TRANSFERABILITY OF ACCIDENT PREDICTION MODELS FOR URBAN INTERSECTIONS

Marchionna A., Perco P., Tavernar M.C.
Department of Civil and Environmental Engineering
University of Trieste, Italy
marchionna@dica.units.it, perco@dica.units.it, mc.tavernar@libero.it

ABSTRACT

More than 75 percent of accidents with casualties and fatalities in Italy occur in urban areas, and more than half of them occur on intersections (826 fatalities and 123.000 injured persons in 2004). Moreover, accidents with only material damage that occur at intersections also cause significant disturbance and delays to traffic flow. Therefore, it is important for highway engineers to know the relationship between accidents and the characteristics of the intersections (traffic data, geometry, regulation) in order to intervene effectively in reducing accidents. For this reason many accident prediction models for urban intersections have been developed and are currently used in many countries. Obviously, the availability of these models can lead to their application in the Italian context. However, the application of these models in a geographic area different from that in which they were developed presents significant problems of transferability because they are based on data (accidents, geometry, regulation, traffic, vehicle characteristics, driver population, etc..) that represent the specific conditions of the original geographic area. Consequently the use of these models in different areas could result in unreliable accident estimations. The transferability of some accident prediction models for urban intersections to a typical Italian context was evaluated using accident and traffic data of intersections in the Trieste urban area. Several GOF statistics were used to assess the performance of the models. The results show a poor transferability of the models tested to the urban intersections of Trieste.

Keywords: accident prediction models, urban intersections, transferability

INTRODUCTION

In order to manage the road safety of urban intersections, highway engineers need to have a good insight into the variables that explain accident occurrence. Accident prediction models (APMs) are useful tools for this task because they relate the annual accident experience of an intersection to its characteristics (geometry, regulation, traffic volume etc.). They can be used, therefore, to compare the actual accident experience of a specific intersection with the expected safety performance of similar intersections. However, an accident prediction model is calibrated using accident data of a sample of intersections representative of a specific geographic area. Therefore, the transferability of the model to other geographic areas requires special attention. In fact, accident frequencies vary across time and space, even between roads that are similar, because of differences in factors such as accident reporting thresholds, accident reporting practices, driver population, law enforcement, animal populations, vehicle characteristics, and climate. Consequently, an accident prediction model cannot be used for sites not included in the geographic area for which it was developed without a transferability evaluation. This problem is well known and has been the object of several studies (Persaud B. et al. 2002; Washington S. et al, 2005). To tackle this problem, Harwood et al. (Harwood et. al., 2000) outlined a procedure to recalibrating "base" accident prediction models to suit the safety conditions present in the states of the United States for which accident data were not used to develop the "base" models. This calibration procedure involves the estimation of a calibration factor appropriate for a particular State that is used to multiply the prediction of the "base" model. Moreover, a second procedure based on the Empirical Bayes approach makes it possible to combine the results of the "base" accident prediction model with accident history data of a specific site. However, to use these calibration procedures is not easy: this task requires an expertise in highway safety analysis procedures, a good knowledge of accident prediction models and the availability of appropriate data. Considering that 75.7% of traffic accidents in Italy occur on urban roads (169,893 accidents) and more than 52% of these accidents occur at urban intersections (826 fatalities and 122.592 injured – ISTAT, 2005), the objective of this study was to evaluate if accident prediction models for urban intersections developed in foreign countries or in a particular Italian region, can be used in other Italian regions without a recalibration procedure.

ACCIDENT PREDICTION MODELS EVALUATED

Currently there are several accident prediction models for urban intersections developed in different countries (Reurings et al., 2005). From all the existing models, the following models were considered to evaluate their transferability to intersections in the urban area of Trieste:

- Bauer and Harwood model (Bauer and Harwood, 2000), United States;
- Canale et al. model (Canale et al., 2005), Italy;
- Summergill et al. model (Summergill et al., 2001), United Kingdom;
- Greibe model (Greibe 2003), Denmark.

