Geospatial Factors Applied to Road Accidents: A Review

Richard B. Watson and Peter J. Ryan *

Ryan Watson Consulting Pty Ltd., Melbourne, Australia
Email: richard@ryanwatsonconsulting.com.au (R.B.W.); peter@ryanwatsonconsulting.com.au (P.J.R.)
*Corresponding author

***Abstract*—Road accidents are a major source of trauma worldwide due to increasing numbers of vehicles and drivers. Considerable data has been collected on road accidents and this is frequently published as open data on the internet. These datasets include parameters such as accident type and number of fatalities as well as environmental variables such as road type, demographics, and area infrastructure. Geospatial analysis provides a means of understanding spatial factors that influence road accidents such as the built infrastructure, natural environment features such as hills and vegetation, traffic volume, and road design and construction. Geospatial visualization techniques can also help identify hotspots and blackspots. The Moran I, Getis-Ord, and Kernel Density Estimation techniques are the most commonly-used geospatial tests and their application is reported in many papers. This paper provides a review of geospatial factors that are relevant to road accidents.**

***Keywords*—road accidents, geospatial analysis, visualization**

I. Introduction

Road accidents are a major source of trauma worldwide caused by increasing numbers of vehicles, drivers, and expanding road infrastructure arising from general population growth and urban encroachment on to previously rural areas. Considerable data has been collected on road accidents and this is frequently published as open data on the internet. These datasets include parameters such as accident type, location, and number of fatalities as well as environmental variables such as light conditions, road type, demographics, and area infrastructure.

Geospatial analysis offers a way to understand spatial environmental factors that influence road accidents such as the built infrastructure, natural environment features such as hills and vegetation, traffic volume, and road design and construction. Geospatial visualization techniques can also help identify traffic blackspots and hotspots.

This paper provides an overview of geospatial analysis techniques applied to road accidents and how this can be used to identify and mitigate issues such as hotspots. Recent analysis from Australia will be provided as an exemplar.

Manuscript November 3, 2023; revised December 7, 2023; accepted January 17, 2024; published March 28, 2024.

The rest of the paper is organized as follows: Section II provides an overview of road accidents; Section III discusses geospatial analysis of road accidents; Section IV describes some use cases for geospatial analysis of road accidents; Section V provides geospatial analysis for a medium-sized Australian city; while Section VI contains the discussion and conclusions.

II. Road Accidents Overview

A. Road Accidents Worldwide

Road accidents are a major cause of death and injury in the modern world. Each year over 1.3 million people are killed in road accidents, and these are a leading cause of death and injury for all age groups. The US Centre for Disease Control and Prevention estimates that road trauma costs the world economy nearly $US 2 trillion annually [1]. While there are more cars per head of population in developed nations, the death rate from such crashes is far higher in less developed nations due to a variety of factors including poor road infrastructure, substandard vehicles and inadequate driver education [1].

Over the years vehicles have incorporated more safety features, such as seat belts, roads have been better designed to reduce accidents, and road safety codes, such as the prohibition of drink driving, have been strengthened. In Victoria, Australia, these measures have contributed to the gradual reduction of deaths and injuries in road accidents [2].

B. Factors Influencing Road Accidents

The key factors that influence the frequency and severity of road accidents are poor transport (road) infrastructure, driver education, natural and built environmental features such as trees and buildings leading to visibility issues, car speed and vehicle overloading (such as trucks carrying too much weight), and poor road safety management (such as lack of accurate speed limits and inadequate numbers of road safety officers) [3].

The influence of the built environment on road accidents has also been the subject of considerable recent research internationally [4–8].

C. Blackspots, Blind Spots, and Hotspots

An accident blackspot or black spot is a place where road traffic collisions have historically been concentrated.

Blind spots in this context refer to situations inside the vehicle where a driver's view is blocked due to mirrors being incorrectly positioned or there is an obstacle blocking the view such as an incorrectly-placed load.

A hotspot differs from a blackspot in that it is defined using statistical analysis rather than historical accident data. GIS techniques can be applied to find hotspots, which can be further categorized as emerging, sporadic, persistent, consecutive, intensifying or diminishing depending on their characteristics [9]. Identification of such hotspots is vital for road planners to reduce the number of potential road accidents [3, 10–12].

III. Geospatial Analysis of Road Accidents

A. *Requirements for Geospatial Analysis*

Geospatial analysis requires data sources, visualization, and appropriate analysis techniques. These are described in the following subsections.

