

Analysis of Spatial and Temporal Distribution of Single and Multiple Vehicle Crash in Western Australia: A Comparison Study

W. Kilamanu^a, J. Xia^a and C. Caulfield^b

^a *Department of Spatial Sciences, Curtin University, Australian*
Email: c.xia@curtin.edu.au

^b *Edith Cowan University, Faculty of Computing, Health and Science, Australia*

Abstract: Forming a major part of any infrastructure, roads are vital in the interplay between the social and economic livelihood of a municipality. The distribution of goods and services is mostly carried out by road. Businesses are provided with the opportunity to expand trading whilst opening up competition. However, when road conditions are directly or indirectly affected by both the physical and natural environment, the likelihood of crashes are inevitable. While there are numerous studies that have investigated these causal factors, only a few have looked into the comparison of single- and multiple-vehicle crashes (SVCs and MVCs) relative to spatial, temporal and spatio-temporal distribution.

This paper applied spatial, temporal and spatio-temporal techniques to investigate patterns of SVCs and MVCs in Western Australia between 1999 to 2008 at different scales. Spider graphs were adapted to identify temporal patterns of vehicle crashes at two different scales: hourly and weekly. The spatial structures of vehicle crashes were analysed using Kernel Density Estimation (KDE) analysis at three different scales: Western Australia, Perth metropolitan area, and Perth Local Government Area (LGA). These are illustrated using spatial zooming theory. Comap was then used to demonstrate the spatio-temporal interaction effect on vehicle crashes.

The results show significant differences in spatio-temporal patterns of SVCs for various crash causes. Furthermore:

- The proportion of SVCs in the Perth LGA compared with MVCs in the same region is not very significant, possibly because of the high density of vehicles in the Perth LGA area.
- It is clearer to illustrate spatial variations of vehicle crashes at the LGA scale level.
- Except for the magnitude differences, six types of hourly and weekly MVCs have similar temporal patterns. Most of them occurred from 8AM to 5PM on weekdays.
- Hit-object SVCs peak between 11PM and 3AM on weekends. In contrast, hit-pedestrian SVCs peak around 8AM, 12 PM, and 3PM to 5PM on workdays when people are travelling between work and home or city workers are out for lunch.
- It is possible to compare hotspots of the SVCs over space and time by integrating spider graphs into Comap and thereby enhancing understanding of location and time dependence and location advantage (or in our case, location disadvantage).

The techniques used here have the potential to help decision makers in developing effective road safety strategies.

Keywords: *Single Vehicle Crashes (SVCs), Multiple Vehicle Crashes (MVCs), kernel density estimation, spider graph, Comap*

1. INTRODUCTION

Over 1.2 million people die each year on the world’s roads and 20 to 50 million suffer non-fatal injuries. Young people constitute the highest number of deaths (WHO, 2009). While there are direct and indirect positive impacts associated with road transportation, increasing road use places a significant burden on people’s health especially relating to traffic injuries. Understanding the spatial and temporal pattern of vehicle crashes is critical in developing location-based road safety strategies to minimise fatalities.

While research into the spatial and temporal analysis of single-vehicle crashes (SVCs) exists, comparison between SVCs and multiple-vehicle crashes (MVCs) is scarce. Ivan et al.(1999) investigated the differences in causality factors for SVCs and MVCs on two-lane roads. SVC crash rates were found to decline when traffic intensity and sight distance increased, which was opposite to MVCs. Probabilistic models of motorcyclists’ injury severities in SVCs and MVCs were estimated by Savolainen and Mannering (2007). They summarised statistics such as rider characteristics and behaviour, roadway and crash characteristics and identified differences between SVCs and MVCs. Geedipally and Lord (2010) discovered that the confidence interval was increased by modelling SVCs and MVCs, which confirmed different characteristics for each.

This paper investigates spatial and temporal distribution of SVCs and MVCs in Western Australia with the aim of geo-visualisation and comparison. The objectives of this study are:

- Investigate temporal variations of SVC and MVC frequencies in Western Australia .
- Investigate spatial distribution of SVCs and MVCs at different scales
- Combine the spatial and temporal analyses to determine how the spatial distribution of SVCs and MVCs varies over time

2. METHODS

2.1. Study area and data collection methods

This paper is based on data obtained from the Main Roads Department in Western Australia. The Main Roads database contains information about all police-reported car crashes in Western Australia. All crashes involving single and multiple vehicles between 1999 and 2008 were analysed. The study area and the distribution of all crashes are illustrated in Figure 1. The crash categories are derived from the “Nature” attribute of the crash data where the values 1 to 6 represent MVCs and values 7 to 10 represent SVCs (see Table 1).