Each of these four models were selected for a specific reason: the Bauer and Harwood model was selected because it is a particularly complete model that considers many geometric characteristics of the intersection as independent variables; the Canale et al. model because it was developed in an Italian city; the model Summergill et al. model because it is based specifically on data from intersections with one or more one-way arms and one-way roads are common in Italian cities, but little is known about how numbers of accidents and their distribution differ from those on two-way roads; finally, the Greibe model was selected because it is a "basic" model that considers only the traffic flow as independent variable and was developed in another European country with right hand driving.

Bauer and Harwood used data from a database provided by the California Department of Transportation. The data included all collision types (i.e., both multiple- and single-vehicle accidents) using 3-year accident frequencies (1990 to 1992) and geometric design, traffic control, and traffic volume data. The set of models developed predict the total accident frequency and the injury accident frequency for the following urban intersection types:

- four-leg, stop -controlled intersections
- three-leg, stop -controlled intersections
- four-leg, signalized intersections

The authors used two general types of statistical models: a lognormal regression model for four-leg intersections (both stop-controlled and signalized) and a loglinear regression model for three-leg, stop-controlled intersections. In the case of lognormal regression the logarithm of the number of accidents is supposed to be normally distributed and the coefficients are estimated by the least-square method. In the case of loglinear regression the number of accidents is supposed to follow a Poisson or a negative binomial distribution and the coefficients are estimated by the maximum likelihood method. Models of both types are in the form:

$$\mu_i = \exp(\beta_0) \cdot (ADT_{major-road})^{\beta_1} \cdot (ADT_{crossroad})^{\beta_2} \cdot \exp(\beta_3 X_{i3}) \cdot ..... \cdot \exp(\beta_q X_{iq})$$

where μ_i is the expected number of accidents at the ith intersection in a 3-year period, X_{ij} are the predictor variables and β_k are the coefficients to be estimated, shown in table 1. The model uses the natural logarithm of the major-road and crossroad average daily traffic (ADT) as variables.

Canale et. al used data from 400 intersections in the urban area of Catania. The data included accidents in a 3-year period, traffic volume, geometric characteristics and traffic control. The set of models developed predict the injury accident frequency in a 3-year period for the following intersection types:

- three-leg, no control intersections
- three–leg, stop - controlled intersections
- four-leg, no control intersections
- four–leg, stop - controlled intersections
- four–leg, signalized intersections.

The authors tested both the types of statistical models: lognormal regression model and loglinear regression model (Poisson regression). However, for all the intersection types the statistical analysis revealed that the loglinear regression model was more appropriate. The model form is the same as Bauer and Harwood's, the coefficients β are shown in table 2.

Table 1 – Accident Prediction models (Bauer and Harwood, 2000)

Variable	Variable level	Coefficients β					
		Four-leg stop		Three-leg stop		Four-leg signalized	
		Total	Injury	Total	Injury	Total	Injury
Intercept		-4.664	-4.693	-5.557	-6.618	-3.428	-5.745
Major road ADT (log)		0.620	0.584	0.683	0.696	0.503	0.574
Crossroad ADT (log)		0.281	0.206	0.245	0.238	0.224	0.215
Major road Left turn	Prohibited	-0.941	-0.747	-0.402	-0.393	-	-
	permitted	0.000	0.000	0.000	0.000	-	-
Major road Left-turn channelization	No left-turn lane	-	-	0.019	-0.057	-	-
	Painted left-turn lane	-	-	0.000	0.000	-	-
	Curbed left-turn lane	-	-	0.210	0.209	-	-
Major road Right turn channelization	No free right turns	-	-	-	-	-0.115	-
	Provision for free right turns	-	-	-	-	0.000	-
Major road Lane width		-0.097	-0.081	-0.037	-0.048	-0.053	-
Major road outside shoulder width		-	-0.020	-	-	-	-
Major road Number of lanes	3 or less	0.401	0.282	-	-	-0.225	-0.163
	4 or 5	0.120	0.049	-	-	-0.130	-0.151
	6 or more	0.000	0.000	-	-	0.000	0.000
Major road Presence of median	Divided	-	-	-0.174	-0.182	-	-
	Undivided	-	-	0.000	0.000	-	-
Major road Access control	None	-0.437	-0.382	-	-	-0.310	-0.290
	Partial	0.000	0.000	-	-	0.000	0.000
Major road Functional class	Principal arterial	0.000	0.000	-	-	-	-
	Minor arterial	-0.153	-0.079	-	-	-	-
	Major collector	-0.229	-0.401	-	-	-	-
Design speed of major road		-	-	-0.006	-	-	0.005
Crossroad Right turn channelization	No free right turns	-0.384	-0.300	-0.559	-0.581	-	-
	Provision for free right turns	0.000	0.000	0.000	0.000	-	-
Crossroad Number of lanes	3 or less	-	-	-	-	-0.130	-0.155
	4 or more	-	-	-	-	0.000	0.000
Lightning	No	-0.160	-	-	0.094	-	-
	Yes	0.000	-	-	0.000	-	-
Signal timing	Pretimed	-	-	-	-	0.063	-0.051
	Semi-actuated	-	-	-	-	0.000	0.000
	Fully actuated	-	-	-	-	0.622	0.400
Signal phasing	Two-phase	-	-	-	-	0.000	0.000
	Multiphase	-	-	-	-	-0.200	-0.240

