SPATIAL NESTED SCALES FOR ROAD ACCIDENTS IN THE PERIPHERY OF BRUSSELS

Nathalie ECKHARDT
Department of Geography, U.C.L
Louvain-la-Neuve, Belgium

Isabelle THOMAS
Department of Geography, U.C.L
Louvain-la-Neuve, Belgium
National Fund for Scientific Research
Brussels, Belgium
Centre of Operation Research and Econometrics
Louvain-la-Neuve, Belgium

(Received October 12, 2004)

This paper tests the usefulness of a multilevel model (MLM) for explaining the spatial occurrence of road accidents; it also shows and confirms how far the characteristics of the geographical environment influence the location and concentration of road accidents at two levels of spatial aggregation. The results are compared to those obtained from a more classical logistic regression. The analysis is performed on the southern periphery of Brussels (Belgium). The main conclusions are: (1) that MLM is a potentially useful technique for modelling road accidents, but that hierarchical levels are not easy to define for spatial data and so MLM are less useful than other regression techniques for modelling spatial occurrences of road accidents; (2) that the characteristics of the environment and the road itself significantly influence the occurrence of road accidents, and changes in these characteristics are quite important elements in the explanation, leading to the suggestion that road users do not adapt their behaviour sufficiently to changes in road conditions. Hence, concentrations of road accidents often correspond to places where improvements could be made in terms of road design, signalling and land-use planning.

Key Words: Multilevel model, Geography, Environment, Accidents, Brussels

1. INTRODUCTION

The main research objective of our team is to analyse the spatial aspects of road-accident occurrence at several scales of analysis[1-4]. The present contribution aims at modelling the spatial occurrence of road accidents by considering different nested levels of spatial data aggregation.

Multilevel analysis is a recent technique that enables the relationships/interactions between variables at several levels of data aggregation to be examined simultaneously[5,6]. It attempts to solve the dilemma of the spatial scale, that is to say the use of scales in analysing the characteristics of accidents at a local level, taking into account the broader spatial context in which these accidents occur. Up to now, applications have mainly been limited to the social and behavioural sciences. Let us mention here the well-known example of school test results, which can be explained by the individual characteristics of the scholars, but also by the characteristics of the class (group) as well as the school or even its environment[7]. Multilevel models are specifically dedicated to the concept of the integration of contextual effects and so to hierarchical models[6,8,9]. In spatial analysis, multilevel models enable the researcher to go beyond the scale defined *a priori* by territorial executives (for example), and to attempt to capture the continuous character of space, taking into account the nested nature of spatial scales[6,10].

In the field of road safety, two recent papers have used multilevel modelling. Jones and Jorgensen[11] consider road accident casualties, and show that the risk of fatality is associated with casualty age and sex, as well as the type of vehicles involved, characteristics of the impact, attributes of the road section on which it took place, time of day, and whether alcohol was suspected. The multilevel analysis shows that 16% of the unexplained variation in casualty outcomes was between accidents, whilst approximately 1% was associated with the area in which each incident occurred. Gee and Takeuchi[12] analyse the cross-sectional relationship between traffic stress and neighbourhood conditions, depression and health status by means of multilevel analyses. They show that perceived traffic stress is associated with both general health status and depression in multilevel models, with people reporting traffic stress having lower health status and more depressive symptoms.

In this paper, we aim at showing the utility of multilevel analysis for understanding the spatial aspects of road safety, and more particularly to show how far the characteristics of space (environment and infrastructure) can influence the location of accidents at different levels of measurements in a periurban environment. By doing so we expect to partly solve some data aggregation problems, well-known in spatial analysis[13,14] as well as in road accident analysis[15,16]. The analysis is here conducted on data from the suburbs of Brussels (southern periphery); the modelling results are interpreted in terms of operational results and are also compared to those obtained by means of a more "classical" logistic regression[3].

This paper is organised as follows. The model is described in Section 2 and choices related to the data are discussed in Section 3. Empirical results are reported in Section 4. Two levels of aggregation are considered: the hectometre (100m) of road and the commune; we aim at understanding why accidents are concentrated in some hectometres and why these hectometres are located in specific communal environments. Our conclusions and discussion are to be found in Section 5.

2. MULTILEVEL ANALYSIS

Let us first justify our modelling choices by briefly defining the type of model (Section 2.1), its advantages in terms of accident modelling (Section 2.2) as well as the methods used for estimating the parameters (Section 2.3). We refer to the literature for further model definition and formulation[6,7,9,17,18].

2.1 Definition

Multilevel modelling (MLM) is a type of regression that has mainly been developed since the 1980s; it is designed to handle hierarchical and clustered data. Such data involve group effects on individuals, which may not be assessed validly by traditional statistical techniques. That is, when grouping is present, observations within a group are often more similar than would be predicted on a pooled-data basis, and hence the assumption of independence of observations is violated. MLM uses variables at several levels of aggregation to adjust the regression of the base-level dependent variables on the base-level independent variables. In our case, for instance, we could predict accident occurrence from the characteristics of the hectometres or from larger environments such as communes. MLM is related to structural equation modelling in that it fits regression equations to the data, then tests alternative models using a likelihood ratio test.

MLM specifies the expected direct effects of variables on each other within any one level, and cross-level interaction effects between variables located at different levels. Hence, mediating mechanisms are postulated; they allow variables at one level to influence variables at another level (e.g. better road infrastructure may influence road-user behaviour and hence prevent accidents at places other than that where the road enhancement has been effected). MLM tests multilevel theories statistically, by simultaneously modelling variables at different levels without having recourse to aggregation or disaggregation. In our case, the global problem is to model the relationship between the place of an accident and the context in which it occurs. We hence aim to detect the amount of context contribution and its effect on the total variation of the "individual behaviour", and to identify which macro characteristics are responsible for the context effect. This means conceptually introducing a multilevel approach in which road accidents are grouped together at different spatial levels; in our case, variables from two levels will be jointly analysed in a unified framework.

The multilevel model is therefore a model with a single dependent variable (Y) measured at the base level (e.g. the accidents are taken individually). As in ordinary least squares (OLS) regression, there may be one or more independent variables collected at the base level. In addition, there will be at least one broader level of aggregation, with at least one explanatory variable (for instance, the environment or the socio-economic characteristics of the ward). In an OLS model, base-level data are analysed for all groups pooled together (e.g. all hectometres, all communes). In a MLM, the regression is performed separately for each group. This produces different regression coefficients and different intercepts (e.g. for each black zone or each commune) and also explains why MLMs are called "random coefficient models". Such models usually use maximum likelihood algorithms to estimate the parameters (coefficients) (see http://www2.chass.ncsu.edu/garson/pa765/multilevel.htm).

