Accident Analysis and Prevention 132 (2019) 105237

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Estimating the occurrence of traffic accidents near school locations: A case study from Valencia (Spain) including several approaches☆

Álvaro Briz-Redón[a,*], Francisco Martínez-Ruiz[b], Francisco Montes[a]

[a] Department of Statistics and Operations Research, University of València, Spain
[b] Statistics Office, City Council of València, Spain

ARTICLE INFO

Keywords:
Traffic safety analysis
School zones
Conditional autoregressive model
Case–control study
Risk in relation to point sources

ABSTRACT

Traffic safety around school locations is a topic of particular interest given the large number of vulnerable users, such as pedestrians or cyclists, that commute to them at certain times of the day.

A dataset of traffic accidents recorded in Valencia (Spain) during 2014 and 2015 is analyzed in order to estimate the effects that school locations produce on traffic risk within their surroundings. The four typologies of school in this city according to the academic levels they offer (All-level, Preschool, Primary, Secondary) are distinguished and taken into consideration for the analysis. Two time windows comprising the starting time in the morning and the evening time once day school has ended are analyzed independently.

Several statistical methods are used, including observed vs expected ratios, macroscopic conditional autoregressive modelling, logistic regression in the context of a case–control study design and risk modelling in relation to several school locations. The distances to each type of school and a set of environmental, traffic-related, demographic and socioeconomic covariates are employed for the analysis.

The macroscopic modelling of accident counts and the modelling of risk as a function of the distance to each type of school serves to confirm that proximity to a school has an effect on the incidence of traffic accidents in particular time windows. Specifically, school types coexisting in Valencia show differential behaviour in this regard. In addition, several covariates have displayed a positive (bus stop density, complex intersections, main road length) and negative (land use entropy) association with accident counts in the time windows investigated.

Finally, the definition of a case–control study design enabled us to observe some differences undetected by the macroscopic approaches that would require further research.

1. Introduction

Guaranteeing safety near school locations is a fundamental objective for the experts in charge of traffic management, especially in order to protect children from traffic accidents. In Spain, high densities of vehicles are usually observed around school locations at starting and ending times on school days. This is a factor that certainly conditions traffic flow and dangerousness at these times, but few scientific research studies are available on this issue in Spain. Hence, the following paragraphs include a description of several studies conducted in the last fifteen years on the subject of traffic safety around schools, the factors involved and the implications. An emphasis is placed on the methodological approaches chosen for these studies.

Most of them have focused on areal zones centred around school locations (buffer zones). First, Abdel-Aty et al. (2007) employed a log-linear model and a set of categorical variables including driver characteristics, pedestrian/cyclist characteristics and other characteristics related to traffic, road and vehicle typologies. These authors showed that most of the traffic accidents involving school-aged children took place close to school locations, considering a buffer zone of 0.5 miles from each school. More specifically, they found higher accident rates for middle and high school children, which was associated with the fact that these schools are frequently situated in the vicinity of a multi-lane high-speed road. Clifton and Kreamer-Fults (2007) modelled five types of aggregated dependent measures (two related to accident occurrence and three to pedestrian exposure) through linear regression on the basis of 0.25 mile buffer zones around schools. Among the set of covariates considered by these authors, percentage of nonwhite residents and population density were associated with more traffic accidents, whereas transit access (percentage of households within 0.25 miles of a transit

☆ Paper has been handled by associate editor Tony Sze.

* Corresponding author.
E-mail address: alvaro.briz@uv.es (Á. Briz-Redón).

https://doi.org/10.1016/j.aap.2019.07.013
Received 15 April 2019; Received in revised form 8 July 2019; Accepted 16 July 2019
Available online 30 August 2019

stop) correlated negatively with accident counts. Warsh et al. (2009) created buffer zones with a radius ranging from 150 to 450 m around schools and confirmed the higher proportion of child pedestrian–vehicular accidents at 7.00–10.00 and 15.00–17.00. They also observed a decrease in the risk for older students, specially those in the 15–17 age group. Yu and Zhu (2016) assessed the presence of modifiable areal unit problems (MAUP) (Amoh-Gyimah et al., 2017; Zhai et al., 2018; Xu et al., 2018) in the specific context of school safety by defining 0.5, 1, 1.5, and 2 mile buffer zones around each school. It was found that highways and interstates, traffic-generating land uses and transit stops were associated with more traffic accidents. On the other hand, higher sidewalk coverage and local roads around schools correlated with fewer traffic accidents.

The use of buffer zones around school locations can be combined with a case–control study design. Indeed, Rothman et al. (2017b) performed a case–control study in Toronto (Canada) by distinguishing school attendance boundaries belonging to the highest quartile of pedestrian–vehicle accident rates (cases) from those in the lowest (controls). It was found through a multivariate logistic regression that some factors such as one way streets, crossing guards, traffic signal density and social disadvantaged areas were associated with a higher incidence of traffic accidents. Moreover, Rothman et al. (2017a) focused their research on traffic safety near schools around risky drop-off behaviours following the same case–control strategy as Rothman et al. (2017b). Observational covariates related to risky behaviours by both drivers and pedestrians in school proximities were obtained, and these were then investigated in combination with a collection of environmental covariates through logistic regressions. Several important findings for safety planning were found, including, for instance, the association between traffic congestion and risky driving and walking behaviours.

Furthermore, in the last few years there has been an increasing number of studies approaching the research from a road segment level perspective. For example, Hwang et al. (2017) considered street segments at a 0.25 mile distance from school locations and a buffer distance of 100 ft around these street segments as the focus of their analysis. Accident counts at the road segment level were recoded into a binary outcome and logistic regression was used for modelling purposes. These authors found a positive correlation between accident rates and block length, proportion of missing sidewalks, crosswalk density and commercial land use. Furthermore, the study revealed some factors specifically affecting students from disadvantaged neighbourhoods. Park et al. (2018) compared negative binomial and Poisson inverse Gaussian model approaches and found that the latter provided better results in terms of forecasting accuracy. They worked at the road segment level, considering roadways connected to a school building or to a nearby area over which school-related activities were taking place. Finally, Yu (2015) made use of a hierarchical logistic model at two spatial levels: neighbourhood and road segment. The findings were similar to those of Yu and Zhu (2016).

The aim of our paper is to present the use of several statistical techniques in order to estimate the effects that school locations and commuting to school may have on the incidence of traffic accidents in specific time windows. These techniques are applied to a dataset of traffic accidents recorded over two years in the city of Valencia (Spain). The objectives are both methodological and practical, as will be highlighted within the text.

Thus, the paper is structured as follows. First, the data used in the analysis is described, including the dataset of traffic accidents, the collection of schools and typologies available in the city, and the set of covariates considered with explanatory objectives. This is followed by a methodological section containing an exploratory analysis that helps to determine the time windows that may be affected by school-related trips, and the explanation of four different statistical methods (observed vs expected ratios, spatial count models, case–control logistic regression and multiple source regression) that allow us to investigate the causal relationship between school locations and traffic accidents occurring during the time windows predefined through the exploratory analysis just mentioned. The results derived from each of the methods are then discussed and compared.

2. Data

2.1. Accident dataset

A total of 18,037 traffic accidents recorded by the Local Police of Valencia (Spain) during the years 2014 and 2015 were used for the analysis. Geographical coordinates for the accidents and information about the date and time of occurrence were provided by the Local Police. The accidents were located with precision on the road network of the city. This road network has a total length of 840.3 km (with a diameter of nearly 11.6 km) and contains 6110 road intersections. Information regarding the type of traffic accident (vehicle–vehicle, vehicle-pedestrian, etc.), its severity (severe, non-severe, etc.) and the age of the people involved was unavailable, although no reporting bias in favour of a specific type of accident should be present.

