An individual risk model in the case of a road accident to support the decisional process in planning

F. A. Marcianò & A. Vitetta
DIMET, Department of Computer Science, Mathematics, Electronics and Transportation, Mediterranea University of Reggio Calabria, Italy

Abstract

In this paper an individual risk model in the case of road accidents is presented. The aim is to estimate the probability of a pedestrian being involved and the probability of injuries under pre-assigned conditions. The model calculates the individual risk of drivers and pedestrians moving in predefined scenarios that can be changed according with the purpose of the analysis. The model evaluates, in a quantitative way, each attribute's share in accident occurrence, and the relation between risk level and attribute. The model can support the decisional process in planning, allowing one to compare hazard levels of various hypothetical scenarios and to analyse the weight and the elasticity of each attribute that characterize the accident scenario. The results are presented of a test performed in the city of Reggio Calabria (Italy) in order to verify the model transferability and reliability.
Keywords: road accident, risk analysis, individual risk, drivers, pedestrians.

1 Introduction

Italian road accident statistics show the gravity for safety on the Italian road network, for the high number of fatalities and injuries. In Italy, every day there are approximately 598 accidents, 13 casualties and 849 injured [1].

One aspect of road safety analysis concerns the protection of "weak categories" of road users, as it is for pedestrians. In Italy 6.6% of accidents with injuries and 13.7% of deaths involve pedestrians. These numbers assume even greater importance if considering that accidents involving pedestrians account

only for 8.5% on the total of accidents [1]. Therefore, we deduct that pedestrian accidents have a lower frequency but at the same time when they occur the risk is much higher.

Current hazard levels could be reduced through some interventions planned on recognized network critical points (black spots). Planning must be preceded by risk analysis in order to support the decisional process through quantitative evaluations.

Risk can be defined as the undesired consequence of an event in relation to its probability of occurring [2] or as the distance between the real context and the absolute safety context. Such definitions indicate that risk (R) can be calculated with the following relations:

$$R = P \cdot V \cdot N \qquad \text{or} \qquad R = F \cdot V \cdot N$$

where

- P is the event probability [0;1] or F is the event observed frequency;
- V is vulnerability, which represents a measure of impact resistance; as it is a probability, it assumes values in the interval [0;1];
- N is exposure, which represents the potential number of people and goods that could be involved [3].

Risk models can be divided into two categories: social risk models, which measure probabilistic (frequentist) collective damage; individual risk models, which measure probabilistic (frequentist) individual damage. They can be: aggregate, when they are defined with an empirical relationship derived from global statistics about an analysed event; disaggregate, when attention is focused on each observation about an analysed event.

All social risk models previously considered are disaggregated [2–4] while individual risk models may be either disaggregated [5–10] or aggregated [10, 11]. Exposure can be estimated using demand models described in papers present in literature [12–15]. In literature different methods and models to plan and to evaluate interventions in order to improve traffic safety are presented. In [16] it was tried to model traffic safety using a driving cost function depending on speed and value of time. In [17] it was investigated the effectiveness of signalling at intersections on the conflict reduction between pedestrians and turning vehicles.

The aim of the paper is to investigate and possibly overcome the main limits of models proposed in the literature. They can be summarised in:

- the lack of completely disaggregate models, because models proposed in literature quantify risk in aggregate way or, even if the quantification of risk is disaggregated, probability and vulnerability is calculated as an aggregate;
- the lack of a model that allows to calculate risk values for drivers and pedestrians moving at the same time on a road in order to identify the scenario where one is more exposed than the other to accident risk.

In this paper an experimental individual risk model is proposed. It allows one to calculate individual risk value for drivers and pedestrians under the hypothesis that a road accident has occurred. Type of accident probability, driver

vulnerability and pedestrian vulnerability models are calibrated. These models have a Multinomial Logit model structure, assumed as a descriptive model, and are completely disaggregate. Moreover, it is constructed an aggregate model that provides the probability of the pedestrian involved having a certain age and sex.

The system of proposed models must be considered in its totality. Each model must not be considered in any way separately.

Model innovation consists in the possibility of calculating the risk value for drivers and pedestrians moving at the same time on roads in different contexts. Contexts can be defined through some external data that characterize the accident scenario. The models also allow one to estimate the weight and the elasticity of each attribute that characterized the accident scenario. Therefore it is possible to evaluate in a quantitative way how each attribute shares in accident occurring and how risk level changes with attribute changing.