Table 2 – Accident Prediction models (Canale et al., 2005)

Variable	Variable level	Coefficients β				
		3 leg		4 leg		
		no control	stop	no control	stop	signalized
Intercept		-7.396	-14.558	-5.515	-5.398	-5.630
Major road ADT		0.731	0.810	0.251	0.275	-
Crossroad ADT		-	0.405	0.244	0.445	1.038
Major road Left turn	Prohibited	0.000	-	-	0.000	-
	Permitted	0.422	-	-	0.819	-
Major road Right turn	Prohibited	-	-	-	0.000	0.000
	Permitted	-	-	-	0.886	1.147
Major road Right turn	Protected	-	-	-	-	0.737
	Not Protected	-	-	-	-	0.000
Major road Side-walk width		-	0.445	-	-	-0.237
Major road Presence of median	Absent	-	1.207	-	-	-
	Present	-	0.000	-	-	-
Major road Operation	One way	0.000	-	0.000	-	0.000
	Two way	0.703	-	0.403	-	1.893
Major road Percent grade	Steep grade	-0.745	-	-	-0.673	0.000
	Level	-0.029	-	-	-0.232	-0.529
	Moderate grade	0.000	-	-	0.000	0.000
Crossroad left turn	Prohibited	-	-	-	0.000	0.000
	Permitted	-	-	-	-0.469	-0.984
Crossroads left turn	Protected	-	-	-	-	-1.470
	Not Protected	-	-	-	-	0.000
Crossroad right turn	Prohibited	-	-	-	0.000	0.000
	Permitted	-	-	-	-0.605	0.342
Crossroads right turn	Protected	-	-	-	-	2.061
	Not Protected	-	-	-	-	0.000
Crossroad lane width		-	0.339	0.317	0.179	-0.293
Crossroad Side-walk width		-	-	0.256	-	-
Crossroad Number of lanes		-	-	-	-	-0.515
Crossroad Operation	One way	-	0.000	0.000	-	-
	Two way	-	0.826	-0.523	-	-
Crossroad Percent grade	Steep grade	1.172	0.152	0.478	-0.604	-
	Level	0.302	-1.253	0.549	-0.153	-
	Moderate grade	0.000	0.000	0.000	0.000	-
Road markings	Present	0.000	0.000	-	-	-
	Absent	-0.707	0.558	-	-	-
Signal phasing	Two-phase	-	-	-	-	-1.669
	Multiphase	-	-	-	-	0.000

The particularity of the set of models developed by Summergill et al. is that it is based on data from intersections with one or more one-way legs. The models are based on the data of injury accidents collected on 433 urban intersections of the following types:

- three-leg, priority intersection

- three-leg, signalised intersection
- four-leg, priority intersection
- four-leg, signalised intersection.

Records of 3,622 personal injury accidents occurring at the junctions were obtained for the period 1987-1994 inclusive. Since the number of accidents in a given period does not follow a normal distribution, the technique of generalised linear modelling was used to develop accident prediction models from the data. Different model forms were developed also considering traffic flow functions, accident types, and geometric factors. This study evaluated only the basic models that estimate the accident frequency at the intersection in function of the total traffic inflow. The form of the basic model is

$A = kQ^{\alpha}$

where A is the expected number of injury accidents per year, Q is the total traffic inflow in thousands of vehicles in a 24-hour period, k and α are the parameters to be estimated shown in table 3.

