp@transport.dtu.dk

[b] Ran Naor Road Safety Research Center
Technion - Israel Institute of Technology
Technion Campus, Haifa 32000, Israel
Telephone: +972.4.8295132, Fax: +972.4.8225716
Email: trivica@technion.ac.il

[c] Faculty of Civil and Environmental Engineering
Technion - Israel Institute of Technology
Technion Campus, Haifa 32000, Israel
Telephone: +972.4.829460, Fax: +972.4.8225716
Email: sbekhor@technion.ac.il

[*] corresponding author

Abstract

This study intends to provide insight into pedestrian accidents by uncovering their patterns in order to design preventive measures and to allocate resources for identified problems. Kohonen neural networks are applied to a database of pedestrian fatal accidents occurred during the four-year period between 2003 and 2006. Results show the existence of five pedestrian accident patterns: (i) elderly pedestrians crossing on crosswalks far from intersections in metropolitan areas; (ii) pedestrians crossing suddenly or from hidden places and colliding with two-wheel vehicles on urban road sections; (iii) male pedestrians crossing at night and being hit by four-wheel vehicles on rural road sections; (iv) young male pedestrians crossing at night wide road sections in both urban and rural areas; (v) children and teenagers crossing road sections in small rural communities. From the perspective of fatality reduction measures, results suggest the necessity of designing education campaigns for parents, promoting information campaigns for road users and allocating resources for infrastructural interventions and law enforcement in order to address the identified major problems.

Keywords: Pedestrian fatalities; Accident patterns; Cluster analysis; Kohonen networks.

1. Introduction

Yearly, over 400 people are killed and thousands more are injured in Israel, with foreseeable costs to the society in terms of human lives, property damages and congestion delays. Statistics reveal that a major problem concerns the most vulnerable road users, since over one third of the traffic fatalities are pedestrians. In fact, in recent years a decrease in the number of accident fatalities has not been accompanied by a proportional decline in the number of pedestrian fatalities (Gitelman et al., 2009).

The literature shows a great interest in comprehending pedestrian accident patterns in order to design preventive measures and to allocate resources for identified problems. In the last thirty years, the main interest has shifted from examining pedestrian accidents involving children (e.g., Bagley, 1992; Brison et al., 1988; Fortenberry and Brown, 1982; Preston, 1972) and elderly (e.g., Sklar et al., 1989; Zegeer et al., 1993) to analyzing determinants of collisions between pedestrian and vehicles (e.g., Al-Ghamdi, 2002; Ballesteros et al., 2004; Kim et al., 2008a; Lee and Abdel-Aty, 2005). In the last decade, the main interest has further shifted toward investigating factors that characterize pedestrian accidents at various spatial aggregation levels (e.g., Beck et al., 2007; Eluru et al., 2008; Fontaine and Gourlet, 1997; Harruff et al., 1998; Kim et al., 2008b; Mabunda et al., 2008; Preusser et al., 2002).

The most common factors used to characterize pedestrian accidents are age and gender of the pedestrians, with emphasis given to the higher vulnerability of children and elderly (Al-Ghamdi, 2002; Eluru et al., 2008; Fontaine and Gourlet, 1997; Harruff et al., 1998; Kim et al., 2008b; Preusser et al., 2002) and to the higher involvement of males (e.g., Al-Madani and Al-Janahi, 2006; Beck et al., 2007; Harruff et al., 1998; Kim et al., 2008a; Mabunda et al., 2008). Other factors frequently utilized to exemplify pedestrian crashes are the location in urban residential areas (e.g., Al-Ghamdi, 2002; Beck et al., 2007; Harruff et al., 1998), the intoxication of pedestrians by alcohol or drugs (e.g., Fontaine and Gourlet, 1997; Mabunda et al., 2008; Kim et al., 2008b; Öström and Eriksson, 2001), the fault of pedestrians and drivers (e.g., Al-Ghamdi, 2002; Preusser et al., 2002) and the type of vehicle involved in the collision (e.g., Ballesteros et al., 2004; Kim et al., 2008a). Generally, the relevance of these factors is evaluated independently rather than being examined jointly to provide a broad picture of concurrent characteristics of pedestrian accidents.

This study contributes to the knowledge about pedestrian accidents by uncovering their patterns and their recurrent underlying characteristics. The implementation of cluster analysis and the retrieval of accident patterns that emerge from the data enable providing an integrative and multi-faceted map of pedestrian accidents concerning the geographical, the demographic and the infrastructural perspectives. Patterns and characteristics of pedestrian accidents in Israel may be compared to existing knowledge about pedestrian accidents in Africa (Mabunda et al., 2008), Asia (Al-Ghamdi, 2002; Al-Madani and Al-Janahi, 2006), Europe (Fontaine and Gourlet, 1997; Öström and Eriksson, 2001), and United States (Ballesteros et al., 2004; Beck et al., 2007; Eluru et al., 2008; Harruff et al., 1998; Kim et al., 2008a; Kim et al., 2008b; Preusser et al., 2002).

Given the interest in reducing the loss of human lives, which motivates the political commitment toward the reduction of the number of fatalities on Israeli roads (in line with the proposition by the European Commission of halving fatalities from 2001 to 2010), this study focuses on 603 pedestrian fatal accidents occurred during the four-year period between 2003 and 2006. Data contain details about modality, location, infrastructure characteristics, environment conditions, vehicles and persons involved in each pedestrian fatal accident.