Let us briefly summarise the formulation. Y_{ij} is the variable to be explained. We consider here a model at two levels of observation (see Section 2.2): the hectometre i (Level 1) localised in a municipality j (Level 2). The general shape of the linear model includes an explanatory variable at Level 1 related to the hectometre (X_i), and a contextual variable at Level 2 related to the commune (municipality) and denoted Z_j. Then

$$\mathrm{Y}_{ij} = \beta_0+\beta_i X_i+\Gamma_j Z_j\ (\mu_{0j}+\mu_{ij}*X_{ij}+\varepsilon_{ij})$$

where $\beta_0+\beta_i X_i+\Gamma_j Z_j$ is the *fixed part* of the model and $(\mu_{0j}+\mu_{ij}*X_{ij}+\varepsilon_{ij})$ is the *random part*. β_0 is the intercept. β_i is the angular coefficient of the right-hand side of the regression; it is the coefficient of the explanatory variables at Level 1 (X_i); Γ_j is the coefficient of the Level 2 explanatory variable (Z_j). *Error terms* (residuals) are associated with β_0 and β_i at the contextual level (μ_{0j} and μ_{ij}); they represent the deviation of the municipality j from the average coefficient. These contextual residuals are assumed to follow a normal distribution law with null averages, variances $\sigma_{0\mu}{}^2$ and $\sigma_{i\mu}{}^2$ and covariance $\sigma_{0i\mu}$. Level 1 residuals (ε_{ij}) have a null-average and equal $\sigma_{0\varepsilon}{}^2$ variances. Hence, this formulation allows every municipality j to have its own constant (β_0) and angular coefficient (β_i). This heterogeneity of the regression coefficients between municipalities can later be tested and explained by Level 1 and Level 2 variables. In other words, the *outputs* of a multilevel model include: (1) a fixed part containing the regression coefficients and the corresponding p-value for the significance levels, at Level 1, Level 2, and for the cross-level interactions; (2) a random part containing regression coefficients and p-values for estimating the variances of Level 1 variables and intercepts; and (3) standardised (beta) coefficients which, as in OLS regression, allow us to compare the relative importance of the independent variables and interactions. Most software packages also produce a measure of deviance.

2.2 Advantages of spatial analysis

Two types of errors are avoided when different levels of aggregation are considered simultaneously[6,19-21]: the *ecological error* which consists of using a global aggregated statistical measure to reveal individual behaviour; and the *atomist error*, which considers the characteristics of the individual but ignores the context in which the human behaviour occurs. This is also true for road accidents: indeed it seems fallacious to isolate the accident from its environment, or the society in which it occurred. The purpose of this paper is to determine the direct effect of the explanatory variables measured at a low level of aggregation and at a higher level of aggregation, and to see if the explanatory variables at the aggregated level moderate the relationships occurring at the individual level, or vice versa.

Levels of analysis are often hierarchically organised: items at one level interact and create a higher homogeneity[8]. This hierarchical structure leads to a correlation of the observations that violates the hypothesis of the independence of residuals (classical regression techniques) and leads to an underestimation of the standard deviations of the regression coefficients. MLMs take this correlation into account in the estimation of the standard deviations of the regression coefficients by including terms of error at the contextual level. With regard to our study, this dependence is spatial and means that the observations supply less information than if they were randomly distributed, as is assumed in the OLS method[22].

Additional advantages of MLM are: (1) regression coefficients are specific to each level of analysis thanks to the contextual residuals; (2) it is possible to test whether the variance in the contextual terms of error is significantly different from 0 (likelihood tests); (3) coefficients of determination (R^2) can be computed for each level of analysis by comparing the residual variances with the variances of an "empty model" (without explanatory variable) for each level; and (4) it is possible to attribute the residual variance of the classical multiple regression to various levels of analysis[6,10]. The assumptions for the application of MLMs are in general less restrictive than for more traditional regression techniques[6,7,9,10]; the main difficulty with MLMs is the definition of the hierarchical levels of observation and the associated variables (see Section 3.2).

2.3 Estimating the parameters

Homoscedasticity (equal variances) is often violated in hierarchical situations: OLS techniques are hence inappropriate. Multilevel linear models are often best estimated by the Newton-Raphson algorithm that is based on maximum likelihood and generalised least squares. The restricted maximum likelihood method is often used. The values of the regression coefficients are first computed on the basis of the first analysis of the matrix of variance-covariance. The matrix is then re-estimated, using the first values of the coefficients. Finally, the estimation of the coefficients is improved by the new variance-covariance matrix until convergence. In this way, the fixed and random parts of the model are more effectively estimated than with OLS.

Several software packages, such as *HLM* and *MlwiN*, are available[23,24]. SAS was used (*Proc Mixed, Proc Nlmixed*) for the research reported in this paper; this package enables the more classic regression models to be considered as well as the hierarchical formulations. Several models can be implemented with *Proc Mixed*: simple random-effects only, simple mixed with a single fixed and random effect, split-plot, multilocation, repeated measures, analysis of covariance, random coefficients, and spatial correlation.

3. METHODOLOGICAL CHOICES

3.1 Road safety in the studied area

Brussels is the capital city of Belgium, located in the centre of the country and containing approximately 1 million inhabitants. As in most urban areas, the city sprawls far beyond its administrative boundaries. Walloon Brabant corresponds to an administrative entity (province) located in the south of the city; it is mainly peri-urban but its landscape results from its historical evolution. Up to the 18th century, it was mainly rural, with many small villages. During the 19th century, industries located in the west (ironworks) and centre (paper mills). Railways and better roads later increased the accessibility. In the first half of the 20th century, industries started closing one after the other. In the sixties, the region was increasingly polarised by Brussels: as in many European cities, people started to move from the centre of Brussels to the countryside, while keeping their jobs in the city. Later on a university was created in Louvain-la-Neuve and several industrial parks were planned all over the area. Hence, the southern periphery of Brussels is now characterised by old villages and small towns, a new town and many allotments, old industrial locations as well as new planned ones (industrial parks), highly urbanised communes close to Brussels as well as more residential areas, woods and agriculture as well as employment and commercial centres. The result is a mosaic of landscapes, a polynuclear structure and quite an interesting spatial pattern[25,26]. Our studied area is hence by definition periurban; its particular structure makes it quite interesting as links between urban sprawl, travel practises and road safety have been little dealt within the literature[27].

In Belgium, any road accident that occurs on a public road and that involves casualties must be officially reported. Location is accurately known on numbered roads: there is a stone marker every hectometre (100 meters); numbered roads are motorways, national and provincial roads linking towns together. On other roads, location is identified by postal addresses that are often less accurate. This analysis is limited to accidents with casualties on numbered roads; the hectometre is the smallest spatial unit for which accident data are spatially and officially available.

Walloon Brabant has 460.4km of numbered roads including 37.7km of motorway. In this paper, a black zone is defined by means of local spatial autocorrelation indices[2,28,29] as a set of contiguous hectometres with a high number of accidents. Black zones vary in length and intensity, and some black zones may include a hectometre with no or very few accidents. These black zones charcterise the same places from year to year[30], despite the many socio-economic changes in this studied area. 47% of the road hectometres did not register any accidents. Black zones represent 38% of the total number of accidents but only 12% of the total number of hectometres. In this paper, the period under study is 1998-2000, a period long enough to minimise random fluctuations, but short enough to limit changes in road traffic conditions. Some 2,363 accidents were registered on numbered roads during this period; in total, 1,388 hectometres out of 4,604 experienced at least one accident with casualties.