Table 3 – Accident Prediction models (Summergill et al., 2001)

Type	3-leg priority	4-leg priority	3-leg signal	4-leg signal
α	0.865	0.430	0.432	0.794
k	0.058	0.270	0.237	0.257

Greibe used data on traffic flow and the geometric characteristics of 1036 urban intersections of the following types:

- three-leg, signalised intersection
- three-leg, non-signalised intersection
- four-leg, signalised intersection
- four-leg, non-signalised intersection.

The accident data included 2534 personal injury and damage only accidents in the period 1987-1991. The traffic flow was considered to be a continuous variable, while all the other variables were converted into class variables. Generalised linear modelling was used to fit the models to the data. The distribution of accidents was supposed to follow a Poisson distribution. The model form is:

$\mu_i = a \cdot N_{pri}^{p1} \cdot N_{\sec}^{p2} \cdot \exp \sum \beta_j \cdot x_{ij}$

where μ_i is the expected number of accidents at the i^{th} intersection per year, N_{pri} is the incoming traffic flow (ADT) from the primary direction, N_{sec} is the incoming traffic flow (ADT) from the secondary direction, x_{ij} the variables describing road geometry or environment of the intersection and α, p, β_j are the parameters to be estimated. The statistical analysis showed that the variables describing road geometry do not increase significantly the percentage explained, therefore the models developed, shown in table 4, use only traffic flows as independent variable.

Table 4 – Accident Prediction models (Greibe, 2003)

Type	3 leg not signalized	3 leg signalized	4 leg not signalized	4 leg signalized
α	1.04E-05	1.34E-05	7.12E-04	1.08E-04
$p1$	0.69	0,88	0.30	0.53
$p2$	0.60	0.33	0.55	0.52

DATA SOURCES

In order to evaluate the transferability of these models, detailed information on accident data, traffic flow and the road layout of urban intersections was required. Therefore this study collected these data for intersections in the urban area of Trieste - a city in the north-east of Italy (population 250,000),.

Crash data

The accident statistics database covering all local police recorded accidents in a 12 year accident period (1990–2001) was available. The database contains more than 38,500 accidents that occurred both on intersections and along streets. 17,445 of them occurred at intersections, causing 36 fatalities and 9,332 injured persons. The data included location, date, time, death, injured, number and type of vehicles involved and pedestrians. All accidents were related to the specific localization using the street names. The intersections were localized with the names of the crossing streets. The data about the nature of accident and the violated road code rules are incomplete and so they are not usable for the purposes of this study. The database included injury accidents and damage only accidents. However, it should be noted that all the injury accidents that occurred were present in the database since the police always intervene in them, while in the case of damage only accidents there are accidents not recorded since the police did not intervene. This element must be considered when the transferability of accident prediction models for all accidents is evaluated.

Traffic data

The traffic data used for this study are based on the report (Rilievi di Traffico, 2002) developed by the Department of Civil and Environmental Engineering of Trieste University that contains the results of an extensive traffic survey campaign commissioned by the Trieste county council to develop the new Traffic Urban Plan. This report contains the traffic volume for the peak hour (7.30-8.30 a.m.) of all the streets in the Trieste urban area, except for minor streets that do not present transit traffic flow but only residential and parking traffic. To calculate the average daily traffic (ADT) starting from the peak hour traffic, the coefficient 0.129 calculated on the basis of the traffic data collected was used. It should be noted that many streets are characterized by one-way traffic reserved for public transport and services. In these cases the traffic flow was not available in the report, therefore the estimation of the ADT was made considering the timetables of all the bus routes, integrated by a field survey to evaluate the other public vehicles that use the reserved routes.

