

# Use of Spatial Analysis Techniques to Identify Statistically Significant Crash Hot Spots in Metropolitan Melbourne

Elizabeth Hovenden<sup>1</sup> and Gang-Jun Liu<sup>2</sup>

<sup>1</sup>*Department of Transport, Melbourne, Australia*

<sup>2</sup>*Royal Melbourne Institute of Technology, Melbourne, Australia*

Corresponding Author: Elizabeth Hovenden, Department of Transport, 60 Denmark Street, Kew Vic 3101, Elizabeth.Hovenden@roads.vic.gov.au, +61 423 597 433.

*This peer-reviewed paper was first presented as an Extended Abstract and Oral Presentation at the 2019 Australasian Road Safety Conference (ARSC2019) held in Adelaide, Australia and first published in the ARSC2019 Proceedings in the form of an Extended Abstract. It was expanded into a 'Full Paper' and underwent further peer-review by three independent experts in the field. It is being reproduced here with the kind permission of the authors and is now only available in this edition of the JACRS.*

## Key Findings

- In metropolitan Melbourne 15.7 per cent of casualty crash locations were hot spots;
- The degree, location and extent of clustering varied for different crash categories;
- Fatal crashes had the lowest clustering level whilst bicycle crashes had the highest;
- There are temporal variations in clustering;
- Hot spot clusters were in locations with a high proportion of commercial land use.

## Abstract

Understanding where, when, what type and why crashes are occurring can help determine the most appropriate initiatives to reduce road trauma. Spatial statistical analysis techniques are better suited to analysing crashes than traditional statistical techniques as they allow for spatial dependency and non-stationarity. For example, crashes tend to cluster at specific locations (spatial dependency) and vary from one location to another (non-stationarity).

Several spatial statistical methods were used to examine crash clustering in metropolitan Melbourne, including Global Moran's I statistic, Kernel Density Estimation and Getis-Ord Gi\* statistic. The Global Moran's I statistic identified statistically significant clustering on a global level. The Kernel Density Estimation method showed clustering but could not identify the statistical significance. The Getis-Ord Gi\* method identified local crash clustering and found that 15.7 per cent of casualty crash locations in metropolitan Melbourne were statistically significant hot spots at the 95 per cent confidence level. The degree, location and extent of clustering was found to vary for different crash categories, with fatal crashes exhibiting the lowest level of clustering and bicycle crashes exhibiting the highest level of clustering. Temporal variations in clustering were also observed.

Overlaying the results with land use and road classification data found that hot spot clusters were in areas with a higher proportion of commercial land use and with a higher proportion of arterial and sub-arterial roads. Further work should investigate network-based hot spot analysis and explore the relationship between crash clusters and influencing factors using spatial techniques such as Geographically Weighted Regression.

## Keywords

Crash Hot Spots, Spatial Statistics, Spatial Analysis, Getis-Ord Gi\*

## Glossary

**Cold spot.** Low values surrounded by other low values. Has a low negative z-score (standard deviation) and a low p-value, indicating that the pattern observed was unlikely to be caused by a random process. The lower the z-score, the more intense the clustering (a z-score near zero means no spatial clustering).

**Crash cluster.** A concentration of crashes.

**Euclidean distance.** Straight line distance between two points.

**Feature Class.** A homogeneous collection of common features (for example, roads), each having the same spatial representation such as points, lines, or polygons, and a common set of attribute columns (Esri, n.d.-a).

**High off-peak period.** The period during the weekday between the AM and PM peak periods. For this research the period was from 10:00 AM to 2:59 PM.

**Hot spot.** High values surrounded by other high values. Has a high positive z-score (standard deviation) and a low p-value, indicating that the pattern observed was unlikely to be caused by a random process. The higher the z-score, the more intense the clustering (a z-score near zero means no spatial clustering).

**Low off-peak period.** The period during the weekday between the PM and AM peak periods. For this research the period was from 7:00 PM to 5:59 AM.

**Modifiable Area Unit Problem (MAUP).** A source of statistical bias and results from the aggregation of data. For example, aggregating crashes to grids of fixed size (such as 100m by 100m) will give a different crash density for a location than if the crashes were aggregated to local government areas.

**Non-stationarity.** Variation from one location to another (spatial non-stationarity), or from one time-period to another (temporal non-stationarity).

**Spatial autocorrelation (spatial dependence).** A measure of the similarity between nearby observations. If spatial autocorrelation exists, then the observations are spatially dependent. This can affect the results from statistical analyses that assume data independence.

## Introduction

Worldwide, more than 1.25 million people die due to road traffic crashes each year and a further 20 to 50 million people suffer injuries, some of which result in disabilities (World Health Organisation, 2018). Road crashes are the ninth leading cause of death globally, are the leading cause of death in people aged between 15 and 29 and are estimated to cost most countries around three percent of their gross domestic product (World Health Organisation, 2018).

Victoria's Towards Zero Road Safety Strategy and Action Plan 2016-2020 (Victoria State Government, 2016, p. 6) states that 'preventable road crashes conservatively cost the Victorian community more than \$3 billion a year'. Between 2012 and 2016, 70,539 casualty crashes occurred on Victorian roads (VicRoads, 2018). Seventy-one per cent of these occurred on metropolitan Melbourne roads, equating to around 10,000 casualty crashes per annum. The resulting cost to the community is massive, both financially and emotionally. Understanding where, when, what type and why the crashes are occurring can help in determining the most appropriate initiatives to reduce this level of road trauma.

Crashes exhibit a spatial and temporal non-random distribution therefore spatial analysis techniques are

well suited to provide this understanding. Only a limited amount of studies has been done in Australia on crashes using spatial analysis techniques. This research project addresses this gap by using spatial analysis and spatial statistics tools to investigate the occurrence of casualty crashes in metropolitan Melbourne.

The purpose of this research was to:

- Investigate where, when and what types of casualty crashes are occurring;
- Determine whether clusters differ temporally and by type of crash and vehicle type;
- Identify some of the factors influencing the location of the clusters.

To address the first two items, all casualty crashes in metropolitan Melbourne between 2012 and 2016 were investigated and clustering was examined by year, by time-period, by severity, by crash type and by vehicle type. To address the third item, the output from the hot spot analysis was used to investigate the relationship between hot spot location and explanatory factors including road classification and land use.

It was hypothesised that there would be locations where crashes are clustered and that there would be both temporal differences and differences related to crash type and vehicle type. There was also an expectation that there would be a statistically significant relationship between hot spot location and road classification, and between hot spot location and land use.

## Literature Review

As crashes are influenced by spatial factors (Loo & Yao, 2012), spatial analysis techniques are well suited to identify crash patterns and the characteristics of those patterns (Nie, Wang, Du, Ren & Tian, 2015; Prasannakumar, Vijith, Charutha & Geetha, 2011). Statistical analysis by itself, without the spatial component, is limited as it does not allow association of crash characteristics with spatial distribution (Gudes, Varhol, Sun & Meuleners, 2017). Traditional statistical techniques assume spatial independence and stationarity, however crashes exhibit spatial autocorrelation (spatial dependence) clustering at specific locations such as on vertical and horizontal curves (Mohaymany, Shahri & Mirbagheri, 2013) and they exhibit non-stationarity, varying from one location to another. Failure to account for spatial autocorrelation can lead to errors when interpreting models (Getis & Ord, 1992).

## Spatial Analysis Techniques

Statistical spatial analysis of crashes can be event-based, where crashes are represented as points, or link- and area-based, where crashes are aggregated to a road segment or area (Loo & Yao, 2012). Clustering can be measured using global or local methods. A global method looks at the whole study area and produces a global statistic indicating

the degree of clustering. A local method looks at individual locations to determine spatial association in the vicinity of the location and produces local statistics for each location indicating the degree of clustering.

### Moran's I Statistic

Whilst there are several techniques to test for spatial autocorrelation, one of the most common ones is the Moran's I statistic (Getis & Ord, 1992; Mohaymany et al., 2013). This method is suitable for area-based analysis and describes how crashes are distributed across the whole study area (Loo & Yao, 2012; Nie et al., 2015). It evaluates whether a pattern is clustered, dispersed or random for a given set of features and associated attributes by comparing the values of the variable at one location with the value at all other locations (Esri, n.d.-b; Prasannakumar et al., 2011).

Being a global index, Moran's I provides statistics based on the entire study area. Although there is a local version of the Moran's I index that can be used to identify clusters, it cannot identify statistically significant clusters of hot spots (high values surrounded by high values) or cold spots (low values surrounded by low values). To do this, the Getis-Ord Gi\* statistic can be used (Nie et al., 2015).

### Kernel Density Estimation

Kernel Density Estimation (KDE) is a common method used to identify spatial patterns for event-based data. It can be used to explore crash clustering and analyse patterns. It creates a density surface, and the darker the colour of the surface, the higher the crash intensity (Mohaymany et al., 2013; Plug, Xia & Caulfield, 2011). Considerations when using KDE is what kernel function should be used and choice of bandwidth. Too large a bandwidth will result in the same density for the area (that is, it will show significant smoothing), too small a bandwidth will just highlight individual points (Plug et al., 2011; Pour, Moridpour, Rajabifard & Tay, 2017). To determine the optimum bandwidth a global Moran's I spatial autocorrelation analysis can be applied, and a visual sensitivity analysis can be undertaken of the results (Lawrence, Stevenson, Oxley & Logan, 2015).

KDE has some limitations. One limitation is that it assumes a two-dimensional (planar) distribution whereas crashes are constrained to the one-dimensional road network. To overcome this limitation, Network-Constrained Kernel Density Estimation (NKDE), in which the distance measure is based on network distance, not Euclidian distance, could be used. Another limitation is that KDE (and NKDE) do not identify which clusters are statistically significant (Nie et al., 2015; Truong & Somenahalli, 2011). However Bíl, Andrášik & Janoška (2013) have developed a method of testing the significance of clusters output from KDE analysis. This method is implemented in the KDE+ application (Transport Research Centre CDV, 2020).

### Getis-Ord Gi\* Statistic

The Getis-Ord Gi\* statistic is a local measure that can be used on link-based or area-based data. It measures the degree of association between weighted points over a given distance (Getis & Ord, 1992). The choice of distance is important. If it is too small or too large, then the assumption of normality is lost (Getis & Ord, 1992). The Getis-Ord Gi\* statistic can be used to identify hot spots. Being a local statistic, it can be used to assess each feature in relation to its neighbours and to compare the local and global situations (Esri, n.d.-c).

A drawback of Getis-Ord Hot Spot analysis is that it requires the aggregation of data, which can lead to the modifiable area unit problem – that is, different results being produced depending on how the data was aggregated (O'Sullivan & Unwin, cited in Plug et al., 2011). There is a network-constrained version of Getis-Ord Gi\* (GLINCS) which can be used to measure spatial autocorrelation and concentration of crashes on a network (Nie et al., 2015).

Although the Getis-Ord Gi\* statistic can be used for local spatial analysis, it cannot diagnose local instability in global measures of association (Anselin, 1995). It is harder to detect local clusters when there is global autocorrelation as this significantly impacts the distribution of local statistics, therefore it is important to consider both the local and global statistics when undertaking spatial analysis (Ord & Getis, 1995). Used in conjunction with global statistics, the Getis-Ord Gi\* statistic gives a better understanding of spatial association and allows local 'pockets' of dependence (clusters or hot spots) to be identified (Getis & Ord, 1992).

### Findings from previous studies using spatial analysis

There have been many studies of crashes undertaken internationally using techniques such as the Moran's I statistic, the Getis-Ord Gi\* statistic and KDE individually and in combination. In some studies the crashes were spatially and/or temporally partitioned (Chance Scott et al., 2016; Prasannakumar et al., 2011). Some studies used severity indices, as well as actual counts, with more severe crashes being given a higher weighting (Choudhary, Ohri & Kumar, 2015; Soltani & Askari, 2017). The results of the studies showed significant clustering of crashes, especially on arterial roads and near urban activity centres, but found that in certain locations the distribution of crashes was random (Chance Scott, Sen Roy & Prasad, 2016; Prasannakumar et al., 2011; Soltani & Askari, 2017). Benedek, Ciobanu and Man (2016) used network-based analysis to identify crash hot spots and found that the number of crashes was directly proportional to the traffic volume, and that the hot spots were in locations where traffic was high.