2.2. School dataset

A total of 372 schools of various age levels located in Valencia were considered for the analysis (Fig. 1). Four main levels of education in Spain can be distinguished (with approximate ages): Preschool I (0–3 years), Preschool II (4–5 years), Primary (6–11 years) and Secondary (12–17 years). For research purposes, the schools were classified into four categories, according to the educational levels offered: All-level (81 schools), Preschool (178 schools, which include centres that offer only Preschool I or II or both), Primary (81 schools) and Secondary (32 schools). Hence, one of our objectives was also to assess whether differential associations with traffic accidents arise depending on the school classification.

2.3. Covariate definition

Several covariates were considered in order to control for baseline effects that may be responsible for the higher incidence of traffic accidents near schools. These covariates were classified into three categories: traffic-related, environmental and socioeconomic/demographic. The values for these covariates were always computed over basic spatial units (BSU) of analysis, which varied depending on the analysis being performed. Table 1 contains a description of all the covariates treated during the analysis, which were conveniently standardized in order to facilitate parameter interpretation and comparison. In particular, the

Fig. 1. School locations in the city of Valencia distinguished by the academic level they offer. The central district of Valencia (city centre) is highlighted with a thicker black line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Description and classification of the covariates defined for the analysis.

Type	Variable
Accidents	No. of traffic accidents
Exposure	Non-pedestrian road length
School-related	No. of schools (all types)
	Distance to the closest All-level school
	Distance to the closest Preschool school
	Distance to the closest Primary school
	Distance to the closest Secondary school
Environmental	No. of education-related services per road km
	No. of services from various sectors (non-educational) per road km
	% of road length with parking spaces available
	No. of parking zones per road km
	No. of bus stops per road km
	Land use entropy
Traffic-related	Average betweenness per road segment
	Complex intersections (four-or-more-leg) per road km
	Main road length per road km
	Traffic lights per road km
Socioeconomic and demographic	No. of school-aged residents (0–18 years) per road km
	Percentage of high-end cars

following lines describe two of the covariates included: betweenness and land use entropy.

Betweenness (BETW) is a measure of network connectivity that was used as a proxy for average vehicle miles travelled per road segment (which was unavailable for most of the streets) following the next formula (Freeman, 1977):

$$\mathrm{BETW}(e) = \sum_{i \sim j} \frac{\sigma_{\mathrm{ij}}(e)}{\sigma_{\mathrm{ij}}}$$

where e is a segment of the road network, i and j are two vertices of the network that are connected by a path ($i \sim j$), σ_{ij} the number of shortest paths between i and j and $\sigma_{ij}(e)$ the number of shortest paths that connect i and j while passing through edge e of the network.

Land use entropy (LUE) was defined as in Rothman et al. (2017b), following the next expression:

$$\mathrm{LUE}(\mathrm{BSU}_i) = -\frac{\sum_j p_{\mathrm{ij}} \log(p_{\mathrm{ij}})}{\log(n)}$$

where j iterates over the indexes associated with the land uses that are present at BSU i (otherwise, the logarithmic expression is not computable), p_{ij} is the proportion of land use of type j at BSU i and n is the total number of land uses considered by the Land Occupancy Information System (SIOSE, by its acronym in Spanish). The LUE index lies in the [0,1] interval, with a value closer to 1 indicating higher land diversity and 0 meaning a unique land use.

3. Methodology

3.1. Software

The R programming language (3.5.2 version, R Development Core

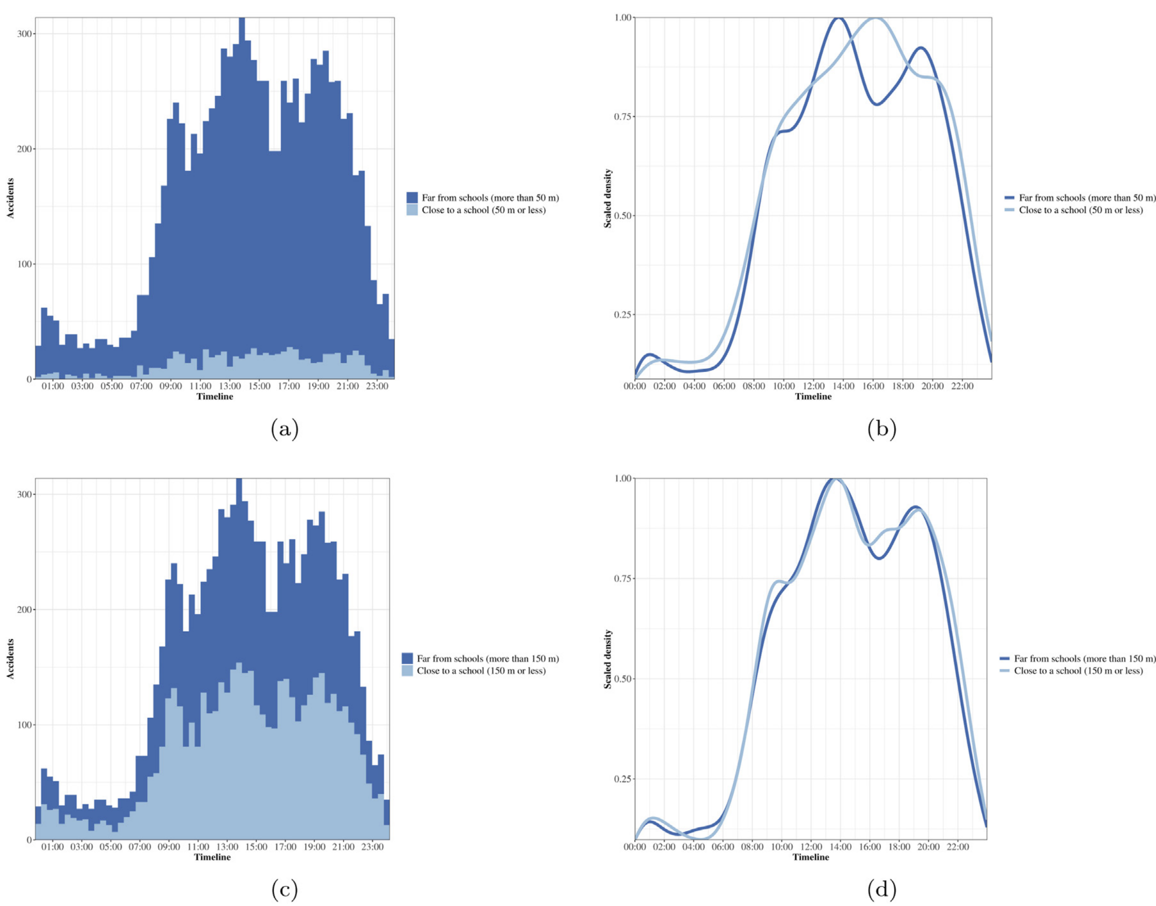


Fig. 2. Histograms and scaled densities along the timeline of school days (in 2014–2015) shown by traffic accidents that occurred close to a school and far from schools for a threshold distance of 50 m (a and b) and 150 m (c and d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Team, Vienna, Austria) (R Core Team, 2018) was used to obtain all the results presented in this work. The R packages *DEoptim* (Mullen et al., 2011), *ggplot2* (Wickham, 2016), *INLA* (Rue et al., 2009; Martins et al., 2013; Lindgren and Rue, 2015), *rgeos* (Bivand and Rundel, 2018), *spatstat* (Baddeley et al., 2015) and *spded* (Bivand and Piras, 2015) were specifically required to perform the complete analysis.