Intersection characteristics

The traffic data were used to determine the intersections for which the traffic flow of all legs were available. Therefore, a reconnaissance survey was undertaken at the intersections to evaluate their characteristics and, in particular, their geometric layout. All the intersections that were not included in the types considered by the accident prediction models to be analysed were eliminated. Moreover, a detailed study was conducted to eliminate all intersections that were modified (geometric layout, regulation, etc..)from the first year of the accident database. Finally, 70 intersections were considered for the following analysis. Subsequently a second reconnaissance survey was undertaken at all intersections collecting the data included in the check list shown in figure 1. This check list was developed on the basis of the geometric variables used by the accident prediction models to be analysed.

1. Intersezione	Strada principale		Starda secondaria	
2. Controllo intersezione	Non controllata ○	Precedenza ○	Stop ○	Semaforo ○
3. Numero corsie strada principale	....			
4. Numero corsie strada trasversale	....			
5. Presenza spartitraffico	Assente ○		Presente ○	
6. Trattamento svolta a sinistra- strada principale	No svolta a sinistra ○	Senza corsia esclusiva per svolta ○		Con corsia esclusiva per svolta ○
7. Tipologia corsia di svolta a sinistra- strada principale	Disegnata ○		Protetta da spartitraffico ○	
8. Trattamento svolta a sinistra- strada secondaria	No svolta a sinistra ○	Senza corsia esclusiva per svolta ○		Con corsia esclusiva per svolta ○
9. Tipologia corsia di svolta a sinistra- strada secondaria	Disegnata ○		Protetta da spartitraffico ○	
10. Trattamento svolta a destra- strada principale	No svolta a destra ○	Senza corsia esclusiva per svolta ○		Con corsia esclusiva per svolta ○
11. Tipologia svolta a destra- strada principale	Non provvisto per svolta libera ○		Provvisto per svolta libera ○	
12. Trattamento svolta a destra- strada secondaria	No svolta a destra ○	Senza corsia esclusiva per svolta ○		Con corsia esclusiva per svolta ○
13. Tipologia svolta a destra- strada secondaria	Non provvisto per svolta libera ○		Provvisto per svolta libera ○	
14. Senso di marcia strada principale	Senso unico ○		Doppio senso ○	
15. Senso di marcia strada secondaria	Senso unico ○		Doppio senso ○	
16. Semaforo	A tempi fissi ○	Semiattuato ○		Attuato ○
17. Fasi semaforiche	2 ○		Multi fase ○	
18. Segnaletica stradale	Presente ○		Assente ○	
19. Pendenza longitudinale strada principale	Piana ○	Leggera pendenza ○		Forte pendenza ○
20. Pendenza longitudinale strada secondaria	Piana ○	Leggera pendenza ○		Forte pendenza ○
21. Larghezza media corsia strada principale	....			
22. Larghezza medi corsia strada trasversale	....			
23. Larghezza marciapiede laterale strada principale	....			
24. Larghezza marciapiede laterale strada secondaria	....			
25. Illuminazione	Assente ○		Presente ○	
26. Accesso controllato sulla strada principale	Nessuno ○		Parziale ○	

Figure 1 – Field survey check list

Table 5 summarizes traffic, accident and control type data of the intersections used for this study. The type “No control” includes the intersections that do not present any horizontal or vertical sign; therefore, the right way depends only on the road code rules. It should be noted that for three types there are only two or three intersections; this limitation must be considered in the following transferability evaluation.

MODEL PERFORMANCE MEASURES

Several goodness-of-fit (GOF) statistics to assess model fit to Trieste data were used. It is important to note that only after an assessment of several GOF criteria has been made, can the performance of a particular model or set of models be assessed. Moreover, it should be noted that the evaluation of the GOFs is subjective, therefore the evaluation of transferability of the models in this study also remains a subjective opinion of the authors. The GOF measures used were:

Pearson’s Product Moment Correlation Coefficients Between Observed and Predicted Accident Frequencies

Pearson’s product moment correlation coefficient, usually denoted by r, reflects the degree of linear relationship between two variables Y_p and Y_o. Pearson’s product moment correlation coefficient is given as:

$$r = \frac{\sum_{i=1}^{n}(Y_{ip} - \overline{Y_p}) \cdot (Y_{io} - \overline{Y_o})}{\left[\sum_{i=1}^{n}(Y_{ip} - \overline{Y_p})^2 \cdot \sum_{i=1}^{n}(Y_{io} - \overline{Y_o})^2\right]^{1/2}}$$

Where

$\overline{Y}_p$ the mean of the Y_{ip} predicted accidents

$\overline{Y}_o$ the mean of the Y_{io} observed accidents

Y_{ip} predicted accidents at *i* intersection

Y_{io} observed accidents at *i* intersection

n validation intersection sample size

Pearson's product moment correlation coefficient ranges from +1 to -1. A correlation of ±1 means that there is a perfect positive/negative linear relationship between predicted (Y_p) and observed (Y_o) values. A correlation of 0 means there is no linear relationship between values. Therefore a low coefficient suggests that the model is not able to predict the observed data well.