Studies into the relationship between crashes and explanatory variables using spatial techniques such as Geographically Weighted Regression found that length of highway had the greatest effect on the number of crashes (Erdogan, 2009) and that speed was an influential factor for single vehicle fatal crash hot spots in rural areas (Liu & Xia, 2015). Other findings were that networks centred on major roads had less pedestrian and cyclist crashes due to less conflicts on local roads between these road users and motorised vehicles (Zhang, Bigham, Ragland & Chen, 2015) and that freeway kilometres and road density were negatively correlated with fatal crashes whilst population density and under 18 population were positively correlated (Li, Wang, Liu, Bigham & Ragland, 2013).

There have only been a limited number of studies of crashes in Australia that have used spatial analysis techniques. Most of the studies used a combination of KDE and Moran's I statistic and some studies used Getis-Ord Gi\* statistics. Lawrence et al. (2015) undertook geospatial analysis using KDE to understand the spatial trend in cycling crashes in Melbourne. Pour et al. (2017) undertook spatial analysis of vehicle-pedestrian crashes in metropolitan Melbourne, using KDE to identify areas with high crash risk and Moran's I to determine spatial dependency, and found a positive spatial correlation between time and location of crashes. Truong and Somenahalli (2011) undertook spatial analysis of pedestrian-vehicle crashes in Adelaide to identify and rank unsafe bus-stops, using the Moran's I statistic to examine spatial patterns and measure spatial autocorrelation, and the Getis-Ord Gi\* statistic to identify clustering. They identified unsafe bus stops by overlaying bus stop data onto the hot spots and ranking the bus stops based on the severity index of the pedestrian-vehicle crashes. Plug et al. (2011) investigated single vehicle crashes in Western Australia using spider-plots for the temporal analysis and KDE for the spatial analysis. By arranging the KDE surfaces sequentially, they illustrated the variation in spatial distribution over the day. Gudes et al. (2017) used the Emerging Hot Spot tool (a space-time implementation of Getis-Ord Gi\*) to identify and categorise concentrations of articulated heavy vehicle crashes in Western Australia. The categories included 'new hot spot' or 'persistent hot spot' and the study used KDE applied to network data to analyse the hot spots.

## Methods

Analysis was undertaken on 2012-2016 crash data for metropolitan Melbourne. The crash data was sourced from the publicly-available VicRoads Crash Stats data extract. Due to incompleteness of 2017 and 2018 data at the time the analysis was undertaken, these years were excluded. As well as spatial clustering of total casualty crashes, the analysis looked at spatial clustering by temporal characteristics (individual year and time period), crash severity, crash type and vehicle type.

Microsoft Excel (Office16) and ESRI ArcGIS 10.3.1 for Desktop were used for the preparation, organisation and analysis of the data. The exploratory and explanatory analyses carried out on the data are briefly described below. For more details, refer to the Appendix.

## Exploratory spatial statistical analysis (spatial autocorrelation and spatial clustering)

The Spatial Autocorrelation (Morans I) tool was used to explore clustering of crashes globally (for the entire study area). As crashes are event-based, an important consideration when identifying local clustering was how to aggregate the data since different aggregation methods can lead to different results. The methods and aggregation levels investigated included the following.

- Kernel Density Estimation (KDE). Two cell sizes were tested (100m and 500m). Each crash was counted only once, the area units were square kilometres, the output values were densities and the method was planar.
- The Getis-Ord Gi\* hot spot analysis method. Alternative aggregation methods were tested, including aggregation to grid of fixed size (100m, 500m and 1000m), aggregation to point with no integration (that is, aggregation of crashes at coincident points), aggregation to point with integration (that is, aggregation of crashes within a specified distance of each other – distances from 10m to 100m were tested).

The 'Integrate' tool (with a 10m tolerance) and the 'Collect Events' tool was used on each of the categories of interest to convert event data (crashes) into weighted point data for further analysis. The 'Spatial Autocorrelation (Morans I)' tool was then used to identify global clustering, with 'Inverse Distance' conceptualisation, Euclidean distance method and a threshold distance of 1000m (other threshold distances, up to 10,000m, were also tested). The 'Hot Spot Analysis (Getis Ord Gi\*)' tool was then used to identify statistically significant hot spots and cold spots, with 'Fixed Distance Band' conceptualisation, Euclidean distance method and a distance band of 1000m. To visualise the results, the 'IDW' tool was used to interpolate a raster surface from the 'GiZScore' from the hot spot analysis point layer for each category and symbology was applied to show the clusters.

## Explanatory spatial statistical analysis

In this part of the analysis, polygons were created (using the 'Aggregate Point' tool) around each cluster of hot spots that were identified in the Getis-Ord Gi\* Hot Spot Analysis. The criteria use in this analysis to define a cluster was an area that had three or more hot spot locations with a confidence level of 95 per cent or above. The polygons were merged to create a single feature class and, after testing various buffer sizes, a 200m buffer was applied to ensure that abutting land use was included in the polygons.

This process was repeated for each cold spot cluster to create a cold spot feature class (with a 200m buffer). The 'Aggregate Point' tool was used to create a feature class of random (non-significant) points, a 200m buffer was applied and the 'Erase' tool was used to remove overlaps with the hot spot and cold spot polygons. The large Melbourne hot spot polygon was split into smaller polygons by the ABS SA2 level.

The length of road for each road classification (based on the CLASS\_CODE field from the TR\_ROAD layer) and the land use area for each land use type (based on MB\_CAT16 field from the MB\_2016\_VIC layer) was calculated for each hot spot, cold spot and random (non-significant) polygons using the 'Intersect' tool, the 'Summary Statistics' tool, and table joins based on attributes. The number of intersections was also calculated for each polygon using the 'Unsplit Line' tool, the 'Intersect' tool (with Output Type set to 'Point'), the 'Integrate' tool (with a 10m tolerance) and the 'Collect Events' tool and a join based on spatial location.

Mesh Block non-spatial attributes were added to the hot spot, cold spot and random (non-significant) polygons. As some Mesh Blocks areas were only partially within the polygons, the 'Clip' tool was used and then the proportion of the clipped area to original area was calculated and the attribute values such as population were recalculated based on the proportional area. For example, if the Mesh Block has a population of 200 and the clipped Mesh Block area is 80 per cent of the original Mesh Block area, then the adjusted population is  $0.8 \times 200 = 160$ . This assumes a uniform distribution of the attribute in the whole area (e.g. Mesh Block) however this may not necessarily be the case. The tables were then joined based on spatial location. The data was exported and comparative statistics for population density, intersection density, land use and road classification were calculated in Excel for the hot spots, cold spots, random (non-significant) locations and the total

metropolitan Melbourne area. As 'Water' made up only 0.5 per cent of the total land use, it was added to 'Other Land Use'.

## Results and Discussion

Metropolitan Melbourne comprises of 31 local government areas covering an area of 8,800 square kilometres. It has over 30,000 kilometres of roads, 5,000 kilometres of which are freeways, highways and arterial roads. The population of metropolitan Melbourne has increased from 4.2 million in 2012 to 4.6 million in 2016 and is now 4.9 million (as at 30 June 2018; Australian Bureau of Statistics, 2019).

Between 2012 and 2016, there were 50,305 casualty crashes in metropolitan Melbourne (VicRoads, 2018). This represents 71 per cent of the total casualty crashes in Victoria for this period. Figure 1 shows the location of the casualty crashes in metropolitan Melbourne. Of the 50,305 casualty crashes:

- 565 were fatal crashes where one or more persons were killed;
- 14,394 were serious injury crashes, where one or more persons required hospitalisation;
- 35,346 were other (minor) injury crashes.

### Exploratory spatial statistical analysis: spatial autocorrelation and spatial clustering

The Spatial Autocorrelation (Moran's I) tool identified statistically significant clustering for 2012-2016 metropolitan Melbourne casualty crashes. Running the tool for each of the crash categories found that crash clustering was statistically significant at a global level for all the crash categories analysed ( $p < .001$ ) except for fatal crashes ( $p = .569$ ) and crashes involving light commercial vehicles ( $p = .286$ ), buses ( $p = .185$ ) and trams ( $p = .224$ ) indicating

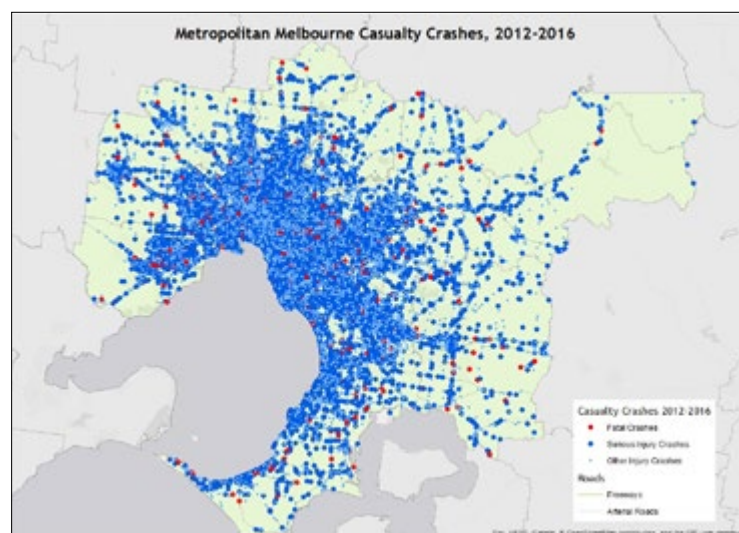


Figure 1. Metropolitan Melbourne casualty crash locations – 2012 to 2016

a random global pattern for these four crash categories. Increasing the distance threshold reduced the p-values for these four crash categories, however even with a distance threshold of 10,000m the p-values were greater than 0.1. This method however could not identify local clustering.

### Comparison of analysis and aggregation methods for local clustering

The KDE method was able to show areas with local crash clustering. The map in Figure 2 shows the results of the KDE analysis method using a 100m cell size (the 500m cell size gave similar results but the output was coarser). Density was calculated using a kernel function and the output (casualty crashes per square kilometre) was classified using 'Natural Breaks (Jenks)' with darker areas representing higher crash densities.

The classification method chosen to symbolise the results was an important consideration as different classification

methods led to different interpretations of the results. The 'Natural Breaks (Jenks)' classification resulted in a high-density area in inner Melbourne, with decreasing density away from inner Melbourne. In comparison, other than for the inner Melbourne area, 'Equal Interval' classification resulted in low density for most of metropolitan Melbourne (refer to Figure 3) whilst 'Quantile' classification resulted in a very extensive high-density area radiating out from the inner Melbourne area, extending into middle Melbourne and decreasing in intensity in outer metropolitan Melbourne (refer to Figure 4).

Whilst the KDE method was able to show areas where there was crash clustering, it was not able to show the statistical significance of the crash clusters (at the time of this analysis KDE+ was not available). In addition, the spatial extent of the density categories in the KDE was large, which led to local patterns being hidden.