3.2. Time window of analysis

A preliminary question was to determine the hours within the day that may be affected by traffic dynamics generated as a consequence of arrivals at (or departures from) schools. School starting and ending times should be the axes of such a temporal interval, but this is hard to define because most schools in Valencia are free to determine their own schedules (subject to some common restrictions). Furthermore, it is quite normal for students in Spain to attend extracurricular activities during weekday evenings (immediately after school), which are usually carried out in the school facilities or in the surrounding area (extracurricular academies, institutions or centres are frequently located in the vicinity of a school).

Exploratory analyses were performed (Fig. 2) and two time windows of interest were finally established: the one around school starting times and the period in the afternoon and evening that includes school ending times and the subsequent hours. Hence, the two time windows 7:30–9:30 and 15:00–19:00 were selected for further analysis. In view of Fig. 2, it seems that school locations do not have an strong effect on the incidence of traffic accidents around starting times, but we still decided to maintain this period of the day for further investigation. The choice of the interval 7:30–9:30 was based on the fact that most schools in Valencia start their classes in the 8:00–9:00 window. In contrast, the period 15:00–19:00 shows a clear differential effect in terms of density of accidents close to and further away from schools, especially for short threshold distances such as 50 m, for example (Fig. 2b). From now on, the time window 7:30–9:30 will be referred to as the Starting Time Window (STW) and 15:00–19:00 will be denoted as the Evening Time Window (ETW).

3.3. Observed vs expected ratios

The first statistical analysis to identify the association between traffic accidents and school locations consisted in computing observed/expected accident ratios at different distance thresholds, σ, from all schools in the city. The sequence of values selected for σ were 25, 50, 75, 100, 150, 200, 250 and 300 m, which allowed us to analyze the effect we are interested in at various spatial scales.

A Monte Carlo approach was taken to assess the statistical significance of the ratios considering the full set of schools and each school type separately for the two time windows established. This process consists in generating 999 datasets that preserve the locations observed for all the accidents recorded in 2014–2015, while permutating their corresponding dates and times of occurrence. Hence, the number of traffic accidents that lie below the threshold distance (σ) from a school location (for both STW and ETW) is kept for each simulation. The average of the 999 simulated values represents an expected value, which is then compared with the real number observed, providing an observed/expected ratio. The 2.5th, 5th, 95th and 97.5th percentiles of the set of simulated distributions of counts are kept to make it possible to assess a significance level for each observed/expected ratio (for a given σ).

3.4. Macroscopic modelling

A conditional autoregressive (CAR) model with negative binomial (NB) response was chosen to fit the accident counts recorded on school days in the STW or ETW (a total of 800 and 2884, respectively, for the period 2014–2015) over a hexagonal grid of 198 BSUs of side length slightly over 250 m and an area of around 0.175 km^2. The use of a hexagonal grid was preferred over the employment of administrative division units (such as boroughs or census tracts) because it provides the possibility of defining a spatial unit of analysis of intermediate scale. This makes it possible to perform an accurate analysis while keeping a good balance between the number of spatial units and the number of covariates being considered. Hexagonal units containing a minimal road structure (mostly located within a green area or in semirural zones along the periphery of the city) were removed from the grid in order to avoid a possible distortion of the results.

If $Y \sim \mathrm{NB}(\mu, \psi)$ (NB distribution of mean μ and shape ψ) then it holds that $E(Y) = \mu$, $V(Y) = \mu + \frac{\mu^2}{\psi}$ and $P(Y = x) = \binom{x+\psi-1}{\psi-1}\left(\frac{\psi}{\mu+\psi}\right)^{\psi}\left(\frac{\mu}{\mu+\psi}\right)^{x}$. Then, assuming a NB distribution for the accident counts, the following spatial model was implemented:

$$Y_i \sim \mathrm{NB}(\mu_i, \psi)$$
$$\log(\mu_i) = \log(E_i) + \lambda_0 + \sum_{m=1}^{p} \lambda_m X_{\mathrm{im}} + \phi_i \quad (1)$$

where Y_i is the number of accidents observed at hexagonal unit i, μ_i and ψ are, respectively, the mean risk (for hexagonal unit i) and overdispersion (shape) values for the NB distribution, the natural logarithm acts as a link function for μ_i, E_i (exposure at hexagonal unit i) is the length of non-pedestrian road at hexagonal unit i (offset of the equation), X_{im} represents the value of the mth covariate at hexagonal unit i, λ_m is the coefficient that controls the effect of the mth covariate and ϕ_i represents a spatial effect for hexagonal unit i, which is derived from the neighbourhood structure formed by the hexagons that are part of the grid.

The spatial effect was modelled through the well-known CAR structure (Besag, 1974; Besag et al., 1991):

$$\phi_i | \phi_j, \quad j \neq i \sim N\left(\alpha \sum_{j=1}^{n} w_{\mathrm{ij}} \phi_j, \tau_i^{-1}\right) \quad (2)$$

where τ_i is a precision parameter that varies with spatial unit i and w_{ij} is an indicator parameter that is 1 if hexagonal units i and j are contiguous and 0 otherwise. The use of a CAR structure for the modelling of accident counts at the macroscopic level is a common practice in traffic safety analysis (Quddus, 2008; Huang et al., 2010).

3.5. Risk modelling in relation to several point sources

Diggle and Rowlingson (1994) proposed a class of regression models inspired by previous works focused on estimating disease risk around one hazardous point source. This kind of model has been successfully applied in many epidemiological studies involving pollution sources that negatively affect human health (Ramis et al., 2011; Reeve et al., 2013). Analogously, this approach could be useful to model the risk triggered by the close presence of a particular type of building within the road network, such as a school. As far as the authors of this paper know, multiple source regression models have not been used before in the field of traffic safety analysis.

Diggle et al. (1997) adapted the work from Diggle and Rowlingson (1994) to enable this class of models to be used with aggregated data. We chose this approach for the available dataset, considering the same hexagonal grid that was employed for the CAR modelling of accident counts. Hence, Eq. (3) displays the mathematical expression of this multiple source regression technique, which accommodates four sources of risk that correspond to the four school types that exist in Valencia.

Table 2
Risk ratios for each combination of distance threshold (σ), school type or combination of types and time window (STW or ETW). Risk ratios in bold were found significant at the 0.1 level. The ratios for $\sigma = 25$ m at STW were unreliable due to the small samples available and are not shown.

STW	σ (m)							
	25	50	75	100	150	200	250	300
All-level	–	0.92	0.91	0.91	1.10	1.05	1.00	1.00
Preschool	–	1.11	1.10	1.05	1.00	0.99	1.01	1.02
Primary	–	0.87	0.86	0.99	0.97	0.94	0.95	0.97
Secondary	–	0.95	1.23	1.12	1.05	0.97	0.90	0.93
All types	–	1.00	1.01	1.00	1.03	1.01	1.02	1.01
ETW	σ (m)							
	25	50	75	100	150	200	250	300
All-level	0.97	0.89	1.04	1.03	1.01	1.01	1.00	0.99
Preschool	**1.23**	**1.14**	1.03	1.03	0.99	1.02	1.01	1.00
Primary	1.25	**1.25**	1.10	1.08	1.06	1.03	1.03	**1.03**
Secondary	0.79	0.83	0.84	1.07	0.99	1.03	1.03	1.03
All types	**1.15**	1.05	1.03	1.04	1.00	1.01	1.00	1.00

$$Y_i \sim \text{Po}(\mu_i)$$
$$\mu_i = \exp\left(\sum_{m=1}^{p} \lambda_m X_{\text{im}}\right) \prod_{j=1}^{4} f(d_{\text{ij}})$$
$$f(d_{\text{ij}}) = 1 + \alpha_j \exp(-(d_{\text{ij}}/\beta_j)^2) \tag{3}$$

where j indicates school type (1 = All-level, 2 = Preschool, 3 = Primary, 4 = Secondary), α_j is the proportional increase (or decrease) in risk produced by school type j, d_{ij} is the distance in km between hexagonal unit i and type j (from the centroid of the hexagon to the closest element of school type j) and β_j measures the rate of decay in risk that occurs as distance from type j increases.