Under the practical point of view the model could support the decisional process in planning [18, 19] allowing both to compare hazard levels of various hypothetical scenarios and to analyse the weight and the elasticity of each attribute that characterizes the accident scenario.

The paper has the following structure: in section 2 the proposed model specification, calibration and validation methods are presented; in section 3 the model is calibrated and validated; in section 4 a test performed on the city of Reggio Calabria (Italy) in order to evaluate model transferability and reliability using quantitative statistics is presented.

2   Proposed model

The proposed model assumes the hypothesis that a road accident has occurred. It is therefore specified with regard to two possible events:
- no pedestrian is involved (the accident involves two or more vehicles or an isolated vehicle), correspondent probability P_{no_ped};
- a pedestrian is involved, correspondent probability P_{ped}.

Given the notation V_d driver vulnerability, V_p pedestrian vulnerability, N_d driver exposure, N_p pedestrian exposure, it is possible to define two social risk functions for drivers, R_d, and pedestrians, R_p:

$$R_d = P_{no\_ped} \, V_d \, N_d$$
$$R_p = P_{ped} \, V_p \, N_p$$

The model is a nested one ant it must be considered in its totality. The output is represented by the risk value for a driver and/or a pedestrian. It means that each component model considered separately doesn't provide any significant solution.

The model allows calculating also the individual risk for drivers ($P_{no\_ped} \, V_d$) and pedestrians ($P_{ped} \, V_p$) under the hypothesis that a road accident has occurred in predefined scenarios that can be changed according with analysis purpose.

Social risk can be obtained multiplying the individual risk by the estimated exposure. However the aim of the model is to develop the individual risk structure through a system of nested models.

In regards to model applications it could be used to support the decisional process in planning. In practice the model should be useful to analyse situations and user' behaviours characterizing the road accident in order to identify the elements on which it is necessary to intervene. The purpose is to estimate the weight and the elasticity of each attribute that characterizes the accident scenario. The model evaluates in a quantitative way how each attribute shares in accident occurring and how risk level changes with attribute changing. In this way it is possible to understand in which cases it is necessary to intervene and above all in which direction. Model results allow one to identify attributes to be changed in order to reduce risk.

The model is defined in term of alternatives, calibration and validation.

The possible alternatives are:

- pedestrians involved
  - with an injured pedestrian of a given age and sex;
  - with a dead pedestrian of a given age and sex;
- no pedestrian involved
  - with an uninjured driver;
  - with an injured driver;
  - with a dead driver.

Probability and user vulnerability are calculated through Multinomial Logit models in relation to the specification accident [13] as a descriptive type. The probability that the pedestrian involved is estimated with a linear regressive model.

In relation to the calibration, the models are calibrated with the Maximum Likelihood method considering Gumbel distribution for residual in the probability and vulnerability models and normal distribution for the pedestrian accident probability model.

Validation consists in verifying parameters reasonableness and significance and model reliability by the t-Student test and ρ^2 statistics.

3 Experimentation

The models are constructed using a road accident database of about 60000 accidents in 188 municipalities of the province of Milan during the two-year period of 2001/2002.

3.1 Attributes identification

Different types of attributes are tested in order to identify the best attribute set for models calibration. The attributes used are those reported in table 1. Population data and municipalities area data are provided by ISTAT (Italian Institute of Statistics). Vehicle data is provided by the ACI (Italian Automobile Club).

Table 1:   Type of attributes.