B. *Data Sources*

There is a vast quantity of data on road accidents worldwide. Rabbani *et al*. [13] reviewed and compared accident data collection systems in developing and developed countries. They noted that developing countries such as Pakistan frequently use manual data recording systems while the developed nations such as Germany and France use automated systems with the latest technology. Further, there is often a lack of official bodies that manage road accident statistics in developing nations. This can lead to erroneous conclusions as to road accident causes.

Data sources are also expanding with technological advances such as onboard electronic equipment built into modern vehicles. Gutierrez-Osario and Pedraza [14] identified five main sources of data as: (1) government, (2) open, (3) onboard from devices installed in vehicles, (4) measurement systems such as radars, cameras, and sensors embedded in roads, and (5) social media. Chand *et al*. [15] discussed how Intelligent Transportation Systems are becoming the greatest source of accident data. They also stated that social media data is untrustworthy since it could be biased and is not dependable.

Road traffic involves millions of vehicles transiting urban and rural areas daily. This leads to an enormous amount of road traffic and accident data. These large accident-related datasets can be considered as big data that require specialized techniques to visualize and analyse.

Besides road traffic and accident data, other datasets are needed for correlation analysis. These can include demographic and census data, built environment data, natural environment data, and weather/climate data as appropriate for the study. Open Street Map (OSM), for example, provides a map of the world that is effectively licence-free [16]. Sophisticated studies may require fusion of several such datasets.

For analysis of Australian road accidents, datasets are available from local, state and federal levels of government. Road accident data for the State of Victoria is published by VicRoads, now within the Department of Transport and Planning (DTP) (https://www.vic.gov.au/department-transport-and-planning). The State Department of Energy, Environment and Climate Action (DEECA) also manages spatial and environmental data for Victoria, and abundant geospatial and demographic environmental data is available from open-source platforms such as AURIN (the Australian Urban Research Infrastructure Network (https://aurin.org.au/portal-retired/), Datashare Victoria (https://datashare.maps.vic.gov.au/), and the City of Melbourne's open data portal (https://data.melbourne.vic.gov.au/pages/home/).

At the federal level, built environment data is available from the Australian Open Data Portal (https://data.gov.au) while Geoscience Australia maintains an extensive portal of environmental data (https://portal.ga.gov.au/). The Australian Bureau of Statistics (https://www.abs.gov.au/) provides census data including Local Government Areas defined by the State and Territory Governments.

Further, there is OSM data available for Victoria. OSM is a free, open geographic database updated and maintained by a community of volunteers via open collaboration (https://www.openstreetmap.org). Google Earth Engine also provides an open data catalogue of satellite imagery and geospatial datasets (https://earthengine.google.com/).

C. *Geospatial Analysis Techniques*

Geospatial analysis refers to the process of collecting, combining and visualizing types of geospatial data. The most commonly used techniques are: Kernel Density Estimation (KDE), Moran I (Global), Moran I_i (Local), and Getis-Ord. These are described in the next paragraphs.

The KDE tool calculates the density of features in a neighborhood around those features. Fig. 1 shows road fatalities in Victoria during the period 2016–2021 from which KDE can be made [17]. The Local Government Areas (LGAs) are also included. As expected, there is a high density of road accidents clustered around the major cities and towns such as the capital city Melbourne surrounding the bay area in Fig. 1.

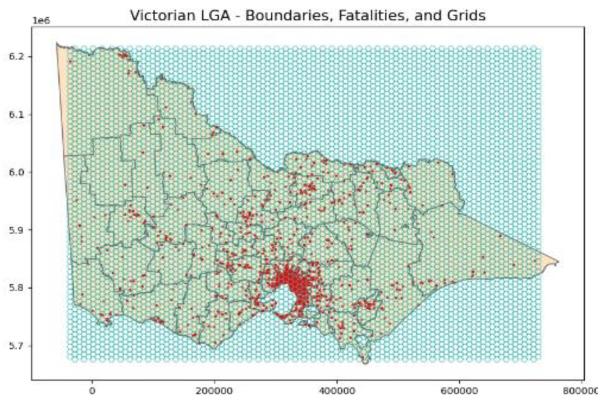

Fig. 1. Map showing Victorian LGAs, 5000 m grid and road fatalities during 2016–2021.

Moran I provides a measure of global spatial autocorrelation [18]. Local Moran I_i values can also be determined for each spatial unit to obtain finer-grained autocorrelation results. The Getis-Ord G test can be used for identification of statistical anomalies [18]. The Space Time Cube approach allows visualization and analysis of geospatial data using time series analysis and spatial and temporal analysis. ESRI's ArcGIS package provides this capability as part of its repertoire [19].