2.2. Kernel Density Estimation (KDE)

Modelling hot spot areas provides insight and understanding into the dependence of events within the spatial processes (Nicholson, 1997). To investigate spatial patterns of SVCs and MVCs, Kernel Density Estimation (KDE) was undertaken using the ArcGIS 10 Spatial Analyst tool (ESRI, 2010). This method calculates density of features (crashes) within the neighbourhood and around those features. Conceptually, a neighbourhood or “kernel” is defined around the points. The number of points that fall within the neighbourhood is totalled and divided by the area of the neighbourhood (Silverman 1986). The KDE parameters used in this study is based on Equation (1) adopted from Silverman (1986).

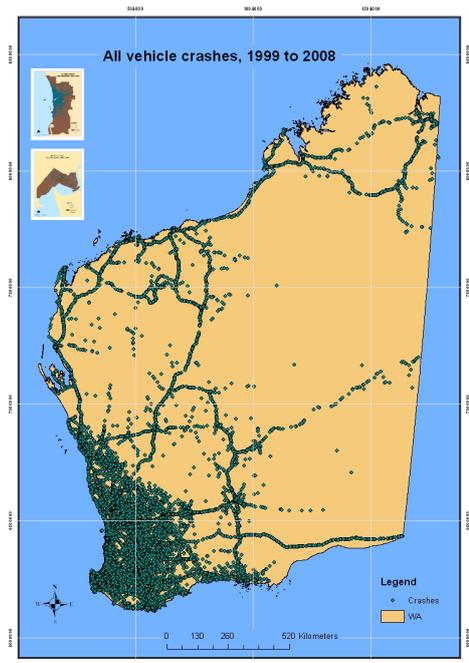


Figure 1. Distribution of all SVC in Western Australia 1999-2008.

This method calculates density of

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (1)$$

- Where K = the kernel function $K(y) = (3\pi)^{-1} (1-y^2)^2$
 h = bandwidth
 x = location of cell centre
 x_i = position of the i -th point (vehicle crash)
 n = total number of crashes

It can be seen that the kernel estimator depends on two parameters – the bandwidth (h) and the kernel density (K) (Silverman, 1986; Flahaut *et al.*, 2003). Flahaut *et al.* (2003) mentions that for any given kernel, the bandwidth or smoothing parameter is “critically important” for the kernel estimator. When h equals the size of the region, density remains the same throughout the region. On the other hand, individual points are highlighted when a very small bandwidth is used (O’Sullivan and Unwin 2003). Silverman (1986) suggests a “subjective” approach for exploratory purposes and an “automatic” choice for subsequent smoothing provided the statistical model and hypothesis are accepted. The bandwidth choice for this study has been based on experimentation or until an appropriate estimation has been produced.

2.3. Spatio- temporal zooming theory

Scale plays a vital role in influencing people’s perception of space and time (Freundschuh and Egenhofer, 1997). People tend to abstract spatial and temporal information from their environment at different scales and conduct mental shifts among these different scales (Hornsby, 2001). Scale is about extent: a large scale shows fine details and high spatial or temporal resolutions while a small scale shows coarse details and low resolution (Longley *et al.*, 2001). Effective visualisation and analysis depend critically on choosing an appropriate scale in order to derive a meaningful outline of the data (Shchaffer *et al.*, 1996).

Different levels of detailed spatial and temporal information can be represented in one map using zooming theory. A single zooming model displays a sequence of layers of gradually simplified representations from fine-grained views (articulations) to coarser-grained views (abstractions) for a given geographical map. Multiple zooming techniques can focus on information of interest with high resolution and present background information with low resolution in one map (Frigioni and Tarantino, 2003; Hornsby and Egenhofer, 2002). This paper applied single zooming theory to the data to illustrate spatial patterns of SVCs at three scales: Western Australia as a whole, the Perth metropolitan area, and the Perth Local Government Area (LGA) (see Figure 5).