Attributes for accident probability	Unit measurement
tot_veh/pop	number of registered vehicles/inhabitants
$X_{day/night}$	1 if day (7-18)
	0 if night (19-6)
$X_{rush-hour}$	1 if accident occurs at rush-hour (7, 8, 12, 13, 18, 19, 20)
	o if accident does not occur at rush-hour
$X_{urb\ road}$	1 if urban road
	0 if extraurban road
$X_{built\ up\ area}$	1 if road inside built-up area
	0 if road outside built-up area
X_{cross}	1 if accident occurs at crossing
	0 if accident does not occur at crossing
$X_{driv\ age}$	years
Attributes for pedestrian accident probability	**Unit measurement**
ped_age/100	years/100
(ped_age /100)^2	(years/100)^2
(ped_age /100)^3	(years/100)^3
sex_m	1 if male pedestrian <10 years
	0 if male pedestrian ≥10 years
sex_f	1 if female pedestrian ≥50 years
	0 if female pedestrian <50 years
Attributes for pedestrian vulnerability	**Unit measurement**
$X_{ped\ age}$	years
$X_{light\ veh}$	1 if light vehicle (car or 2-wheeled vehicle)
	0 if heavy vehicle
$X_{resp\ speed-limit}$	1 if speed limit is observed
	0 otherwise
Attributes for driver vulnerability	**Unit measurement**
X_{wheels}	1 if 4 (or plus)-wheeled vehicle
	0 if 2-wheeled vehicle
$X_{resp\ speed-limit}$	1 if speed limit is observed
	0 otherwise
$X_{veh\ years}$	years since vehicle registration year
$X_{driv\ age}$	years
$X_{driv\ sex}$	1 if male
	0 if female
$X_{prot\ syst}$	1 if driver used protection systems (helmet, belts)
	0 if driver did not use protection systems

3.2 Specification

Specification is done defining different impedances given by different sets of attributes which are calibrated and validated.

The model of accident probability provides the two alternative probabilities conditioned by the occurrence of a road accident:

1. no pedestrian involved (the accident involves two or more vehicles or an isolated vehicle);
2. pedestrian involved.

In the first alternative only a shadow variable, termed X_{no_ped}, is included while in the second alternative the attributes defined at section 3.1 are included.

The model of pedestrians' accident probability is regressive and the attributes defined in section 3.1 are included.

The model of pedestrian vulnerability provides the two alternative probabilities conditioned by pedestrian involved occurring:

1. injured pedestrian;
2. dead pedestrian.

In the first alternative only a shadow variable, termed $X_{injured}$, is included while in the second alternative the attributes defined at section 3.1 are included.

The driver vulnerability model is split into two levels.

The first level provides the two alternative probabilities conditioned by no pedestrian involved occurring:

1. uninjured driver;
2. hurt driver (injured or dead).

In the first alternative only a shadow variable, termed $X_{uninjured}$, is included while in the second alternative the attributes defined at section 3.1 are included.

The second level provides the two alternative probabilities conditioned by hurt driver occurring:

1. injured driver;
2. dead driver.

In the first alternative only a shadow variable, termed $X_{injured}$, is included while in the second alternative the attributes defined at section 3.1 are included.

3.3 Calibration

Various calibrations are performed to reduce the objective function for each model. In the pedestrian accident probability model each pedestrian category is replaced with its barycentric value of age (5, 15, 25, 35, 45, 55, 65 and 75 years) in order to apply the Least Squares method. The model has a step function trend where each step represents accident probability for a pedestrian belonging to that age category; a continuous trend (the function passes through estimated probability values) has only a qualitative meaning.

3.4 Validation

Calibrated models are validated in order to verify each attribute weight and to discard attributes with a wrong parameter sign and with a low t-Student value. The best calibration and validation results are reported in table 2.

4   Test on a real context

The test is performed on the city of Reggio Calabria (Italy) in order to verify model transferability and reliability using quantitative statistics. The main results are presented.

4.1 Targets and methods used

The test on a real context is executed in order to:

-   verify model transferability and reliability when applied to a context other than that on which model was calibrated;

Table 2:   Calibrated parameters and validation.

Accident probability (Observations = 60623)

	1	2							ρ²	log L_in	log L_fin
Attrib.	X_{no_ped}	$T_{ot\_veh/pop}$ vehic/inhabit	$X_{day/night}$ 1 if day	$X_{rush\text{-}hour}$ 1 if rush-hour	X_{urb_road} 1 if urb road	$X_{built\text{-}un\_area}$ 1 inside built-up area	X_{cross} 1 if at crossing	X_{driv_age} years			
estim. β	8.838	4.352	0.556	0.205	0.927	1.383	-0.574	0.0127	0.64	-42019.9	-15049.4
t-Stud.	35.6	16.2	13.8	6.4	9.4	8.6	-18.1	11.9			

Pedestrian accident probability (5195)