D. Machine Learning

Machine Learning (ML) provides a means of identifying mapping relationships among spatial datasets [20]. Previously the authors have used predictive ML to visualize and analyse Victorian road accident event data [21]. ML can also be used for hotspot analysis. ML is the most recent addition to the analysis toolset and can be used to predict trends and determine the significant features that lead to road accidents [22]. Spatial data has unique properties such as spatial dependence, spatial heterogeneity and scale which need to be explicitly handled by the ML method used [23]. Many ML models can be used to predict road accident hotspots, including Support Vector Machine, Random Forest and Artificial Neural Networks [10].

E. Geospatial Visualization

A key feature of road accident investigation is the requirement to visualize complex datasets so that features can be interpreted and trends detected. While visualization is often considered as part of analysis, it is the separate process of presenting information and data in a pictorial or graphical format. Numerous different techniques have been applied for accident statistics. These include standard bar and column charts as well as pie charts and other graph types. Rabbani *et al*. [24] showed how cartographic presentation enhances the understanding of accidents for the lay person by showing them on a map projection where road accidents occur. Hotspots can easily be seen on such maps.

A further visualization technique is the infographic that can combine standard charts with text and imagery to inform both the layperson and also decision makers [25]. Such infographics have been shown to be more effective at conveying information about road accidents than other techniques.

F. Python Library Approach

Key analysis systems for geospatial analysis include those developed by ESRI (ArcGIS) [19] and CARTO [26]. An alternative and cheaper approach is to use the vast array of Python libraries that can perform geospatial data analysis. These include data analysis packages such as *pandas* and *geopandas*, geospatial analysis packages including *shapely*, *geocube*, and *pysal* and visualization tools such as *matplotlib*, *seaborn*, *plotly*, *contextily*, and *splot*.

Specialized packages for accessing and graphing Open Street Map data include *pyrosm* and *OSMNX*. Python ML packages include *scikit-learn*, also known as *sklearn*. These libraries are generally open source and are regularly updated and enhanced by their user communities to provide the latest features [24].

G. Limitations and Challenges of Current Approaches

Data availability is a key challenge for geospatial analysis. In the developed countries it has been noted that there is generally data available for analysis collected by government organizations that manage road accidents, whereas developing nations may not have the same quantity, quality, and reliability of comparable data [13]. Further, in more recent work, there is a requirement to include street view imagery in accident analysis [8]. Such data is not available for many areas of interest. Finally, new data types may be required for future analysis such as Point Cloud Data (PCD) that is now starting to become available.

Researchers are also developing approaches to improve traffic management and thus increase road safety. Monica *et al*. [27] presented a study on the role of AI, Internet of Things and GPS in traffic management, and proposed a traffic signal management system that reduces congestion, prioritizes emergency vehicles, reduces commuting time, and tracks vehicles. Li *et al.* [28] described a novel collision warning system that uses the visual road environment. Earlier, Nadarajan *et al*. [29] applied ML to estimate the probabilistic space-time representation of traffic scenarios.

Other challenges for geospatial analysis include the integration of various technologies, complexity in algorithms, scalability concerns, security, and privacy. It is also anticipated that there will be refinements in analysis techniques, improvements in existing ML algorithms and development of new algorithms, for example for hotspot detection, and application of enhanced visualization for accident zones.

IV. CASE STUDIES

Numerous case studies of road accident investigations have been published. A comprehensive set of such use cases is provided by the International Road Assessment Programme (IRAP) [30]. Case studies tend to be restricted to small urban areas where there is high traffic flow and road accidents are most common. Most case studies are from the US and Europe where there is detailed data on road accident statistics and also land use, demographics, and the built and natural environment. Selected case studies are presented from Europe, China, the US, the Middle East, and India providing a contrast between highly developed and developing countries.

Mesquitela *et al*. [31] presented a case study for Lisbon in Portugal using the ArcGIS Pro system to apply Kernel Density and hotspot analysis to identify traffic accident blackspots and determine the factors that influence accident frequency and severity. The authors required six datasets for analysis—wind information, meteorology, rain observations, accident statistics, emergency occurrences, and historical traffic statistics. Global Moran I was used for spatial autocorrelation to determine if accidents are clustered, scattered, or random. Kernel Density Estimation was used for spatial pattern analysis and Getis-Ord was applied to identify hotspots. The analysis identified 12 streets where traffic accidents were high and the locations of 14 hotspots.