Given the interest in providing a multi-faceted map of pedestrian fatal accidents by extracting patterns from the data, this study applies Kohonen neural networks for clustering analysis. Unsupervised self-organizing maps are constructed by a competitive learning algorithm and each element of the map corresponds to a cluster of pedestrian fatal accidents that is interpreted according to the categories of the most relevant variables. Existing methods for the individuation of accident patterns include frequency analysis combined with odd ratio calculation (e.g., Ballesteros et al., 2004; Beck et al., 2007; Harruff et al., 1998; Zegeer et al., 1993), multiple correspondence analysis (e.g., Al-Ghamdi, 2002; Al-Madani and Al-Janahi, 2006; Mabunda et al., 2008), and generalized linear or discrete choice modeling (e.g., Eluru et al., 2008; Kim et al., 2008b; Lee and Abdel-Aty, 2005). The advantage of implementing Kohonen networks consists in the ability of unraveling accident patterns from a large number of accident characteristics without the need of analyzing independently accident factors as in frequency analyses, restricting significantly the examined number of variables and categories as in the multi-correspondence analyses, or introducing restrictive assumptions about relevant accident factors to be considered in the model. Moreover, the advantages of implementing one-dimensional Kohonen networks (i.e., $1 \times J$ maps where J is the number of clusters) are the stability of the solutions and the convergence to a global optimum (e.g., Cottrell and Rousset, 1997; Erwin et al., 1988; Ritter and Schulten, 1992), as demonstrated by their application to operations research for the solution of the Traveling Salesman Problem (e.g., Aras et al., 1999; Burke, 1996) and to oncology research for the discovery of anticancer drugs (Augen, 2004).

The remainder of the paper is structured as follows. Section 2 describes the pedestrian fatal accident data. Section 3 describes the methods applied to uncover accident patterns. Section 4 presents the results of the cluster analysis. Section 5 discusses the major findings of this study and introduces some measures recommendations.

2. Data

The present study focuses on the 603 pedestrian fatal accidents occurred in Israel during the four-year period between 2003 and 2006, and extracted from the general accident database provided by the Central Bureau of Statistics (CBS). The four-year period is long enough to limit random fluctuations in the accident counts and short enough to control for changes in road and traffic conditions.

Thousands of fatal and non-fatal accidents are reported every year by the police and three different files are constructed by the CBS on the basis of the collected information: (i) accident file, (ii) vehicle and driver file, and (iii) injured person file.

The accident file reports several details concerning each accident, including time, date, type, modality, cause, whereabouts in terms of police district and location related characteristics (i.e., urban or rural area, intersection or road section), level of severity (i.e., fatal, severe and light), infrastructure and environmental characteristics (e.g., allowed speed, presence and condition of a median barrier, traffic control and signal system, condition of the surface, weather), and details about pedestrians and objects for specific accident typologies. The vehicle and driver file describes each vehicle and driver involved in the accident, with each record listing generic vehicle information (e.g., type, age, motor, weight and direction of travel of the vehicle) as well as drivers' socio-economic characteristics (e.g., gender, age, year of licensure). The injured person file provides information about each person injured, with each record registering socio-economic information (e.g., gender, age, nationality, place of residence and type of injury sustained).

Data quality is verified while preparing the pedestrian fatal accident dataset. Firstly, the consistency of the coding system is extremely important since variations in the definition of the variables by CBS throughout the years could lead to a bias in the application of data mining methods. The analysis focuses on the period between 2003 and 2006 for which the coding system is consistent and allows avoiding potential problems related to heterogeneity in the definition of the variables. Secondly, merging the files requires attention because the accident file contains one accident for each record, but the other two files contain multiple records corresponding to the same accident and equivalent to the number of involved drivers, vehicles and injured persons. The constructed pedestrian fatal accident dataset contains as many records as accidents, each with its characteristics and its number of vehicles, drivers and injured persons. Last, missing values are encountered in the accident database with possible consequences on the pattern recognition process. The selection of fatal accidents as the focus of the study allows avoiding problems related to data quality, as most of the errors are reported in accidents that result into light injuries to one person involved in the accident, while fatal accidents are more scrupulously described and almost error free.

The pedestrian fatal accident dataset consists of 603 records and the categorical variables considered for data analysis, after merging the three files and verifying the data quality, are summarized in table 1.

[Insert Table 1 about here]

3. Methodology

The current study intends to recognize accident patterns from the analysis of a rich dataset of pedestrian fatal accidents, and not to restrict the analysis to factors individually and arbitrarily chosen prior to the implementation of any method. Kohonen neural networks are unsupervised learning methods that answer the requirements of treating large numbers of variables, assessing their importance in order to refine the analysis, and obtaining correlation patterns without any predefined assumption (Cottrell and Rousset, 1997; Kohonen, 2001). Kohonen networks are preferable to alternative unsupervised learning methods for several reasons. Firstly, they allow treating large samples, unlike k-means clustering and hierarchical clustering that are difficult to interpret when more than 200-300 cases are considered (Augen, 2004). Secondly, they enable providing intelligible solutions while considering a large number of variables, unlike multi-correspondence analysis and principal component analysis that require a selection of a limited number of variables that de facto removes the advantage of applying unsupervised learning (Augen, 2004). Last, they allow reaching stable optimal solutions under the one-dimensional configuration (e.g., Cottrell and Rousset, 1997; Erwin et al., 1988; Ritter and Schulten, 1992), as shown in applications to operations research (e.g., Aras et al., 1999; Burke, 1996) and oncology research (Augen, 2004).