3.2 Levels of analysis and dependent variables

The definition of the hierarchical levels is not as easy as for some topics in human sciences. It is a compromise between the significance and the availability of the data.

Ideally, accidents should be considered as the first level of analysis. However, accident data are collected by hectometre (no GPS was available). Moreover, most environmental variables are also only available for hectometres or segments of roads rather than for pinpoint locations on the road. Let us here add that some environmental characteristics are collected at the time and place of an accident and are reported on the statistical form filled in for each accident with casualties. These could have been used for characterising the places at which accidents took place, but their quality is doubtful[31] and we do not have any comparable information about places where no accidents occurred. Hence the lowest aggregated level of analysis used here is the hectometre.

Initially, we decided to take the road segment (several hectometres long) as the second level of analysis. However, it was not easy to choose the "best" length for this segment and to justify it. Thresholds would have been quite artificial and such segmentation does not correspond to any official definition; explanatory data are not available at this level, and they would have had to be a combination of data available at the hectometre level. Hence, the municipality j (commune) was chosen as the second level of analysis. It is not directly related to the road network itself, but to its global environment. A lot of data are available at this level of administrative aggregation, which corresponds quite well to mobility patterns in Belgium.

Three dependent variables are used and modelled separately: $Y1$ takes the value 1 when a hectometre belongs to a black zone and 0 otherwise. It enables us to understand why some hectometres are more dangerous than others. *Y2* is the total number of accidents observed in each hectometre of road, and *Y3* is a measure of the risk of accident, roughly estimated as the total number of accidents divided by the average daily traffic intensity. *Y2* has the advantage of giving the real number of accidents: some hectometres can belong to a black zone (*Y1*) without any accidents being recording[2], and, on the contrary, a hectometre with several accidents can be surrounded by hectometres with no accidents and hence not belong to a black zone. The absolute number of accidents (*Y2*) is interesting for some public authorities (such as the emergency services), while for others, the relative number (*Y3*) is important. For each *Y* variable, the situation on motorways is modelled separately because traffic has a different structure (two separate lanes, etc.) on motorways.

3.3 Explanatory variables

The purpose here is to identify environmental conditions associated with road safety/danger at two nested levels of analysis. Most explanatory variables are selected from official data bases (Belgian National Institute of Statistics, Ministry of Equipment and Transport) or constructed by means of G.I.S. techniques from official IGN maps[32]; some of these variables are identical with (or close to) those used by Flahaut[3]. They are briefly described below.

Let us remind that our objective is not here to get an exhaustive list of all the potential explanatory variables. Our choice was guided by former published studies and constrained by data availability as well as the structure of the studied area. We are conscious that this choice biases the results of the analysis and limits the conclusions. However, most results presented hereunder are consistent with former published results. There is indeed a large number of papers in the literature justifying the choice of roadway geometrics or the characteristics of the adjacent environment of roads for explaining road accidents occurrences. Roadway geometrics include variables such as the road type (functional or physical), the number of lanes and their width, the horizontal curvature and the vertical grade[33-36]. Spatial environment can be apprehended in various ways : land use[37], roadside features (guardrail, bridges, etc.)[38], residential development[27,39], human activities[39,40], social disparities[41,42].

Most explanatory variables used at Level 1 describe the environment of the hectometre itself: the physical characteristics of the road, land use or natural environment of the hectometre (see Table 1). Changes in environmental conditions are also considered by means of so-called “transition variables” which pinpoint road and environmental discontinuities that could influence drivers’ ability (or inability) to cope with changes. These variables are referenced by the absolute distance to the spatial discontinuity and are recorded as *_dist*. If the change is within the hectometre, the distance is recorded as 0 metres. A maximum of 200m is considered on numbered roads and 300m on motorways. Information on traffic, speed limits and the physical characteristics of the roads (type of road, type of surface, adherence of the road, presence of rutting or obstacles) were made available by the Ministry of Equipment and Transport (MET). Traffic density is measured by the 1999 average daily number of vehicles in both directions, including all types of vehicle, between 6 a.m. and 10 p.m. Traffic density is measured for well-defined hectometres and extrapolated to larger road segments, but is never known at the time and place of the accident. It is a daily average. As in many other countries, traffic is macroscopically measured[43]. Land use variables refer to human activities or to measurable characteristics of the physical environment (e.g. % of area built-up, level of afforestation along the road at a distance of 50m). These are constructed on the basis of 1:50,000 digital maps provided by the National Geographic Institute. Finally, the orientation, bends and slopes of the roads are obtained from measurements made on digitalised topographic maps by means of appropriate GIS techniques. These characteristics of the roads can potentially influence the visibility and the behaviour of road users. Most variables are (0,1) variables.

At level 2, explanatory variables mainly pertain to the socio-economic characteristics of the population living, working or shopping in the commune (Census data)[41]. Table 2 contains the definitions of the variables used in this paper. As in most human geography studies, it is difficult to estimate the population at risk who is really present in the municipality or traversing it. We limited ourselves to proxies that are officially available. Structural variables were also taken into account: the level of urbanisation[44] and the rurality (as expressed by the ratio of agricultural areas to the total land area of a municipality). We also expected that the morphology of the built-up environment could influence accident occurrence; morphology was estimated by fractal dimension[25,26]. Let us remind that the fractal dimension (D) describes the extent to which a mass (here the built-up area) is concentrated within a zone. Fractal dimension is not equal to density[26]. Thus for spatial mass distributions, *D* can be interpreted as a measure of mass concentration in a given area. It can be shown that a value of *D* close