2.4. Spider Plots

To visualise and explore temporal patterns of both SVC and MVCs, spider plots are employed. Temporal analyses are often shown in line charts and circular plots since they illustrate continuity and chronological order (Asgary, 2010). Subsets of the crash dataset are used to explore the temporal nature and distribution of SVCs and MVCs. Spider plots are provided and compared for SVCs and MVCs for Perth LGA. Figure 4 illustrates an example spider plot used in this analysis.

2.5. Comap

Comap is a method of exploratory data analysis to visualise changes in a pair of variables over time (Asgary *et al.*, 2010). This method can investigate the relationship between the location of crashes and when they occurred. The steps involved in performing a Comap analysis are outlined by Asgary *et al.* (2010) as (1) Subset vehicle crashes into different classes based on time interval in which they occur such as 1AM to 3AM. (2) Run a KDE analysis for each subset to generate hot-spot maps. (3) Illustrate hot spot maps according to the ordered time intervals. Asgary *et al.* (2010) states that the process of classification may affect the result of the Comap analysis. For example, creating too small a subset may prevent a pattern from being easily discerned. Hence, he suggests that class boundaries should overlap and each class must contain the same number of events.

3. RESULTS

3.1. General comparison of SVCs and MVCs

As mentioned in section 2, the crash nature values 1 to 6 represent MVCs and the values 7 to 10 represent SVCs.

Table 1 shows comparison of crash nature with (a) showing distribution of crashes by metropolitan area and LGA while

Table 1. Summary of Vehicle Crashes by nature

Value	Crashes by Nature	Perth LGA		Perth Metropolitan	
		Freq	% Crash Nature	Freq	% Crash Nature
0		2089	0.09	19083	0.06
1	Rear End	9278	0.41	129932	0.41
2	Head On	67	0.00	2720	0.01
3	Sideswipe Opposite Direction		0.00		0.00
4	Sideswipe Same Direction	4800	0.21	34184	0.11
5	Right Angle	3438	0.15	65422	0.21
6	Right Turn Thru	1416	0.06	23725	0.08
7	Hit Pedestrian	724	0.03	4244	0.01
8	Hit Animal	1	0.00	1250	0.00
9	Hit Obj	848	0.04	28903	0.09
10	Non Collision	181	0.01	4502	0.01
	Total	22842	1.0000	313965	0.94

(a)

Overall Totals	AVC	SVC	% AVC	MVC	% AVC
WA	381294	68570	0.18	288268	0.76
Metro	313965	38899	0.12	255983	0.82
LGA	22842	1754	0.08	18999	0.83

(b)

(b) shows comparison between the overall total (WA), metropolitan area and LGA area as percentages. The column “% Crash Nature” in (a) is the percentage of crash frequency to the total number of crashes within the area. Major MVCs are Rear End (41%), Sideswipe Same Direction (21%) and Right Angle (15%). Major SVCs are hit Objects (4%) in Perth LGA. The proportion of the SVCs in the Perth LGA compared the MVCs within the same area is relatively small probably due to the high density of vehicles in the Perth LGA area.

3.2. Spatial analysis

KDE provides us with an overview of the spatial variations of crashes. A spatial window or kernel is passed over the areas of interest. The result is a continuous surface of density estimates with high values depicted as peaks. It also reveals hot spot areas of the phenomena under study. More detailed methods can be found in (O’Sullivan & Unwin; 2003; Silverman 1986). Spatial analysis performed for the three areas of interest is shown in Figure 2, which illustrates how spatial scale affects the level of information exhibited. For instance, density estimation for all crashes for Western Australia shows Perth is the major hot spot area— more than 80 percent of the total crashes in WA occur in the Perth metropolitan area and LGA. When zooming into these areas, more detailed vehicle crashes patterns were revealed. Figure 3 shows that it is possibly clearer to illustrate spatial variations of vehicle crashes at Perth LGA scale level. The differences of spatial patterns between SVCs and MVCs will be discussed in section 3.3.