Kohonen neural networks are based upon the idea of self-organized learning (Kohonen, 1982; Kohonen, 2001). A single layer of neurons is arranged in a one- or two- dimensional map through a competitive learning algorithm that connects all neurons in an input layer with all neurons in an output layer and assigns weights to each connection. The input layer is composed by neurons that correspond to each pedestrian fatal accident with its characteristics in terms of categorical values of the considered variables. The output layer is composed by neurons that correspond to pedestrian accident patterns with shared characteristics unraveled by the competitive learning algorithm. Since this algorithm does not attempt to predict values of target variables, Kohonen networks are suitable for cluster analysis.

Initially, the weights of the connections between input and output neurons are randomly drawn. During the training phase, the competitive learning algorithm presents an input vector x_s to the network (i.e., a pedestrian fatal accident) and only the neuron whose weight vector w_s is most similar to this input is stimulated. Mathematically, the "winning" neuron *c* is selected as the output neuron providing the minimal Euclidean distance to the input vector x_s:

$$c \leftarrow \min_j \left\{ \sum_i \left(x_{si} - w_{ji} \right)^2 \right\} \qquad (1)$$

where x_{si} is the i-th coordinate of the input vector x_s (i.e., a characteristic of the accident), w_{ji} is the i-th weight level of neuron j, and the number of neurons j depends on the structure of the self-organizing map. After the winning neuron is selected, the weights w_{ji} of each neuron j in the output layer are updated on the basis of the difference between their old value and the values of the correspondent coordinates x_{si} of the input vector. The corrections depend on the topological distance d_r from the winner:

$$\Delta w_{ji} = \eta \left(1 - \frac{d_r}{d_{\max} + 1}\right)\left(x_{si} - w_{ji}^{old}\right) \quad for\, dr_r = 0,1,...,d_{max} \tag{2}$$

where Δw_{ji} is the weight update, w_{ji}^{old} is the numerical value of the weight w_{ji} at the previous iteration, η is the learning rate, d_r is the topological distance defined as the number of neurons separating the neuron j from the winning neuron. The size of the neighborhood d_{max}, which at the beginning of the learning covers the entire network, decreases during the training phase and is eventually limited only to the winning neuron. The learning rate η is also changing during the training phase, and a clustering solution is obtained at convergence when either the learning parameter η arrives to a null value or the number n_{epoch} of training cycles equals a pre-specified maximum value n_{tot}:

$$\eta = \left(\eta^{start} - \eta^{final}\right)\left(1 - \frac{n_{epoch}}{n_{tot}}\right) + \eta^{final} \tag{3}$$

where η^{start} and η^{final} are the values of the learning rate η at the beginning and at the end of the n_{epoch} cycle. The algorithm consists of two stages for which learning rate η and number n_{tot} of cycles are defined by searching for the values that guarantee the fastest convergence to a stable and meaningful solution. In the current study, this search shows that a stable solution is obtained for different combinations of parameter values, and the fastest convergence is obtained for a first stage with learning rate η equal to 0.4 and number of affected neighbors equal to 2 to train the network faster, and a second stage with learning rate η reduced to 0.1 and number of neighbors reduced to 1 to fine-tune the map. Also, the fastest convergence is obtained for number n_{tot} of cycles equal to 20 in the first stage and to 100 in the second.

The number of clusters in the one-dimensional map is determined through a trial and error process, which exploits the property by Kohonen neural networks of assigning progressively the data to different clusters by dividing bigger clusters into smaller ones. In general, the search for the optimal number of clusters terminates when the addition of one cluster does not produce an additional relevant group and also does not decrease the error in the network.

4. Results

In the four-year period between 2003 and 2006, the 603 pedestrian fatalities represent around one third of the 1,793 total traffic fatalities.

General characteristics of the accidents describe that most pedestrian fatalities occur in urban areas (72.1%), in road sections (70.6%) and in the center of the country where the two major metropolitan areas are located (56.7%). Pedestrian fatalities are mainly registered during day (57.5%), generally during the morning and the afternoon off-peak periods, and prevalently concern crossing pedestrians (77.1%) who in over 50% of the cases are reported to have either crossed suddenly or from a hidden place. Even though this might suggest that pedestrians are at fault in the majority of the accidents, records show that the cause of the accident is the offense of the driver in 58.7% of the cases. Pedestrian fatalities are mainly male (60.7%), elderly (36.7%), children and teenagers (18.7%), and the share of non-Jewish pedestrian fatalities (33%) is higher than the share of non-Jewish population (20%).

Frequency analysis of the characteristics of the accidents offers only a blurred picture of pedestrian fatal accidents, while the implementation of Kohonen neural networks provides further insight by recognizing patterns of concurrent accident characteristics. After searching for the number of clusters with the described iterative process and testing several self-organizing maps, the optimal Kohonen network is a one-dimensional map composed by five clusters. The interpretation of the clusters depends on the frequency of the categories of the most relevant variables as suggested by the learning algorithm and represented in figure 1.

[Insert Figure 1 about here]

Relevant variables for the first cluster concern the location of the accident, the condition of the infrastructure and the median, the age of the pedestrian, and the behavior of the pedestrian. This first cluster includes accidents where elderly pedestrians cross narrow roads on crosswalks far from intersections, mainly in urban areas and more specifically in the Tel Aviv metropolitan area.

