Identification of Hazardous Road Locations in Urban Areas through Spatial Analysis: Application to Vila Real

Inês Franco Moreira

Instituto Superior Técnico, Lisbon, Portugal, ines.moreira@ist.utl.pt

Abstract

In Portugal, statistics show that injury road accidents causing considerable human and material losses are more frequent (163% higher in 2010) in urban areas than outside them. An effective approach to reduce the urban road accidents is the site action, which treats sites where these tend to be aggregated: the hazardous road locations (HRL).

This work analyses the injury road accidents that occurred during the 2004-2008 period in Vila Real. Road accident data was collected from the national database maintained by Autoridade Nacional de Segurança Rodoviária, the Portuguese national institution for road safety, and completed with information kept by the local police forces. The aim of this research project is to compare and discuss the applicability of different methods and definitions for HRL identification and their respective results. The choices of methods and definitions follow different criteria and consider the current availability of data and resources of Portuguese municipalities.

Three types of methods were applied. These were based on: (i) definitions applied in European countries; (ii) Euclidean distance between points; and (iii) density functions. Analyses were carried out using ArcGIS™ and CrimeStat® softwares. Different methods and parameterisation led to the identification of different HRL both in number and location; however, a considerable number of HRL was regularly identified using different methods and/or parameter values. The adequacy and effectiveness of the diverse methods is discussed, as well as their advantages and limitations.

Finally, one HRL identified in Vila Real was selected and analysed in greater detail to inspect site-related problems, which could be mitigated through the implementation of engineering measures.

1. Introduction

Statistics show that the number of road accidents in Portugal is still unacceptable, especially when compared with other countries (WHO, 2004; EC, 2011). In 2010, 35 426 injury road accidents were recorded in Portuguese roads causing more than 47 000 casualties (ANSR, 2011). About 69% of those casualties occurred in urban areas (ANSR, 2011). Over the last decade in Portugal, a reduction of the casualties was observed which was larger outside urban areas than inside them (ETSC, 2010; OECD/ITF, 2011).

An efficient approach to reduce road accidents as well as their casualties and severity, in particular inside urban areas, is the treatment of locations(s) where road accidents tend to aggregate (the hazardous road locations - HRL) through the implementation of corrective road safety measures (IHT, 1990), which often include low cost engineering measures. Despite its effectiveness, this approach has yet hardly been implemented in Portugal.

This work intends to contribute with a survey on methods and definitions that enable the identification of HRL in Portuguese urban areas. The identified locations can be then analysed, treated and monitored helping the achievement of the National Strategy targets and the European commitments (ANSR, 2008; EC, 2010).

The main objective of this research project is the comparison of different methods of identification of HRL in urban areas with spatial analysis software taking into consideration the current conditions of available data and resources in the Portuguese local authorities – Câmaras Municipais (CM). It focuses on the comparison of different methods to identify HRL from a practical point of view, and explores their possibilities for application in other Portuguese cities as well as the city where the case study was undertaken.

2. Analysis of urban road accidents

One of the goals of urban road accident analysis is to provide working elements (including diagrams, maps and identification of factors) to the technicians with responsibilities in the management of road safety, in order to support interventions to reduce the number of such accidents and their severity.

2.1. Road accidents

The definition of road accident is not universal (OECD, 1998). In this study, only injury accidents are considered, as they are mandatorily reported to the authorities and are recorded in the Portuguese national database. Road accidents are rare events on public roads involving at least one moving vehicle, registered by the authorities, which leads at least to a personal injury occurrence. Such events can be divided into fatal, serious or slight according to the severity of the most severe casualty.

Road accidents are considered as random, rare and independent events. It is impossible to predict under which circumstances a single accident will occur (Brijs et al., 2007) and their occurrence in a specific location does not influence
the occurrence of other accidents in other places or moments. However, the concentration of accidents in the same site might indicate a spatial relation between accidents and the environment and/or road conditions (Levine et al., 1995a; Elsenaar and Abouraad, 2005; Geurts et al., 2005a; Elvik, 2006).