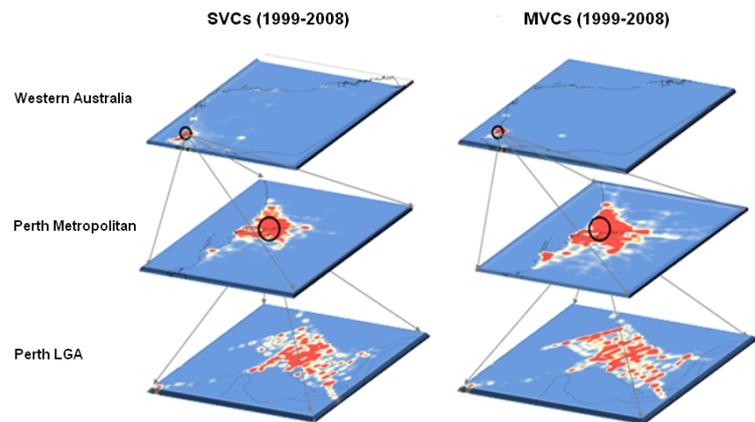


Figure 2. KDE results for SVCs and MVCs at different scales.

3.3. Temporal analysis

Figure 3 shows the temporal distributions for single and multiple vehicle crashes by hours of the day within the metropolitan and LGA for the period 1999 to 2008. As illustrated, there is a distinct difference in the temporal distribution for both crash categories. For MVCs, distribution by hour shows fluctuations in crash numbers between 8AM and 5PM. The crash frequency throughout the day shows a sharp increase between 5AM to 8AM and 2PM to 3PM and a gradual increase between 3PM and 5PM (peaking at 5PM). A sharp

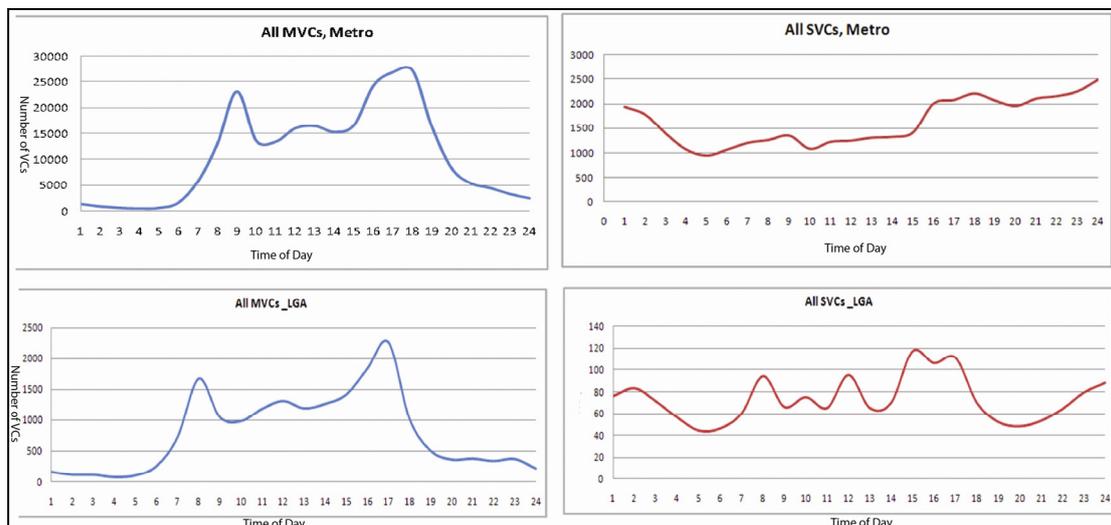


Figure 3. Number of SVCs and MVCs at the Perth Metro and LGA, 1999-2008

decrease is also seen at 9AM and between 5PM and 8PM. These patterns suggest that the MVCs are more likely to occur at peak hours.

In contrast to MVCs, temporal analysis of SVCs shows a relatively steady distribution throughout the day for the metropolitan area. Figure 3 shows a steady but gradual increase throughout the day and reveals high incidence between 2PM and 12AM, inferring that a higher number of SVCs occur within this time period. While MVC frequency variation for the metropolitan and LGA areas are not distinct, as can be seen from Figure 3, the SVC plot shows an intriguing pattern. The frequency of SVCs in the metropolitan area is generally distributed uniformly whereas those in the LGA seem to fluctuate distinctly throughout the day. Further investigation of the factors might shed some light on this.