Table 1 Explanatory variables at Level 1

	Variables	Description
Road use	TRAFFIC	Logarithm of the average daily volume of traffic (Source: MET)
	VMAX	Maximum speed limit (Source: MET)
	VMAX_*dist*	Distance to a change in speed limit (in meters): >200; 200; 100; 0 m and 300 m on motorways. Written as VMAX(>200)_*dist* etc.
Physical characteristics of the road	LANES	Number of lanes: (1) 1+1; (2) 2+2; (3) 2; (4) 3
	LANES_*dist*	Distance to a change in road type in terms of number of lanes: >200; 200; 100; 0 m and 300 m on motorways. Written as LANES(100m)_*dist* etc.
	SURFACE	Type of surfacing (1) BE: concrete; (2) HFN: conventional asphalt; (3) HOM: thin asphalt; (4) HON: draining asphalt
	SURFACE_*dist*	Distance to a change in surfacing: >200; 200; 100; 0 m for numbered roads and 300 m on motorways
	RUT	Presence (1) / absence (0) of ruts
	JUNCT_*dist*	Distance to major crossroad: >200; 200; 100; 0 m on numbered roads and >300 ; 300 m on motorways
	ADHERENCE	Adherence of the road surface: (0) good or normal; (1) bad or very bad
	PROX_ACCES	Proximity (<300 m) to entry/exit. (0,1) On motorways only
	BZONE	Hectometre belonging (1) or not (0) to a black zone
Land-use	BUILT	Estimated % of built-up area (from the road to 50 m from it): <20; 20; 25; 30; 40; 50. Written as Built30 etc.
	BUILT_*dist*	Distance to transition in density of built-up area (BUILT): >200; 200; 100; 0 m (>300; 300 m on motorways)
	WOODS	Estimated % of wooded area along the road (at 50 m): < 20; 20; 25; 30; 40; 50. Written as Woods30 etc.
	FIRMS	Proximity of firms or large supermarkets (50 m) (0,1)
	OBSTA _*dist*	Distance to an obstacle such as a bridge pillar: >200; 200; 100; 0 m (>300; 300 m on motorways)
Landscape	AGGLO	Inside/outside an urban agglomeration (F1/F3 road sign)
	DIRECTION	Road segment direction: (1) Others (N-S); (2) E-W≤22.5°; (3) 22.5°<E-W45°
	RELIEF	At the top of a slope; at 100 m from the top; at the bottom of a slope; at 100m from the bottom; other

Table 2 Explanatory variables at Level 2

Variables	Description
MIXITY	Number of jobs/number of inhabitants in a commune in 2001 (= level of mixing of the activities and hence traffic).
ATTRACT	(Numbers of jobs + number of inhabitants)/surface of the commune in 2001 (= attractivity of the commune)
EMPLOYDENS	(Working population residing in a commune + population working in the commune) / total surface of the commune
MOBILITY	Working population residing in commune *i* / Total resident population in i
ROADLENGTH	Total number of hm of roads in the commune (logarithm)
URBE	Level of urbanisation
DCORR	Fractal dimension of the built-up area obtained by correlation
RURALITE	Area devoted to agriculture in 2001 / total area (%)

to 2.0 describes a fairly homogeneous distribution. The lower the value of D, the more the mass is concentrated: thus a dimension close to 0.0 corresponds to a concentration of the mass in one isolated point (one farm in the middle of fields), while the value 1.0 corresponds to a line (houses along a road), but also characterises a hierarchical spatial distribution of masses. No value smaller that 1.0 was obtained in this study. Such values would refer to structures composed of a disconnected set of points. It is possible that road safety is affected by whether built-up areas are distributed homogeneously or heterogeneously throughout the neighbourhood. D is here measured by correlation[25] and denoted DCORR.

4. MODELLING RESULTS

Section 4.1 and Table 3 contain the results for accidents occurring on numbered roads (regional roads) with the exception of motorways, which are reported in Section 4.2 and Table 4. The three dependant variables are analysed separately. The exploratory data analysis conducted on the explanatory variables is not reported here; it was mainly based on correlation coefficients and odd ratios, and enabled the number of explanatory variables to be reduced and their "best" formulation to be selected.

4.1 Accidents on numbered roads

Proc Nlmixed uses a non-linear formulation of the model. It requires initial values for the parameters to be specified, so that the model converges. Convergence is due to the fact that maximum likelihood estimation is an iterative algorithm which may require many runs before reaching stable coefficient estimates. Our first estimate was β_0 = -2.05; this initial value was introduced into the "empty" model (Table 3, Column 2). The intra-municipal variance (Level 1) is an indicator of the variability between hectometres, whereas the inter-municipalities variance (Level 2) concerns the variability between municipalities. The total residual variance of the empty model was 7.77. The intra-class correlation $\rho_\mu = \frac{\sigma_{\mu 0}{}^2}{\sigma_{\mu 0}{}^2 + \sigma_{\varepsilon 0}{}^2}$ indicates the proportion of the variance due to the communes (Level 2); ρ here is equal to 24.95% ($\sigma_{0\mu}{}^2$ = 1.94). For the hectometres (Level 1) $\rho_\varepsilon = \frac{\sigma_{\varepsilon 0}{}^2}{\sigma_{\mu 0}{}^2 + \sigma_{\varepsilon 0}{}^2}$ = 75.05% ($\sigma_{0\varepsilon}{}^2$ = 5.83). The smaller the intra-class correlation, the better the standard errors of the parameters are estimated[11]. Let us also mention the pseudo-coefficient of determination (R^2), which indicates the weight of the explanatory variables at every level by comparing the residual variances to those of the empty model ($R_1^2 = 1 - \frac{\sigma_{\varepsilon 0 actuel}{}^2}{\sigma_{\varepsilon 0 ancien}{}^2}$ and $R_2^2 = 1 - \frac{\sigma_{\mu 0 actuel}{}^2}{\sigma_{\mu 0 ancien}{}^2}$).

Pseudo R-squared (denoted R^2) should not be compared to the R^2s obtained by OLS. There are analogous, but not equivalent.

The empty model contains no explanatory variables. It simply results in a partition of the total variation between the intra- and inter-level constituents. The purpose is then to reduce these variances by introducing explanatory variables. The explanatory variables that make a significant contribution to the equation related to *Y1* (being or not being part of a black zone) only explain 5.15% of the variance (Table 3, Column 3). However the decomposition by level of observation leads to a larger R^2 at the municipal level ($R^2{}_2$ = 23.19%). Let us here note that, quite surprisingly, the variance observed at the hectometre level (5.88) is almost the same than that observed in the empty model (5.83). At this stage of the analysis, this is not easy to explain. MLM models are perhaps not the best way to model *Y1*. Interpretation of the coefficients may be meaningless.

Y2 is the total number of accidents observed on each hectometre of road. Table 3 shows that its variability between communes is small ($\sigma_{0\mu}{}^2$= 0.0001; ρ_μ= 0.01%); communes can therefore be considered as homogeneous in this respect. Compared to the empty model, the *X* variables have a greater explanatory power at the municipal level ($R^2{}_2$ = 66.30%) than at the hectometre level ($R^2{}_1$ = 9.07%). Several variables explain the variation of *Y2*, including the quality of the adherence of the road surface (ADHERENCE). We suspect the role of this variable to be associated with users' behaviour; given the many recent technical improvements to motor vehicles, drivers may take greater risks because they feel more confident when driving vehicles equipped with these safety features even in difficult conditions[45]. When adherence is not good, road conditions seem not to be manageable by road users for one reason or another (weather conditions, infrastructural or behavioural circumstances), leading to an accident. As expected[46,47], *Y2* is also high when the hectometre is located in the vicinity of a junction (JUNCT_*dist*). We know that better signalled intersections are a short-term solution for avoiding black zones, with roundabouts being a solution in the longer term. Traffic density (TRAFFIC) also has a significant positive influence on *Y2*, confirming former results such as those recently published by Hiselius[48], Golob[43] or Davis[15]. As expected, the total number of accidents (*Y2*) is higher in urban agglomerations (AGGLO), which also confirms the effect of density and/or mobility[27,43]. We see from Table 3 that *Y2* is once again higher close to places where driving conditions change (LANES 0, 100_*dist*; VMAX 0_*dist*). These latter are quite interesting because they identify road discontinuities that are potentially dangerous: narrowing of the road, agglomerations, etc. We know from experience that better warning signs are an easy short-term solution for reducing risks at these transition places[47]. However, in the longer term, they should be accompanied by changes in infrastructure and environment.