Important variables for the second cluster relate to the location of the accident, the involvement of vehicles and the modality of crossing. The second cluster contains accidents where pedestrians cross roads suddenly or from hidden places and collide generally with two-wheel vehicles in urban sections, prevalently inside one of the three main metropolitan areas (i.e., Tel Aviv, Jerusalem and Haifa).

Significant variables for the third cluster involve the location of the accident, the characteristics of the infrastructure, the time of the accident, the gender of the pedestrian, and the modality of occurrence of the collision. The third cluster comprises accidents occurring prevalently in the evening or at night, where male pedestrians cross rural road sections outside crosswalks and are hit by cars after exiting suddenly from a hidden place.

Relevant variables for the fourth cluster relate to the location of the accident, the width of the road crossed, the gender and the age of the pedestrian, and the modality of the accident. The fourth cluster contains accidents where prevalently male pedestrians under the age of 35 cross wide road sections in both urban and rural areas at night, while coming out of hidden places.

Important variables for the fifth cluster include the location of the accident, the condition of the infrastructure, the age of the pedestrian and the crossing location. The fifth cluster consists of accidents where mostly young children and teenagers cross roads without a median separation that are located in small rural communities that are prevalently located in the north of the country.

Figure 2 summarizes the Kohonen map and provides a definition of the clusters of pedestrian accidents. The most important variables revealed by the Kohonen networks are the accident location, the width and the presence of median separation in road sections, the modality of the accident, the period of the accident, the age and the place of residence of the pedestrians. Considering similarities between neighboring clusters, major differences are characterized in terms of (i) age of the pedestrian, as the first cluster contains accidents with elderly while the last cluster includes accidents with children and teenagers, and (ii) geographical location, as from the metropolitan areas of the first cluster the map moves through the urban areas of the second to the small rural communities of the last one.

[Insert Figure 2 about here]

5. Discussion and conclusions

This study contributes to the body of literature focusing on pedestrian accidents at various spatial aggregation levels by applying Kohonen neural networks to unravel patterns and recurrent characteristics of pedestrian fatal accidents in Israel. Five clusters emerge from the data to provide an integrative and multi-faceted map of pedestrian accidents : (i) elderly pedestrians crossing on crosswalks far from intersection in metropolitan areas; (ii) pedestrians crossing suddenly or from hidden places and colliding with two-wheel vehicles on urban road sections; (iii) male pedestrians crossing at night and being hit by four-wheel vehicles on rural road sections; (iv) young male pedestrians crossing at night wide road sections in both urban and rural areas; (v) children and teenagers crossing road sections in small rural communities.

The results of the current study confirm general findings presented in the literature with respect to locations, circumstances and demographic characteristics of pedestrian accidents. Among accident characteristics, relevant factors are (i) accident location, with prevalence for urban areas as observed in the U.S. (e.g., Beck et al., 2007; Harruff et al., 1998) and Saudi Arabia (Al-Ghamdi, 2002), and (ii) accident cause, with almost equal repartition of the fault between pedestrians and drivers as viewed in the U.S. (Preusser et al., 2002) and Saudi

Arabia (Al-Ghamdi, 2002). Among socio-demographic characteristics, important factors are (i) age of the pedestrians, with stress on the higher vulnerability of children and elderly as examined in the U.S. (e.g., Eluru et al., 2008; Harruff et al., 1998; Kim et al., 2008b; Preusser et al., 2002), France (Fontaine and Gourlet, 1997) and Saudi Arabia (Al-Ghamdi, 2002), and (ii) gender of the pedestrians, with emphasis on the larger involvement of males as noticed in the U.S. (Beck et al., 2007; Harruff et al., 1998; Kim et al., 2008a), Bahrain (Al-Madani and Al-Janahi, 2006) and South Africa (Mabunda et al., 2008).

In addition to confirming general findings, the results of the current study emphasize the relevance of different factors such as the width of the road combined with the presence of a median, the involvement of other vulnerable road users such as motorcyclists and bicyclists, and the geographical location. The latter aspect appears extremely important, since the differentiation between large metropolitan areas and small rural communities along the Kohonen map suggests that pedestrians living and moving in different regions may be exposed to risk related to land use and activity patterns (Elias et al., 2010). The difference across population groups confirms recent findings about the personal and social background of pedestrians affecting the likelihood of being involved in accidents (Al-Madani and Al-Janahi, 2006).

Notably, the recognition of the five accident patterns may help designing preventive measures and allocating resources for the identified problems. After recognizing the most compelling problems with pattern analysis, a system approach (see, e.g., Wegman and Aarts, 2006) could propose policy interventions initially focusing on these priorities identified by the Kohonen networks.

As pattern analysis emphasized the relevance of pedestrian behavior, a system approach may introduce the necessity for education and information campaigns . An education campaign could target insufficient road safety problem awareness (Wegman and Aarts, 2006), in particular in rural communities, and dangerous behavioral habits (Wegman and Aarts, 2006), in particular among children and teenagers. For example, education in hazard perception could help children and teenagers in recognizing risks, and contextually support the existing graduated driver licensing program to improve road safety once teenagers reach the licensing age. An information campaign could target dangerous behavioral habits (Wegman and Aarts, 2006), typically distract behavior related to the uncovered accident patterns: pedestrians should patiently approach crosswalks and cautiously check incoming traffic rather than crossing suddenly from hidden places and possibly being distracted by cellular phones or music players; bicyclists should carefully ride their bicycles when urban residential areas do not have reserved lanes and should not consider large sidewalks as a natural biking location; motorcyclists and car drivers should carefully concentrate on the task at hand rather than being constantly distracted by using cellular phones, reading newspapers and sometimes chatting animatedly with the other occupants of the vehicle.