Asadi *et al*. [4] carried out a comprehensive analysis of the relationships between the built environment and traffic safety in selected Netherlands urban areas. The authors investigated the correlation of these factors with vehicle-bicycle and vehicle-vehicle Property Damage Only (PDO) and Killed and Severe Injury (KSI) crashes in urban areas. A Spatial Hurdle Negative Binomial regression model was employed for the Netherlands-Randstad Area where major land-use developments have occurred since the 1970s. The study was conducted by developing a rich dataset composed of various national and local databases. The analysis showed that environmental factors and land-use policies have substantial impacts on safety. Furthermore, low socioeconomic levels are associated with a higher frequency of traffic crashes.

Increasing population and urbanization have led to rising accident rates in China, A recent paper investigated the effects of Points of Interest (POI) on traffic accidents in a busy industrial park [5]. These authors applied a variety of geospatial analysis techniques and found that local models such as Geographically Weighted Regression performed better than global models such as Negative Binomial Regression. Land use was determined from POI data for the selected area. The local regression results showed that population density and road length are positively correlated with the frequency of traffic accidents. Land use for commercial purposes, green land for leisure, and land for transport all negatively impact on road safety. Such analysis can help city managers to build roads, plan open and commercial spaces and develop public transport in ways to reduce road accident frequency.

Huang *et al.* [6] examined the relationships between the built environment and crashes using a geographically weighted regression approach for the Detroit (Michigan, US) urban area. Eight datasets were used in the study including those for census data and land use. The Geographically Weighted Regression (GWR) approach was shown to provide greater accuracy than Ordinary Least Squares (OLS). The findings showed that most built environment parameters have some correlation with crashes. Commercial land use and four-way intersections were found to generally indicate a higher crash rate.

Mohammed *et al.* [9] carried out a similar study for the small, oil-rich Middle East state of Qatar that has an area less than 12, 000 km^2 and a population less than 3 million. Accordingly, road traffic is concentrated in the capital city Doha metro area that has 40% of the Qatari population. Space Time Cube analysis, GWR, Moran's I and Getis-Ord G were employed to determine the geospatial characteristics of road accidents. Hotspots of type sporadic, persistent, consecutive, intensifying and diminishing were identified. The findings of this study can inform the development of more effective methods for preventing crashes and improving safety on Qatar's roads.

Singh and Katiyar discussed how GIS analysis can inform Indian government bodies to help with planning a safer urban road network, for example by reducing the number of potential accident blackspots [3]. An interesting statistic noted here is that the US has 5 times as many road accidents as India but less than a quarter of the fatalities indicating that traffic accidents are a major issue for India and that its road system needs to be considerably improved to meet global safety standards. India is a rapidly developing nation whose transport infrastructure and driver training has failed to keep up with its industrial and domestic infrastructure.

V. Application to Geelong, Australia

A. *Choice of Geelong as a Testbed for Analysis*

The city of Geelong in the state of Victoria, Australia was chosen as the testbed for this analysis as it is a medium-sized city with many road accidents occurring in the 2006–2020 recording period. It has a well-defined Central Business District (CBD) where road accidents are most numerous, and is thus similar to other cities which have been studied for geospatial road accident analysis such as Barcelona, Spain [11], Harbin, China [32], and Tabriz City, Iran [33].

The built environment of the Geelong CBD is catalogued in the OSM database, and there are other sources of environmental data that we are currently investigating. Population data of both residents and commuting workers at the 500 m grid level is also available from Australian census records (https://www.abs.gov.au/census/find-census-data), although this is not included in the analysis described here.

The city of Geelong is also the administrative centre of the City of Greater Geelong Local Government Area (LGA), which covers an area of 1,248 km^2 and has a population of 271,000 as of the 2021 Australian census. The analysis described here covers both the City of Greater Geelong LGA and the Geelong CBD. A 500 m hexagonal grid was used to analyse the data using an algorithm described in Ref. [34]. This is comparable to the grids used in other studies of this kind ([11]—150 m, [32]—230 m, and [6]—1000 m). The hexagonal grid covering the Geelong CBD is shown in Fig. 2.

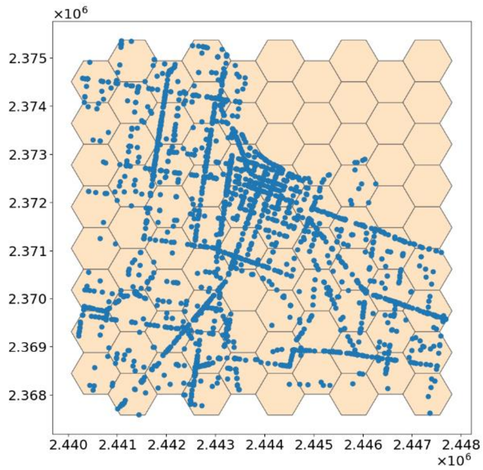

Fig. 2. Map of the Geelong CBD showing 500 m grid and locations of road accidents of all kinds.