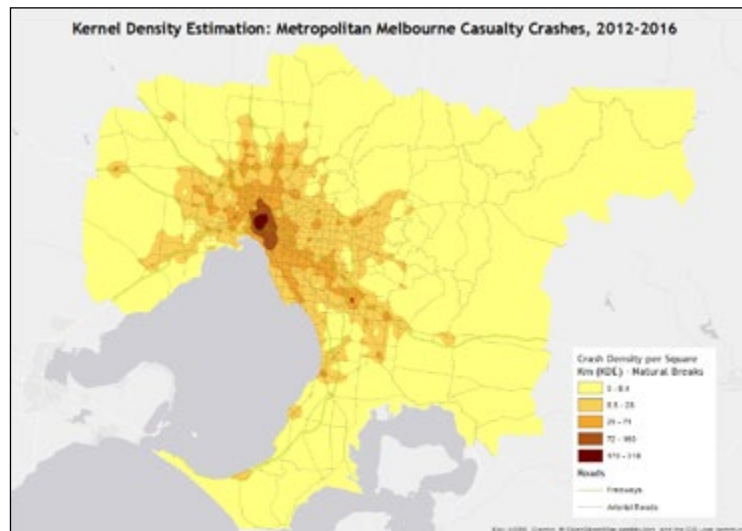


Figure 2. Kernel Density Estimation method, using Natural Breaks for classification

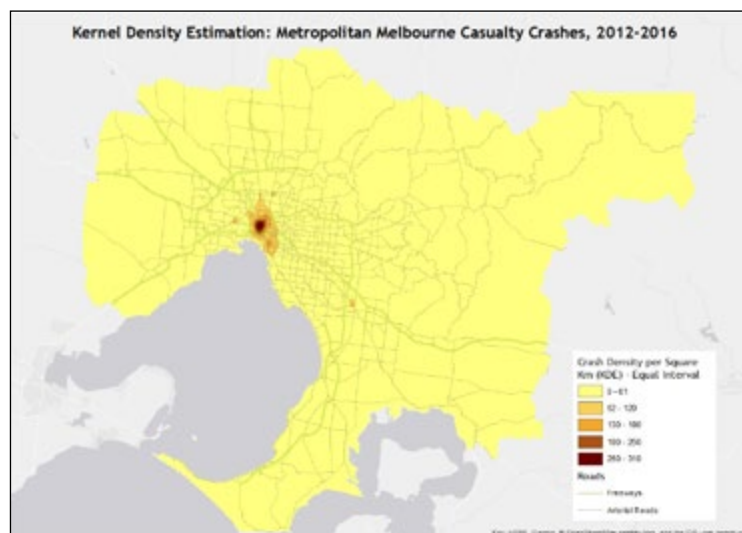


Figure 3. Kernel Density Estimation method, using Equal Intervals for classification

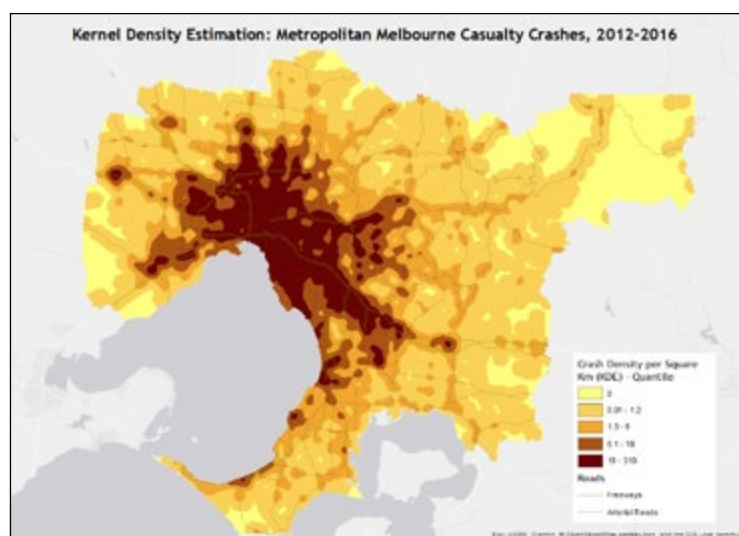


Figure 4. Kernel Density Estimation method, using Quantiles for classification

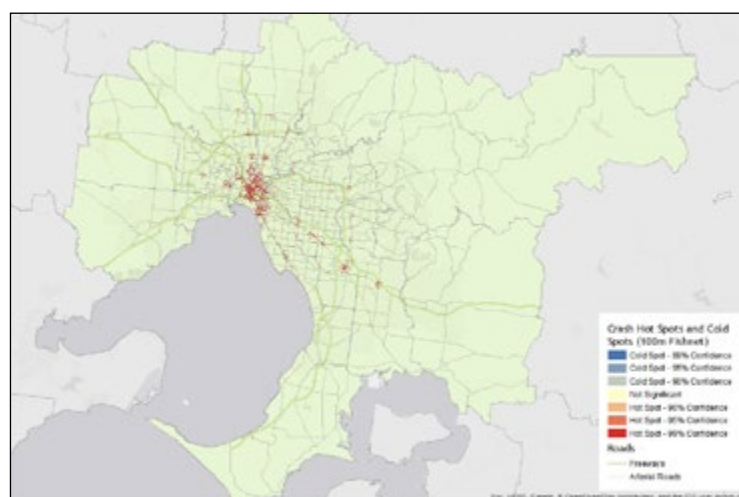


Figure 5. Metropolitan Melbourne casualty crash hot spots and cold spots, 2012-2016 - based on Getis-Ord Gi\* analysis using 100m fishnet aggregation

The Getis-Ord Gi\* hot spot analysis method was able to identify statistically significant crash clusters. As this method works with aggregated data but crash data is event-based, an important consideration was how to aggregate the data since different aggregation methods can lead to different results. Aggregation to fishnet (a grid of fixed size) and aggregation to point were investigated. The aggregation to administrative boundaries, such as the ABS SA1 areas, was not pursued due to problems in determining which area to allocate boundary crashes (the administrative boundaries generally coincide with the road network and the majority of crashes are assigned to the centreline of the road). Both aggregation to 1000m fishnet and aggregation to 500m fishnet were too coarse, especially when normalised by kilometre of road. Aggregation to 100m fishnet gave good results but took a long time to process. Point aggregation with no integration gave reasonable results, however the point aggregation with 10m integration identified more statistically significant hot spots.

The maps in Figures 5 and 6 show the output of the Getis-Ord Gi\* hot spot analysis method with aggregation to 100m fishnet (grids). The map in Figure 7 shows the output of the Getis-Ord Gi\* hot spot analysis method with point aggregation and 10m integration. Both methods gave similar results in terms of locations and extent, especially for the larger clusters (see Figures 8 and 9). Although both methods gave good results in terms of identifying areas where there were statistically significant hot spot and cold spot clusters, the latter analysis ran much faster and had the advantage in that it identified actual hot spot and cold spot points as distinct to areas. This was the method chosen for the subsequent crash analysis, using a distance band of 1km.

A limitation with this analysis is that all the methods tested assumed a planar (two-dimensional) distribution of crashes however crashes are constrained to a one-dimensional road network. When spatial analysis methods designed for planar surfaces are applied to network-constrained

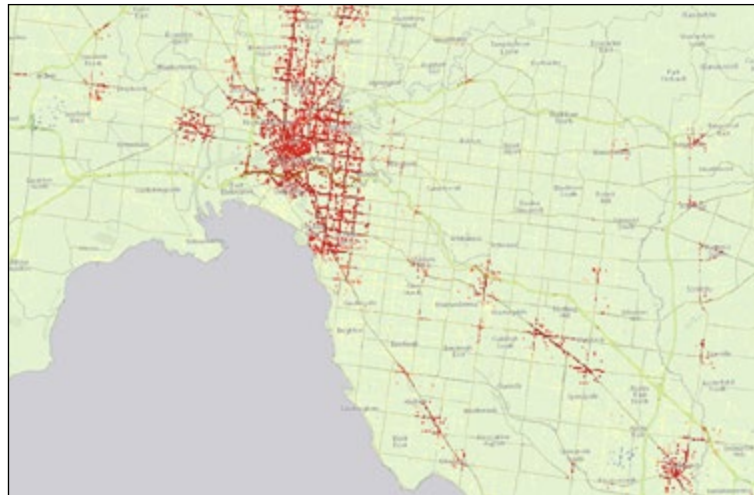


Figure 6. Metropolitan Melbourne casualty crash hot spots and cold spots, 2012-2016 - based on Getis-Ord Gi\* analysis using 100m fishnet aggregation – Melbourne zoom-in

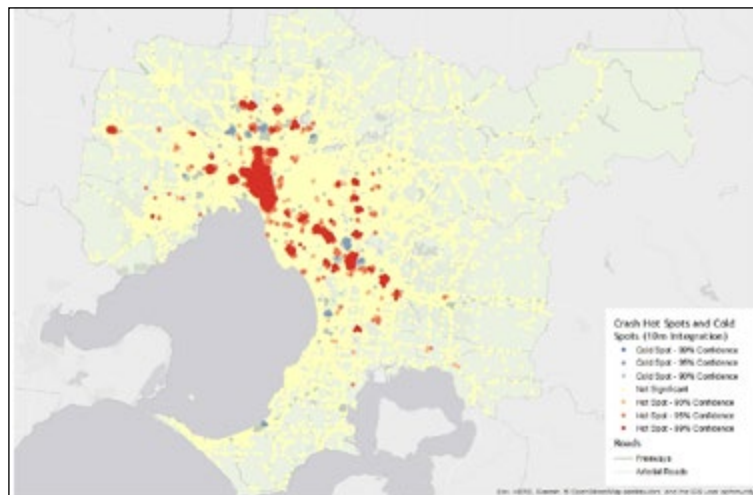


Figure 7. Metropolitan Melbourne casualty crash hot spots and cold spots, 2012-2016 - based on Getis-Ord Gi\* analysis and point aggregation with 10m integration

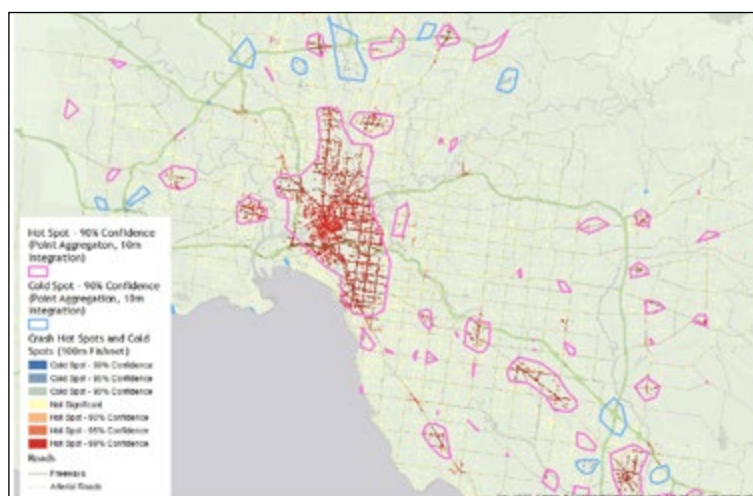


Figure 8. Comparison of aggregation to 100m fishnet with point aggregation with 10m integration for identification of hot spots and cold spots for metropolitan Melbourne 2012-2016 casualty crashes



**Figure 9. Comparison of aggregation to 100m fishnet (shown as areas) with point aggregation with 10m integration (shown as dots) for identification of hot spots and cold spots for metropolitan Melbourne 2012-2016 casualty crashes – Dandenong zoom**

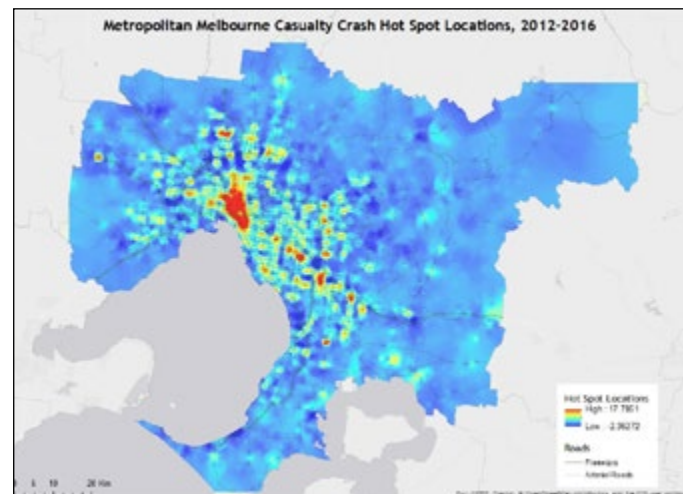


**Figure 10. Casualty crash hot spots in metropolitan Melbourne, 2012 to 2016**

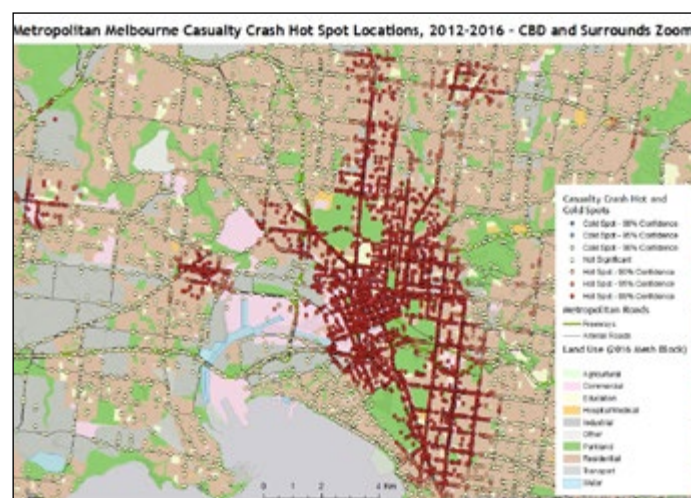


**Figure 11. CBD and surrounds, Footscray, Sunshine and Preston zoom: casualty crash hot spots in relation to land use and arterial road network, 2012 to 2016**

phenomena such as crashes, it is possible for a crash cluster to be identified just because the network is clustered (Yamada & Thill, 2007) therefore network-based analysis may be useful for identifying crash hot spots and cold spots. One network-based analysis tool is the SANET (Spatial Analysis on a NETwork) tool (Okabe, Okunuki, & Shiode, 2005).