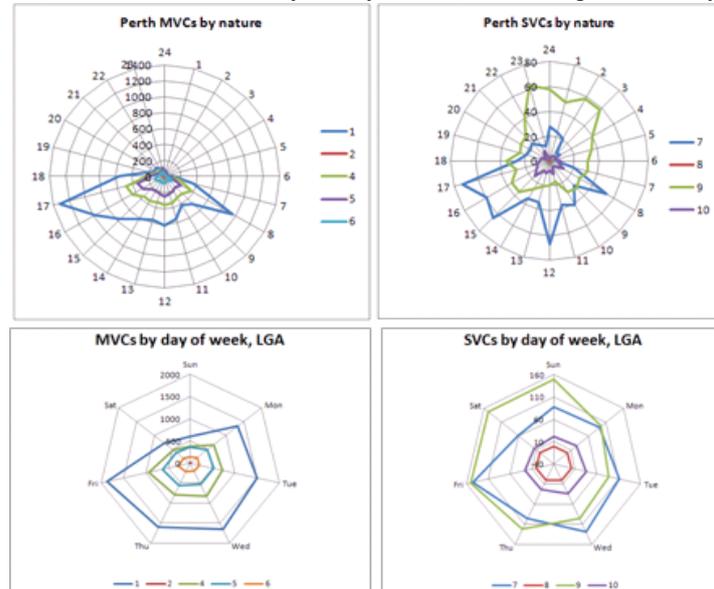


Figure 4. Spider graph of VCs by nature.

Spider graphs were used to illustrate the temporal pattern of SVC and MVCs by nature (Figure 4 and Table 1).

Except for the magnitude differences, six types of hourly and weekly MVCs have similar temporal patterns. Most of them occurred during between 8AM to 5PM on weekdays. While four types of SVCs show different distributions. Hit-object SVCs peak between 11PM and 3AM on weekends. In contrast, hit-pedestrian SVCs peak around 8AM, 12PM, and 3PM to 5PM on workdays when people travel between work and home or city workers are out for lunch. Non-collision SVCs peak between 8AM and 3PM.

3.4. Spatial and temporal analysis

In the Comap (Figure 5), each time interval (such as midnight to 3AM) is labelled with a number from 1 to 8 so that we can identify temporal and spatial patterns of vehicle crashes. Analyses show that there are variations in crash distribution throughout the day. For example, it is evident the MVCs increase gradually from 8 AM to 5PM, and clusters of MVCs shift from the Central Business District (CBD) to the Northbridge entertainment district from day time to night time. This is suggestive of the influence of social factors of Northbridge at night.

SVCs in Figure 5 are a little more dynamic than MVCs. The hot spots are moving around city centre (St George's Terrace, Barrack Street, William Street and Wellington Street), the Northbridge entertainment district and the west (Subiaco). Hit-pedestrian vehicle crashes at night are evidently clustered in the northern part of the suburb (the Northbridge entertainment district). During the day pedestrian crashes are more common in the city centre (St George's Terrace, Barrack Street, William Street and Wellington Street) and some clustering occurs in the area to the west (Subiaco). Object crashes do not exhibit the same degree of clustering as pedestrian crashes. However, they do appear to be most common on the freeways and causeways. In the afternoon and evening many occur on Barrack Street, while late at night they tend to be located on Wellington Street between the Barrack and William Street intersections. William and Barrack streets feed into the bus and train stations, which might explain why there are so many pedestrians around.

Kilamanu *et al.*, Analysis of Spatial and Temporal Distribution of Single and Multiple Vehicle Crash in Western Australia: A Comparison Study