Table 3 MLM for accidents on numbered roads

	Y1		*Y2*		*Y3*	
Variables	Empty Model	MLM	Empty Model	MLM	Empty Model	MLM
β_0	–2.05***	–13.71***	0.21***	–1.97***	0.47***	–1.47**
Level 1 variables						
AGGLO				0.38***		0.18***
TRAFFIC		1.52**		0.53***		
VMAX (0m) *_dist*				0.19***		
LANES(3)		–0.75***				
LANES (0m) *_dist*				0.34**		
LANES(100m)*_dist*				0.31***		
RUT						–0.31**
JUNCT (0m) *_dist*		1.33***		0.32**		2.15***
JUNCT(100m)*_dist*				0.35***		
ADHERENCE		0.75***		0.23***		0.35***
BZONE						0.75***
BUILT30		0.32***				0.54***
WOODS25						0.50*
FIRMS		0.47**				
DIRECTION (3)		–0.29**				
Level 2 variables						
EMPLOYDENS		2.64***				–0.97***
DCORR						2.82***
Variance $\sigma_{0\varepsilon}^2$ *(level 1)*	5.83	5.88	1.11	1.01	3.79	2.98
Variance $\sigma_{0\mu}^2$ *(level 2)*	1.94	1.49	0.0003	0.0001	0.0002	0.0002
Total variance	7.77	7.37	1.11	1.0088	3.79	2.98
ρ_ε	75.05%	79.78%	99.98%	99.99%	99.99%	99.99%
ρ_μ	24.95%	20.21%	0.02%	0.01%	0.01%	0.01%
R^2_1	–	–	–	9.05%	–	21.34%
R^2_2	–	23.19%	–	66.3%	–	–
R^2_{total}	–	5.15%	–	9.07%	–	21.34%

***significant at 99.9%; ** significant at 99%; * significant at 95%

Y3 is a simple measure of risk: the total number of accidents observed in a hectometre of road divided by the average daily traffic intensity. The scale effect is here due to hectometres only. Some 21.34% of the variance is explained. This means that infrastructure and environment play a limited but significant role in accident risk. *Y3* is also explained by distance to crossroads (JUNCT*_dist*), location within urban agglomeration (AGGLO), bad adherence of the road surface (ADHERENCE), built-up (BUILT30) and wooded (WOODS25) environments. There is also an unexpected negative effect of rutting (RUT) (β = -0.31): the risk of accidents is small where rutting occurs. This might be explained by the fact that, in the area we studied, rutted roads often correspond to roads with dense traffic and hence congestion. We know that congestion leads to more damage-only accidents (fewer casualties). But rutted roads may also correspond to small roads between hamlets with little traffic, where vehicles do not necessarily adapt their speed. Variables measured at the municipality level have a very small but significant explanatory power for the risk of accidents. However DCORR, which is an index of urban morphology, plays an interesting role: the greater the uniformity in the built space (no hierarchy), the greater the risk of accidents. This is quite an important and novel finding in terms of planning. Specific spatial organisations (morphologies) may affect the relative speed of vehicles. Moreover, the risk of accidents also decreases when population density increases:

the greater the density, the greater the activity and traffic and hence congestion or speed limits. These latter reduce the risk of accidents; as already discussed[49], this should be taken into account when considering mobility problems within urban areas.

We can here conclude that the utility of MLM for modelling the occurrence of accidents is limited: effects of scale are small. In all cases, one level of analysis is predominant, although this may differ for different dependent variables. In the model for the number of accidents (*Y2*), the significant variables explain 66.3% of the variance at the level of the commune, whereas in the model for the risk of accidents (*Y3*), the explanatory variables account for 21.34% of the variance at the level of the hectometre. In the first model, for the occurrence of black spots (*Y1*), the variation is mainly at the level of the hectometre, but the explanatory variables account for 23.19% of the variance at the municipal level. The effect of scale in this case is shared between hectometres (79.78%) and municipalities (20.21%). Although MLM is known to be helpful in revealing differences in variance among units of analysis at different levels, it is less interesting than expected in this frame of application (see Section 5).

The explanatory variables that are significant in all three models are proximity to crossroads (JUNCT_*dist*), the adherence of the road (ADHERENCE), and measures of traffic density (TRAFFIC or EMPLOYDENS). This confirms results of other published studies (see references mentioned). The quality of the road surface is a risk factor that deserves particular attention in the maintenance of roads, especially in a peri-urban context. At the municipal level the risk of accidents increases with the urban morphological indicator (DCORR) and decreases with population density.

4.2 Accidents on motorways

Given the characteristics of motorway traffic (separate lanes, few entries/exists, minimum/maximum speed conditions, etc.) and the specificities of road accidents on this kind of road, models were computed separately for this type of road accidents. The results are given in Table 4.

The first analysis, for the existence of a black spot (*Y1*), shows an increase in the total residual variance for the MLM compared to the empty model (from 17.81 to 55.46) (Table 4, Columns 2 and 3); this means that, overall, the significant explanatory variables do not improve the explanation of whether or not a particular hectometre is part of a black zones of the motorway. The fact that a hectometre belongs to a black zone is not to be explained by the here used explanatory environmental variables.

The model related to the number of accidents by hectometre (*Y2*) is reported in Columns 4 and 5 of Table 4. The introduction of explanatory variables in the model now reduces the total variance (from 1.43 to 1.23); this is exclusively due to Level 1 variables (R^2_1 = 14.24%). Hence, the effect of scale is here mainly due to the hectometres. This is quite obvious for motorways where entries/exists are sparse compared to the scale of the communes. Let us have a look at the explanatory variables. Traffic density has a positive relationship to accidents: the larger the traffic volume, the larger the number of accidents. We know that his relationship is true on average: traffic varies with the time of the day leading to congestion at some periods. Let us remember that accidents on motorways often only involve one road user and are often associated with a loss of control and/or travelling in excess of the speed limit[38,50,51], which occur when traffic is not dense (often at night). Distance to a change in surface (SURFACE200m_*dist*), distance to a change in maximum speed limit (VMAX 0 and 100m_*dist*) are once again sources of changes in road behaviour; they should be the focus of planners' attention, as they normally correspond to roadworks on motorways or places close to cities. Better road signs avoid these situations; speed limits signs are indeed a safety measure and their use has already often been debated[52]. The nearness of firms/department stores (FIRMS) is also a risk factor at the entrance of cities; this risk factor could be avoided by better land use/infrastructure planning. Surfacing of type HFN and the variable BUILT<20 have negative coefficients: asphalt with the conventional texture and sparsely built-up environments are factors of road safety rather than danger.