#### **Local clustering (Getis-Ord Gi\* statistical method): metropolitan Melbourne casualty crashes**

The Hot Spot Analysis (Getis-Ord Gi\*) tool identified 4,523 locations (15.7% of total crash locations) as hot spots at the 0.05 significance level. Hot spots were located in areas where there is high trip attraction, such as the Melbourne CBD and its surrounds, and major activity centres such as Dandenong and Footscray (refer to Figures 10 and 11). For more examples of hot spot locations for the different crash categories, refer to the Appendix.

#### **Local clustering (Getis-Ord Gi\* statistical method): temporal analysis**

Temporal differences were found in the level, location and extent of clustering (see Figure 12). Some locations such as the CBD have remained hot spots from year to year, whilst other locations were hot spots in a particular year, or years, only (e.g. Bundoora), some hot spots decreased over time (e.g. Campbellfield) and other hot spots emerged over time (e.g. Caroline Springs, Carrum Downs).

Understanding which hot spots are persistent or emerging, and why, is important in determining which sites to treat and what treatments to apply. Similarly, understanding the characteristics of the declining hot spots is beneficial as the learnings could be applied to the persistent and emerging hot spots.

Weekdays were found to have a higher percentage of locations that exhibited statistically significant clustering compared to weekends (refer to Figure 13). The extent of clustering in the CBD and in Dandenong (both commercial locations) was greater on weekdays than it

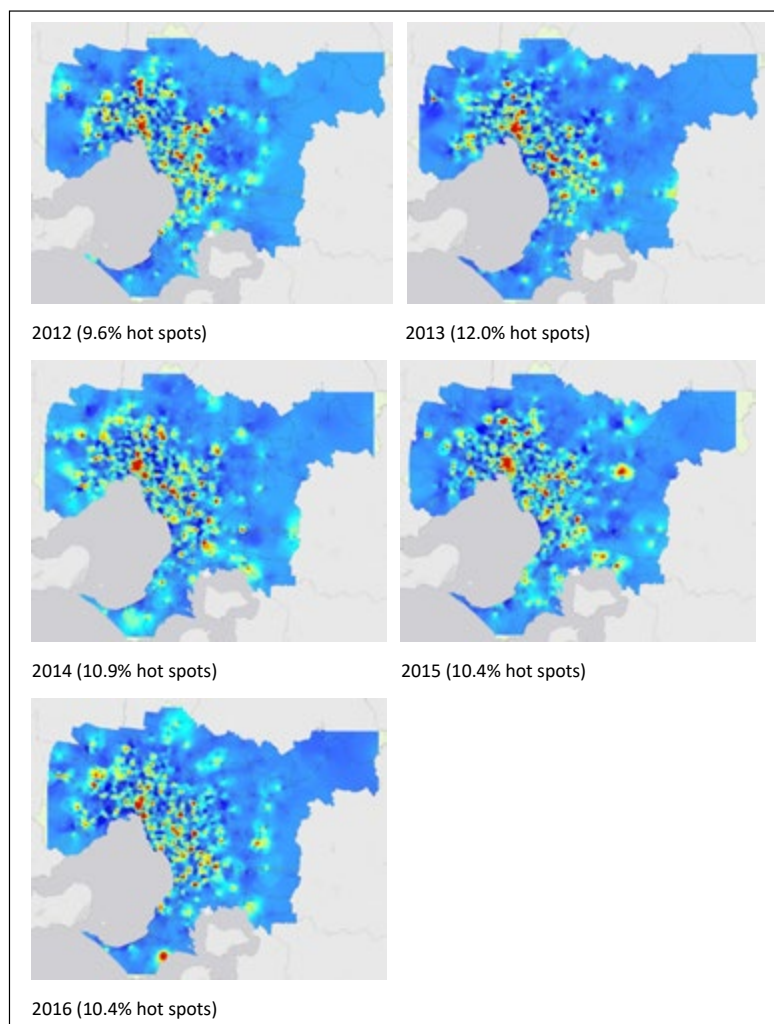


**Figure 12. Comparison of metropolitan Melbourne casualty crash hot spots by year: percentage of locations that are hot spots (95% confidence level)**

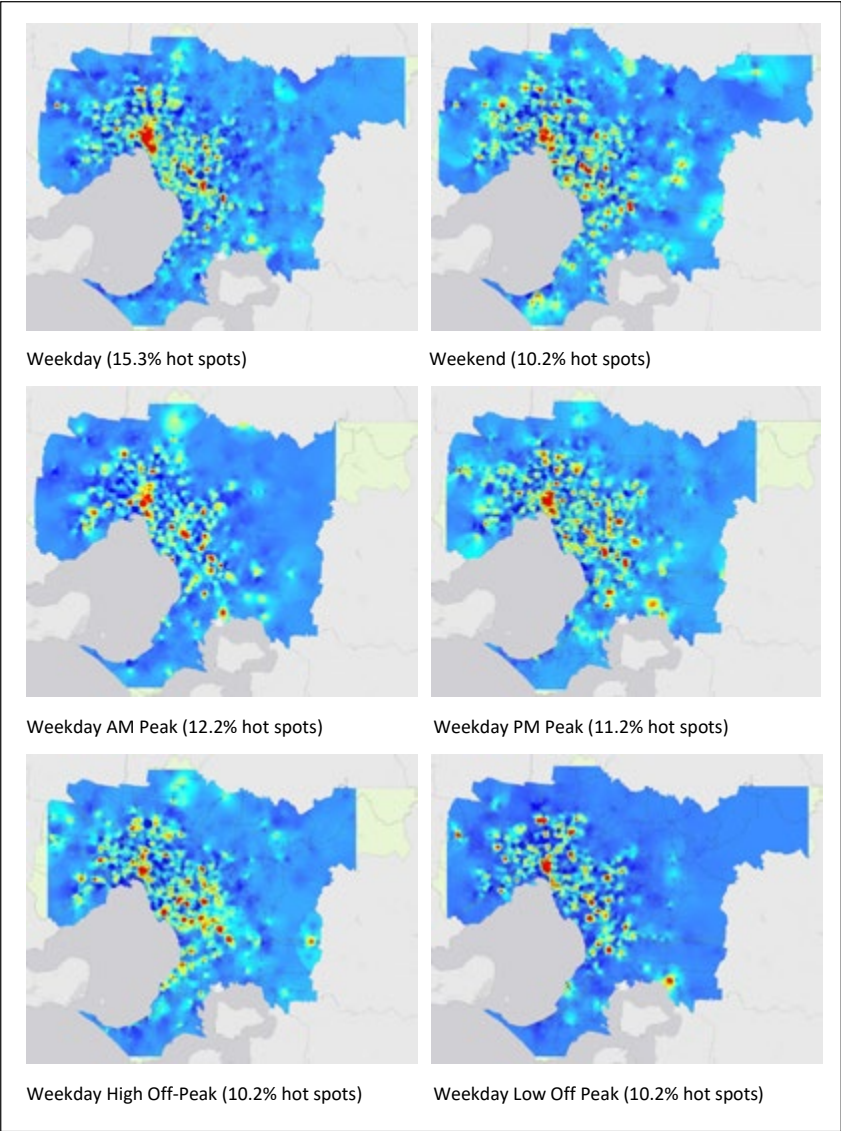


Figure 13. Comparison of metropolitan Melbourne 2012-2016 casualty crash hot spots by weekday, weekend and weekday time-periods (95% confidence level)

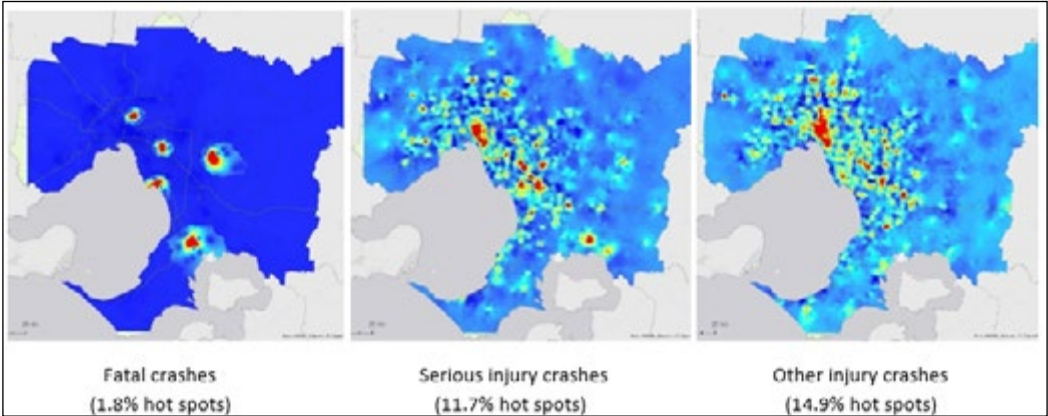


Figure 14. Comparison of metropolitan Melbourne 2012-2016 casualty crash hot spots by crash severity (95% confidence level)

was on weekends. Some locations, such as the Mornington Peninsula and Brighton (coastal locations popular for recreational activities and tourism), were hot spots on weekends but not on weekdays whilst others, such as Footscray and Wantirna (commercial locations), were hot spots on weekdays but not on weekends. Some locations had clustering in some of the periods but not in the others. For example, Melton had crash clustering in the weekday PM peak period and low off-peak period but not during the other periods.

#### Local clustering (Getis-Ord Gi\* statistical method): crash severity

The analysis showed differences in the hot spot locations and the level of clustering between fatal crashes, serious injury crashes and other (minor) injury crashes (refer to Figure 14). Fatal crashes had the lowest level of clustering.

For serious injury crashes and other injury crashes, some locations were hot spots for both severities, whilst other locations were hot spots for only one of these severities.

Narre Warren is an example of a location that had no clustering for serious injury crashes but did have clustering for other injury crashes. Caroline Springs is an example of a location that had clustering for serious injury crashes but no clustering for other injury crashes.

#### Local clustering (Getis-Ord Gi\* statistical method): crash type

Differences were found in the hot spot locations and level of clustering for the different crash types (refer to Figure 15). The pedestrian crash type has the highest clustering, more than double that of run off road crashes, head on not overtaking crashes and rear-end crashes and nearly double that of side impact at intersection crashes. Lane change / side swipe crashes had the next highest clustering.

- Pedestrian crash hot spots were located mainly in the inner and middle suburbs of Melbourne and included the CBD and other high pedestrian activity areas such as St Kilda and Richmond.