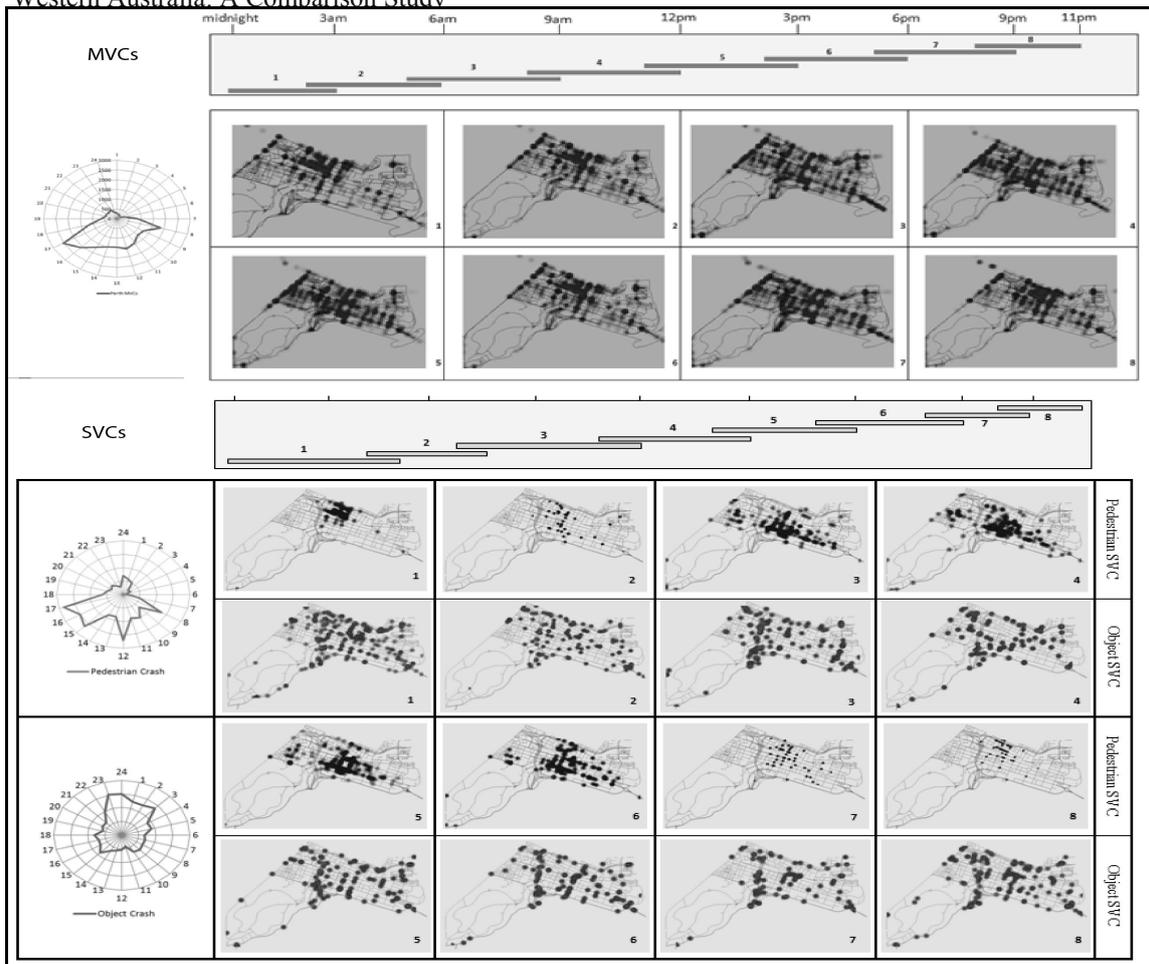


Figure 5. Comap of VCs by hours of a day

4. DISCUSSIONS AND CONCLUSIONS

The main objective of this analysis was to investigate temporal, spatial and spatial-temporal structure of single and multiple vehicle crashes. To achieve this, temporal and spatial analysis was performed for different spatial scales, especially metropolitan and LGA areas, and then combined using Comap.

Persaud & Mucsi (1995) found that MVCs were more likely to occur during the daytime while the light conditions were good, whereas SVCs mostly occur at night. Our findings confirmed the MVCs pattern. However, the temporal distribution of SVCs varied between the metropolitan area and Perth LGA area. SVCs occurred more frequently after sunset than during the day in the metropolitan area. But, more SVCs were likely to occur during the day in Perth LGA area, especially from 2PM to 5PM. This might be due to the high percentage of workers or shoppers in the city, which causes more frequent pedestrian vehicle crashes. We also discovered that hit-pedestrian vehicle crashes mostly occurred during the day, while hit-object vehicle crashes are more likely to occur at night.

To demonstrate the clustering of crash frequency, thereby revealing areas of spatial hotspot, KDE was applied. While analysing at a larger spatial scale provides a general perception of the nature of crashes, it does not reveal underlying trends and patterns that might be beneficial in understanding the characteristics of vehicle crashes. For instance, KDE analysis of crashes at an LGA level revealed that most crashes occur at intersections and midblock. Therefore, this suggests some further analysis is needed concerning the causal factors at hot spot areas.