2.2. Approaches to decrease urban road accidents

The spatial distribution of accidents is not uniform along the road network. According to their spatial distribution, different approaches to decrease their number as well as their severity have been conceived considering the spatial distribution of road accidents. According to the Chartered Institute of Highways and Transportation (1990) there are four main types of approaches: (i) **Site action** – treatment of a specific location at which road accidents are clustered; (ii) **Mass action** – application of a road safety treatment procedure to various locations which share common accident factors; (iii) **Route action** – application of road safety measures to a particular type of road or road section (usually considering 0.5 to 1.5 km sections) where accidents are especially accumulated; and (iv) **Area action** – application of a zoning in areas where the most hazardous zones (e.g. with higher accident rates) are thoroughly analysed; road measures are subsequently proposed and implemented taking into consideration the entire zone.

This research project focuses on the **Site action** approach only. It is divided into the following stages: (i) establishment of the study objectives; (ii) identification of HRL (such as road sections or intersections); (iii) ranking and selection of HRL for subsequent analysis; (iv) accident analysis within the selected hazardous road elements in order to understand and identify the contributing factors to road accident occurrence; and (v) proposal, implementation and evaluation of adequate corrective road safety measures.

Although this approach is considered to be effective and has produced large reductions in the number of accidents (IHT, 1990; Heydecker and Wu, 2001; Elsenaar and Abouraad, 2005), it presents some problems such as the regression to the mean effect, a phenomenon related to the fluctuation on the number of accidents explained by the accident random effect, a phenomenon related to the fluctuation on

2.3. Data

Within this analytical process, information that characterises the local environment and traffic as well as accident data play an important role. Indeed, according to PIARC (2007) three levels of data can be considered:

- Road accident data (e.g. accident identification, time, location, accident type, number of casualties). This data is key for the analysis and should, when available, be complemented with;
- Road and traffic data (e.g. road description, location of crosswalks, parking places, road surface, roadside obstacles and traffic flows); and
- Other data (e.g. driver, pedestrian and passenger information: gender, age, alcohol rate).

2.4. Identification of Hazardous Road Locations

The investigation of factors which lead to accidents’ clusters is an important topic in the road safety activity (Moons et al., 2009). According to Elsenaar and Abouraad (2005) a high concentration of accidents at the same location is a strong indicator of a road related problem that should be investigated.

While there is no internationally accepted definition of HRL (Geurts and Wets, 2003; Geurts et al., 2005a; Brimicombe 2004 cit. in Anderson, 2007; Cheng and Washington, 2008; Anderson, 2009) several different methods have been applied to identify these locations (Retting et al., 2001; Hauer et al., 2002a; Fiafaut et al., 2003; Brijs et al., 2007; Elvik, 2008b; Pirdavani et al., 2010).

Elvik (2007), based on an OECD report, proposed a distinction between the different definitions of HRL into **numerical, statistical and model-based definitions**. Generally, most of the identification through **numerical definitions** can be applied with information only about the road network and the road accidents, in particular their location. In general, those definitions identify the places where the recorded accident number or rate (e.g. accidents per traffic flow) during a period, is higher than a threshold. Examples of numerical definitions are presented in Table 2.1. The remaining types of definitions require complementing information, such as traffic data, which is often not available, especially in countries where road safety programmes have not been recurrently applied (Sørensen, 2007). For this reason, statistical and model-based definitions are out of the scope of this work.

Besides the abovementioned definitions, other practical methods have been used with the purpose of identifying HRL (Retting et al., 2001; Mungnimit et al.2009; Xie and Yan, 2008). Examples are the **Kernel Density Estimation** and the nearest-neighbour methods (Anderson, 2009; Nicholson, 1998).

To apply these definitions and/or methods, the use of spatial analysis software in the HRL identification is recommended (Elsenaar and Abouraad, 2005). From the literature and among the several options available (TRL, 2011; TES, 2011; Jarvis and Kamal, 2009; BioMedware, 2010), two possible options are:

- **ArcGIS™** - A widely used multipurpose proprietary geographic information system (GIS) with a
comprehensive range of spatial analysis tools and useful data processing functionalities. Examples of the implementation of this software for HRL identification are described by Anderson (2007) and Xie and Yan (2008);

- CrimeStat® – A spatial statistics software that was created to analyze crime incidents (Levine, 2010), but it has been applied to other domains, in particular in road safety. According to Erdogan et al. (2008), it is an extensive tool for HRL identification and an attractive alternative to “traditional” GIS. It is freely downloadable from the Internet.