Diggle et al. (1997) approximated the log-likelihood function associated with the model in Eq. (3) through the next expression:

$$L(\boldsymbol{\Lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = -\sum_i \mu_i + \sum_i O_i \log(\mu_i) \tag{4}$$

where $\boldsymbol{\Lambda}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ represent the three vectors of coefficients that include, respectively, the λ_m's, α_j's and β_j's, μ_i follows Eq. (3) and O_i is the number of accident counts observed in hexagonal unit i. The expression in Eq. (4) was maximized with the aid of the R package *DEoptim* (Mullen et al., 2011). The method implemented in *DEoptim* relies on the theory of differential evolution algorithms for global optimization (Price et al., 2006).

3.6. Case–control study

The calculation of observed/expected ratios in buffer zones of various radii from each school location enabled us to detect those with higher ratios. Thus, the same Monte Carlo procedure described in Section 3.3 was applied around each of the schools for STW and ETW, separately, to assess statistical significance. Schools with a significantly (at the 0.10 level) high (over 1) observed/expected ratio within a certain buffer zone built around them were considered as cases for establishing of a case–control study. On the other hand, schools presenting a ratio lower than 1 (not necessarily showing statistical significance for such a low ratio) were declared as control schools. Hence, a sample of cases and controls was obtained, constraining buffer zones around school locations to not overlap (avoiding an excess of multicollinearity). A distance of $\sigma = 100$ m was found optimal for defining the buffer zones according to the number of cases and controls provided (choosing this distance allowed having a case:control ratio close to 1:4) and the spatial accuracy achieved. Then, a multivariate logistic regression model was specified:

$$\log\left(\frac{pi}{1 - p_i}\right) = \log(E_i) + \sum_{m=1}^{p} \lambda_m X_{\text{im}} + \phi_i + \text{Type}_i \tag{5}$$

where $\frac{pi}{1-p_i}$ represents the odds ratio for a school being a case. E_i, λ_m, X_{im} and ϕ_i are as in Eq. (1), but now for the buffer zone associated with school i. Neighbourhood relationships for estimating ϕ_i were now established between schools (either cases or controls) if they were closer than 500 m. For each of the few isolated schools in the city (under a threshold distance of 500 m), the closest school from those available was defined as the unique neighbour. Finally, a factor representing school type was added to the model.

4. Results and discussion

This section contains one subsection for each of the statistical approaches chosen for this study. Each part begins by reporting the results obtained for the corresponding approach, indicating some technical issues, and then sets out the interpretations and implications in terms of traffic safety analysis that may be drawn from them.

4.1. Risk ratios by school type

The observed vs expected ratios strategy was used with buffer zones of various radii centred at all the school locations available in Valencia. Table 2 indicates the estimated risk ratios (observed/expected) that were overall achieved for each school type and for the complete set of schools, along with the statistical significance derived for them from 999 Monte Carlo simulations. No significant associations were found (at the 0.1 level) for the STW, although the estimated ratios are higher for Preschool and Secondary. On the other hand, several associations were determined to be significant for ETW, including the one considering all school typologies for the lowest value of σ. The lower sample of traffic accidents at STW probably reduced the statistical power of the test, and therefore the chances of observing statistically significant ratios.

Therefore, it can be concluded in the light of Table 2 that the contribution of school locations to traffic accidents in Valencia is not high in magnitude, although some of the associations found for ETW are quite notable. In this regard, Fig. 2b already suggested the high presence of traffic accidents within 50 m of school locations from 15:00 to 19:00.

4.2. Spatial count models

The analysis through CAR models revealed several associations between accident counts at the macroscopic level (hexagonal units) and the set of covariates, with little difference shown by the two time windows considered (see Fig. 3 for a graphical summary of the data). When the distances to each school type were considered, the differences between STW and ETW increased. Hence, the CAR model for STW revealed a negative association for the Primary type, although for ETW the same model indicated a positive association for the All-level type and a negative one for Preschool and Primary (Table 3). Here the signs of the coefficients estimated need to be understood differently given the nature of these distance-based covariates. Thus, higher distances to Preschool and Primary schools correlated with fewer traffic accidents, whereas the same situation was associated with more traffic accidents when considering All-level schools. In other words, a decrease in accident counts was attributed to the proximity of an All-level school (given these results), while an increase arose with the close presence of Preschools and Primary schools.

Main roads and higher bus stop density correlated with more traffic accidents for STW and ETW, as shown in Table 3. On the other hand, higher land use entropy was associated with fewer accidents (Table 3), possibly suggesting calming effects in traffic produced by the coexistence of different types of facilities. The positive association of traffic

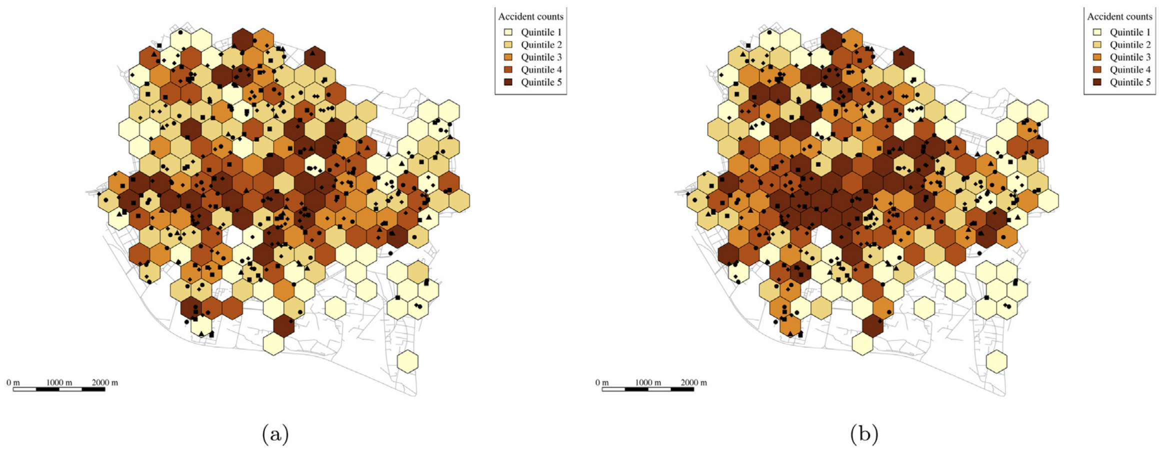


Fig. 3. Hexagonal grid coloured according to accident counts observed at STW (a) and ETW (b). The school locations are represented with black squares (All-level), circles (Primary), diamonds (Preschool) and triangles (Secondary). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

accidents with main roads and areas dense in bus stops is consistent with other studies (Yu, 2015; Yu and Zhu, 2016). However, the result for land use entropy is far more unexpected according to previous literature (Rothman et al., 2017b). Moreover, complex intersections were associated with more traffic accidents at STW (also showing a positive estimate for ETW), which seems entirely plausible given the natural increase in risk that these road entities produce (Miaou and Lord, 2003; Huang et al., 2017; Lee et al., 2017). The analysis of intersection characteristics and their relationship with actual and perceived risk for students is another topic of interest (Lee et al., 2016).