For comparison, Fig. 3 shows the choropleth map of road accidents.

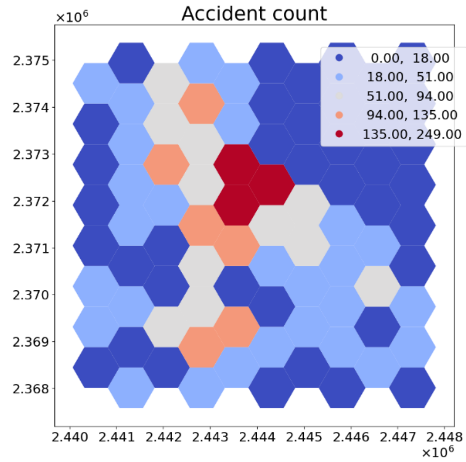

Fig. 3. Choropleth map of the Geelong CBD 500 m grid road accident counts.

B. *Moran I (Global) Analysis*

The most commonly used statistic for spatial autocorrelation is Moran's I developed by Patrick Moran [35].

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} \tag{1}$$

where N is the number of observations, z_i is the standardized value of the variable of interest at location i, and w_{ij} is the cell corresponding to the i-th row and j-th column of a W spatial weights matrix.

The Moran I plot for the Geelong CBD is shown as Fig. 4(b). This demonstrates that accidents have strong spatial autocorrelation since high values of accident counts cluster near high values and low values near other low values resulting in a roughly linear plot [36]. Here the Moran index is 0.71 and the p-value 0.001. The reference distribution shown in Fig. 4(a) was generated by simulating 999 random maps with the values of the count variable and then calculating Moran's I for each of these maps. It is clear that the observed spatial pattern is not random.

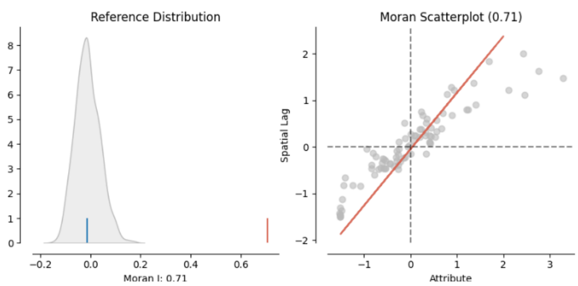

Fig. 4. (a) Moran I reference distribution and (b) Scatterplot for the Geelong CBD accident counts in a 500 m hexagon grid.

C. *Moran Ii (Local) Analysis*

Moran local analysis was applied to the Geelong LGA by adapting the python code given in [36], which uses the *splot* library. A "cluster map" of results, which shows clusters where high-high and low-low values are located. is shown in Fig. 5.

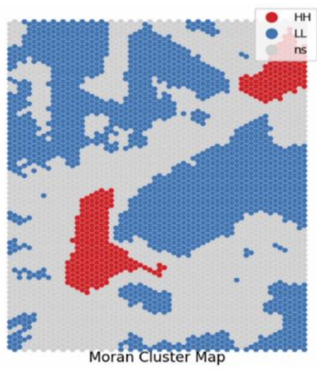

Fig. 5. Moran I_i Cluster Map for the Greater Geelong LGA accident counts in 500 m grid.

D. *Regression Analysis Using Ordinary Least Squares*

Regression analysis using Ordinary Least Squares (OLS) was also applied to the Geelong CBD for accident count related to the number of adjacent buildings and accident count related to the distance from the centre of Geelong. The Python code given in [36] was adapted for this application. No clear trends are visible here although it appears there are more accidents in built-up areas and there is a higher accident probability close to the Geelong Post Office (GPO). These results are shown as Fig. 6.