Table 5 – Characteristics of Intersection types

Type	Number	Accident Severity	accident data per year			ADT Major Road		ADT Minor Road	
			Min	Max	Mean	Min	Max	Min	Max
3 leg									
No control	12	Total	0.25	2.25	10.08	1428	8415	193	3777
		Injury	0.08	0.75	3.92				
Give way	10	Total	0.33	7.75	23.75	2322	22378	77	9830
		Injury	0.08	3.50	9.08				
Stop	3	Total	0.67	2.58	4.17	5235	12615	766	3468
		Injury	0.17	1.42	2.00				
Signal	8	Total	0.25	5.42	24.00	1855	23854	1498	11099
		Injury	0.25	1.42	6.75				
4 leg									
No control	2	Total	1.08	7.50	8.58	3171	5133	2806	3089
		Injury	0.58	4.00	4.58				
Give way	2	Total	8.25	11.92	20.17	3874	10391	882	1792
		Injury	4.42	7.00	11.42				
Stop	8	Total	1.25	9.50	42.25	2493	8732	573	2102
		Injury	0.58	5.00	22.33				
Signal	25	Total	1.42	12.42	118.7	2460	24521	604	12355
		Injury	0.67	6.75	60.42				

Mean Prediction Bias (MPB)
The MPB is the sum of predicted values minus observed values in the validation data set, divided by the number of validation sample size. The MPB can be positive or negative, and is given by:

$$MPB = \frac{\sum_{i=1}^{n} (Y_{io} - Y_{ip})}{n}$$

This statistic provides a measure of the magnitude and direction of the average model bias as compared to observed data. The smaller the MPB, the better the model is at predicting observed validation data. A positive MPB suggests that on average the model overpredicts the observed data. Conversely, a negative value suggests systematic underprediction. The magnitude of MPB provides the magnitude of the average bias.

Mean Absolute Deviation (MAD)
MAD is the sum of the absolute value of the differences between the predicted values minus observed values, divided by the number of validation sample size. The MAD can only be positive and is given by:

$$MAD = \frac{\sum_{i=1}^{n} \left| Y_{io} - Y_{ip} \right|}{n}$$

The MAD gives a measure of the average magnitude of variability of prediction. A large MAD suggests that the model is not able to predict the observed data well.

Mean Absolute Percentage Error (MAPE)
MAPE is the sum of the absolute value of the differences between the predicted values and observed values divided by the observed values again and by the number of validation sample size. MAPE is a measure of accuracy and it is expressed as a percentage.

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{\left| Y_{ip} - Y_{io} \right|}{Y_{io}} \cdot 100$$

The magnitude of MAPE provides the magnitude of the average error in respect of the observed value. A large MAPE suggests that the model is not able to predict the observed data well. It should be noted that the observed accidents in this study are calculated as an average per year and in no case is Y_{io} equal to zero.

Washington et al. (Washington S. et al, 2005) also used the comparison between the Mean Squared Prediction Error (MSPE) and the Mean Squared Error (MSE) to reveal potential overfitting or underfitting of the models to estimation data. However, this comparison was not possible in the present study because the MSE of all the models evaluated were not available.

To normalize the GOF statistics to compensate for the different number of years associated with different data sets, GOF statistics were computed on a yearly basis.

EVALUATION OF TRANSFERABILITY

Depending on the number of legs and the type of control, the available intersections from table 5 were selected for each model which had to be evaluated. In particular, the intersections with both the controls "give-way" and "stop" and with at least one one-way leg were considered for the priority intersection models developed by Summergill et al., whereas the intersections with all controls except "signal" were considered for the not signalized intersection model developed by Greibe.