Comap was used to understand the spatial variation of crashes over a certain time interval. Comap methods can be either univariate, (Figure 5) showing how crash density varies over time, or bivariate (Figure 6) to compare hit-pedestrian or hit-object vehicle crashes. This information is vital in risk analysis of road

Kilamanu *et al.*, Analysis of Spatial and Temporal Distribution of Single and Multiple Vehicle Crash in Western Australia: A Comparison Study

segments or road category, consequently leading to improved road safety. Comap is also able to compare hotspots of SVCs over space and time by integrating spider graphs, and thereby enhancing the understanding of location and time dependence and location advantage (or in our case, location disadvantage).

Classifications inherent in the Comap method must be considered when using this technique. Because time point data are grouped into time intervals, the intervals selected will affect the results of the density estimation. Natural breaks were used for purposes of classification as it better highlighted overall temporal distribution.

Appropriate bandwidth choice will affect the results of kernel density estimation. Suitable bandwidth was selected relative to the spatial scales used in the estimation. The use of unsuitable bandwidths may lead to either more or less smoothing of the discrete data points, producing different maps. Thus, care should be taken in determining bandwidths as well as the interpretation of resulting maps.

It is evident that the nature and structure of SVCs and MVCs in Western Australia varies with location and time. Analyses reveal clustering at intersections and midblock for both crash categories and therefore would be ideal for further research. This would shed light on the dynamics of crashes at these locations and provide vital information on allocation of road safety measures.

ACKNOWLEDGMENTS

We wish to acknowledge the assistance of Main Roads WA in providing access to the WA road crash data files needed for this study and two reviewers for their comments that help improve the manuscript.

REFERENCES

- Asgary, A., Ghaffari, A., Levy, J. (2010). "Spatial and temporal analysis of structural fire incidents and their causes: A case of Toronto, Canada." *Fire Safety Journal* 45 (1), 44 -54.
- ESRI (2010) ArcGIS 9.3. New York, ESRI.
- Flahaut, B., Mouchart, M., Martin, E. S., & Thomas, I. (2003). The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. *Accident Analysis & Prevention*, 35(6), 991-1004.
- Freundschuh, S. & Egenhofer, M. (1997). Human Conceptions of Spaces: Implications for Geographic Information Systems, *Transactions in GIS*, 2(4), 361-375.
- Frigioni, D. & Tarantino, L. (2003). Multiple zooming in geographic maps, *Data & Knowledge Engineering*, 47(2), 207-236.
- Geedipally, S. R., & Lord, D. (2010). Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson-gamma models. *Accident Analysis & Prevention*, 42(4), 1273-1282.
- Hornsby, K. (2001). Temporal zooming, *Transactions in GIS*, 5(3), 255-272.
- Hornsby, K. & Egenhofer, M. (2002). Modelling Moving Objects over Multiple Granularities, *Annals of Mathematics and Artificial Intelligence*, 36, 177-194.
- Ivan, J. N., Pasupathy, R. K., & Ossenbruggen, P. J. (1999). Differences in causality factors for single and multi-vehicle crashes on two-lane roads. *Accident Analysis & Prevention*, 31(6), 695-704.
- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2001). *Geographic information systems and science* Chichester, New York : Wiley.
- Nicholson, A. (1997). Analysis of spatial distributions of accidents. *Safety Science* 31(1), 71 - 91.
- O'Sullivan, D., D.J. Unwin. *Geographic Information Analysis*. Hoboken: John Wiley & Sons, Inc., 2003.
- Persaud, B., & Mucsi, K. (1995). Microscopic accident potential models for two-lane rural roads. *Transportation Research Record 1485*, TRB, National Research Council, Washington, DC, 134-139.
- Savolainen, P., & Mannering, F. (2007). Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. *Accident Analysis & Prevention*, 39(5), 955-963.
- Shchaffer, D., Zuo, Z., Greenberg, S., Bartram, L., Dill, J., Dubs, S. & Roseman, M. (1996) Navigating Hierarchically Clustered Networks through Fisheye and Full-Zoom Methods, *ACM Transactions on Computer-Human Interaction*, 3(2), 162-188.
- Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.
- WHO. (2009). Global status report on road safety. Retrieved 26 June, 2011, from http://www.who.int/gho/road_safety/reports/en/index.html