Attrib.	ped_age/100 years/100	(ped_age)² (years/100)²	(ped_age)³ (years/100)³	sex_f 1 if female	sex_m 1 if male	ρ²
estim. β	0.595	-1.609	1.274	0.021	0.015	0.99
t-Stud.	13.4	-9.2	7.9	1.8	3.6	

Pedestrian vulnerability (Observations = 4964)

	1	2			ρ²	log L_iniz	log L_fin
Attrib.	$X_{injured}$	X_{ped_age} years	X_{light_vehic} 1 if light vehicle	$X_{resp\_speed\text{-}limit}$ 1 if speed limit observed			
estim. β	3.526	0.02118	-1.020	-0.951	0.88	-3440.1	-396.1
t–Stud.	8.2	4.1	-4.0	-3.5			

Driver vulnerability level I (Observations = 26542)

	1	2						ρ²	log L_iniz	log L_fin
Attrib.	$X_{uninjured}$	X_{wheels} 1 if 4 (or +) wheels	$X_{resp\_speed\text{-}limit}$ 1 if speed limit observed	X_{vehic_years} years since veh. registration year	X_{driv_age} years	X_{driv_sex} 1 if male	X_{prot_syst} 1 if prot systems used			
estim. β	-4.883	-3.365	-0.1259	0.01227	-0.01252	-0.4542	-1.079	0.24	-18396.8	-13957.9
t–Stud.	-48.9	-59.4	-2.5	3.7	-12.0	-13.5	-21.2			

Driver vulnerability model, level II (Observations = 15279)

	1	2				ρ²	log L_iniz	log L_fin
Attrib.	$X_{injured}$	$X_{resp\_speed\text{-}limit}$ 1 if speed limit observed	X_{driv_age} years	X_{driv_sex} 1 if male	X_{prot_syst} 1 if used prot systems			
estim. β	4.315	-1.0520	0.0099	1.0010	-1.6050	0.96	-10590.6	-458.5
t–Stud.	7.9	-3.7	1.3	2.5	-6.8			

- plot probability and vulnerability curves;
- plot individual risk curves as regards predefined accident scenarios.

Simulated accident scenarios are:
- injured male pedestrian hit by a light vehicle (car or 2-wheeled vehicle);
- injured male pedestrian hit by a heavy vehicle;
- no pedestrian involved accident with the following characteristics
  - night / no rush-hour / 2-wheeled vehicle / hurt driver (injured or dead);
- no pedestrian involved in accident with the following characteristics
  - day / rush-hour / 4 or more wheeled vehicle / hurt driver (injured or dead).

The test is carried out in the city of Reggio Calabria (Italy). The 2005 road accident database is used, consisting of a series of records, in all 1146. Each record refers to a single accident and contains information on vehicles, drivers, pedestrians and environmental conditions.

   Model transferability and reliability are verified using quantitative statistics. In particular the Kolmogorov-Smirnov test is used to verify that the model, calibrated on the Milan road accident data set, fits the data of Reggio Calabria road accident data set. The Kolmogorov-Smirnov test is verified by each model for significance levels of 90 and 95%, which indicates a good model reliability.

4.2 Test on Reggio Calabria

The purpose of the test is to plot the probability and vulnerability curves, and hence individual risk curves. Curves are plotted as regards predefined accident scenarios according to user age. In [18] accident scenario is defined as the prototype of an accident event that simulates a set of similar accidents in terms of events concatenation and causes interrelationship.

4.2.1 Test on type of accident probability model
The test is carried out for six scenarios described in table 3. The purpose is to point out that pedestrian involved and no pedestrian involved probabilities change with: driver age and time band (scenarios I and II); location (scenarios III, IV, V, VI).
   The curves obtained for the scenarios I and II are reported in figure 1. In the same way, the figures for the other scenarios can be obtained.

Table 3:   Accident scenarios.