Modelling of the risk of accidents (*Y3*) is reported in Columns 6 and 7 of Table 4. It turns out that the effect of scale is once again only due to hectometres, and that 13.14% of the variance is explained by the explanatory variables introduced. The MLM adds very little in this case because the variance associated with one of the levels (here Level 2) is almost equal to 0. Significant explanatory variables are the distance to a maximum speed (VMAX 0 or 100_*dist*), membership of the hectometre of a black zone (BZONE) and the type of surfacing (HFN).

Hence, on motorways, Level 2 is useless in modelling *Y2* and *Y3* because variables describing the municipal level only apply to those sections of the motorways that are located near Brussels. It might have been better to choose another contextual environment because the municipal effect only represents a small part of the effects of scale on motorways. For *Y1*, the effects of scale

Table 4 MLM for accidents on motorways

	Y1		*Y2*		*Y3*	
Variables	Empty Model	MLM	Empty Model	MLM	Empty Model	MLM
β_0	–5.27**	92.91**	0.75***	–11.22***	1.54***	4.85*
Level 1 variables						
TRAFFIC		–21.47**		2.55***		
VMAX (0m)*_dist*				2.38***		4.07***
VMAX (100m) *_dist*				1.83**		3.71**
LANES(4)		–2.44**		3.74***		
SURFACE (HFN)				–0.41***		–0.49*
SURFACE(200m)*_dist*				1.06**		
PROX_ACCES (300m)		1.29**				
BZONE						1.70***
BUILT (20%)				–2.63**		
FIRMS		2.46**		0.48*		
Level 2 variables						
EMPLOYDENS						7.19*
MIXITY						–5.31**
ATTRACT						–6.88*
Variance $\sigma_{0\varepsilon}^2$ *(niv 1)*	12.02	30.14	1.43	1.23	4.78	4.15
Variance $\sigma_{0\mu}^2$ *(niv 2)*	5.79	25.32	0.0002	0.0002	0.0002	0
Total variance	17.81	55.46	1.43	1.23	4.78	4.15
ρ_ε	67.47%	54.35%	99.98%	99.98%	99.96%	100%
ρ_μ	32.51%	45.65%	0.02%	0.02%	0.04%	–
R^2_1	–	–	–	14.24%	–	13.14%
R^2_2	–	–	–	–	–	–
R^2_{total}	–	–	–	14.24%	–	13.14%

***significant at 99.9%; ** significant at 99%; * significant at 95%

are shared between hectometres (54.35%) and municipalities (45.65%), but no variable explains the total variation, which increases when explanatory variables are introduced. None of the explanatory variables recurs in each model, but each is useful in defining sensitive places on motorways (such as, for example, the zones with different speed limits, the changes in type of road, the surfacing of roads, etc.). To sum up, when modelling the presence of a hectometre in a black zone on motorways (*Y1*), variations are shared in a more or less equivalent way between hectometres (ρ_ε = 54.35%) and municipalities (ρ_μ = 45.65%). The introduction of explanatory variables does not improve the predictive power of the model. When modelling the number of accidents (*Y2*), the variation occurs at the level of the hectometres and 14.24% (R^2_{total}) is explained. When the dependent variable is the number of accidents divided by the average traffic intensity (*Y3*), the effect of scale is only due to hectometres but only 13.14% of the total variation is explained.

4.3 Comparing MLM to logistic regression results

Multilevel results are not really comparable to other regression results[13,14]. MLM relies on complex, particular distributions of relationships across and within levels. MLM outcomes are hence less general since each best-fitting model may be very specific to the dataset used. Let us here roughly compare *Y1* modelling results obtained on numbered roads with those obtained by Flahaut[3] by means of a more classical logistic regression (Table 5) using the same area of study and almost (but not exactly) the same explanatory variables. We see that (1) infrastructure and land-use have a smaller power of explanation in the MLM model (smaller pseudo R^2), and that (2) the explanatory effect of the variables is only slightly different. Similar effects are to be found for traffic density, crossroads, adherence, built-up areas, and proximity of firms/department stores. MLM modelling adds variables such as the orientation of the roads and the number of lanes; at the municipal level, MLM adds population density, which confirms other density effects measured at the hectometre level and other published papers already mentioned. In the logistic model, distances to changes in speed limit or in the number of lanes as well as the type of road surface play a more determining role also confirming the importance of speed variance.

Table 5 Comparing multilevel and logistic modelling of *Y1* on numbered roads

Variables	Multilevel model	Logistic model
TRAFFIC	++	+++
VMAX (0m) _*dist*		+++
LANES	---	
LANES (0m) _*dist*		+
SURFACE		++
JUNCT (0m) _*dist*	+++	+++
ADHERENCE	+++	+
BUILT (30%)	+++	+++
FIRMS	++	+++
DIRECTION	--	
EMPLOYDENS (level 2)	+++	

(+ a positive relationship; - a negative relationship.
+++/---significant at 99.9%; ++/-- significant at 99%; +/- significant at 95%).

Operationally, both modelling approaches lead to specific results, but, on average, road accidents data in Brabant Walloon seem to have a strong spatial structure that comes through in both modelling procedures. Results are quite stable, confirm common practise and should hence be better integrated into land use planning and road infrastructure enhancement policies.

5. CONCLUSIONS

The importance of MLM in understanding contextual effects on road safety lies in its ability to meaningfully specify the latent structure of relationships, which involve individuals and their environments. This paper considered the spatial occurrence of road accidents by means of MLM. Three dependant variables were studied separately. Explanatory components were limited to infrastructure and environment; the choice of the data was constrained by the availability of the data, the characteristics of the studied area as well as recent published results. Due to the nature of the road accident and data limitations, only two levels of analysis were taken into account: the hectometre and the commune. In spite of these many limitations, we can conclude that:

(1) MLM enables the relative importance of spatial levels in the explanatory process to be assessed. In our case, the commune has, on average, less importance in the explanation than the hectometre: road accidents occur at micro-locations (hm) which can be analysed in a broader spatial context, but this context does not seem to correspond to the commune. If communes are useful official statistical and administrative units, they vary in size and shape (modifiable areal unit problem) and are not suitable for explaining the locations and spatial concentrations of road accidents. This level is not appropriate for taking into account the complex relationships in which a car/road user is involved in an accident. Unfortunately, in road accident analysis, the choice of the spatial level of observation is not as straightforward as in human and behavioural sciences. In both road accident and mobility analyses, there is often a compromise between data availability and meaning. In this case hectometres and communes were the only possible choices. MLM for road accidents is not straightforward to apply for spatial analysis.