The accident database was used to calculate the Y_{io} of each intersection, whereas the accident prediction models were used to calculate the Y_{ip} of each intersection. Therefore, the average number of predicted accidents, the average number of observed accidents and the GOF statistics were calculated for each model evaluated. Table 6 shows the numbers of intersections used for each model evaluated and the corresponding statistics on a yearly basis.

Table 6 – Transferability evaluation results (one-year basis)

Model	Number of intersections	Average accidents predicted	Average accidents observed	*MAPE*	*r*	*MPB*	*MAD*
Bauer and Harwood – all accidents							
3-leg stop	3	3.89	1.39	247.0	0.36	2.50	2.50
4-leg stop	8	3.04	5.28	57.62	-0.03	-2.24	2.88
4-leg signalized	25	4.70	4.75	42.91	0.39	-0.04	1.50
Bauer and Harwood – injury accidents							
3-leg stop	3	1.65	0.67	234.1	0.54	0.98	0.98
4-leg stop	8	0.93	2.79	63.93	0.11	-1.86	1.94
4-leg signalized	25	1.37	2.42	38.80	0.35	-1.05	1.16
Canale et al. – injury accidents							
3-leg no control	12	0.15	0.32	65.22	-0.01	-0.17	0.25
3-leg stop	3	0.32	0.67	40.09	0.61	-0.35	0.35
4-leg no control	2	0.76	2.29	68.50	-0.60	-1.53	1.84
4-leg stop	8	0.62	2.79	69.56	-0.05	-2.18	2.20
4-leg signalized	25	1.47	2.42	107.5	-0.03	-0.95	2.28
Summergill et al. – injury accidents							
3 leg priority	8	0.46	2.24	70.07	0.65	-1.78	1.78
4 leg priority	10	0.53	6.24	88.57	0.12	-5.71	5.71
3 leg signalized	6	1.39	0.85	124.6	0.01	0.55	0.55
4 leg signalized	24	1.32	2.24	39.98	0.25	-0.92	1.04
Greibe – all accidents							
3 leg not signalized	25	0.40	1.52	69.00	0.34	-1.12	1.14
3 leg signalized	8	0.94	2.38	102.2	0.03	-1.43	1.67
4 leg not signalized	12	0.47	5.92	86.62	0.08	-5.45	5.45
4 leg signalized	25	1.17	4.75	71.04	0.38	-3.57	3.57

Table 6 shows significant differences between the average accidents observed and the average accidents predicted by the models evaluated. Not considering the four cases with few intersections (less than 6), in five cases out of fifteen the average accidents predicted are more than three times lower than the average accidents observed. The MAPE, that gives a measure of the accuracy of the prediction models evaluated, shows values that achieve 247 percent, with a minimum value equal to 40 percent. Three cases have MAPE>100% whereas twelve cases have MAPE>50%. Pearson's Product Moment Correlation Coefficients *r* has low values, with some exceptions only for models evaluated using few intersections. The MPB is negative, with only a few exceptions, showing that the models evaluated underpredict the observed accident frequencies. Considering that the accident database does not contain all the damage only accidents, it is significant to note that the models that predict total accidents (Bauer and Harwood, Greibe) with only one exception, also underpredict the observed accident frequencies. The MAD, that gives a measure of the average magnitude of variability of prediction, shows significant values if compared with the total number of accidents observed.

Figure 2 shows an example of the observed versus predicted accident frequency for 4-leg signalized intersections. It is evident from the figure that the models evaluated do not fit the Trieste data well.