Accident probability

Scenario	Time band	Rush-hour	Road type	Location	Position
I, III, V	day	Yes	urban	inside built-up area	straight
II	night	No	urban	inside built-up area	straight
IV	day	yes	urban	inside built-up area	crossing
VI	day	yes	urban	outside built-up area	straight

Pedestrian vulnerability

Scenario	Type of vehicle	Speed limit observed
VII	light vehicle (car; 2-wheeled vehicle)	yes
VIII	heavy vehicle	yes

Driver vulnerability model (level I)

Scenario	Type of vehicle	Speed limit observed	Vehicle age	Driver sex	Protection systems
IX	4 (or +) wheels	yes	3	male	no
X	2 wheels	yes	3	male	no

Pedestrian risk functions

Scenario	Type of vehicle	Effects on pedestrian
A	heavy vehicle	injured
B	light vehicle	injured

Driver risk functions

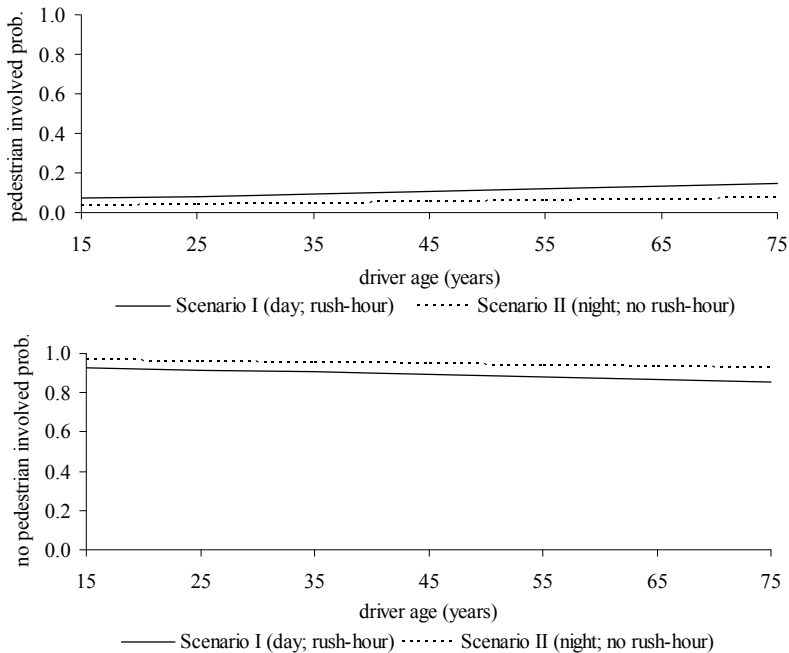| Scenario	Time band	Rush hour	Type of vehicle	Effects on driver
C	night	no	2 wheels	injured or dead
D	day	yes	4 (or +) wheels	injured or dead

Figure 1:   Comparison between pedestrian involved and no pedestrian involved probabilities at changes in driver age and proposed scenarios.

The results indicate that the pedestrian involved probability is higher during the day and at rush-hour; hence the no pedestrian involved probability is higher at night and at non rush-hour.

This may be explained simply: pedestrians move by day and at rush-hour mainly with an accident concentration in this time band. Hence most accidents at night are statistically those where no pedestrian involved. Moreover, pedestrian involved probability is lower at crossings because the presence of traffic lights and/or signs makes crossing safer. There is a higher probability of pedestrians not being involved outside built-up areas where traffic is not congested and speeds are higher.

4.2.2 Tests on vulnerability models

Tests are carried out on the four scenarios described in table 3. The purpose is to show how injured pedestrian and dead pedestrian probabilities and hurt driver and dead driver probabilities change with:

- pedestrian and driver age;
- type of vehicle.

The curves obtained for the scenarios 7 and 8 are reported in figure 2. In the same way, the figures for the other scenarios can be obtained.

The results show that the dead pedestrian probability is higher for heavy vehicles and that the dead driver probability is higher for 2-wheel vehicles.

4.2.3 Risk curves for drivers and pedestrians
Risk curves for drivers and pedestrians are plotted for the higher and lower danger scenarios. Risk scenarios are described in table 3. The results obtained are reported in figures 3 and 4.