(2) If the level of spatial explanation is not high, it is however significant and corroborates former published results. Environment and infrastructure explain between 5% and 21% of the total observed variation in road accidents in Brabant Walloon. We are conscious that our models are mis-specified: we didn't take into account the many other factors that could interact (user behaviour, mobility patterns, etc.). Given these results, the physical characteristics of the road, as well as its environment, should be better and further integrated into safety and land-use policies. Let us here however add that the assessment of environmental accident factors can, however, been discussed: circumstances can be present and regarded as risk factors, but not necessarily as accident factors and causes[36].

(3) Three different dependent variables were analysed: whether or not a hectometre belongs to a black zone; the number of accidents per hectometre; and the risk of accidents, defined as the number of accidents divided by the average traffic volume. Each *Y* variable has a specific meaning for police forces, emergency services or road engineers/planners. In our analysis,

each model has also a specific form, with a difference combination of independent variables.

(4) Most explanatory variables are associated with hectometres. Many are related to changes in road conditions. The importance of these changes in road conditions in the explanation reveals the inability of the road user to adapt his or her behaviour to changes in road conditions and road infrastructure[53]. In-depth analysis of each sub-type of accident should increase our understanding of each type of circumstance, but this is far beyond the scope of this paper. All these associations show that spatial concentrations of road accidents often correspond to places where improvements could be made in terms of road design, signalling and land-use planning. This corroborates previous results on road accidents and road geometry[33,34,47,54].

(5) The introduction of a morphological index (fractal dimension) is quite novel in measuring land use and more particularly in explaining road safety. Further analyses will be performed with this variable. In this paper we showed that homogeneity in texture leads to greater danger. Specific spatial organisation (morphology) may affect the relative speed of vehicles as well as the mobility patterns. This should be taken into consideration when considering safety and mobility problems within urbanised areas.

(6) Multilevel results are not really comparable to other regression results. MLM outcomes are less general since each best-fitting model may be very specific for the dataset used. In our case, logistic regression seems to be easier to use and could be extended to cope with autocorrelation[3]. Other analytical methods, such as a weighted geographic regression[21] (which is based on the hypothesis that the variations between variables measured in different places cannot be constant in the space), may however also be interesting to use. Instead of considering the local variations as averages and as unobservable, weighted regression allows us to measure local variations and to map them. However, the quality of the spatial data for this type of analysis is a strong preliminary condition.

The findings of this paper are both suggestive and limited in that they are based on only one data set, and only consider the environment and infrastructure as explanatory variables. Our modelling results depend strongly upon the many choices made, and these are strongly related to data availability. Neither the social characteristics of the road users nor the technical characteristics of vehicles are considered here. The interactions between social, technical and spatial variables are not taken into account. Our paper shows the importance of the hectometre as a basic spatial unit and the limited usefulness of multilevel models in analysing road accident locations. Other statistical techniques may be better suited to this task. Spatial concentrations of accidents are characterised by specific accident circumstances, which require different counter-measures to reduce their number (e.g. improvements in terms of road design, signalling, and local environment). There is no unique combination of characteristics associated with road accident locations: it is a complex phenomenon of which only a very few aspects have been considered here.

REFERENCES

1. Thomas, I. Spatial data aggregation: exploratory analysis of road accidents. "ACCIDENT ANALYSIS AND PREVENTION" 28(2): pp.251-264. (1996).
2. Flahaut, B., Mouchart, M., San Martin, E. and Thomas, I. Identifying black zones with a local spatial autocorrelation index and a kernel method: a comparative approach. "ACCIDENT ANALYSIS AND PREVENTION" 35(6): pp.991-1004. (2003).
3. Flahaut, B. Impact of infrastructure and local environment on road insecurity. Logistic modelling with spatial autocorrelation, "ACCIDENT ANALYSIS AND PREVENTION" 36(6): pp.1055-1066. (2004).
4. Lassarre, S. and Thomas, I. Exploring road mortality ratios in Europe: national versus regional realities. "JOURNAL OF THE ROYAL STATISTICAL SOCIETY" A. (in press - January 2005).
5. Courgeau, D. and Baccaïni, B. Analyse multiniveau en sciences sociales. "POPULATION" 4: pp.831-864. (1997).
6. Mathian, H. and Piron, M. Echelles géographiques et méthodes statistiques multidimensionnelles. in Sanders L. (ed.) Modèles en analyse spatiale. Paris, Hermès Science Publications. pp.62-103. (2000).
7. Goldstein, H. Multilevel Statistical Models. London, Arnold. 178p. (1995).
8. Jones, K. Concepts and Techniques in Modern Geography: Multi-level Models for Geographical Research. Norwich, Environmental Publications, University of East Anglia. 50p. (1991).
9. Bryk, A. and Raudenbsuch, S. Hierarchical Linear Models: Applications and Data Analysis Methods. Second edition, California, Sage Publications. 485p. (2002).
10. Jones, K. and Duncan, C. People and places: the multilevel model as a general framework for the quantitative analysis of geographical data. in Longley P. and Batty M. (eds). Spatial Analysis: modelling in GIS environment. John Wiley and Sons, London. pp.79-104. (1996).
11. Jones, A. and Jorgensen, S. The use of multilevel models for the prediction of road accident outcomes. "ACCIDENT ANALYSIS AND PREVENTION" 35(1): pp.59-69. (2003).
12. Gee, G. and Takeuchi, D. Traffic stress, vehicular burden and well-being: a multilevel analysis. "SOCIAL SCIENCE & MEDICINE" 59: pp.405-414. (2004).
13. Openshaw, S. and Taylor, P. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. in Wrigley N. (ed). Statistical applications in the spatial sciences (London: Pion). pp.127-144. (1979).
14. Robinson, W. Ecological correlations and the behaviour of individuals. "AMERICAN SOCIOLOGICAL REVIEW" 15: pp.351-357. (1950).
15. Davis, G. Is the claim that variance kills an ecological fallacy ? "ACCIDENT ANALYSIS AND PREVENTION" 34(3): pp.343-346. (2002).