<table>
<thead>
<tr>
<th>Definition</th>
<th>ANSR</th>
<th>Belgium</th>
<th>Flanders</th>
<th>Hungary</th>
<th>Germany</th>
<th>Norway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period of records</td>
<td>1 year</td>
<td>Over a year</td>
<td>3 years</td>
<td>3 years</td>
<td>3 years</td>
<td>5 years</td>
</tr>
<tr>
<td>Location (road extension)</td>
<td>100 m</td>
<td>50 m</td>
<td>50 m</td>
<td>50 m</td>
<td>50 m</td>
<td>50 m</td>
</tr>
<tr>
<td>Severity index (SI)</td>
<td>SI=100F+10SerI+3Sill</td>
<td>SI=5F+3SerI+Sill</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Criterion 1: No. of accidents</td>
<td>≥ 5</td>
<td>≥ 3 serious</td>
<td>≥ 3</td>
<td>≥ 4</td>
<td>≥ 5 with casualties or ≥ 3 serious</td>
<td>≥ 4</td>
</tr>
<tr>
<td>Criterion 2: Severity index</td>
<td>$\sum$ SI &gt; 20</td>
<td>-</td>
<td>$\sum$ SI ≥ 15</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Source: ANSR, 2011; Flahaut et al., 2003; Elvik, 2007; Geurts et al., 2006; Elvik, 2007; Elvik, 2007; 2008b; Elvik, 2007; 2008b

Legend: F – fatal injury, SerI – seriously injured, Sill – Slightly injured

2.5. Ranking and selection of a HRL

HRL ranking, usually according to technical criteria, has the purpose to prioritize places for further analysis and implementation of measures. Some of the possible ranking criteria for HRL are accident frequency, accident rate, proportion of accident types considered susceptible for treatment and severity of casualties.

Such ranking is followed by the selection of HRL, often a decision made by local authorities. In this case, the selection can be based on different criteria (apart from technical criteria), such as political, economical or logistical aspects.

2.6. Analysis of road accidents

Accident analysis within each HRL has the purpose of understanding what causes are in the origin of road accidents, which might be suitable for treatment. Several inspections or analyses can be carried out in this stage, within a site action approach: examples are the site visits or the production and analysis of stick diagrams which identify common patterns among accidents that occurred in the same place, supporting the search for common problems between accidents and often contributing to the elaboration of countermeasures. It is, however, important to state that accident analysis may not lead to the finding of common contributory factors to road accidents in a HRL. This might be due to the intrinsic nature of accidents, most probably from the existing fluctuation of variables taken from events considered as random.

2.7. Proposal, implementation and evaluation of mitigating measures

Following the analysis of road accidents registered in each HRL, corrective measures can be proposed. Different causes of road accidents require different road safety measure types, including enforcement, education, encouragement and engineering measures. A comprehensive description of several measures can be found in Elvik and Vaa (2004).

2.8. Urban road accidents in Portugal

The current National Road Safety Strategy in Portugal recommends a special focus on particular groups, such as accidents in urban areas (ANSR, 2008), aiming to achieve European target levels. The Portuguese local authorities, Câmaras Municipais (CM), include in their competences the analysis of road accidents and the subsequent implementation of measures to decrease the number and severity of urban road accidents.

2.8.1. Road accident data collection in urban roads

According to the Portuguese law, all injury road accidents must be recorded by the police (Portuguese Government, 2001). There are two police forces responsible for this role: Polícia de Segurança Pública (PSP), Guarda Nacional Republicana (GNR). Collection of road accident data is compulsory for injury road accidents. For each injury road accident two written reports are filled by the police officer:

- Boletim estatístico de acidentes de viação (BEAV) – Contains generic information about the road accident, (e.g. type of accident, moment of occurrence, street name, vehicles involved, information of drivers, passengers and...
pedestrians without private data). BEAVs are subsequently sent to Autoridade Nacional de Segurança Rodoviária (ANSR), the Portuguese national authority for road safety responsible for supporting the governmental road safety policies; and

- Participação de acidentes de viação (PAV) – mainly contains information about location and moment of accident occurrence, drivers’ personal information specific information about the vehicle an accident description, possible causes and a sketch of the accident (which is a crucial information to determine the exact location of the road accident). These reports are kept by the police forces which filled them. To access this information a request should be sent to the Comissão de Acesso aos Documentos Administrativos (CADA), an independent public entity which regulates access to administrative documents.