As pattern analysis emphasized the relevance of the location of the pedestrian accident, a system approach may introduce the necessity for allocating resources to infrastructural interventions and law enforcement (Wegman and Aarts, 2006). A typical accident pattern concerns elderly on crosswalks in metropolitan areas, and a system approach may propose infrastructures that forgive pedestrians such as graded crossings to compel drivers to slow down and decrease of the distance between crosswalks, or law enforcement of parking restrictions nearby the crosswalks. Another typical accident pattern involves children and teenagers crossing urban roads in rural communities, and a system approach may propose infrastructures that create physical separation such as safer crossings of existing roads or bypass arterials that would reduce the number of major roads crossing small rural communities. A third typical accident pattern relates to collisions between pedestrians and two-wheel vehicle users, and a system approach may propose infrastructures that enforce speed management to slow down motorcyclists when approaching crosswalks and intersections, and constructing reserved bike lanes for cyclists who would not invade sidewalks.

Summarizing, cluster analysis contributes to the general knowledge about pedestrian accidents by presenting findings that on the one hand relate to general results obtained when analyzing realities that are different from the geographical and the social perspectives, and on the other hand provide new insight into concurrent characteristics of pedestrian accidents. Moreover, cluster analysis may contribute to the design of policies within a system approach that could benefit from the identification of priority problems through the Kohonen networks implemented to the rich database of pedestrian fatal accidents. Further research could benefit from an extension of the available data to information concerning intoxication by alcohol or drugs and containing additional details about the persons involved in terms of income and education level, in order to reach a better profiling of both pedestrians and drivers in light of the expected importance of social aspects other than the place of residence in the recognition of accident patterns.

Acknowledgements

The authors would like to thank the Or Yarok no-profit organization for the financial support to the Ran Naor Road Safety Research Center where the study has been conducted, Tsippy Lotan and Ayala Cohen for their helpful comments during the study, and two anonymous reviewers for their insightful remarks.

References

Al-Ghamdi, A.S., 2002. Pedestrian-vehicle crashes and analytical techniques for stratified contingency tables. Accident Analysis and Prevention 34 (2), 205-214.

Al-Madani, H., Al-Janahi, A., 2006. Personal exposure risk factors in pedestrian accidents in Bahrain. Safety Science 44 (4), 335-347.

Aras, N., Oommen, B.J., Altmel, I.K., 1999. The Kohonen network incorporating explicit statistics and its application to the Traveling Salesman Problem. Neural Networks 12, 1273-1284.

Augen, J., 2004. Bioinformatics in the post-genomic era: genome, transcriptome, proteome, and information-based medicine. Addison Wesley, Calgary, Canada.

Bagley, C., 1992. The urban setting of juvenile pedestrian injuries: a study of behavioural ecology and social disadvantage. Accident Analysis and Prevention 24 (6), 673-678.

Ballesteros, M.F., Dischinger, P.C., Langenberg, P., 2004. Pedestrian injuries and vehicle type in Maryland 1995-1999. Accident Analysis and Prevention 36 (1), 73-81.

Beck, L.F., Paulozzi, L.J., Davidson, S.C., 2007. Pedestrian fatalities, Atlanta metropolitan statistical area and United States, 2000-2004. Journal of Safety Research 38 (6), 613-616.

Brison, R.J., Wicklund, K., Mueller, B.A., 1988. Fatal pedestrian injuries to young children: a different pattern of injury. American Journal of Public Health 78, 793-795.

Burke, L.I., 1996. Neural methods for the Travelling Salesman Problem: insights from operations research. Neural Networks, 7, 681-690.

Cottrell, M., Rousset, P., 1997. The Kohonen algorithm: a powerful tool for analyzing and representing quantitative and qualitative data. Lecture Notes in Computer Science, 1240, 861-871.

Elias, W., Toledo, T., Shiftan, Y., 2010. The effect of daily-activity patterns on crash involvement. Accident Analysis and Prevention, 42 (6), 1682-1688.

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention, 40 (3), 1033-1054.

Erwin, E., Obermayer, K., Schulten, K., 1992. Self-organizing maps: ordering, convergence properties and energy functions. Biological Cybernetics, 67, 47-55.

Fontaine, H., Gourlet, Y., 1997. Fatal pedestrian accidents in France: a typological analysis. Accident Analysis and Prevention 29 (3), 303-312.

Fortenberry, J.C., Brown, D.B., 1982. Problem identification, implementation and evaluation of a pedestrian safety program. Accident Analysis and Prevention 14 (4), 315-322.

Gitelman, V., Balasha, D., Carmel, R., 2009. Examination of infrastructure solutions for improving pedestrian safety in Israel. Proceedings of the European Transport Conference, Leeuwenhorst, the Netherlands.

Harruff, R., Avery, A., Alter-Pandya, A., 1998. Analysis of circumstances and injuries in 217 pedestrian traffic fatalities. Accident Analysis and Prevention 30 (1), 11-20.

Kim, J.K., Ulfarsson, G.F., Shankar, V.N., Kim, S., 2008a. Age and pedestrian injury severity in motor-vehicle crashes: a heteroskedastic logit analysis. Accident Analysis and Prevention 40 (5), 1695-1702.