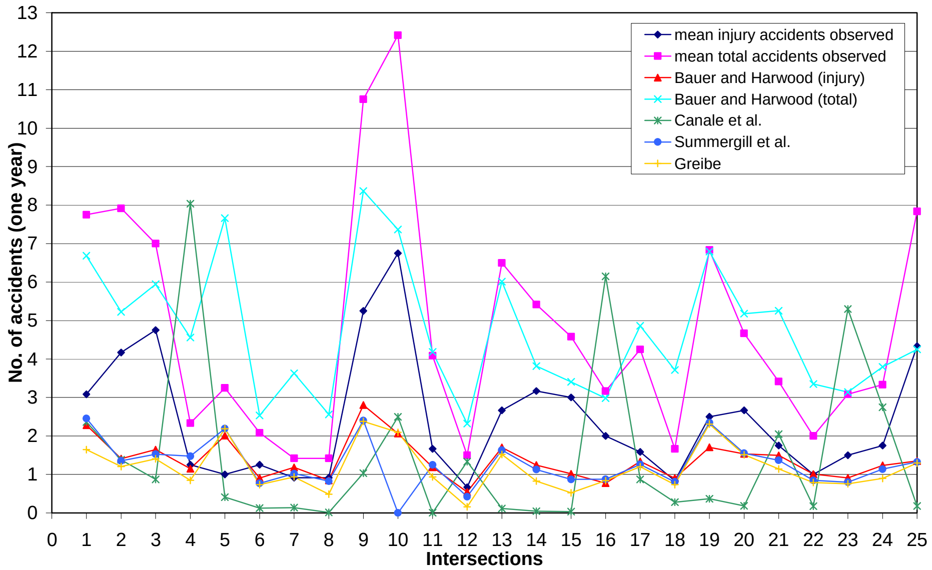

Figure 2 – Observed versus predicted accident frequency (4-leg signalized)

In all, the analysis of the statistics in table 6 and of the graphics of the observed versus predicted accident frequencies, of figure 2, makes it possible to state that all the models evaluated are not able to predict accident frequency on the intersections of Trieste with sufficient reliability.

It is significant to note that none of the four sets of models fits the Trieste data. Therefore, there is no geographic area for which the accident data can be considered similar to the Trieste accident data, not even the models of Canale et al. developed in another Italian city.

CONCLUSIONS

The result of the transferability evaluation of accident prediction models shows that the models tested are not able to predict the accidents observed on the intersections of the Trieste urban area. This result makes clear that the use of existing accident prediction models for intersections not included in the same geographic area of those used to calibrate the model requires careful attention. This is true not only if the model was developed in foreign countries, but also if it was developed in Italy. Therefore, it is necessary to adjust the accident prediction of the existing model to suit the safety conditions present in the specific context. This adjustment could be made estimating a calibration factor or using the Empirical Bayes approach that consider the accident history data of the specific site. Therefore, both these procedures will require special attention on the part of Italian researchers in this field in the near future.

REFERENCES

1. Bauer K.M., D.W. Harwood (2000). *Statistical Model of At-Grade Intersection Accidents – Addendum*. Report FHWA-RD-99-094, U.S. Department of Transportation.
2. Canale S., S. Leonardi, G. Pappalardo (2005). The reliability of the urban road network: Accident forecast models, *Proceedings of IIIrd International SIIV Congress*, Bari, Italy.
3. Greibe P. (2003). Accident Prediction models for Urban Roads, *Accident Analysis and Prevention*, n°35, p. 273-285.
4. Harwood D.W., F.M. Council, E. Hauer, W.E. Hughes, and A. Vogt (2000). *Prediction of the expected safety performance of rural two-lane highways*. Report FHWA-RD-99-207, U.S. Department of Transportation.
5. ISTAT (2005). *Statistica degli incidenti stradali – anni 2003 2004 – edizione provvisoria*. Istituto Nazionale di Statistica, Rome, Italy.
6. Reurings M., T. Janssen, R. Eenink, R. Elvik, J. Cardoso, C. Stefan (2005). *Accident Prediction Models and Road safety Impact Assessment: a state-of-the-art*. Project Ripcord Iserest – 6^th framework Programme EU.
7. *Rilievi di Traffico e Monitoraggio delle Condizioni di Circolazione - Comune di Trieste* (2002). Department of Civil and Environmental Engineering, University of Trieste, Italy.
8. Summergill I., J.V. Kennedy, R.D. Hall, A. Hickford, S.R. Barnard (2001). *Accidents at Junctions on one-way urban roads*. TRL Report 510, UK.
9. Washington S., B. Persaud, C. Lyon, J. Oh (2005). *Validation of Accident Models for Intersections*. Report FHWA-RD-03-037, U.S. Department of Transportation.