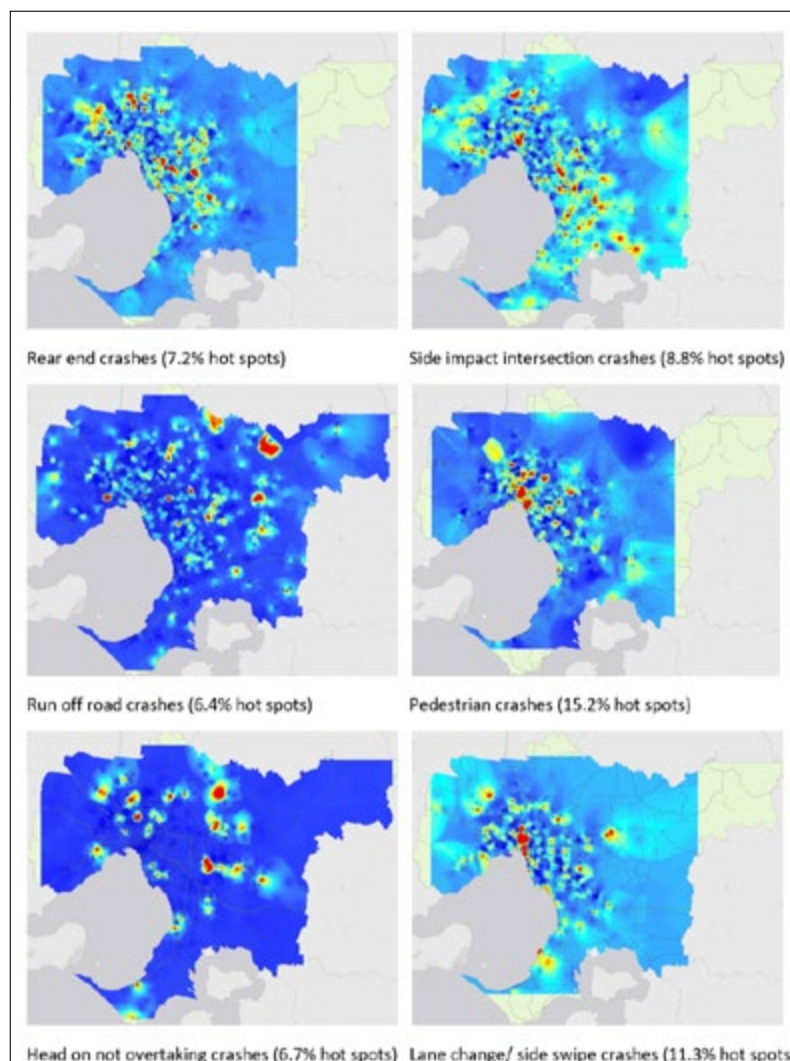


Figure 15. Comparison of metropolitan Melbourne 2012-2016 casualty crash hot spots by crash type (95% confidence level)

- Run off road crash hotspots were clustered in the outer suburbs and some middle suburbs of Melbourne and included areas such as Melton and Belgrave.
- Similarly head on not overtaking crash hot spot clusters were mainly located in the middle and outer suburbs. The head on not overtaking hot spot clusters were often located on sections of roads rather than over an area and included roads like Belgrave-Hallam Road and Maroondah Highway. Many of these roads were curved, although some like Barry Road were straight.
- Rear end crash hot spot clusters are mainly located around major intersections and on motorways (freeways and tollways) within a radius of 20km north and west and a radius of 45km south and east from the CBD. Rear end crash hot spot clusters included roads such as Dandenong Road, the Monash Freeway, the West Gate Freeway and Hume Highway.

- Side impact at intersection hot spot clusters included intersections in the CBD and surrounds, and in middle and outer metropolitan Melbourne areas such as Clayton and Mornington.
- Lane change and side swipe crash hot spot clusters are mainly located in the inner metropolitan Melbourne area including the CBD, and areas such as South Melbourne. There was also a cluster at Brighton.

#### Local clustering (Getis-Ord Gi\* statistical method): vehicle type

Differences were found in the hot spot locations and level of clustering for the different vehicle types (refer to Figure 16). Crashes involving bicycles had the highest clustering, nearly double that of crashes involving passenger cars. Crashes involving light commercial vehicles and crashes involving buses / coaches had the lowest clustering of the vehicle types.

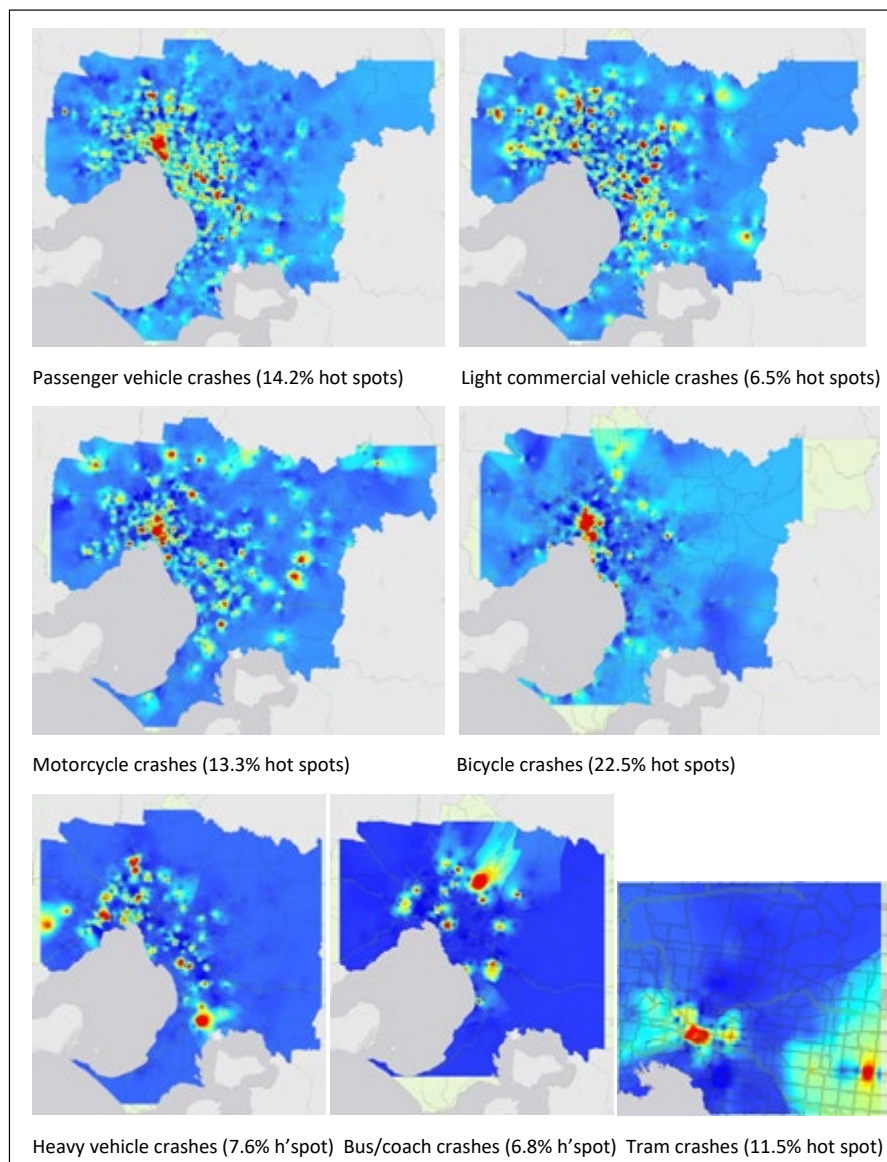


Figure 16. Comparison of metropolitan Melbourne 2012-2016 casualty crash hot spots by vehicle type (95% confidence level)

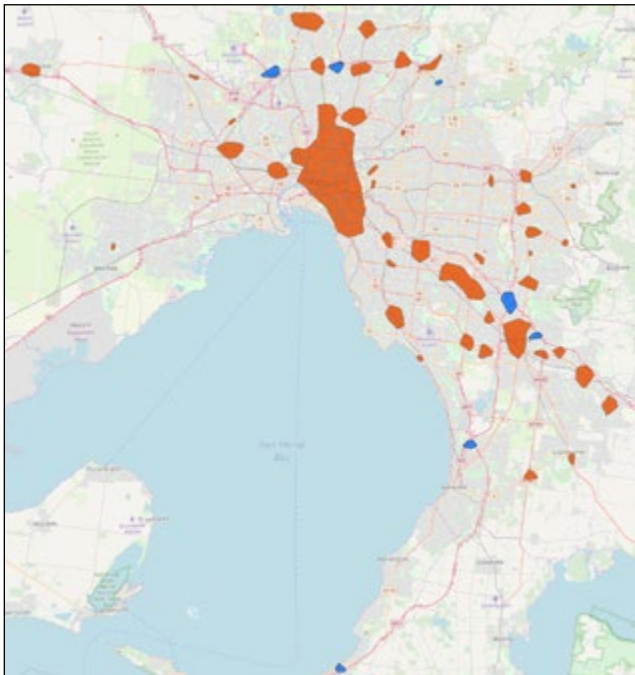


**Figure 17. Polygons of statistically significant (95% confidence level) hot spot crash clusters (shown in brown) and cold spot crash clusters (shown in blue) for metropolitan Melbourne 2012-2016 casualty crashes**

- As passenger vehicles were involved in 84 per cent of casualty crashes, the hot spot clusters for crashes involving this vehicle type were almost identical to that for total casualty crashes. One exception was Box Hill which was identified as a hot spot crash cluster for crashes involving passenger vehicles but not for total casualty crashes.
  - Light commercial vehicle hot spot crash clusters were mainly located in North Melbourne and Docklands and in middle and outer metropolitan suburbs. They were mainly located on freeways, the approaches to freeways, along major arterial roads (including highways) and around industrial and commercial areas.
  - Only 34 locations were identified as hot spots for bus crashes, with most clusters comprising of one or two sites only. Dandenong, Wantirna South, Mordialloc, Toorak, Preston and Epping had hot spot clusters of three or more sites.
  - For tram crashes, the CBD, St Kilda Road between Flinders Street and Southbank Boulevard and Carlton were locations that had a cluster of hot spots. Burwood East had one site that was a tram hot spot, as did East Melbourne.
- Bicycle crash hot spot clusters were mainly located along popular bicycle routes in inner Melbourne (including the CBD) and along St Kilda Road, Chapel Street and Beach Road.
  - Heavy vehicle crash hot spot clusters were mainly located around industrial areas such as Laverton, and Campbellfield. The clusters were also located on roads leading to ports, such as Western Port Highway, or on roads connecting ports with industrial areas such as the West Gate Bridge.
  - Motorcycle crash hot spot clusters were mainly located in inner and middle metropolitan areas such as the CBD, North Melbourne, Brunswick and Clayton. There were also some clusters in the outer metropolitan Melbourne areas, particularly on windy roads.

### Explanatory spatial statistical analysis

There were 42 clusters with three or more statistically significant hot spot locations (at the 95% confidence level) and seven clusters of three or more statistically significant cold spot locations (at the 95% confidence level) identified from the Getis-Ord Gi\* analysis of metropolitan Melbourne 2012-2016 casualty crashes. Figure 17 shows the clusters of hot spots (brown) and cold spots (blue).

A comparison of the characteristics of the hot spot and cold spot clusters is shown Table 1. Both hot spot and cold spot clusters had a higher density of population, road network, intersection and casualty crashes than the total metropolitan Melbourne area. The lower density in the total metropolitan Melbourne area is not surprising as a large proportion of Melbourne is comprised of parkland and agricultural areas. Although road density and, to a

**Table 1. Comparison of characteristics of hot spot clusters, cold spot clusters and the total metropolitan Melbourne area**

|                                                                           | Hot Spot Clusters | Cold Spot Clusters | Total Metropolitan Melbourne |
|---------------------------------------------------------------------------|-------------------|--------------------|------------------------------|
| Population density: persons per square kilometre (based on 2016 ABS data) | 3,427             | 1,833              | 527                          |
| Road density: kilometre road per square kilometre                         | 14                | 14                 | 3.8                          |
| Intersection density: number of intersections per kilometre of road       | 4.6               | 3.7                | 3.3                          |
| 2012-2016 casualty crashes per kilometre of road                          | 5.6               | 2.0                | 1.5                          |
| 2012-2016 casualty crashes per square kilometre                           | 78                | 28                 | 5.7                          |

lesser extent, intersection density were similar between the hot spot and cold spot clusters, hot spot clusters had a much higher population density (nearly double) and a much higher crash density (nearly three times higher) than cold spot clusters.