2.8.2. Road safety policies in Portugal

In the last decade, the number of road accidents and casualties has decreased in Portugal due to the implementation of several complementary measures and policies, as well as improvements on national road infrastructure and technological advances in terms of vehicle safety (Simão, 2010).

The current National Road Safety Strategy recommends the creation of Municipal Road Safety Plans emphasising the responsibility of local authorities in the treatment of urban road safety problems. However, few municipalities have so far created those plans. Despite recommendations from the National Road Safety Strategy, specific urban road safety actions and programmes at a local level have not been widely put into practice in urban areas and only sporadic road safety actions in urban areas were carried out. Other initiatives have also been promoted as a way for stimulating works in the road safety field, such as road safety campaigns.

Some urban road safety research projects have been developed in or about Portugal (Rodrigues, 1997; Costa, 2002; Costa, 2005; Rebelo, 2007): concerning specifically HRL identification, some research studies have been carried out, particularly on motorways and national roads (Guerreiro, 2008; Cunha 2009; Jordão 2008). Marrana (1996) proposed a methodology for HRL identification based on the total number of accidents in each road segment using a GIS.

3. Methodology

The proposed methodology has the main purpose of identifying HRL, using the limited resources and information usually available to local authorities. Moreover, it intends to be easily applicable to any urban area in Portugal in short term, using generically available data (accident and road data) and software (GIS or other programs with analytical capabilities for spatial data).

3.1. Delimitation of the urban area and data collection

The first step of this methodology is the delimitation of the area under analysis. Such area might correspond to the following examples: (i) the entire municipal urban area; (ii) the consolidated urban area; (iii) the corresponding urban area in the road accidents data collection. For this, there is an assumption on the availability of administrative boundaries, soil use or other spatial data that applies.

Besides the crucial accident information, several types of information are suggested in the process of road accident analysis (LAA, 1989), such as road traffic data and geographic and demographic data. However, even in the case of unavailability of these datasets, the methodology may be applied using only basic resources, as road accident data (from BEAV and PAV) and the road network map are the truly indispensable datasets. To collect this information requests to ANSR and CADA are necessary (section 2.8.1). This information is then treated and georeferenced.

Road network should be analysed in order to establish its functional road hierarchy, as it might have an important role in the problem diagnosis and in the recommendation of solutions related with the enhancement of road safety. Usually, this procedure is defined based on the traffic and pedestrian information and also the knowledge of experts, independently of the georeferencing process. However, if the traffic and pedestrian figures are not available, it can be established based only on experts’ knowledge.

3.2. Descriptive analysis

A descriptive statistical analysis of injury road accidents has the purpose to characterise the main trends of data, such as the main conditions of occurrence of most road accidents and the main characteristics of the individuals involved on them. This analysis identifies trends that might be utilised in the spatial analysis in greater detail.

3.3. Identification of HRL

In this methodology, the identification of HRL consists in the implementation of three types of methods on the datasets: (i) methods based on definitions applied in some European countries; (ii) methods based on the Euclidean distance between points; (iii) methods based on the point density.

3.3.1. Methods based on definitions applied in some European countries

These methods have the purpose to reproduce numerical definitions (see section 2.4) applied in some European countries to identify hazardous road locations. Chosen definitions were those from: Belgium, Flanders, Germany, Hungary, Norway and Portugal (ANSR), as these are not based on traffic flows (as they might not be available in Portuguese municipalities), consider the universe of injury road accidents and provide a noteworthy differentiation between them (see Table 2.1). To replicate these definitions, two methods were created, here named I and II.
On Method I the number of recorded injury road accident data in each definition is previously selected; for each road accident a severity index (SI) is calculated (according to the definition); a radius is fixed, corresponding to half of the road section length present in the definition; the centre of the circle is applied to the entire urban area according to the cells of a grid, with a specific spatial resolution. Then, two calculations are executed: (i) the number of accidents in each circle and (ii) the sum of the severity index in each circle, resulting in two grids. These are then aggregated and a reclassification is made according to the criteria of the definition in practice to identify the HRL.

On Method II the number of recorded injury road accident data in each definition is previously selected; for each road accident a severity index (SI) is calculated (according to the definition); over the line segments of the road network points regularly distanced are placed; a fixed radius, corresponding to half of the section length of road present in the definition, is considered; a circle with this radius is centred over the regularly distanced road network points; in each circle road accidents are counted and the values of the severity index are summarised; finally the areas within the circles which verify the correspondent criteria of minimum number of road accidents and sum of the severity index given by the definition in use are selected as being HRL.