4.3. Multiple source regression

The model described in Eq. (3) was fitted for the four types of schools, which were taking the role of sources of putative risk, and the set of covariates considered. The optimization of the log-likelihood function associated with this model (Eq. (4)) required choosing some constraints for the parameters involved to reach convergence. The addition of covariates to the model was done successively in order to avoid overfitted models and to guide the choice of the constraints. The baseline model with no covariates was also tested. Finally, the constraints defined were $-1 \leq \lambda_m \leq 1$, $-1 \leq \alpha_j \leq 5$ and $0 \leq \beta_j \leq 2$. Furthermore, only the four covariates that showed a greater effect through the CAR modelling of accident counts were used in the final model: the number of bus stops, land use entropy, the number of complex intersections and main road length. Hence, a parsimonious criterion was followed, as the inclusion of more covariates barely increased the value of the log-likelihood function. The parameter estimates obtained for the multiple source regression model are shown in Table 4. The relative risk curves (as a function of the distance to each school type) derived from the estimates of the α and β parameters are shown in Fig. 4.

All the coefficients found for the multiple source regression are coherent overall with those obtained through the macroscopic CAR modelling. In particular, the differential effects produced by the two factors, school type and time window of analysis, appear again, as is evident from Fig. 4. The use of this modelling technique enables the risk to be characterized easily as a function of the distance to the source. For instance, according to the results obtained, the relative risk starts at a

Table 3
Estimates (Est.) and 5th, 10th, 90th and 95th percentiles of the posterior distributions of the parameters involved in the CAR model for both time windows considered (STW and ETW). Estimates shown in bold are significant with 90% credibility.

Covariate	STW					ETW				
	Est.	p_5	p_{10}	p_{90}	p_{95}	Est.	p_5	p_{10}	p_{90}	p_{95}
(Intercept)	**1.29**	1.19	1.21	1.36	1.38	**2.58**	2.51	2.53	2.63	2.65
No. of schools	−0.00	−0.11	−0.09	0.08	0.11	−0.06	−0.14	−0.13	0.01	0.03
Distance to closest All-level	0.13	−0.07	−0.03	0.28	0.33	**0.20**	0.06	0.09	0.31	0.34
Distance to closest Preschool	−0.10	−0.27	−0.23	0.03	0.07	**−0.17**	−0.29	−0.26	−0.07	−0.04
Distance to closest Primary	**−0.22**	−0.44	−0.39	−0.04	0.01	**−0.32**	−0.48	−0.44	−0.20	−0.16
Distance to closest Secondary	−0.04	−0.19	−0.16	0.08	0.11	−0.01	−0.12	−0.10	0.08	0.11
No. of educational services	−0.04	−0.20	−0.16	0.08	0.12	−0.01	−0.13	−0.10	0.09	0.11
No. of services (non-educational)	0.01	−0.18	−0.14	0.16	0.21	0.06	−0.09	−0.05	0.18	0.22
% of road with parking slots	0.01	−0.11	−0.08	0.10	0.12	0.06	−0.02	−0.01	0.13	0.15
No. of parking zones	0.01	−0.13	−0.10	0.12	0.16	0.04	−0.07	−0.05	0.12	0.14
No. of bus stops	**0.13**	0.01	0.04	0.22	0.25	**0.08**	−0.00	0.02	0.15	0.17
Land use entropy	**−0.11**	−0.22	−0.20	−0.03	−0.01	**−0.17**	−0.25	−0.23	−0.10	−0.09
Betweenness	−0.01	−0.12	−0.10	0.07	0.10	−0.01	−0.10	−0.08	0.05	0.07
No. of complex intersections	**0.13**	0.01	0.04	0.23	0.26	0.05	−0.05	−0.03	0.12	0.14
Main road length	**0.34**	0.20	0.23	0.45	0.48	**0.23**	0.12	0.15	0.31	0.33
No. of traffic lights	−0.04	−0.20	−0.16	0.08	0.12	−0.01	−0.13	−0.10	0.09	0.11
No. of school-aged residents	0.01	−0.12	−0.09	0.11	0.14	−0.07	−0.17	−0.15	0.00	0.03
Percentage of high-end cars	0.06	−0.06	−0.04	0.17	0.19	0.07	−0.03	−0.01	0.14	0.17
ψ	**3.34**	2.35	2.52	4.30	4.66	**4.15**	3.23	3.41	4.95	5.21

Table 4
Parameter estimates obtained for the multiple source regression models in the two time windows, STW and ETW. Parameters λ_1, λ_2, λ_3 and λ_4 were associated with the number of complex intersections, main road length, number of bus stops and land use entropy, respectively. The indexes for the α's and β's represent: 1 = All-level, 2 = Preschool, 3 = Primary, 4 = Secondary.

	λ_1	λ_2	λ_3	λ_4	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
STW	0.05	0.33	0.09	−0.13	−0.56	0.43	4.24	0.45	1.15	1.24	0.95	0.83
ETW	0.09	0.29	0.08	−0.18	−0.49	3.38	2.99	0.71	1.96	1.73	1.97	1.39

value slightly over 4 in the immediate vicinity of a Primary school for the ETW, and is then progressively reduced to 2 as the distance from the source reaches a value of 1.5 km.

4.4. Multivariate logistic regression

A total of 40 cases and 160 controls were established for the STW (Fig. 5a), and 45 cases and 155 controls for the ETW (Fig. 5b), following the Monte Carlo procedure explained in Section 3.3. Hence, a fair ratio of around 4 control schools for each school defined as a case were included in the multivariate logistic model. As indicated in Section 3.5, the choice of a radius $\sigma = 100$ m for the construction of the buffer zones was based on the availability of a controls/cases ratio that allows us to detect the differences between the two conditions. The consideration of a higher value of σ made the non-overlapping condition between buffer zones too restrictive to obtain a sufficient sample of both cases and controls, whereas a smaller value complicated particularly the identification of cases given the loss of sample size and statistical power.

Table 5 shows the results for the multivariate logistic regression fitted to the samples of cases and controls (for STW and ETW) and covariates available in a radius of 100 m around them. Contrary to the CAR modelling of accident counts over the hexagonal grid, the results from the case–control approach presented multiple differences between STW and ETW. Indeed, the number of complex intersections, the percentage of road length with parking slots, the number of parking zones and the percentage of high-end cars decreased the probability of a school being a case in STW, whereas the number of non-educational services and the number of bus stops had the same effect for the ETW. On the other side, the number of traffic lights was the only covariate that increased the probability of case in the STW, while main road length, the number of educational services and the percentage of high-end cars showed this behaviour for the ETW.

These findings will now be discussed in relation to a school location being unsafe (case) or not (control). The facility of parking (represented by both the percentage of road providing parking slots and the number of parking zones) indicated a protective effect in the STW. This is a reasonable result, as a lack of parking opportunities usually generates complicated and competitive traffic movements that may be responsible for an increase in traffic accidents. On the other hand, the number of parking zones seems to increase the risk in the ETW. The proximity of most of the parking zones in Valencia to shopping areas may be producing this effect in the ETW. In any case, one should not overlook the fact that establishing a correlation between parking difficulties and the number of vehicles arriving at a school at starting and ending times (possibly causing complex traffic dynamics that may lead to more accidents) is a hard task. Indeed, the scarcity of parking facilities leads many students' parents to reduce the use of their private vehicles to commute to school, but such reductions can also be seen as an incentive for other parents, generating some kind of equilibrium situation (Black et al., 2001).