### Road classification

Local roads make up the highest proportion of the road network in metropolitan Melbourne, and in the hot spot and cold spot clusters (refer to Figure 18). Highways, arterial roads and sub-arterial roads made up only 16 per cent of metropolitan road network, yet they made up 22 per cent of the road network in the hot spot clusters. This is not surprising as these road types carry high volumes of traffic. Normalising crashes by traffic volume could be undertaken prior to running the spatial analyses

to identify hazardous locations without the influence of traffic volumes and is a recommendation for future research. Although freeways carry even higher volumes of traffic, they only made up 3.6 per cent of the road network in the hot spot clusters, slightly higher than in the total metropolitan area (3%). However, in the cold spot clusters they represented 30 per cent of the road network. One potential reason for that is that freeways are built to high safety standards, are divided, and have restricted entry which limits some vehicle-to-vehicle interactions and vehicle-to-pedestrian (and cyclist) interactions, and as a result are safer (in terms of crash rates per vehicle kilometre travel) compared to other road types. Freeway crashes tend to be more evenly dispersed as they are often related to the traffic conditions rather than to the infrastructure itself (Hovenden, Zurlinden & Gaffney, 2020).

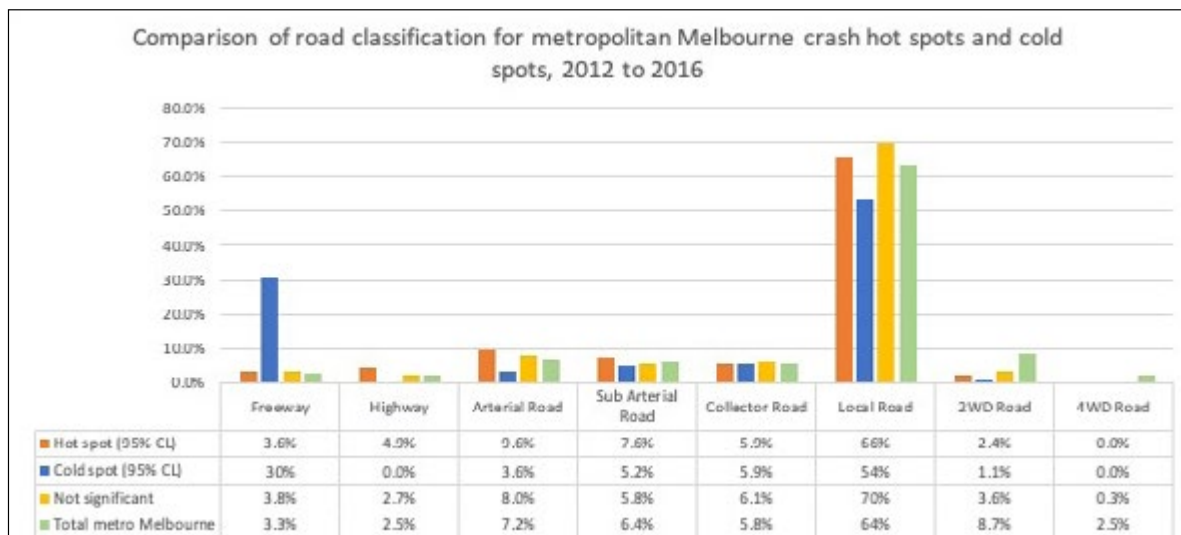


Figure 18. Comparison of percentage of roads in casualty crash hot spot clusters, cold spot clusters, non-significant (random) crash locations and total metropolitan Melbourne area by road classification, 2012-2016

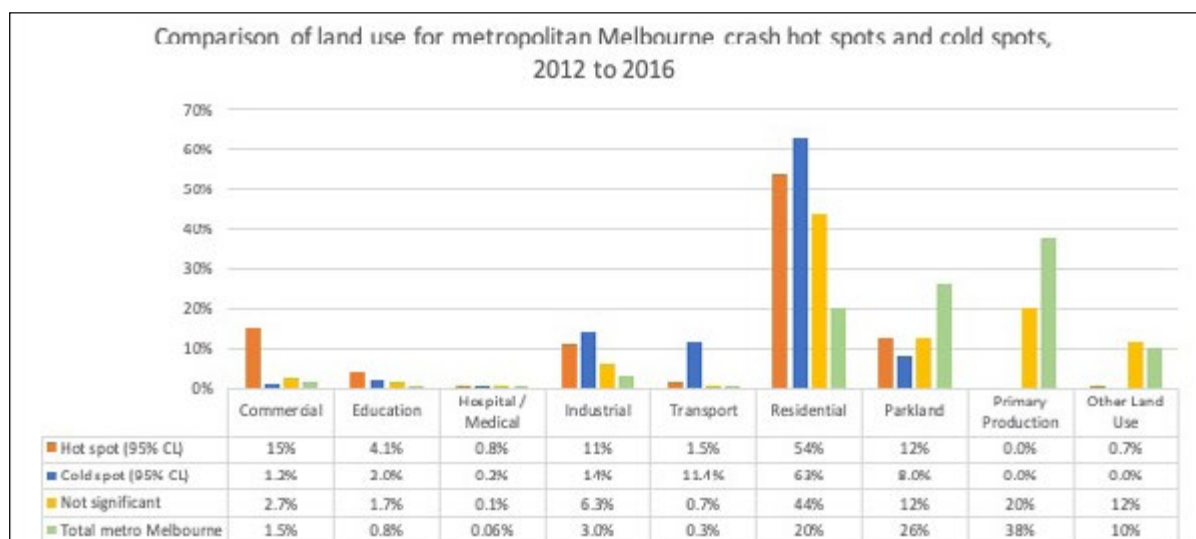


Figure 19. Comparison of percentage of roads in casualty crash hot spot clusters, cold spot clusters, non-significant (random) crash locations and total metropolitan Melbourne area by land use, 2012-2016

## Land use

Metropolitan Melbourne is primarily made up of primary production (agricultural) land use (38%), parkland (26%) and residential land use (20%). Only three per cent of Melbourne's land use is industrial, one per cent is commercial and less than one per cent each is education, medical or transport related (refer to Figure 19).

Exploratory analysis of the relationship between statistically significant crash clusters and land use found that hot spot clusters were in areas with a higher percentage of commercial land use (15%) compared to both cold spot clusters (1.2%) and to the total metropolitan Melbourne area (1.5%), suggesting a relationship between crash cluster location and level of activity. Although numbers are small, hot spot clusters also had a higher percentage of education land use and hospital / medical land use, further supporting this observation. As mentioned earlier, normalising crashes by traffic volume could be undertaken prior to running the spatial analyses to identify hazardous locations whilst taking into account the level of activity.

Both hot spot clusters and cold spot clusters had a higher percentage of industrial land use compared to the total metropolitan Melbourne area (11% and 14% respectively compared to 3%). As industrial areas have a strong temporal component, the percentage of this land use in hot spot clusters could be even higher during certain times and this is an area for further investigation.

Cold spot clusters had a much higher percentage of transport (rail and road) land use (11%) compared to both the total metropolitan Melbourne area (0.3%) and to hot spot clusters (1.5%). It should be noted that transport land use includes freeways, which are discussed in the Road Classification section above.

Further work is required to test the statistical significance of the relationship between the clusters, road classification and land use. It is recommended that spatial techniques including Exploratory Regression analysis, Ordinary Least Squares and Geographically Weighted Regression analysis be undertaken to identify the factors influencing the location of the clusters and explore the relationships between these factors and the clusters.

## Conclusions

Metropolitan Melbourne is a rapidly growing area in terms of population and has a sizeable casualty crash problem. Whilst much research of metropolitan Melbourne crashes has been conducted using traditional statistical methods, there has been limited analyses of these crashes using spatial techniques. Crashes are a spatial phenomenon and patterns can be missed using traditional techniques. Without spatial analysis, it is not possible to determine whether crashes are randomly dispersed or clustered which can lead to errors in interpreting results. This research

addressed this gap by providing an understanding of the influence of spatial factors on metropolitan Melbourne crashes. The techniques used in this research allowed identification of locations in metropolitan Melbourne with statistically significant crash clusters. Visualisation of the results through mapping provided a better understanding of the spatial and temporal distribution of crashes and allowed identification of persistent and emerging hot spots.

The Hot Spot Analysis (Getis-Ord  $G_i^*$ ) statistic method identified 15.7 per cent of casualty crash locations in metropolitan Melbourne as statistically significant hot spots at the 95 per cent confidence level. Running this method on the different crash categories found that the level (that is, percentage) of crash clustering and the location and extent of the clusters is dependent on temporal factors such as year, time of day and day of week, and on crash severity, type of crash and type of vehicle involved in the crash. Analysis of the clusters with three or more statistically significant hot spot and cold spot locations (at the 95% confidence level) found that hot spot clusters were located in areas with a higher proportion of commercial land use compared to cold spots and to the total metropolitan Melbourne area, suggesting a relationship between crashes and level of activity. Cold spot clusters were located in areas with a higher proportion of freeways, which could be explained by the high safety standards of freeway infrastructure and controlled access.

The results of this analysis can be used to inform the development of strategies, initiatives and treatments to improve road safety on metropolitan Melbourne roads. The spatial techniques used in this research have been applied to systematically identify those locations where the crash clusters are statistically significant. A targeted, evidence-based approach can be adopted to further investigate these locations, find common factors and to develop treatments. This approach can be expanded to the whole of Victoria. Future work in this area includes network-based hot spot analysis, normalisation by traffic volume, and exploration of the relationship between crash clusters and the factors influencing the location of those clusters using spatial techniques such as Exploratory Regression analysis and Geographically Weighted Regression.

## Acknowledgements

The authors would like to thank John Gaffney, Amy Stebbing and Hendrik Zurlinden for their feedback.