The definitions which use the definition of serious road accident were applied according to the definition adopted by ANSR (ANSR, 2011). Relatively to the Belgian definition, two periods of records were considered: one and five years as that corresponds to the smallest period allowed by this definition and the larger period, taking into consideration the available dataset.

3.3.2. Methods based on the Euclidean distance between points

This type of methods encompasses those that only require the Euclidean distance between points (road accidents) to identify locations where points are spatially aggregated. Examples of these methods are:

- **Count coincident points** – when there are more than a specified number of (road accident) points spatially coincident, such location is considered as a HRL;
- **Count points within a neighbourhood of each point** – (road accident) points within a neighbourhood area of a road accident location (point) are counted; areas where the number exceed threshold are selected as HRL; the procedure is repeated for all road accident locations.
- **Identify point groups which are spatially close** – the Euclidean distance between points (road accidents) is compared to a threshold distance; pairs of points closer than this distance are aggregated and define a spatial cluster); point aggregations are then selected according to an established minimum number of elements. This minimum value does not interfere with the shape or size of point aggregations which are dependent on the value of the threshold distance. The threshold distance can be either defined by the analyst or based on a random distribution of points for which the area and number of points are equal to the case study.

3.3.3. Methods based in density functions

These methods are based on a point density function, which is calculated for all the cells of a grid (after setting a spatial resolution value) covering the entire area under analysis. Cells with point density values higher than a cut-off point density value, are identified as HRL. This density value is chosen by the analyst and can be adopted according to the size of the identified HRL (e.g., it can be determined according with the maximum distance between the HRL boundary from a road network centreline).

To represent this type of methods, this work adopted the **Kernel density estimation** (KDE). Its application requires the following steps: (i) definition of the extension (area under study) and spatial resolution of the grid; (ii) selection of a density function (e.g.: based on the Normal distribution) and of a bandwidth value; (iii) “application” of a symmetrical surface (defined by the density function selected) centred on each injury road accident location point which is dependent from the bandwidth value; (iv) calculation, for each cell on the grid, of the sum of the values of the abovementioned density surfaces (placed over each point) producing a single surface (kernel density surface) covering the area under study. After this calculation the cells with density over a cut-off value are selected as being HRL.

3.4. Choice of hazardous road location for detailed analysis

After the identification of HRL, these can be ranked in order to prioritise the locations for detailed analysis and subsequently for treatment. Some of the adopted technical criteria to rank the identified HRL could be: (i) the coincident identification by different methods; (ii) the number of road accidents that each HRL contained (accident frequency).

3.5. Detailed analysis of a hazardous road location

Within the framework of this study, stick diagrams analysis are carried out. These charts present the road accident information in a diagram form which allows the identification of common characteristics between the accidents and are useful analytical tools supporting the search for common characteristics among the diverse accidents.

4. Case Study

4.1. Choice of Vila Real

Vila Real was selected for the application of this research work for two main reasons. First, the case study required an amount of data that could be manageable and analysed compatible with available resources allocated to the work leading to a MSc dissertation thesis. Secondly, the municipality of Vila Real has shown interest to participate on
road safety projects, and there was the expectation of availability of some resources that would facilitate the data collection, which indeed happened.

4.2. Data collection

Injury road accident data collected for Vila Real from 2004 to 2008 had two sources: the non spatial information (included in BEAV) was obtained in the scope of a protocol between Instituto Superior Técnico (IST) and ANSR, whereas spatial information was obtained via direct PAV access on Vila Real police departments (after the permission of CADA) – see section 2.8.1.

To minimise the duration of the PAV data collection procedure, a database in Microsoft Access was created, with a set of user-friendly forms specially designed for this purpose. The task of collecting accident data took a long time and some of the records in the ANSR database were not found in the records of the police forces, and vice-versa. Road accidents from motorways were excluded as well as those that occurred outside the Vila Real urban area. In total it was not possible to determine their locations of about 50 road accidents with the used sources, remaining 381 injury road accidents, the set that was considered in this work.

The collected information was homogenised in terms of road accident attributes and data validation. Additional geographic and cartographic data, such as the road network datasets was supplied by the municipality.