5 Conclusions

In this paper an individual risk model in the case of a road accident is presented. The purpose is to estimate the probability of a pedestrian being involved besides
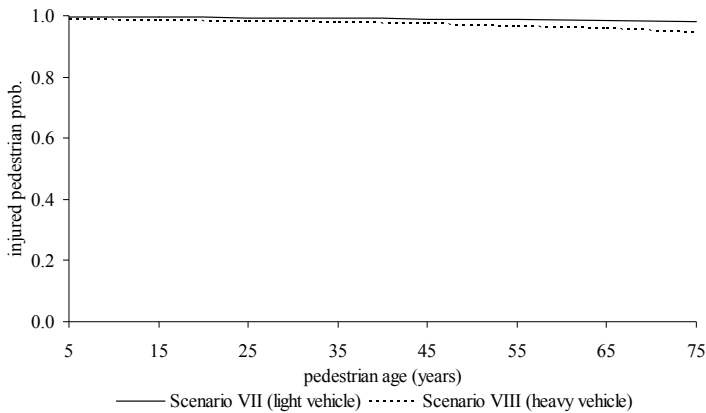
Figure 2: Comparison between injured pedestrian probabilities at changes in pedestrian age and proposed scenario.
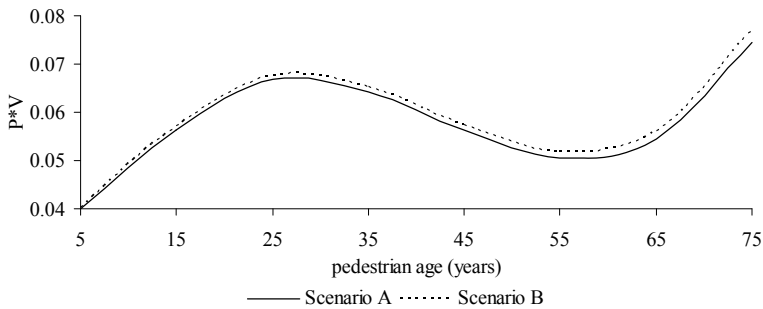

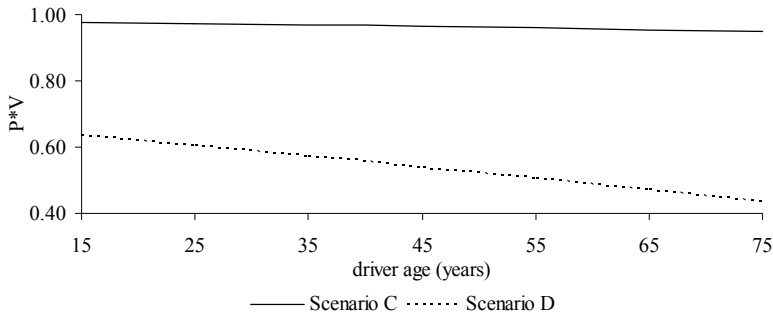Figure 3: Comparison between pedestrian risk functions at changes in proposed scenario.

Figure 4:   Comparison between driver risk functions at changes in proposed scenario.

the driver and the probability of their being injured under pre-assigned conditions. The model also allows one to analyse the weight and the elasticity of each attribute that characterized the accident scenario.

Although the definition of risk has been established for a long time now, in the scientific literature there is substantial heterogeneity in approaches and methods for risk analysis. Therefore in this paper a risk model whose formal expression derived directly from risk definition is presented. Type of accident probability, driver vulnerability and pedestrian vulnerability models are calibrated. These models have a Multinomial Logit model structure assumed as a descriptive model. Importantly, the models are of a completely disaggregated type.

Pedestrian accident probability analysis is also investigated in depth. An aggregated model is proposed which provides the probability of the pedestrian involved being of a certain age and sex.

Finally, a test is carried out on the city of Reggio Calabria (Italy) to verify model transferability and reliability using quantitative statistics. In particular the Kolmogorov-Smirnov test is used to verify that the model, calibrated on one data set, fits the data of another data set. Probability, vulnerability and individual risk curves are plotted for predefined accident scenarios. Model results are validated using accident data recorded in Reggio Calabria in 2005 through the Kolmogorov-Smirnov test. This test is verified by each model for significance levels of 90 and 95%.

The proposed model allows one to compare hazard levels in various hypothetical situations. It is shown that the risk of pedestrian injury is almost the same whether he/she is hit by a light or heavy vehicle (figure 3). This means that the consequences of pedestrian accidents are linked much more closely to the absence of any protection system. By contrast, the driver's risk of injury is higher in the case of accidents on 2-wheeled vehicles than on 4 (or more)-wheeled vehicles (figure 4).