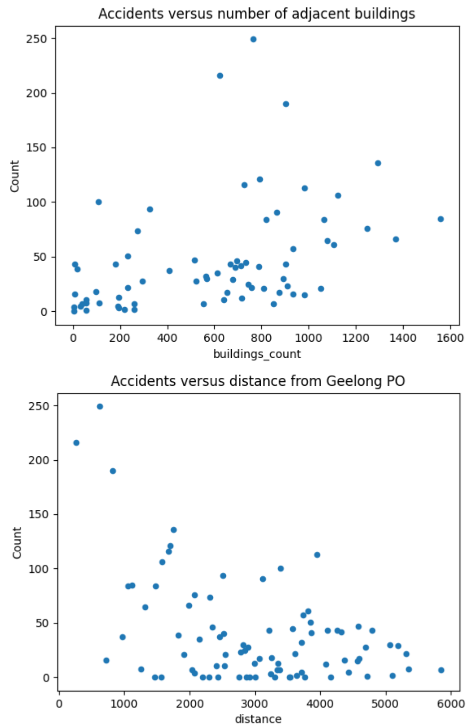

Fig. 6. Results of OLS Regression Analysis of accident counts in Geelong CBD 500 m grid, for two independent variables. (a) (top): Accident count versus number of adjacent buildings in a 500 m buffer zone from grid centroids and (b) (lower) Accident count versus distances of grid centroids from Geelong Post Office (GPO).

We are planning to consider other spatial and demographic features and apply predictive ML to identify accident hotspots.

E. Multiscale Geographically Weighted Regression Analysis

Geographically Weighted Regression (GWR) is a spatial statistical technique that, like spatial local regression, recognizes that traditional 'global' regression models may be limited when processes vary by context. GWR captures a process's spatial heterogeneity (i.e., process variation by spatial context) via an operationalization of Tobler's first law of geography: "everything is related to everything else, but near things are more related than distant things" [37]. We note that much recent work has been done applying GWR to road accident modelling [5, 38].

A GWR model may be specified as:

$$y_i = \beta_{i0} + \sum_{k=1}^{r} \beta_{ik}\, x_{ik} + \epsilon_i, \qquad i = 1, \ldots, n \qquad (2)$$

where y_i is the dependent variable at location i, β_{i0} is the intercept coefficient at location i, x_{ik} is the k-th explanatory variable at location i, β_{ik} is the k-th local regression coefficient for the k-th explanatory variable at location i, and ε_i is the random error term associated with location i.

We have adapted the Python code of Mendez [39] to analyse the road accident counts in all Victorian LGAs. Fig. 7 gives some indicative results of this analysis. It is planned to investigate Victorian urban areas like Geelong to determine if spatial heterogeneity is significant in the context of road accident geospatial analysis.

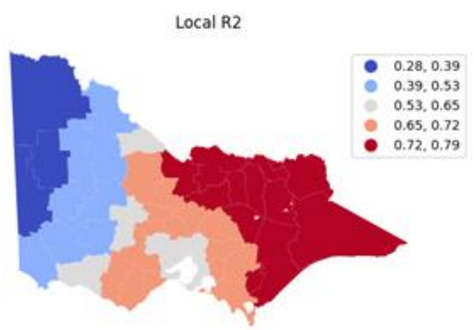

Fig. 7. Geographically Weighted Regression analysis of accidents by LGA for Victoria, Australia.

VI. Conclusions

Geospatial data analysis provides a means of understanding road accident severity and frequency. This complements traditional statistical analysis techniques and is an emerging tool for managing the global problem of road accidents. It can establish the significance of the influence of spatial factors such as the built infrastructure, natural environment features, and road construction on road accidents. Techniques such as Moran I, Getis-Ord and Kernel Density Estimation can be applied to identify where road traffic is likely to lead to accidents and where hotspots are located. Machine learning is the most recent addition to the analysis toolset and can be used to predict trends and determine the significant features that lead to road accidents. Selected results for analysis of road accidents in the City of Greater Geelong, Australia were provided to illustrate these techniques.

Further challenges for the researcher include data availability, refinements in analysis techniques, improvements in existing ML algorithms and development of new algorithms, and application of enhanced visualization for accident zones.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

RBW initiated the road accident research project; PJR wrote the initial version of the paper; RBW and PJR jointly developed the code and analyzed the data. Both authors approved the final version.

Acknowledgment

The project was carried out with the assistance of student interns from Swinburne University of Technology (Victoria, Australia) who did some of the coding and analysis.

References

[1] Center for Disease Control and Prevention (US). Global Road Safety. (2023). Available: https://www.cdc.gov/injury/features/global-road-safety/index.html#:~:text=1%20Each%20year%2C%201.35%20million%20people%20are%20killed,people%205%E2%80%9329%20years%20of%20age.%20...%20More%20items

[2] Australian Government, Road Trauma Australia, *2021 Statistical Summary, Bureau of Infrastructure and Transport Research Economics*, 2021.