4.3. Descriptive statistical analysis

The 381 road accidents caused 464 casualties (most of which were slightly injured individuals – 433) of which 53% were drivers, 25% pedestrians and 22% passengers. Among these, pedestrians were the type of casualty that suffered more serious injuries or deaths and were mainly (54.4% of pedestrians) run over on crosswalks. About 96% of drivers had the driver licence appropriate to the vehicle.

Looking to the types of road accidents, the larger number were running over pedestrians (28%), followed by sidewalk collisions (22%), screenings (19%), frontal collisions (12%) and rear-end collisions (10%). The vehicles involved in the injury road accidents were mostly cars (78.8%) followed by two wheeled vehicles (15.4%).

Injury road accidents generally occurred more often during the working weekend and at daylight (67%) and at night with lighting (28%). Concerning to weather conditions, the injury road accidents mostly occurred with good weather (76.4%) or rain (21.3%).

Most accidents occurred on roads with pavements with good or regular conditions of maintenance, and at clean and dry road surface. According to the considered road hierarchy, most road accidents occurred on main distributor roads (72.4%).

4.4. Identification of HRL

4.4.1. Methods based on definitions applied in some European countries

Some parameters were adopted in Method I and II, both applied using ArcGIS™. Method I was applied using ArcGIS™ Point Density function to produce a density grid with 10 m-sided cells whereas in Method II equidistant points (distanced of 20 m) were placed over the centrelines of the entire road network.

Firstly, Method I was applied to reproduce the definition of ANSR (Portugal), leading to the identification of 1 to 3 HRL per year (from 2004 to 2008). The same definition when reproduced by Method II led to the identification of 3 HRL (a single HRL in 2004, 2005 and 2007). The comparison between these two results and taking into consideration the computation process of Methods I and II allowed realising that the latter allows a more accurate reproduction of the definition as explained in section 0. For his reason, for the definitions of Belgium, Hungary, Flanders, Germany and Norway, only Method II was applied. Table 4.1 presents the results. Figure 4.1 illustrates the results of the Norwegian definition applied with Method II.

Table 4.1 – Number of identified HRL by applying Method II for each applied definition

<table>
<thead>
<tr>
<th>Year</th>
<th>ANSR</th>
<th>Belgium</th>
<th>Flanders</th>
<th>Hungary</th>
<th>Germany</th>
<th>Norway</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>2005</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2006</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2007</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 4.1 – HRL identified with Method II applying the definition of Norway in the period of 2004-2008 in Vila Real

4.4.2. Methods based on the Euclidean distance between points

The analysis of HRL based on the Euclidian distance between road accident location points was carried out using CrimeStat®. This software includes three routines based on
the Euclidean distance between points, usable to identify point agglomerations: (1) Mode routine – counting coincident points; (2) Fuzzy Mode routine – counting points within a neighbourhood of each point, and (3) Nearest Neighbor Hierarchical Spatial Clustering (NNH) – which identifies point groups which are spatially close based on a random distribution of points. This random distribution has the same area and number of points (road accidents) as the case study. The calculation of these routines is described in section 3.3.2. All these routines were applied to the dataset of injury road accidents that occurred between 2004 and 2008 in Vila Real.

In the Mode routine only the locations with at least three road accidents were considered as hazardous. This criteria was adopted as the occurrence in the exactly same place of three accidents in five years, might indicate road safety problems; on the other hand this value, although not isolated, is also used in other definitions (see Table 2.1). Hence, this routine identified a single location with four coincident accidents and 4 with three accidents.

In Fuzzy Mode routine, a minimum of 4 road accidents were adopted as the selection criterion to the HRL identification and two values of searching radius were applied: 50 m and 100 m These values were considered adequate for this purpose and it has been applied for the same purpose, as Table 2.1 shows. For example, for the radius of 50 m, this routine led to the identification of 19 HRL.

The NNH routine was computed considering several values for the minimum number of points as presented in Table 4.2. One of these solutions is illustrated in Figure 4.2.