16. Greenland, S. and Robins, J. Invited commentary: ecologic studies – biases, misconceptions, and counterexamples. "AMERICAN JOURNAL OF EPIDEMIOLOGY" 139: pp.747-760. (1994).
17. Kreft, I. and De Leeuw, J. Introducing Multilevel Modelling. London: Sage Publications. (1998).
18. Snijders, T. and Bosker, R. An Introduction to Basic and Advanced Multilevel Modelling. London, Sage Publications. 266p. (1999).
19. Hox, J. Applied Multilevel Analysis. Amsterdam, TT-Publikaties. 119p. (1995).
20. Jones, K. and Duncan, C. Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. "HEALTH AND PLACE" 1(1): pp.27-40. (1995).
21. Fotheringham, A., Brunsdon, C. and Charlton, M. Multilevel modelling in Fotheringham A. (ed). Quantitative Geography: Perspectives on Spatial Data Analysis. London, Sage Publications. pp.103-106. (2000).
22. Polsky, C. and Easterling, W. Adaptation to climate variability and change in the US Great Plains: a multi-scale analysis of Ricardian climate sensitivities. "AGRICULTURE, ECOSYSTEMS AND ENVIRONMENT" 85: pp.133-144. (2001).
23. Singer, J. Fitting multilevel models using SAS Proc Mixed. "MULTILEVEL MODELLING NEWSLETTER" 10(2): pp.5-9. (1998).
24. Singer, J. Using SAS Pproc Mixed to fit multilevel models, hierarchical models, and individual growth models. "JOURNAL OF EDUCATIONAL AND BEHAVIORAL STATISTICS" 24(4): pp.323-355. (1998).
25. De Keersmaecker, M.-L., Frankhauser, P. and Thomas, I. Analyse de la réalité fractale p_riurbaine: l'exemple de Bruxelles. "L'ESPACE GÉOGRAPHIQUE" 2004 (3): pp.219-240. (2004).
26. Thomas, I., Frankhauser, P. and De Keersmaecker, M-L. Fractal dimension versus density of the built-up surfaces in the periphery of Brussels. Paper presented at the ERSA 2004 Conference in Porto, refereed session.
27. Millot, M. Urban growth, travel practises and evolution of road safety. "JOURNAL OF TRANSPORT GEOGRAPHY" 12: pp.207-218. (2004).
28. Flahaut, B. and Thomas, I. Identifier les zones noires d'un réseau routier par l'autocorrélation spatiale locale. Analyses de sensibilité et aspects opérationnels, "REVUE INTERNATIONALE DE GEOMATIQUE" 12(2): pp.245-261. (2002).
29. Steengergen, T., Dufays, T., Thomas, I. and Flahaut, B. Intra-urban location of road accident black zones: a Belgian example. "INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE" 18(2) : pp.169-181. (2004).
30. Eckhardt, N., Flahaut, B. and Thomas, I. Spatio-temporalité des accidents de la route en périphérie urbaine. L'exemple de Bruxelles. "RECHERCHE TRANSPORTS SECURITE" 82: pp.35-46. (2004).
31. Casaer, F., Eckhardt, N., Steenberghen, T., Thomas, I., Wets, G. and Wijnants, J. Een onderzoek naar de kwaliteit van de Belgische ongevallendata. Working paper, LUC. (Diepenbeek). (2002).
32. Institut Géographique National (IGN) Banques de données et cartes topographiques. (Top50r), Bruxelles, IGN. (2002).
33. Agent, K. and Deen, R. Relationship between roadway geometrics and accidents. TRANSPORTATION RESEARCH RECORD 541: pp.1-11. (1975).
34. Wong Y. and Nicholson, A. Driver behaviour at horizontal curves: risk compensation and the margin of safety. "ACCIDENT ANALYSIS AND PREVENTION" 24(4): pp.425-436. (1992).
35. Okamoto, H. and Koshi, M. A method to cope with the random errors of observed accidents rates in regression analysis. "ACCIDENT ANALYSIS AND PREVENTION" 21(4): pp.317-332. (1989).
36. Garber, N., and Ehrhart, A. The effect of speed, flow, and geometric characteristics on crash frequency for two-lane highways. TRANSPORTATION RESEARCH RECORD 1717: pp.76-83. (2000). (see also full report on http://www.virginiadot.org/vtrc/main/online_reports/pdf/00-r15.pdf.)
37. Ossenbruggen, P., Pendharkar, J. and Ivan, J. Roadway safety in rural and small urbanized areas. "ACCIDENT ANALYSIS AND PREVENTION" 33(4): pp.485-498. (2001).
38. Lee, J. and Mannering, F. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. "ACCIDENT ANALYSIS AND PREVENTION" 34(2): pp.149-161. (2002).
39. Henning-Hager, U. Urban development and road safety. "ACCIDENT ANALYSIS AND PREVENTION" 18(2): pp.135-145. (1986).
40. Vasconcellos, E. Urban development and traffic accidents and Brazil. "ACCIDENT ANALYSIS AND PREVENTION" 31(4): pp.319-328. (1999).
41. Noland, R. and Quddus, M. A spatially disaggregate analysis of raod causalties in England. "ACCIDENT ANALYSIS AND PREVENTION" 36(6): pp.973-984. (2004).
42. Abdalla, I., Raeside, R., Barker, D. and McGuigan, D. An investigation into the relationships between area social characteristics and road accident casualties. "ACCIDENT ANALYSIS AND PREVENTION" 29(5): pp.583-593. (1997).
43. Golob, T., Recker, W. and Alvarez, V. Freeway safety as a function of traffic flow. "ACCIDENT ANALYSIS AND PREVENTION" 36(6): pp.933-946. (2004).
44. Halleux, J., Derwael, F. and Merenne, B. Urbanisation. Monographie 11A. Recensement Général de la Population et des Logements au 1er mars 1991. (1998).
45. Edwards, J. Weather-related road accidents in England and Wales: a spatial analysis. "ACCIDENT ANALYSIS AND PREVENTION" 4(3): pp.201-212. (1996).
46. Larsen, L. and Klines, P. Multidisciplinary in-depth investigations of head-on and left-turn road collisions. "ACCIDENT ANALYSIS AND PREVENTION" 34(3): pp.367-380. (2002).
47. Geurts, K., Thomas, I. and Wets, G. Understanding accidents at black zones using frequent item sets. "ACCIDENT ANALYSIS AND PREVENTION" (submitted).
48. Hiselius, L. Estimating the relationship between accident frequency and homogeneous and inhomogeneous traffic flows. "ACCIDENT ANALYSIS AND PREVENTION" 36(6): pp.985-992. (2004).
49. Shefer, D. and Rietveld, P. Congestion and safety on motorways: towards an analytical model. "URBAN STUDIES" 34: pp.679-692. (1997).
50. Navon, D. The paradox of driving speed: two adverse effects on motorway accident rate. "ACCIDENT ANALYSIS AND PREVENTION" 35(3): pp.361-367. (2003).
51. Thiffault, P. and Bergeron, J. Monotony of road environment and driver fatigue: a simulator study. "ACCIDENT ANALYSIS AND PREVENTION" 35(3): pp.381-391. (2003).
52. Aljanahi, A., Rhodes, A. Metcalfe, A. Speed, speed limits and road traffic accidents under free flow conditions. "ACCIDENT ANALYSIS AND PREVENTION" 31(1)-(2): pp.161-168. (1999).
53. Elvik, R. To what extent can theory account for the findings of road safety evaluation studies? "ACCIDENT ANALYSIS AND PREVENTION" 36(5): pp.841-849. (2004).
54. Greibe, P. Accident prediction models for urban roads. "ACCIDENT ANALYSIS AND PREVENTION" 35(2): pp.273-285. (2003).

ACKNOWLEDGEMENTS

This work was funded by the Federal Scientific Policy in Belgium (project CP/10/342). We would like to thank Benoît Flahaut and the anonymous referees for helpful comments on earlier versions of this paper.