[3] N. Singh and S. K. Katiyar, "Application of Geographical Information System (GIS) in reducing accident blackspots and in planning of a safer urban road network: A review," *Ecological Informatics*, vol. 66, 101436, 2021. https://doi.org/10.1016/j.ecoinf.2021.101436

[4] M. Asadi, M. B. Ulak, K. T. Geurs, W. Weijermars, and P. Schepers, "A comprehensive analysis of the relationships between the built environment and traffic safety in the Dutch urban areas," *Accident Analysis & Prevention*, vol. 172, 106683, 2022. https://doi.org/10.1016/j.aap.2022.106683

[5] H. Chung, Q. Duan, Z. Chen, and Y. Yang, "Investigating the effects of POI-based land use on traffic accidents in Suzhou Industrial Park, China," *Case Studies on Transport Policy*, vol. 12, 100933, 2023. https://doi.org/10.1016/j.cstp.2022.100933

[6] Y. Huang, X. Wang, and D. Patton, "Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach," *Journal of Transport Geography*, vol. 69, pp. 221–233, 2018. https://doi.org/10.1016/j.jtrangeo.2018.04.027

[7] M. Obelheiro, A. Silva, C. Nodari, H. Cybis, and L. Lindau, "A new zone system to analyze the spatial relationships between the built environment and traffic safety," *Journal of Transport Geography*, vol. 84, 102699, 2020. doi: 10.1016/j.jtrangeo.2020.102699

[8] S. Hu, H. Xing, W. Luo, L. Wu, Y. Xu, W. Huang, W. Liu, and T. Li, "Uncovering the association between traffic crashes and street-level built-environment features using street view images," *International Journal of Geographical Information Science*, p. 1–25, 2023. doi: 10.1080/13658816.2023.2254362

[9] S. Mohammed, A. H. Alkhereibi, A. Abulibdeh, R. N. Jawarneh, and P. Balakrishnan, "GIS-based spatiotemporal analysis for road traffic crashes; In support of sustainable transportation planning," *Transportation Research Interdisciplinary Perspectives*, vol. 20, 100836, 2023. https://doi.org/10.1016/j.trip.2023.100836

[10] B. D. S. P. Amorim, A. A. Firmino, C. D. S. Baptista, G. B. Júnior, A. C. D. Paiva, and F. E. D. A. Júnior, "A machine learning