### 4.4.3. Methods based in density functions

Kernel Density Estimation (KDE) was the adopted to identify HRL. This routine requires a set of parameters, displayed on Table 4.3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Adopted option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpolation function</td>
<td>based on Normal (or Gaussian) distribution</td>
</tr>
<tr>
<td>Bandwidth type</td>
<td>adaptive</td>
</tr>
<tr>
<td>Bandwidth criterion</td>
<td>minimum number of road accidents of 10, 5 and 3</td>
</tr>
<tr>
<td>Cutoff value (density)</td>
<td>0.0002 points/m² (corresponds to a 50 m limit away from the road network centreline)</td>
</tr>
<tr>
<td>Cell size</td>
<td>10 m</td>
</tr>
</tbody>
</table>

Erro! Auto-referência de marcador inválida. presents the number of HRL according to the tested bandwidth values. One of the results is illustrated in Figure 4.3.

<table>
<thead>
<tr>
<th>Minimum no. of points to include in the bandwidth</th>
<th>Number of HRL</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
</tr>
</tbody>
</table>

### 4.4.4. Discussion of results

Results suggest that Method II allows a more accurate reproduction of the definition because it considers only an area around the road network avoiding the identification of HRL outside road network, while Method I applies a grid into the entire area, regardless of the road network location. Despite Method II being considered to perform better than Method I, it presents some approximations to the strict application of definitions. The results of Method II show that the identified HRL varied according to the definition applied; furthermore, some definitions identified no HRL at all (Flemish, Belgian) whereas the Norwegian led to the identification of 25 HRL (see Table 4.1). This might indicate that definitions can be more suitable according to the case study. For cities with a larger number of injury road accidents, the Belgian or Flemish definition might be more
suitable to identify locations with more severe road accidents, than the remaining applied definitions.

Also the results produced by the application of these definitions might be a signal of the performance in road safety of each country. For instance, the Norwegian definition applied by Method II, identified more HRL than the remaining definitions, evidencing the good performance of this country at this level (EC, 2011). This analysis must have into consideration that part of the obtained results also depends on the parameters adopted in the Method II, as well as the method itself, which is not an exact reproduction of the definitions.

The Mode routine revealed a low suitability for HRL identification, at least in this case study, because it only identifies locations where road accidents are exactly in the same place (point). Furthermore, its results might be considerably influenced by the georeferencing options that are considered (e.g., snap the road accident points to the centroid of the road network and minimum spatial unit used).

As one of the essays of Fuzzy Mode routine adopted similar parameters as the Norway definition, Method II and Fuzzy Mode routine were compared. It was possible to conclude that (for the adopted values of parameters) Method II identified more HRL than the routine (for the adopted parameters), as it focus the HRL identification process over each point regularly distanced over the entire road network, whereas the routine only focus over the location of road accidents.

The NNH routine might lead to the identification of HRL that form chains along a road, without any extension limit, which is particularly suitable for the route action approach (see section 2.2).

The results of KDE routine show that the specification of the bandwidth has a large influence on the number of identified HRL. Smaller bandwidths might have a tendency to identify more locations.

In general, all methods are relatively simple to apply and require only few calculation minutes; however, other characteristics, capabilities and consequences might influence the choice of a method for identifying HRL and the interpretation of its results. For this reason, a brief comparison was performed between Method II and the Fuzzy Mode, NNH and KDE routines, summarised in Table 5.1. Method I and the Mode routine were not included in this comparison, as they were considered to perform worse than other essayed methods, mentioned in this section.

4.5. Detailed analysis of a hazardous road location

Several locations were identified as hazardous: some of them are consistently recognised by different methods. One of them had the highest number of injury road accidents and therefore was chosen for detailed analysis. It corresponds of a road extension of about 250 m and includes a roundabout. The selected HRL contained 18 road accidents, mostly running over pedestrians (mainly located over crosswalks).

The analysis of this location was based on stick diagrams with the purpose of finding possible problems susceptible to be treated through engineering road safety measures. The analysis of road accidents involving only one vehicle (running over pedestrians and screenings), indicated a possible problem related with exceeding speed. The analysis of the remaining road accidents did not enable the finding of a common pattern related to the probable causes for their occurrences.

5. Conclusions and Discussion

The results showed that the number and location of the identified HRL in Vila Real in the 2004-2008 period varied according to the applied method and definition or routine. Among all the identified HRL, a considerable number was systematically produced by most methods and/or routines. Results indicate the identification of HRL from 0 for the definitions of Belgium and Flanders (reproduced according with Method II) to 28 for the KDE (considering adaptive bandwidth including at least 3 road accidents, the Normal distribution as interpolation function, 10 m cells and a cut-off density value of 0.0002 points/m²). In general, all the methods identified HRL of different sizes and shapes, mostly located on intersections of the highest road functional hierarchy levels.