Regarding the increase in risk suggested by main road length and the number of educational services in the ETW, both were expected. The first, because it usually correlates with more traffic accidents, regardless of the context being studied. The second, because these activities attract high numbers of commuters in the ETW, although the tendency to locate educational services near schools may be increasing this effect (which may also be responsible for the positive association of educational activities with traffic accidents in the STW).

The rest of the associations revealed by the case–control analysis are harder to interpret and in many cases require us to consider confounding factors that have not been included in the models due to their unavailability. For example, the number of complex intersections was

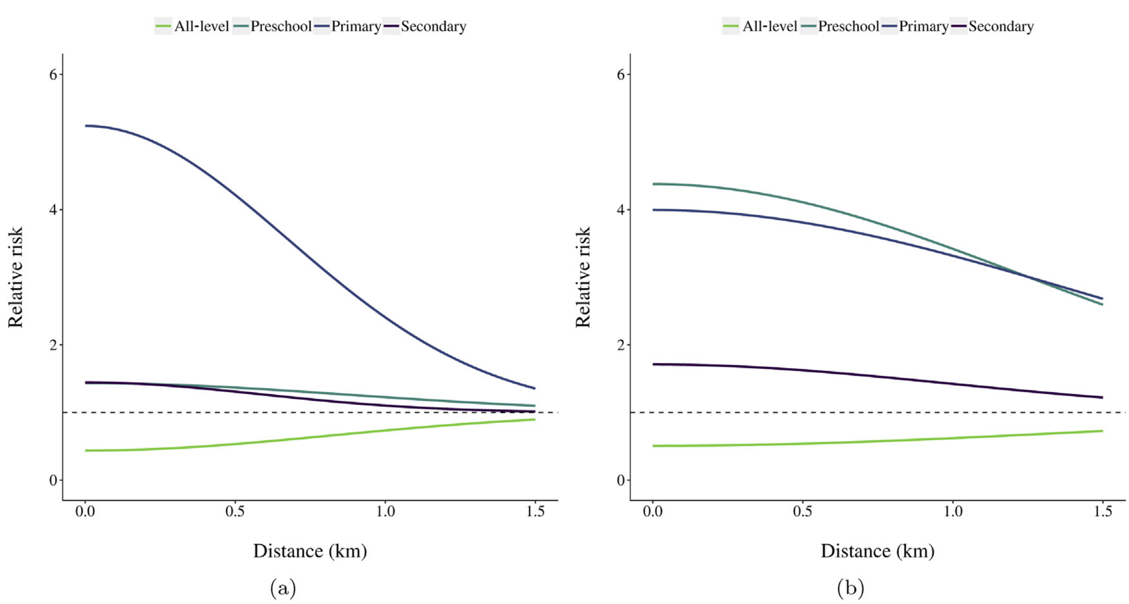
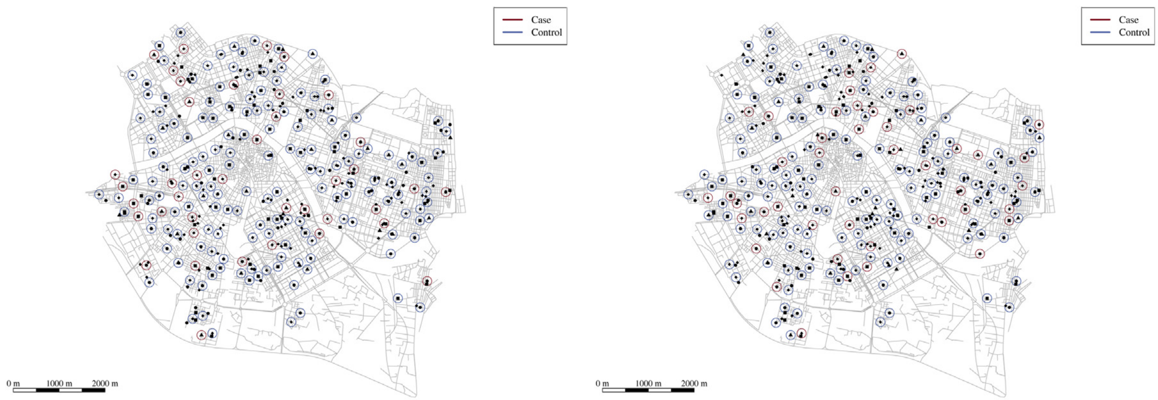


Fig. 4. Relative risk curves derived from the source regression model with covariates for the four types of schools available in Valencia in STW (a) and ETW (b). The black dashed line indicates a relative risk of 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) (b)

Fig. 5. Case and control buffer zones (100 m radius) defined around school locations in STW (a) and in ETW (b). The school locations are represented by black squares (All-level), circles (Primary), diamonds (Preschool) and triangles (Secondary). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

associated with a decrease in the probability of being a risky school zone, the opposite result to that found with the CAR model. It is logical that users drive more carefully in a small area (100 m buffer zone) which is dense in complex intersections, a fact that may explain this result. However, when a higher level of spatial aggregation is considered, as in the macroscopic modelling described in Section 3.4, a high number of complex intersections are more likely to be associated with an increase in traffic accidents.

The number of bus stops, which showed a positive correlation with traffic accidents according to the macroscopic modelling, was associated with a decrease in risk at the ETW. The fact of being a school located in a zone of the city endowed with several bus stops should reduce the number of commuters by private vehicle and therefore traffic congestion, although a higher number of bus stops does not always imply a higher level of connectivity within the public bus network. As in the case of complex intersections, at a higher level of spatial aggregation, it is more plausible that bus stop density should be associated with an increase in traffic dangerousness.

The behaviour of other covariates, such as the number of traffic lights or the number of non-educational services, may be a consequence of the presence of confounding effects. The number of traffic lights in the STW is possibly absorbing the more expected effect of main road length, as these two covariates are moderately correlated. On the other hand, the number of non-educational services is highly concentrated in the southern zone of the city centre and its contiguous district in this direction, which mostly presented control schools in the ETW, even though the two facts do not seem to be related.

Another covariate whose behaviour showed a strong dependence on the time window being considered was the percentage of high-end cars registered in the zone. This covariate was chosen as an approximation to socioeconomic status, given the unavailability of other sources of information (such as income level or housing price), yielding a positive association with traffic accidents only in ETW. Participating in extracurricular activities is much more frequent among students belonging to wealthier families (Leung et al., 2019), as most of these activities are not publicly financed. Consequently, more vehicles may be arriving at schools located in wealthier areas and afterwards transporting students to the location of the extracurricular activities. In contrast, students enrolled in schools situated in more economically depressed areas of the city are more likely to go home by their own means (under the plausible assumption that students' homes are on average closer to their school than extracurricular activities that do not take place at the same school). Nevertheless, these findings would require a deeper and more specific investigation to confirm this effect.

Table 5
Estimates (Est.) and 5th, 10th, 90th and 95th percentiles of the posterior distributions of the parameters involved in the logistic model for both time windows considered (STW and ETW). Estimates shown in bold are significant with 90% credibility.