approach for classifying road accident hotspots," *ISPRS International Journal of Geo-Information*, vol. 12, no. 6, 227, 2023.
[11] M. Alvarez. (2020). Predicting traffic accident hotspots with spatial data science. [Online]. Available: https://carto.com/blog/predicting-traffic-accident-hotspots-with-spatial-data-science
[12] K. Hazaymeh, A. Almagbile, and A. H. Alomari, "Spatiotemporal analysis of traffic accidents hotspots based on geospatial techniques," *ISPRS International Journal of Geo-Information*, vol. 11, no. 4, 260, 2022. https://doi.org/10.3390/ijgi11040260
[13] M. Rabbani, M. A. Musarat, W. Alaloul, S. Ayub, H. Bukhari, and M. Altaf, "Road accident data collection systems in developing and developed countries: A review," *International Journal of Integrated Engineering*, vol. 14, pp. 336–352, 2022. doi: 10.30880/ijie.2022.14.01.031
[14] C. Gutierrez-Osorio and C. Pedraza, "Modern data sources and techniques for analysis and forecast of road accidents: A review," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 7, no. 4, pp. 432–446, 2020. https://doi.org/10.1016/j.jtte.2020.05.002.
[15] A. Chand, S. Jayesh, and A. B. Bhasi, "Road traffic accidents: An overview of data sources, analysis techniques and contributing factors," *Materials Today*, vol. 47, pp. 5135–5141, 2021. https://doi.org/10.1016/j.matpr.2021.05.415
[16] Open Street Map. (2023). [Online]. Available: https://www.openstreetmap.org
[17] S. M. A. Kazmi, M. Ahmed, R. Mumtaz, and Z. Anwar, "Spatiotemporal clustering and analysis of road accident hotspots by exploiting GIS technology and kernel density estimation," *The Computer Journal*, vol. 65, no. 2, pp. 155–176, 2020. doi: https://doi.org/10.1093/comjnl/bxz158
[18] R. Satria and M. Castro, "GIS tools for analyzing accidents and road design: A review," *Transportation Research Procedia*, vol. 18, pp. 242–247, 2016. https://doi.org/10.1016/j.trpro.2016.12.033
[19] ESRI. ESRI—ArcGIS for Desktop. 2023. [Online]. Available: https://www.esri.com/en-us/arcgis/about-arcgis/overview
[20] W. Yang, M. Deng, J. Tang, and L. Luo, "Geographically weighted regression with the integration of machine learning for spatial prediction," *Journal of Geographical Systems*, vol. 25, no. 2, pp. 213–236, 2023. doi: 10.1007/s10109-022-00387-5
[21] R. Watson and P. Ryan, "Big data analytics for Australian local government," *Smart Cities*, vol. 3, no. 3, pp. 657–675, 2020. https://doi.org/10.3390/smartcities3030034
[22] M. Megnidio-Tchoukouegno and J. A. Adedeji, "Machine learning for road traffic accident improvement and environmental resource management in the transportation sector," *Sustainability*, vol. 15, no. 3, 2014, 2023.
[23] B. Nikparvar and J.-C. Thill, "Machine learning of spatial data," *ISPRS International Journal of Geo-Information*, vol. 10, no. 9, 600, 2021.
[24] M. Rabbani, M. A. Musarat, W. Alaloul, A. Maqsoom, H. Bukhari, and W. Rafiq, "Road Traffic accident data analysis and its visualization," *Civil Engineering and Architecture*, vol. 9, pp. 1603–1614, 2021. doi: 10.13189/cea.2021.090530
[25] J. Steinhardt, "The role of numeric and statistical content on risk perception in infographics about road safety," *Journal of Risk Research*, vol. 23, no. 5, pp. 613–625, 2020. doi: 10.1080/13669877.2019.1596147
[26] CARTO. (2023). Spatial analytics for the modern data stack. [Online]. Available: https://carto.com/
[27] C. Monica, B. Jyothi, A. Ramagiri, S. Gottipati, V. Jahnavi, S. A. Akther, and R. Chinnaiyan, "Intelligent traffic monitoring, prioritizing and controlling model based on GPS," in *Proc. 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, 2023.
[28] Z. Li, B. Yu, Y. Wang, Y. Chen, Y. Kong, and Y. Xu, "A novel collision warning system based on the visual road environment schema: An examination from vehicle and driver characteristics," *Accident Analysis & Prevention*, vol. 190, 107154, 2023. https://doi.org/10.1016/j.aap.2023.107154
[29] P. Nadarajan, M. Botsch, and S. Sardiña, "Machine learning architectures for the estimation of predicted occupancy grids in road traffic," *Journal of Advances in Information Technology*, vol. 9, pp. 1–9, 2018.
[30] International Road Assessment Programme. Case Studies—Road Safety Toolkit—iRAP. [Online]. Available: https://toolkit.irap.org/case-studies/
[31] J. Mesquitela, L. B. Elvas, J. C. Ferreira, and L. Nunes, "Data analytics process over road accidents data—A case study of Lisbon city," *ISPRS International Journal of Geo-Information*, vol. 11, no. 2, 143, 2022. https://doi.org/10.3390/ijgi11020143
[32] M. Wang, J. Yi, X. Chen, W. Zhang, and T. Qiang, "Spatial and temporal distribution analysis of traffic accidents using GIS-based data in Harbin," *Journal of Advanced Transportation*, 9207500, 2021. doi: 10.1155/2021/9207500
[33] B. Feizizadeh, D. Omarzadeh, A. Sharifi, A. Rahmani, T. Lakes, and T. Blaschke, "A GIS-based spatiotemporal modelling of urban traffic accidents in Tabriz City during the COVID-19 pandemic," *Sustainability*, 2022.
[34] M. Mann, S. Chao, J. Graesser, and N. Feldman. (2022). Python open source spatial programming & remote sensing. [Online]. Available: https://pygis.io/docs/a_intro.html#
[35] C. Heyde, "Patrick Alfred Pierce Moran 1917–1988," *Hist. Rec. Aust. Sci*., vol. 9, pp. 17–30, 1992.
[36] S. J. Rey, D. Arribas-Bel, and L. J. Wolf. *Geographic Data Science with Python*. [Online]. Available: https://geographicdata.science/book/intro.html#geographic-data-science-with-python
[37] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Economic Geography*, vol. 46, sup. 1, pp. 234–240, 1970. doi: 10.2307/143141
[38] X. Qu, X. Zhu, X. Xiao, H. Wu, B. Guo, and D. Li, "Exploring the influences of point-of-interest on traffic crashes during weekdays and weekends via multi-scale geographically weighted regression," *International Journal of Geo-Information*, vol. 10, 791, 2021. doi: 10.3390/ijgi10110791
[39] C. Mendez. Introduction to GWR and MGWR. (2020). [Online]. Available: https://deepnote.com/@carlos-mendez/PYTHON-GWR-and-MGWR-71dd8ba9-a3ea-4d28-9b20-41cc8a282b7a