The comparison between methods and their results highlight that Mode routine seemed to be inadequate for the HRL identification and Method I showed to perform worse when compared to Method II to reproduce the applied European definitions (see section 0).

Some of the applied methods and/or routines (Method I and II, Fuzzy Mode, NNH and KDE routines) considered that space is assumed to be contiguous and homogeneous. This might lead to the identification of HRL which contain road accidents that occurred in non-contiguous roads, which is undesirable, as local characteristics might be different from case to case, and the possible underlying problems of road accidents might also be different.

The NNH and KDE routines can be also useful to be applied in other approaches such as the route action (both) or area action (KDE routine). The NNH routine was the least dependent from parameters established by the analyst as it is has a statistical basis of spatial randomness. However, its calculation does not consider characteristics such as the shape of the area under analysis and relies on the assumption of a homogeneous space.

The suggested methods are relatively simple and easy to understand and can be implemented by the municipalities in a short term with current available resources (even if no traffic data is available), achieving one of the objectives of this work and contributing to the improvement of road safety through the identification of HRL. The identified sites can...
subsequently be analysed and corrected through
engineering measures in the scope of road safety. However,
several aspects need to be considered in the
implementation process, such as the establishment of the
parameters used in the several methods or routines which
might influence considerably the identification of HRL.

<table>
<thead>
<tr>
<th>Parameters Analysis dependence</th>
<th>Method II</th>
<th>Fuzzy Mode routine</th>
<th>NNH routine</th>
<th>KDE routine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Searching radius; Selection criterion; Period of data records; Step of sliding window.</td>
<td>Searching radius; Selection criterion (no. of accidents); Period of data records.</td>
<td>Value of study area; Minimum number of points per agglomeration; Period of data records.</td>
<td>Cell size; Minimum number of bandwidth; Cut-off value; Interpolation function; Period of data records.</td>
<td></td>
</tr>
</tbody>
</table>

| Consideration of accident severity | Yes, according to the definition | No | No | No (although possible, not applied in this study) |

| Some adopted parameters | Time period of accident records; Radius and selection criteria according the applied definition; Step of sliding window: 20 m | 5 years of accident records; Radius: 50 m Selection criterion: ≥ 4 accidents with casualties | 5 years of accident records; Threshold distance based on a random distribution; Minimum of 5 accidents per agglomeration | 5 years of accident records; Cell size: 10 m; Adaptive bandwidth; Cut-off density value: 0.0002 points/m²; Interpolation function: Normal distribution |

| Shape of HRL | Buffers along the road network centrelines | Buffers along the road network centrelines | Convex hulls | Grid cells with density above the cut-off value |

| Main advantages | Allows the definition application; Applies a sliding window; which in the same conditions of Fuzzy Mode routine identifies more HRL | Identification HRL centred on road accidents | Requires calibration of few parameters; Identification based on statistical evidence; HRL size dependent of area and number of accidents in the case study; Might be adopted in route or area action | Promote the HRL identification on areas where more road accidents occur (adequate to areas wide analysis where HRL needed to be identified) |

| Main drawbacks | Results varied according to the applied definition; Implementation proceeding more susceptible to errors: Measurements through Euclidian distance | Choice of search radius and criterion for HRL identification; Measurements through Euclidian distance | Delimitation of the study urban area without considering its shape; Choice of minimum number of accidents in the HRL; Measurements through Euclidian distance | Choice of cut-off value which interferes with the number and shape of HRL; Need of bandwidth determination; Isolated aggregations of accidents have less density value than aggregations surrounded by areas with more road accidents. |

The main difficulty found in this work was the process of collecting accident data from paper documents with imprecise information on the event locations. This process required a considerable time that could be significantly decreased if GPS devices were used *in situ* during accident registration, facilitating the georeferencing process and the removal of location ambiguities and incoherencies.

With the application to a single city, for future studies it is recommended the implementation of the same methods and routines to other cities with different contexts (e.g.: a wider city or a city with a similar size with many more road accidents of with considerably less road accidents) in order to test their influence on the results, with a sensibility analysis to the remaining parameters not tested in this work.

**References**