Kim, K., Brunner, I.M., Yamashita, E., 2008b. Modeling fault among accident-involved pedestrians and motorists in Hawaii. Accident Analysis and Prevention 40 (6), 2043-2049.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biological Cybernetics 43, 59-69.

Kohonen, T. (2001). Self-Organizing Maps (3rd extended edition). Springer & Verlag, Berlin Heldelberg, Germany.

Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37 (4), 775-786.

Mabunda, M.M., Swart, L.A., Seedat, M., 2008. Magnitude and categories of pedestrian fatalities in South Africa. Accident Analysis and Prevention 40 (2), 586-593.

Öström, M., Eriksson, A., 2001. Pedestrian fatalities and alcohol. Accident Analysis and Prevention 33 (2), 173-180.

Preston, B., 1972. Statistical analysis of child pedestrian accidents in Manchester and Salford Accident Analysis and Prevention 4 (4), 323-332.

Preusser, D.F., Wells, J.K., Williams, A.F., Weinstein, H.B., 2002. Pedestrian crashes in Washington, DC and Baltimore. Accident Analysis and Prevention 34 (5), 703-710.

Ritter, H., Schulten, K., 1988. Convergence properties of Kohonen's topology conserving maps: fluctuations, stability and dimension selection. Biological Cybernetics, 60, 59-71.

Sklar, D.P., Demarest, G.B., McFeeley, P., 1989. Increased pedestrian mortality among the elderly. American Journal of Emergency Medicine 7 (4), 387-390.

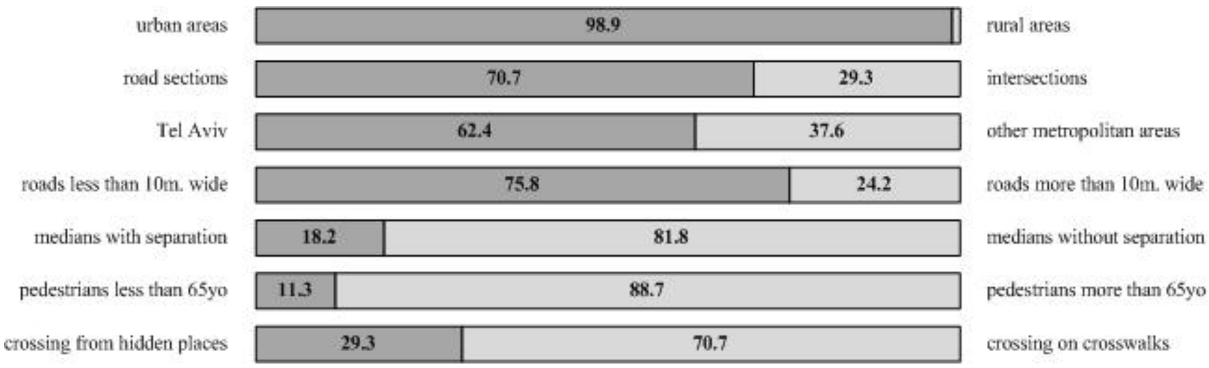
Wegman, F., Aarts, L., 2006. Advancing sustainable safety: national road safety outlook for 2005-2020. SWOW Institute for Safety Research, Leidschendam, The Netherlands (available at http://www.swov.nl/rapport/dmdv/Advancing_sustainable_safety.pdf).

Zegeer, C.V., Stutts, J.C., Huang, H., Zhou, M., Rodgman, E., 1993. Analysis of elderly pedestrian accidents and recommended countermeasures. Transportation Research Record 1405, 56-63.