Starting from these surveys it is possible to define types of intervention to be pursued: in the first case it is advisable to adopt preventive measures in order to reduce accident probability (such as seeking to separate pedestrian flows from

vehicular flows); in the second case mitigation measures should be adopted in order to improve 2-wheeled vehicles safety, as accident probability may be considered constant during planning.

References

[1]   ISTAT - Italian Institute of Statistics, *Incidenti stradali - anno 2008*. On www.istat.it.

[2]   Center for Chemical Process Safety, *Guidelines for chemical transportation risk analysis*. American Institute of Chemical Engineers, New York, 1995.

[3]   Russo F. & Vitetta A., Risk evaluation in a transportation system. *International Journal of Sustainable Development and Planning*, 1 (2), pp. 170-191, 2006.

[4]   Russo F. & Vitetta A., Safety of users in road evacuation: General methodology and main results. *Proc. of Urban Transport XIII, Urban Transport and the Environment in the 21st century*, Brebbia C. A. (ed.), WIT Press, Southampton, pp. 763-772, 2007.

[5]   Keall, M.D., Pedestrian exposure to risk of road accident in New Zealand. In *Accident Analysis*, 27 pp. 729-740, 1995.

[6]   Lee, C., Abdel-Aty, M., Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. In *Accident Analysis*, 37 pp. 775-786, 2005.

[7]   Vitetta A., Musolino G. & Marcianò F.A., Safety of users in road evacuation: Supply and demand-supply interaction models for users. *Proc. of Urban Transport XIII, Urban Transport and the Environment in the 21st century*, Brebbia C. A. (ed.), WIT Press, Southampton, 2007.

[8]   Vitetta A., Musolino G. & Marcianò F. A., Safety of users in road evacuation: calibration of cost functions and simulation. *Proc. of Urban Transport XIV, Urban Transport and the Environment in the 21st century*, Brebbia C. A. (ed.), WIT Press, Southampton, pp. 715-725, 2008.

[9]   Vitetta A., Quattrone A. & Polimeni A., Safety of users in road evacuation: design of path choice models for emergency vehicles. *Proc. of Urban Transport XIII, Urban Transport and the Environment in the 21st century*, Brebbia C. A. (ed.), WIT Press, Southampton, pp. 803-812, 2007.

[10] Delfino, G., Rindone, C., Russo, F., Vitetta, A., Risk analysis in road safety: a model and an experimental application for pedestrians. In *Proceedings of the European Transport Conference*, Strasbourg, France, 2005.

[11] Evans, L., *Safety and the driver*. Van Nostrand Reinhold, New York, 1991.

[12] Ben Akiva M., Lerman S.R., *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press Cambridge, Massachusetts, 1985.

[13] Cascetta, E., *Transportation systems engineering: theory and methods*, Kluwer, 2001.

[14] Domencich, T.A., Mc Fadden, D., *Urban travel demand: a behavioural analysis*. American Elsevier, New York, 1975.

[15] Ortuzar, J., de, D., Willumsen, L.G., *Modelling Transport*. John Wiley & Sons, Second Edition, 1996.
[16] Delhaye, E., Traffic safety: Speed limits, strict liability and a km tax. In *Transportation Research Part A*, Vol. 40, Issue: 3, pp. 205-226, 2006.
[17] Al-Kaisy, H. N. A., Investigation of signing to reduce conflicts between pedestrians and turning vehicles at signalized intersections. In *Transportation Research Part A*, Vol. 30, Issue: 1, pp. 64-65, 1996.
[18] Russo F. & Rindone C., Safety of users in road evacuation: planning internal processes and guidelines. *Proc. of Urban Transport XIII, Urban Transport and the Environment in the 21st century*, Brebbia C. A. (ed.), WIT Press, Southampton, pp. 825-834, 2007.
[19] Russo F. & Vitetta A., An assignment model with modified Logit, which obviates enumeration and overlapping problems. *Transportation*, 30 (3), pp. 177-201, 2003.
[20] Brenac, T., Megherby, B., Diagnostic de sécurité routière sur une ville: intérêt de l'analyse fine de procédures d'accidents tirées aléatoirement. In *Recherche Transports Sécurité*, 52 pp. 59-71, 1996.