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

## References

- Anselin, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27(2), 93-115. doi:10.1111/j.1538-4632.1995.tb00338.x
- Australian Bureau of Statistics. (2018). *Regional Population Growth, Australia, 2016-17, Population Estimates by Local Government Area (ASGS 2017), 2016 to 2017* (No. 3218.0). Canberra, Australia. Retrieved from <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02016-17?OpenDocument>
- Australian Bureau of Statistics. (2019). *Regional Population Growth, Australia, 2017-18, Population Estimates by Local Government Area (ASGS 2018), 2017 to 2018* (No. 3218.0). Canberra, Australia. Retrieved from <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02017-18?OpenDocument>
- Benedek, J., Ciobanu, S.M. & Man, T.C. (2016). Hotspots and social background of urban traffic crashes: A case study in Cluj-Napoca (Romania). *Accident Analysis and Prevention*, 87, 117-126. doi:10.1016/j.aap.2015.11.0260001-4575
- Bíl, M., Andrášik, R., & Janoška, Z. (2013). Identification of hazardous road locations of traffic accidents by means of kernel density estimation and cluster significance evaluation. *Accident Analysis & Prevention*, 55, 265-273. doi:10.1016/j.aap.2013.03.003
- Chance Scott, M., Sen Roy, S. & Prasad, S. (2016). Spatial patterns of off-the-system traffic crashes in Miami-Dade County, Florida, during 2005-2010. *Traffic Injury Prevention*, 17(7), 729-735. doi:10.1080/15389588.2016.1144878
- Choudhary, J., Ohri, A. & Kumar, B. (2015, December). *Spatial and statistical analysis of road accidents hot spots using GIS*. Paper presented at the 3rd Conference of Transportation Research Group of India (3rd CTRG). Kolkata, India. Retrieved from [https://www.researchgate.net/publication/291973317\\_Spatial\\_and\\_statistical\\_analysis\\_of\\_road\\_accidents\\_hot\\_spots\\_using\\_GIS](https://www.researchgate.net/publication/291973317_Spatial_and_statistical_analysis_of_road_accidents_hot_spots_using_GIS)
- Erdogan, S. (2009). Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, 40(5), 341-351. doi:10.1016/j.jsr.2009.07.006
- Esri. (n.d.-a). Feature class basics. Retrieved from <http://desktop.arcgis.com/en/arcmap/10.3/manage-data/geodatabases/feature-class-basics.htm>
- Esri. (n.d.-b). How Spatial Autocorrelation (Global Moran's I) works. Retrieved from <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm>
- Esri. (n.d.-c). How Hot Spot Analysis (Getis-Ord Gi\*) works. Retrieved from <http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm>
- Getis, A. & Ord, J.K. (1992). The Analysis of Spatial Association by Use of Distance Statistic. *Geographical Analysis*, 24(3), 189-206. doi:10.1111/j.1538-4632.1992.tb00261.x
- Gudes, O., Varhol, R., Sun, Q.C. & Meuleners, L. (2017). Investigating articulated heavy-vehicle crashes in Western Australia using a spatial approach. *Accident Analysis and Prevention*, 106, 243-253. doi:10.1016/j.aap.2017.05.026
- Hovenden, E., Zurlinden, H. & Gaffney, J. (2020). Safety on Heavily Trafficked Urban Motorways in Relation to Traffic State. *Journal of Road Safety*, 31(1), 51-65. doi: 10.33492/JRS-D-19-00247
- Lawrence, B.M., Stevenson, M.R., Oxley, J.A. & Logan, D.B. (2015). Geospatial Analysis of Cyclist Injury Trends: An Investigation in Melbourne, Australia. *Traffic Injury Prevention*, 16(5), 513-518. doi:10.1080/15389588.2014.973947
- Li, Z., Wang, W., Liu, P., Bigham, J.M. & Ragland, D.R. (2013). Using Geographically Weighted Poisson Regression for county-level crash modeling in California. *Safety Science*, 58, 89-97. doi:10.1016/j.ssci.2013.04.005
- Liu, X. & Xia, J. (2015). Locally analysing the risk factors for fatal single vehicle crashes hot spots in Western Australia. *International journal of crashworthiness*, 20(6), 524-534. doi:10.1080/13588265.2015.1055649
- Loo, B.P.Y. & Yao, S. (2012). Geographical information systems. In G. Li & S.P. Baker (Eds), *Injury Research: Theories, Methods, and Approaches*. (pp. 447-463). New York: Springer.
- Mohaymany, A.S., Shahri, M. & Mirbagheri, B. (2013). GIS-based method for detecting high-crash-risk road segments using network kernel density estimation. *Geo-spatial Information Science*, 16 (2), 113-119. doi:10.1080/10095020.2013.766396
- Nie, K., Wang, Z., Du, Q., Ren, F. & Tian, Q. (2015). A Network-Constrained Integrated Method for Detecting Spatial Cluster and Risk Location of Traffic Crash: A Case Study from Wuhan, China. *Sustainability*, 7(3), 2662-2677. doi:10.3390/su7032662
- Okabe, A., Okunuki, K. & Shiode, S. (2005). SANET: A Toolbox for Spatial Analysis on a Network. *Geographical Analysis*, 38(1), 57-66. doi:10.1111/j.0016-7363.2005.00674.x
- Ord, J.K. & Getis, A. (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, 27(4), 286-306. doi:10.1111/j.1538-4632.1995.tb00912.x
- Plug, C., Xia, J.C. & Caulfield, C. (2011). Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis and Prevention*, 43(6), 1937-1946. doi:10.1016/j.aap.2011.05.007
- Pour, A.T., Moridpour, S., Rajabifard, A. & Tay, R. (2017). Spatial and temporal distribution of pedestrian crashes in Melbourne metropolitan area. *Road and Transport Research*, 26(1), 4-20.
- Prasannakumar, V., Vijith, H., Charutha, R., & Geetha, N. (2011). Spatio-Temporal Clustering of Road Accidents: GIS Based Analysis and Assessment. *Procedia - Social and Behavioral Sciences*, 21, 317-325. doi:10.1016/j.sbspro.2011.07.020

- Soltani, A. & Askari, S. (2017). Exploring spatial autocorrelation of traffic crashes based on severity. *Injury*, 48(3), 637-647. doi:10.1016/j.injury.2017.01.032
- Transport Research Centre (CDV). (2020). *KDE+*. Retrieved from <http://www.kdeplus.cz/en/>
- Truong, L.T. & Somenahalli, S.V.C. (2011). Using GIS to Identify Pedestrian-Vehicle Crash Hot Spots and Unsafe Bus Stops. *Journal of Public Transportation*, 14(1), 99-114. doi:10.5038/2375-0901.14.1.6
- Victoria State Government. (2016). *Towards Zero 2016-2020 Victoria's Road Safety Strategy & Action Plan*. Retrieved from [https://www.towardszero.vic.gov.au/\\_\\_data/assets/pdf\\_file/0010/183556/STU\\_0206\\_RS\\_STRATEGY\\_2016\\_web.pdf](https://www.towardszero.vic.gov.au/__data/assets/pdf_file/0010/183556/STU_0206_RS_STRATEGY_2016_web.pdf)
- VicRoads. (2018). Crash Stats – Data Extract. Retrieved from <https://www.data.vic.gov.au/data/dataset/crash-stats-data-extract>
- World Health Organisation. (2018). *Road traffic injuries Fact sheet*. Retrieved from <http://www.who.int/en/news-room/fact-sheets/detail/road-traffic-injuries>
- Yamada, I. & Thill, J.-C. (2007). Local Indicators of Network-Constrained Clusters in Spatial Point Patterns. *Geographical Analysis*, 39, 268-292. doi:10.1111/j.1538-4632.2007.00704.x
- Zhang, Y., Bigham, J., Ragland, D. & Chen, X. (2015). Investigating the associations between road network structure and non-motorist accidents. *Journal of Transport Geography*, 42, 34-47. doi:10.1016/j.jtrangeo.2014.10.010

## Appendix

### Supplementary information to methodology section

**Table A1. List of the 31 local government areas used to define metropolitan Melbourne**

|               |                   |                      |             |
|---------------|-------------------|----------------------|-------------|
| Banyule       | Bayside           | Boroondara           | Brimbank    |
| Cardinia      | Casey             | Darebin              | Frankston   |
| Glen Eira     | Greater Dandenong | Hobsons Bay          | Hume        |
| Kingston      | Knox              | Manningham           | Maribyrnong |
| Maroondah     | Melbourne         | Melton               | Monash      |
| Moonee Valley | Moreland          | Mornington Peninsula | Nillumbik   |
| Port Phillip  | Stonnington       | Whitehorse           | Whittlesea  |
| Wyndham       | Yarra             | Yarra Ranges         |             |

**Table A2. Categories used for crash analysis**

| Category     | Group                       | Definition                                                                                                      |
|--------------|-----------------------------|-----------------------------------------------------------------------------------------------------------------|
| Time-Period  | AM peak                     | Weekday, 06:00 to 09:59                                                                                         |
| Time-Period  | PM peak                     | Weekday, 15:00 to 18:59                                                                                         |
| Time-Period  | High off-peak               | Weekday, 10:00 to 14:59                                                                                         |
| Time-Period  | Low off-peak                | Weekday, 19:00 to 23:59 and 00:00 to 05:59                                                                      |
| Crash Type   | Rear end                    | DCA_CODE between 130 and 132                                                                                    |
| Crash Type   | Side impact at intersection | DCA_CODE in (110, 111, 113, 114, 116, 121, 122) and NODE_TYPE = 'I' (intersection)                              |
| Crash Type   | Run off road                | DCA_CODE in (151, 170, 171, 172, 173, 180, 181, 182, 183)                                                       |
| Crash Type   | Pedestrian                  | DCA_CODE between 100 and 109                                                                                    |
| Crash Type   | Head on not overtaking      | DCA_CODE = 120                                                                                                  |
| Crash Type   | Lane change or side swipe   | DCA_CODE between 133 and 137                                                                                    |
| Vehicle Type | Heavy vehicle               | VEHICLE_TYPE in (6, 7, 60, 61, 62, 63, 72)<br>(includes rigid trucks > 4.5 tonnes GVM and articulated vehicles) |
| Vehicle Type | Motorcycle                  | VEHICLE_TYPE in (10, 11, 12)                                                                                    |
| Vehicle Type | Bicycle                     | VEHICLE_TYPE = 13                                                                                               |
| Vehicle Type | Light commercial vehicle    | VEHICLE_TYPE in (4, 5, 71)<br>(includes rigid trucks <= 4.5 tonnes GVM, utilities and panel vans)               |
| Vehicle Type | Passenger vehicle           | VEHICLE_TYPE in (1, 2, 3, 9)                                                                                    |
| Vehicle Type | Bus                         | VEHICLE_TYPE = 8                                                                                                |
| Vehicle Type | Tram                        | VEHICLE_TYPE = 15                                                                                               |

**Table A3. Datasets used in the analysis**

| Dataset                                                                                                                                      | Source                                                                                                                                     | Description                                    | Comments                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|----------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| LGA_POLYGON                                                                                                                                  | <a href="http://services.land.vic.gov.au/SpatialDatamart">http://services.land.vic.gov.au/SpatialDatamart</a><br>(VICMAP_ADMIN)            | LGA boundaries                                 | A metropolitan Melbourne feature class was created using this dataset, based on the 31 local government areas listed in Table A1.<br><br>This feature class was then used to clip other spatial datasets to the metropolitan Melbourne boundary.                                                                                                                                                                                                                                                                                                                                                       |
| Crash Stats – Data Extract:<br><ul style="list-style-type: none"><li>• NODE table</li><li>• ACCIDENT table</li><li>• VEHICLE table</li></ul> | <a href="https://www.data.vic.gov.au">https://www.data.vic.gov.au</a>                                                                      | Crash data                                     | The NODE, ACCIDENT and VEHICLE tables were used to create a master crash dataset.<br><br>Crash locations within metropolitan Melbourne were selected from the NODE table using the ‘Intersect’ tool and the metropolitan Melbourne feature class. The ACCIDENT table was then filtered using the NODE table.<br><br>Variables in the ACCIDENT table were grouped to make querying easier.<br><br>The VEHICLE table was used to select specific vehicle categories (e.g ‘heavy vehicles’) and this was applied to the ACCIDENT table to identify crashes that involved the particular vehicle category. |
| TR_ROAD                                                                                                                                      | <a href="http://services.land.vic.gov.au/SpatialDatamart">http://services.land.vic.gov.au/SpatialDatamart</a><br>(VICMAP_TRANSPORT)        | Road network                                   | To exclude roads not open to the public and paths, the following definition query was used:<br>“CLASS_CODE” IN (0, 1, 2, 3, 4, 5, 6, 7) AND “RESTRICTN” NOT IN (‘1’, ‘4’, ‘5’).                                                                                                                                                                                                                                                                                                                                                                                                                        |
| MB_2016_VIC                                                                                                                                  | <a href="http://www.abs.gov.au">http://www.abs.gov.au</a>                                                                                  | Mesh block data – includes land use (MB_CAT16) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 2016 Census Mesh Block Counts                                                                                                                | <a href="http://www.abs.gov.au">http://www.abs.gov.au</a><br>2074.0 - Census of Population and Housing: Mesh Block Counts, Australia, 2016 | Mesh block population and household data       | The 2016 Census Mesh Block Counts (aspatial) dataset was joined to the MB_2016_VIC (spatial) dataset. The attributes used in the analysis included ‘MB_CAT16’ (land use category), ‘Dwelling’ and ‘Person’.                                                                                                                                                                                                                                                                                                                                                                                            |