Covariate	STW					ETW				
	Est.	p_5	p_{10}	p_{90}	p_{95}	Est.	p_5	p_{10}	p_{90}	p_{95}
(Intercept)	**−2.13**	−2.90	−2.71	−1.57	−1.43	**−0.95**	−1.64	−1.47	−0.44	−0.30
No. of educational services	**0.30**	−0.04	0.03	0.57	0.65	**0.66**	0.24	0.33	0.99	1.09
No. of services (non-educational)	0.16	−0.51	−0.35	0.66	0.79	**−1.04**	−1.80	−1.61	−0.51	−0.38
% of road with parking slots	**−0.36**	−0.75	−0.66	−0.08	−0.00	0.07	−0.29	−0.21	0.36	0.44
No. of parking zones	**−0.47**	−0.98	−0.85	−0.12	−0.05	**0.30**	−0.01	0.05	0.56	0.64
No. of bus stops	0.14	−0.20	−0.12	0.40	0.48	−0.23	−0.61	−0.52	0.05	0.12
Land use entropy	**−0.40**	−0.81	−0.71	−0.09	−0.00	0.01	−0.38	−0.29	0.32	0.40
Betweenness	0.05	−0.30	−0.22	0.31	0.37	0.08	−0.31	−0.21	0.37	0.44
No. of complex intersections	**−0.49**	−0.92	−0.82	−0.17	−0.09	0.09	−0.24	−0.17	0.34	0.41
Main road length	0.02	−0.39	−0.29	0.33	0.42	**0.35**	−0.04	0.05	0.65	0.74
No. of traffic lights	**0.58**	0.17	0.25	0.92	1.01	0.25	−0.16	−0.07	0.57	0.66
No. of school-aged residents	0.09	−0.32	−0.23	0.40	0.48	0.11	−0.33	−0.24	0.45	0.54
Percentage of high-end cars	−0.16	−0.59	−0.48	0.14	0.22	0.25	−0.11	−0.02	0.52	0.59
School type (Preschool)	0.17	−0.71	−0.52	0.87	1.08	−0.51	−1.36	−1.17	0.17	0.36
School type (Primary)	**1.11**	0.15	0.36	1.87	2.10	−0.48	−1.40	−1.20	0.23	0.43
School type (Secondary)	**1.14**	−0.01	0.24	2.05	2.31	−0.63	−1.94	−1.63	0.36	0.62

Finally, the categorical covariate adding the effect of school type presented different behaviour between the STW and the ETW, considering that in Table 5 the estimates are computed in relation to the All-level type (which is hidden because it is the reference category). In the case of the STW, Primary and Secondary types showed a significant increase in the odds ratios for being a case, in agreement with the other models tested (especially in the case of Primary schools). However, all the estimates for the ETW were not significant with 90% credibility, a surprising result in view of the previous models.

5. Conclusions

In this paper, we have followed several approaches to measure the effect of school locations on traffic accidents in Valencia. From a methodological perspective, the main conclusion is the desirability of applying different methods to obtain consistent knowledge about the phenomenon of interest. All the statistical techniques chosen for accomplishing the research purposes established have provided information about the question of interest from different perspectives, providing several points of agreement between them. However, each technique has offered a particular view of the phenomenon, highlighting the necessity of carefully choosing a specific approach and even the advisability of using more than one in order to strengthen our findings.

First, observed vs expected ratios of traffic accidents were computed at a range of distances from school locations. Although this approach has the disadvantage of not allowing the use of any auxiliary information (covariates), it serves to get an overall idea of the magnitude of the effect being analyzed. In this study, it was a starting point which suggested that school locations have an impact on traffic accidents, albeit only a moderate impact. Furthermore, the observed vs expected ratios also made it possible to define case and control schools for the subsequent performance of a case–control study design. In the absence of external information (if no particular school zone has been declared dangerous by traffic experts), this simple method can fill the gap.

Besides observed vs expected ratios, three other statistical models were used in the analysis: macroscopic CAR modelling, multiple source regression and a logistic regression under a case–control design. The macroscopic modelling of accident counts indicated that some environmental and traffic-related factors, such as the number of bus stops, land use entropy, main road length and the number of complex intersections, have an effect on the incidence of traffic accidents in the two time windows investigated within school-days. Furthermore, this model enabled us to confirm that in the city of Valencia the type of school entails a specific risk with respect to traffic accident occurrence. The last statement was also confirmed through the multiple source regression model. According to both modelling approaches, the All-level school type shows an overall protective effect against traffic accidents in the city of Valencia. On the other hand, proximity to Primary (in both time windows studied) and Preschool schools (in the ETW) is subject to more risk.

Regarding the case–control study, the logistic models showed substantially different results for the STW and the ETW. For instance, this modelling approach does not detect any significant effect from school types for the ETW, a result that is inconsistent with the other methods employed. It is worth of noting that the case–control study design may be too sensitive to the choice of cases and controls and the results should be interpreted with caution, especially if the methodology for selecting both cases and controls is not supported by other evidence apart from the not large sample of traffic accidents available. Furthermore, it needs to be remembered that each model type relies on a different type of spatial unit of analysis (arbitrary hexagonal units covering the city vs buffer zones around schools), a fact that is likely to give rise to MAUP effects. Hence, the scope and extent of each approach is different and direct comparisons are not completely suitable. In any case, the results suggest that this approach could be powerful for detecting small effects that a macroscopic modelling of accident counts may otherwise miss. One drawback of this approach, in comparison to the former, is the likely possibility of dealing with small samples of school locations that can make difficult, or even impossible, the achievement of reliable parameter estimates.

From a practical perspective, within the context of traffic safety around school locations, we can highlight two important results that are now discussed: the need to distinguish between coexisting school types and the consideration of time windows other than those that strictly correspond to school starting and ending times. With regard to school types, and getting back to the fact that All-level schools have been shown to be a safer context for vehicle traffic, it is worth noting that the differences between school types coexisting in Valencia go beyond the age levels. Indeed, most of the All-level schools in Valencia are semi-private, and parents of students enrolled in this kind of school have, on average, higher income levels and educational attainment than parents of students registered at public institutions (Llera and Pérez, 2012), a category that includes the vast majority of Primary and Secondary schools in Valencia. Furthermore, it needs to be remarked that it is a very challenging task to disentangle the origin of the protective effect found for All-level schools, meaning that it is unclear whether commuting time is particularly dangerous for certain school surroundings given the travel mode choices or travel behaviours of their students (or of the adults in charge) or, alternatively, whether some schools (or school types) are located in intrinsically riskier/safer areas for driving. Although we are aware that All-level schools in Valencia correlate with higher socioeconomic status of students, we were not able to extract more conclusions in this regard, in the absence of more precise information on commuters' travel mode choices and attitudes towards safe driving, walking or cycling. We could only guess that more educated parents will be more predisposed towards safety (Murray, 1998). The place of residence and socioeconomic characteristics of students belonging to each school, which are sometimes used to make approximations about travel modes (Wilson et al., 2010; Kelly and Fu, 2014), were not available either. In any case, even if we had been in possession of this information, the inference of travel modes from students' characteristics could have been highly uncertain, as Valencia is quite different from the cities where these studies were developed.

On the other hand, it is remarkable that the evening time window is often overlooked in the study of traffic safety around school locations. Many of the research papers available focus on the first hours in the morning, but we think that traffic generated in the evening as a consequence of collecting students and taking them to leisure activities deserves more attention. Indeed, several papers that have compared school and non-school trips (those related to extracurricular or leisure activities) have coincided in pointing out that the use of private vehicles is much more likely in non-school trips (Hjorthol and Fyhri, 2009; Fyhri and Hjorthol, 2009; Park et al., 2018), probably augmenting traffic volume. The initial exploratory analysis and the computation of risk ratios near schools already suggested this situation, which was confirmed via the macroscopic modelling and the multiple source regression with the detection of significant effects for most of the school types.