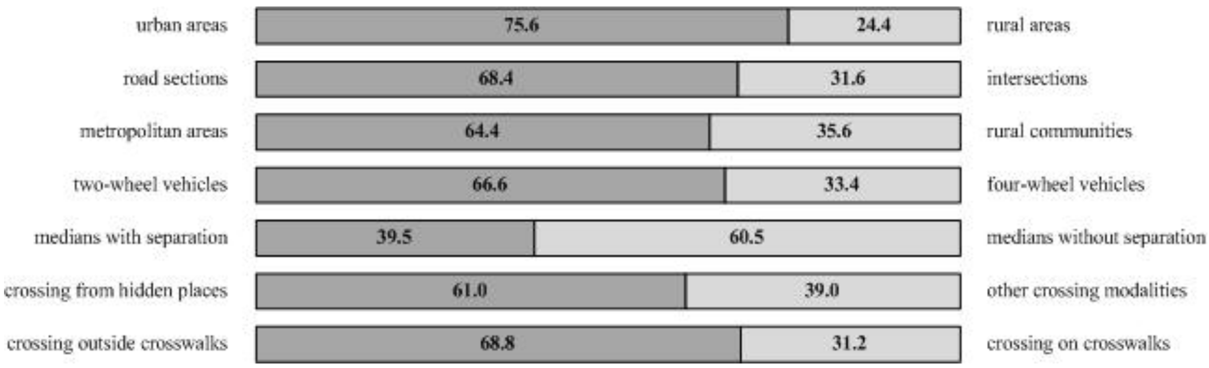
Table 1

Categorical variables for pedestrian fatal accidents

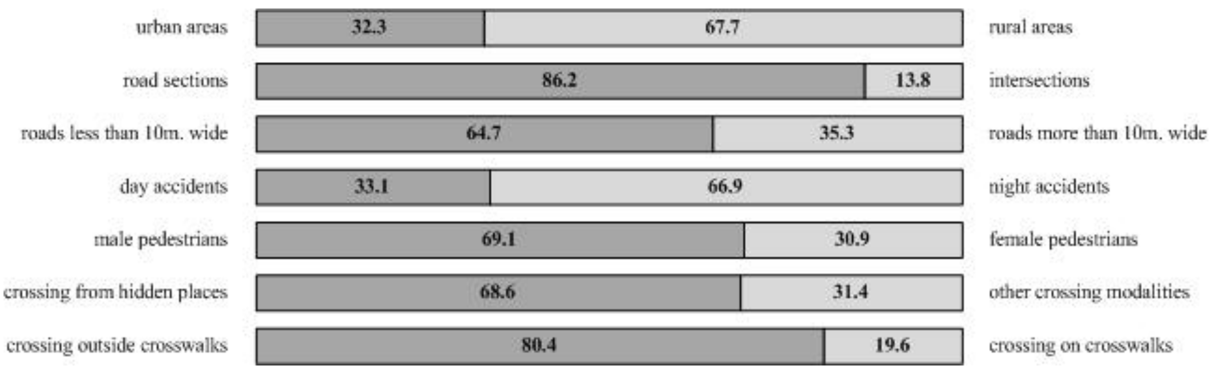
Variables	Categories
year of the accident	2003 – 2004 – 2005 – 2006
season of the accident	spring – summer – autumn – winter
day of the week	Sunday – Monday – Tuesday –Wednesday – Thursday – Friday – Saturday
day / night	day – night
period of the day	morning peak – off-peak – afternoon peak – evening/night
type of day	normal – pre-festive – festive – inner-festive
cause of the accident	driver behavior – passenger behavior – pedestrian behavior – motorcyclist behavior – cyclist behavior – vehicle malfunctioning
location of the accident	urban intersections – urban sections – rural intersections – rural sections
allowed speed	50 km/h – 60 km/h – 70 km/h – 80 km/h – 90 km/h – 100 km/h
weather conditions	clear – rainy – hot – foggy – not specified
road surface conditions	dry – wet from water – wet from slippery material – covered with mud – covered with sand – not specified
number of ways	one-way road – two-way road with separation – two-way without separation – not specified
median	with separation fence – without separation fence – not specified
shoulders of the road	good conditions – bad conditions – rough road
width of the road	up to 5 m. – 5 to 7 m. – 7 to 10 m. – 10 to 14 m. – over 14 m.
illumination on the road	daylight – night without illumination – night with illumination
traffic control	no control – working traffic light – failing traffic light – blinking yellow – stop sign – priority sign – not specified
crossing location	on crosswalk with traffic light – on crosswalk without traffic light – outside crosswalk next to intersection – outside crosswalk far from an intersection
crossing modality	suddenly – from hidden places – normally – not specified
age of the pedestrians	less than 14 – 15 to 19 – 20 to 24 – 25 to 34 – 35 to 44 – 45 to 54 – 55 to 64 – more than 65
gender of the pedestrians	male – female
social group of the pedestrians	Jewish – non Jewish –not specified or others
type of vehicles involved	car – light truck – heavy truck – public transport –motorcycle –bicycle – not specified
age of the vehicles	up to 2 years – 2 to 5 – 6 to 10 – 11 to 15 – more than 15
safety on board	seatbelts – helmets – child-seat – not used
property of the vehicles	private – army – police – other
age of the drivers	17 to 19 – 20 to 24 - 25 to 34 –35 to 44 –45 to 54 –55 to 64 – more than 65
licensure years for the drivers	up to 2 –2 to 5 –6 to 10 –11 to 20 – more than 20
gender of the drivers	male – female
social group of the drivers	Jewish – non Jewish – not specified or other