## Results of Getis-Ord Gi\* statistic analysis of metropolitan Melbourne casualty crashes

| Category                     | Hot Spots<br>99%CL | Hot Spots<br>95%CL | Hot Spots<br>90%CL | Not<br>Significant | Cold Spots<br>90%CL | Cold Spots<br>95%CL | Cold Spots<br>99%CL | Total<br>Locations | Percentage<br>Hotspots<br>>=95%CL | Percentage<br>Hotspots<br>>=90%CL |
|------------------------------|--------------------|--------------------|--------------------|--------------------|---------------------|---------------------|---------------------|--------------------|-----------------------------------|-----------------------------------|
| Casualty Crashes 2012-2016   | 3456               | 1067               | 768                | 23030              | 392                 | 125                 | 0                   | 28838              | 15.7%                             | 18.3%                             |
| Casualty Crashes 2016        | 512                | 358                | 178                | 7264               | 17                  | 0                   | 0                   | 8329               | 10.4%                             | 12.6%                             |
| Casualty Crashes 2015        | 548                | 320                | 238                | 7264               | 3                   | 0                   | 0                   | 8373               | 10.4%                             | 13.2%                             |
| Casualty Crashes 2014        | 612                | 292                | 203                | 7165               | 10                  | 0                   | 0                   | 8282               | 10.9%                             | 13.4%                             |
| Casualty Crashes 2013        | 617                | 332                | 219                | 6732               | 5                   | 0                   | 0                   | 7905               | 12.0%                             | 14.8%                             |
| Casualty Crashes 2012        | 431                | 333                | 302                | 6880               | 6                   | 0                   | 0                   | 7952               | 9.6%                              | 13.4%                             |
| Weekday Crashes              | 2672               | 911                | 659                | 18696              | 315                 | 114                 | 0                   | 23367              | 15.3%                             | 18.2%                             |
| AM Peak Crashes              | 608                | 253                | 236                | 5930               | 15                  | 1                   | 0                   | 7043               | 12.2%                             | 15.6%                             |
| PM Peak Crashes              | 733                | 394                | 347                | 8537               | 55                  | 0                   | 0                   | 10066              | 11.2%                             | 14.6%                             |
| High Off Peak Crashes        | 487                | 290                | 210                | 6599               | 7                   | 1                   | 0                   | 7594               | 10.2%                             | 13.0%                             |
| Low Off Peak Crashes         | 442                | 172                | 164                | 5213               | 0                   | 0                   | 0                   | 5991               | 10.2%                             | 13.0%                             |
| Weekend Crashes              | 613                | 366                | 293                | 8349               | 17                  | 0                   | 0                   | 9638               | 10.2%                             | 13.2%                             |
| Fatal Crashes                | 10                 | 0                  | 0                  | 550                | 0                   | 0                   | 0                   | 560                | 1.8%                              | 1.8%                              |
| Serious Injury Crashes       | 959                | 381                | 321                | 9812               | 25                  | 0                   | 0                   | 11498              | 11.7%                             | 14.4%                             |
| Other Injury Crashes         | 2322               | 890                | 564                | 17515              | 242                 | 74                  | 1                   | 21608              | 14.9%                             | 17.5%                             |
| Rear End Crashes             | 359                | 257                | 252                | 7679               | 7                   | 0                   | 0                   | 8554               | 7.2%                              | 10.1%                             |
| Side Impact Crashes          | 359                | 183                | 149                | 5377               | 63                  | 9                   | 0                   | 6140               | 8.8%                              | 11.3%                             |
| Run off Road Crashes         | 262                | 116                | 74                 | 5418               | 0                   | 0                   | 0                   | 5870               | 6.4%                              | 7.7%                              |
| Pedestrian Crashes           | 460                | 159                | 110                | 3331               | 0                   | 0                   | 0                   | 4060               | 15.2%                             | 18.0%                             |
| Head On Crashes              | 50                 | 35                 | 24                 | 1151               | 0                   | 0                   | 0                   | 1260               | 6.7%                              | 8.7%                              |
| Lane Change Crashes          | 259                | 38                 | 56                 | 2275               | 0                   | 0                   | 0                   | 2628               | 11.3%                             | 13.4%                             |
| Heavy Vehicle Crashes        | 96                 | 42                 | 61                 | 1610               | 0                   | 0                   | 0                   | 1809               | 7.6%                              | 11.0%                             |
| Motorcycle Crashes           | 532                | 219                | 146                | 4735               | 0                   | 0                   | 0                   | 5632               | 13.3%                             | 15.9%                             |
| Bicycle Crashes              | 925                | 116                | 76                 | 3496               | 12                  | 0                   | 0                   | 4625               | 22.5%                             | 24.2%                             |
| Light Commercial Veh Crashes | 267                | 187                | 167                | 6382               | 11                  | 0                   | 0                   | 7014               | 6.5%                              | 8.9%                              |
| Passenger Vehicle Crashes    | 2503               | 1016               | 737                | 20258              | 291                 | 51                  | 0                   | 24856              | 14.2%                             | 17.1%                             |
| Bus Crashes                  | 15                 | 19                 | 4                  | 461                | 0                   | 0                   | 0                   | 499                | 6.8%                              | 7.6%                              |
| Tram Crashes                 | 13                 | 13                 | 7                  | 194                | 0                   | 0                   | 0                   | 227                | 11.5%                             | 14.5%                             |

Figure A1. Number of hot spots, cold spots and non-significant location resulting from Getis-Ord Gi\* analysis of metropolitan Melbourne 2012-2016 crashes

Example of locations (suburbs and/or roads) in metropolitan Melbourne with the highest number of hot spots (at the 95% confidence level) for various crash categories

Casualty crash hot spots:

- Melbourne CBD and its surrounds (from Coburg in the north to St Kilda in the south, and Flemington in the west to Toorak in the east);
- Dandenong, Clayton, Springvale, Footscray, Preston, Sunshine, Narre Warren, Roxburgh Park, Broadmeadows, Cheltenham, Bundoora, Chadstone, Epping, Melton, and Ringwood.

Fatal crash hot spots:

- Coburg East, Ashwood, Mentone, Upper Ferntree Gully and Pearcedale.

Pedestrian crash hot spots:

- Mainly in the inner and middle suburbs of Melbourne;
- Included areas such as the CBD, Docklands, North

Melbourne, Carlton, Fitzroy, Brunswick, Coburg, Preston, Glenroy, Heidelberg, St Kilda, Prahran, Windsor, South Yarra, Kensington, Flemington, East Melbourne, Abbotsford, Richmond, Mitcham, Nunawading, Burwood, Ashburton, Chadstone, Clayton and Seaford.

Run off road crash hot spots:

- Outer suburbs and some middle suburbs of Melbourne;
- Included areas such as Melton, Laverton, Roxburgh Park, Mitcham, Ringwood, Belgrave, Lysterfield, Ferny Creek, Gembrook and Cranbourne North.

Head on not overtaking crash hot spots:

- Mainly in the middle and outer suburbs of Melbourne;
- Often located on sections of roads rather than over an area;
- Included roads like Belgrave-Hallam Road in Belgrave, Barry Road in Coolaroo, Maroondah Highway in Lilydale, Yan Yean Road in Plenty and Sydney Road in Coburg.

**Rear end crash hot spots:**

- Mainly located around major intersections and on motorways (that is, freeways and tollways) within a radius of 20km north and west and a radius of 45km south and east from the CBD;
- Included roads such as Dandenong Road, North Road and Clayton Road in Clayton, North Road and Wellington Road in Rowville, Monash Freeway, Warrigal Road and Dandenong Road in Chadstone, Springvale Road and Burwood Highway in Burwood East, Monash Freeway, Narre Warren North Road, Cranbourne Road and Princes Freeway in Narre Warren, Dandenong Road and Burke Road in Malvern East, West Gate Freeway and Western Link Tollway in Port Melbourne, Ballarat Road between Anderson Street and Duke Street in Sunshine, and Western Ring Road, Hume Highway, Camp Road, Barry Road, Somerton Road and Pascoe Vale Road in Campbellfield-Coolaroo;
- Also included areas such as Mentone, Taylors Lakes, Epping, Thomastown, Ringwood, and Ferntree Gully.

**Side impact at intersection hot spots:**

- Included intersections in the CBD and surrounds;
- Also located in middle and outer metropolitan Melbourne areas such as Clayton, Springvale, Noble Park, Dandenong, Cranbourne, Narre Warren, Wantirna, Wantirna South, Preston, Somerton, Meadow Heights, Burnside Heights, Melton, Sunshine, Truganina (Leakes Road), Heatherton, Highett (Bay Road), Springvale South, Carrum Downs, Langwarrin and Mornington.

**Lane change and side swipe crash hot spots:**

- Mainly located in the inner metropolitan Melbourne area including the CBD, Carlton, North Melbourne, Brunswick, Collingwood, South Melbourne, South Yarra and St Kilda;
- There is also a cluster at Brighton.

**Bicycle crash hot spots:**

- Mainly located in the inner metropolitan Melbourne areas including the CBD and extending from Brunswick in the north to South Melbourne in the south, and from North Melbourne in the west to Collingwood in the East;
- There was also a cluster on St Kilda Road (from the CBD to Alma Road in St Kilda) and on Chapel Street and surrounding areas (from East Richmond in the north to Windsor in the South);
- There were three clusters on Beach Road, a well-known riding spot for cyclists. These were at Brighton Beach, Black Rock and Mordialloc.

**Heavy vehicle crash hot spots:**

- Mainly located around industrial areas such as Laverton, Sunshine West, Altona North, the Melbourne Docks area, Sunshine North, Campbellfield, Somerton, Thomastown and Keysborough.
- The clusters were also located on Western Port Highway which leads to the Port of Hastings and on the West Gate Bridge between Williamstown Road and Todd Road which connect two industrial areas and leads to the Port of Melbourne.
- Other locations included the South Melbourne industrial area and the Brunswick industrial area.

**Motorcycle crash hot spots:**

- Mainly located in the inner and middle metropolitan Melbourne suburbs;
- There were some clusters in the outer metropolitan Melbourne areas such as Gembrook Road and Gembrook-Launching Place Road in Gembrook (both windy roads), Mount Dandenong Tourist Road in Upper Ferntree Gully (windy road), Warburton-Woods Point road in Upper Yarra Valley (windy road), Diamond Creek Road in Greensborough, Gap Road and Elizabeth Drive in Sunbury, and Heaths Road in Werribee.
- Clusters in the middle metropolitan suburbs included Clayton, Gardenvale, Hawthorn, Camberwell, Kew, Footscray, Ascot Vale and Brunswick.
- The inner metropolitan hot spot clusters included the CBD and extended from Brunswick in the north to St Kilda in the south and from North Melbourne in the west to Richmond in the east.

**Passenger vehicle crash hot spots:**

- The hot spot cluster locations were almost identical to that for total casualty crashes.
- One exception was that Box Hill was identified as a hot spot crash cluster for crashes involving passenger vehicles but not for total casualty crashes.

**Light commercial vehicle crash hot spots:**

- Clusters were located in North Melbourne and Docklands and in the middle and outer suburbs of metropolitan Melbourne, including Melton, Taylors Lakes, Caroline Springs, Truganina, Campbellfield, Pascoe Vale, Thomastown, Epping, Watsonia, Preston, Greensborough, Ringwood, Bayswater, Chirnside Park, Ferntree Gully, Chadstone, Clayton, Rowville, Highett, Springvale South, Noble Park and Keysborough.
- They were mainly located on freeways and approaches to freeways, on highways and on major arterial roads including Western Freeway, Melton Highway, Hume Highway, Western Ring Road, Burwood Highway, Dandenong Road and Nepean Highway and around industrial areas and commercial areas.

Bus crash hot spots:

- Most of the bus crash hot spot clusters comprise of only one or two sites only;
- Dandenong, Wantirna South, Mordialloc, Toorak, Preston and Epping had hot spot clusters of three or more sites. Carlton;
- Niddrie and Kingsbury (La Trobe University) had hot spot clusters of two sites.

Tram crash hot spots:

- The CBD, St Kilda Road between Flinders Street and Southbank Boulevard and Carlton were locations that had a cluster of hot spots;
  - Burwood East had one site that was a tram hot spot, as did East Melbourne.
-