Finally, we need to mention the main limitations of this study, which are a consequence of data unavailability. Besides the absence of information regarding travel mode choices, there was also a lack of data regarding the trip destination of vehicles involved in the traffic accidents employed for the analysis and of the age of the people travelling in them. In other words, an unproved causal relationship had to be established: traffic accidents that occurred in proximity to schools in the STW or ETW were assumed to be a consequence of the traffic dynamics generated by people commuting to school. Of course, these accidents could involve the commuters themselves or other traffic users with a non-school destination not influenced by school commuters. Furthermore, one cannot overlook the possibility that traffic accidents temporally associated with school starting and ending times do not occur as close to school locations as one might expect, a possibility suggested by Kingham et al. (2011), which makes this kind of research very challenging.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the Local Police Department and the Statistics Office of València for providing the data that was used in the study.

References

Abdel-Aty, M., Chundi, S.S., Lee, C., 2007. Geo-spatial and log-linear analysis of pedestrian and bicyclist crashes involving school-aged children. J. Saf. Res. 38 (5), 571–579.
Amoh-Gyimah, R., Saberi, M., Sarvi, M., 2017. The effect of variations in spatial units on unobserved heterogeneity in macroscopic crash models. Anal. Methods Accid. Res. 13, 28–51.
Baddeley, A., Rubak, E., Turner, R., 2015. Spatial Point Patterns: Methodology and Applications with R. CRC Press.
Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Stat. Soc. Ser. B (Methodol.) 192–236.
Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Stat. Math. 43 (1), 1–20.
Bivand, R., Piras, G., 2015. Comparing implementations of estimation methods for spatial econometrics. J. Stat. Softw. 63 (18), 1–36.
Bivand, R., Rundel, C., 2018. rgeos: Interface to Geometry Engine – Open Source (‘GEOS’). R Package Version 0.3-28.
Black, C., Collins, A., Snell, M., 2001. Encouraging walking: the case of journey-to-school trips in compact urban areas. Urban Stud. 38 (7), 1121–1141.
Clifton, K.J., Kreamer-Fults, K., 2007. An examination of the environmental attributes associated with pedestrian–vehicular crashes near public schools. Accid. Anal. Prev. 39 (4), 708–715.
Diggle, P., Morris, S., Elliott, P., Shaddick, G., 1997. Regression modelling of disease risk in relation to point sources. J. R. Stat. Soc. Ser. A (Stat. Soc.) 160 (3), 491–505.
Diggle, P.J., Rowlingson, B.S., 1994. A conditional approach to point process modelling of elevated risk. J. R. Stat. Soc. Ser. A (Stat. Soc.) 157 (3), 433–440.
Freeman, L.C., 1977. A set of measures of centrality based on betweenness. Sociometry 35–41.
Fyhri, A., Hjorthol, R., 2009. Children's independent mobility to school, friends and leisure activities. J. Transp. Geogr. 17 (5), 377–384.
Hjorthol, R., Fyhri, A., 2009. Do organized leisure activities for children encourage car-use? Transp. Res. Part A: Policy Pract. 43 (2), 209–218.
Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-level crash risk analysis in Florida: Bayesian spatial modeling. Transp. Res. Rec.: J. Transp. Res. Board 2148, 27–37.
Huang, H., Zhou, H., Wang, J., Chang, F., Ma, M., 2017. A multivariate spatial model of crash frequency by transportation modes for urban intersections. Anal. Methods Accid. Res. 14, 10–21.
Hwang, J., Joh, K., Woo, A., 2017. Social inequalities in child pedestrian traffic injuries: differences in neighborhood built environments near schools in Austin, TX, USA. J. Transp. Health 6, 40–49.
Kelly, J.A., Fu, M., 2014. Sustainable school commuting-understanding choices and identifying opportunities: a case study in Dublin, Ireland. J. Transp. Geogr. 34, 221–230.
Kingham, S., Sabel, C.E., Bartie, P., 2011. The impact of the ‘school run’on road traffic accidents: a spatio-temporal analysis. J. Transp. Geogr. 19 (4), 705–711.
Lee, G., Park, Y., Kim, J., Cho, G.-H., 2016. Association between intersection characteristics and perceived crash risk among school-aged children. Accid. Anal. Prev. 97, 111–121.
Lee, J., Abdel-Aty, M., Cai, Q., 2017. Intersection crash prediction modeling with macro-level data from various geographic units. Accid. Anal. Prev. 102, 213–226.
Leung, K.Y., Astroza, S., Loo, B.P., Bhat, C.R., 2019. An environment–people interactions framework for analysing children's extra-curricular activities and active transport. J. Transp. Geogr. 74, 341–358.
Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. 63 (19).
Llera, R.F., Pérez, M.M., 2012. Colegios concertados y selección de escuela en España: un círculo vicioso. Presup. Gasto Público 67, 97–118.
Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: new features. Comput. Stat. Data Anal. 67, 68–83.
Miaou, S.-P., Lord, D., 2003. Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. Transp. Res. Rec.: J. Transp. Res. Board (1840), 31–40.
Mullen, K.M., Ardia, D., Gil, D.L., Windover, D., Cline, J., et al., 2011. DEoptim: an R package for global optimization by differential evolution. J. Stat. Softw. 40 (i06).
Murray, Å., 1998. The home and school background of young drivers involved in traffic accidents. Accid. Anal. Prev. 30 (2), 169–182.
Park, J., Abdel-Aty, M., Lee, J., 2018. School zone safety modeling in countermeasure evaluation and decision. Transportmetrica A: Transp. Sci. 1–16.
Price, K., Storn, R.M., Lampinen, J.A., 2006. Differential Evolution: A Practical Approach to Global Optimization. Springer Science & Business Media.
Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accid. Anal. Prev. 40 (4), 1486–1497.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Ramis, R., Diggle, P., Cambra, K., López-Abente, G., 2011. Prostate cancer and industrial pollution: risk around putative focus in a multi-source scenario. Environ. Int. 37 (3), 577–585.
Reeve, N.F., Fanshawe, T.R., Keegan, T.J., Stewart, A.G., Diggle, P.J., 2013. Spatial analysis of health effects of large industrial incinerators in England, 1998–2008: a study using matched case–control areas. BMJ Open 3 (1), e001847.
Rothman, L., Buliung, R., Howard, A., Macarthur, C., Macpherson, A., 2017a. The school environment and student car drop-off at elementary schools. Travel Behav. Soc. 9, 50–57.
Rothman, L., Howard, A., Buliung, R., Macarthur, C., Richmond, S.A., Macpherson, A., 2017b. School environments and social risk factors for child pedestrian-motor vehicle collisions: a case–control study. Accid. Anal. Prev. 98, 252–258.
Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 71 (2), 319–392.
Warsh, J., Rothman, L., Slater, M., Steverango, C., Howard, A., 2009. Are school zones effective? An examination of motor vehicle versus child pedestrian crashes near schools. Inj. Prev. 15 (4), 226–229.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
Wilson, E.J., Marshall, J., Wilson, R., Krizek, K.J., 2010. By foot, bus or car: children's school travel and school choice policy. Environ. Plan. A 42 (9), 2168–2185.
Xu, P., Huang, H., Dong, N., 2018. The modifiable areal unit problem in traffic safety: basic issue, potential solutions and future research. J. Traffic Transp. Eng. (Engl. Ed.) 5 (1), 73–82.
Yu, C.-Y., 2015. How differences in roadways affect school travel safety. J. Am. Plan. Assoc. 81 (3), 203–220.
Yu, C.-Y., Zhu, X., 2016. Planning for safe schools: impacts of school siting and surrounding environments on traffic safety. J. Plan. Educ. Res. 36 (4), 476–486.
Zhai, X., Huang, H., Xu, P., Sze, N., 2018. The influence of zonal configurations on macro-level crash modeling. Transportmetrica A: Transp. Sci. 1–18.