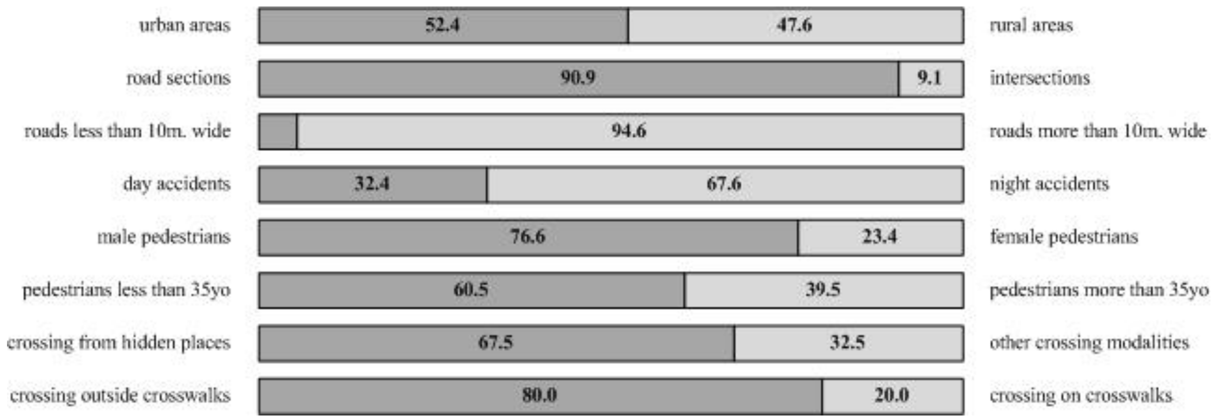
(a) first cluster



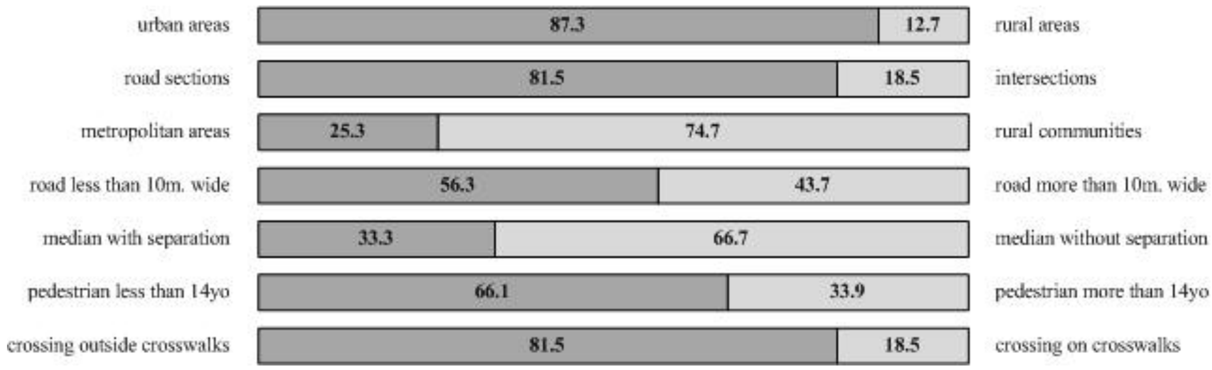
(b) second cluster



(c) third cluster



(d) fourth cluster



(e) fifth cluster

Figure 1. Relevant variables and frequent characteristics of the clusters of pedestrian accidents

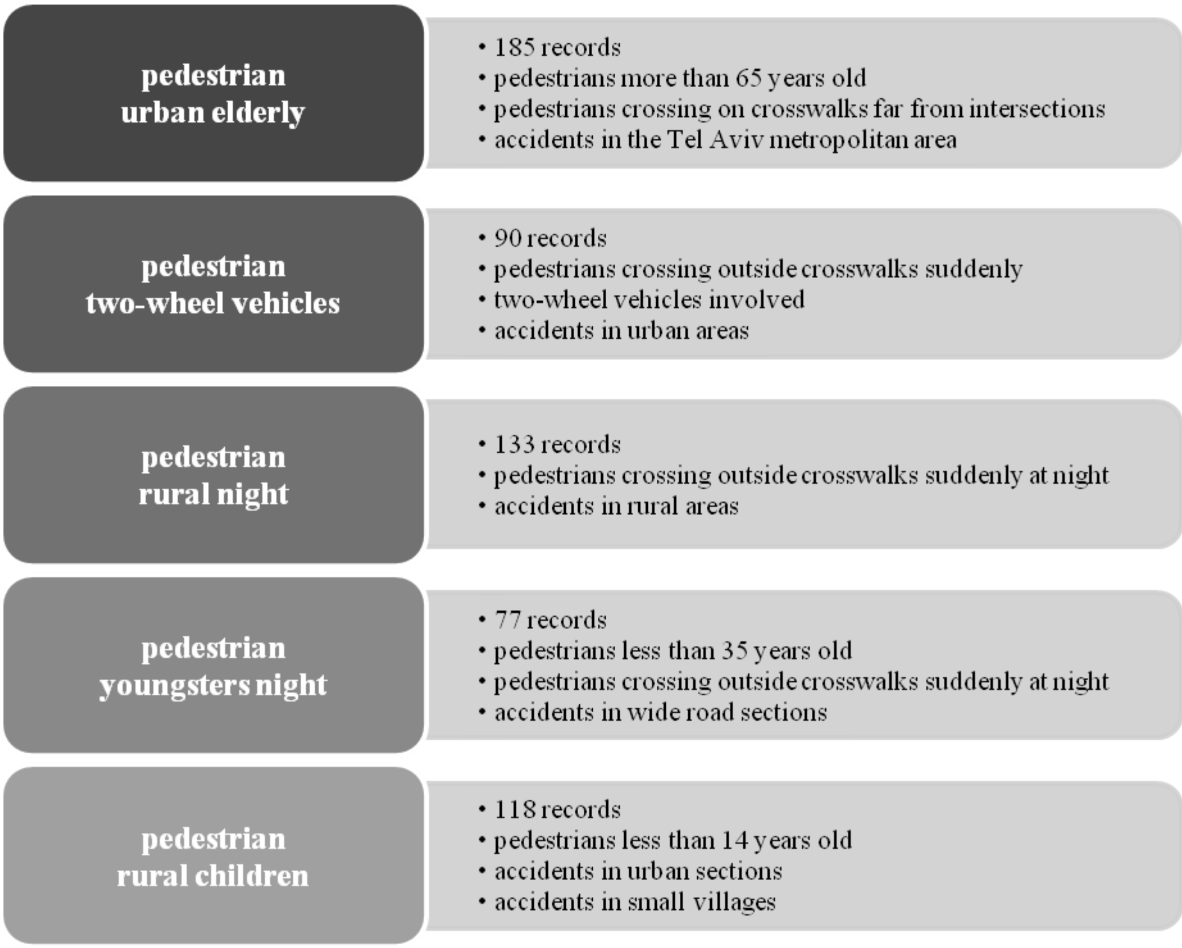


Figure 2. Definition of the clusters of pedestrian accidents