Practice Research Paper on 1st/2nd Order & Network Constrained Spatial Point Pattern Analysis on Melbourne City, Australia

Harith Oh Yee Choon
Singapore Management University
harith.oh.2020@scis.smu.edu.sg

Teo Ren Jie
Singapore Management University
renjie.teo2021@scis.smu.edu.sg

Kwek Ming Rong
Singapore Management University
mrkwek.2020@scis.smu.edu.sg

ABSTRACT

Traveling is many people's favorite form of temporal escapism from life's demands, dramas and deadlines and it enables us to clear our mind. Upon research, Melbourne city is ranked the 3rd top trending traveling destination in 2023 (Glynn, 2023).

Hence, Spatial Bros is created to assist non technologically savvy users in performing geographical point pattern analysis. This application aims to assist users in 2 types of analysis, particularly in performing 1st/2nd Order & Network Constrained Spatial Point Pattern Analysis. For each of the analyses, the application will provide users with statistical functions, kernel density heat map estimation, various mappings and G&K function results. The application will cover an array of spatial points located in Melbourne City such as childcare centres, business establishments such as banking, clothing, famous landmarks including places of interest such as schools, theaters, health service, sports facilities, drinking fountains and public toilets. The spatial points will work in conjunction to cover areas of the city's road, pedestrian and tram network. From this application, users would be able to perform types of hypothesis testing that allow them to generate insights towards their conclusion on the distribution along the spatial points along the network.

1. MOTIVATION OF THE APPLICATION

In today's technological advancing world, there are many useful and interesting spatial data sources that exist in the forms of Geospatial and Aspatial format. Geographical Geospatial data sets the foundation based on the geographical boundary locations and the Aspatial data are the records of observation that can be further prepared to be used to derive meaningful insights.

Despite all the data sources out on the internet, there are not many people who are knowledgeable and trained to perform such analysis. Without the fundamental knowledge and training involved, users are not able to interpret the results..

Our group aim is to mainly focus on performing analysis and develop a website based geographical spatial tool. R Shiny tool will be used with regards to developing the 1st/2nd Order & Network Constrained Point Pattern Analysis of Melbourne City, Australia. We hope this application can achieve the below objectives of educating and empowering users to:

- Have a better understanding of 1st/2nd Order Kernel Density Estimation (KDE)
- Have a better understanding of the capabilities of Network Constrained Point Pattern Analysis
- Conduct Network Constrained Point Pattern Analysis Density Estimation, 1st Order (KDE & Clark Evans) and 2nd Order (G & K) function analysis
- Have the ability to generate their insights and conclusions based on data and user guide provided

2. REVIEW AND CRITIC ON PAST WORKS

2.1 Case Study 1: Identifying factors of influence in the spatial distribution of domestic fires

2.1.1 Objective

To gain an understanding of how domestic fires arise based on different factors such as time of day, building types and income groups.

2.1.2 Methodology Used

- Statistical point pattern analysis
- Nearest Neighbour Analysis
- Non-stationary Strauss process
- G function

2.1.3 Learning Point

The use of the Strauss process for inferring different variables such as building types, income groups is an interesting way to generate random patterns of points where it stimulates the realization of the Strauss process. Developed by Strauss, Kelly and Ripley, it is a model for spatial inhibition that ranges from a strong 'hard core' inhibition to a completely random pattern according to gamma and beta value. The results also show that there are significant differences in factors that influence the spatial distribution of fires with the above variables listed.

2.2 Case Study 2: Network-constrained and category-based point pattern analysis for Suguo retail stores in Nanjing, China

2.2.1 Objective

To gain an understanding of the retail service hot-spot areas and the spatial clustering patterns of local retail giant, Suguo. There is also the study towards how the road network and stores are categorized to investigate the weighting influence for different categories of point events.

2.2.2 Methodology Used

- Point pattern analysis
- Network Kernel Density Estimation
- Network K Function

2.2.3 Learning Point

The bandwidth for generating the Network Constrained Kernel Density maps would produce representations of distributions for Suguo retail stores in Nanjing, China. Learning from this, we can explore various network KDE bandwidths to find suitable bandwidth for Melbourne City's road network and its respective spatial points. We could also try to explore possible options for users to adjust the network bandwidth.

3. DESIGN FRAMEWORK

3.1 User Interface Design

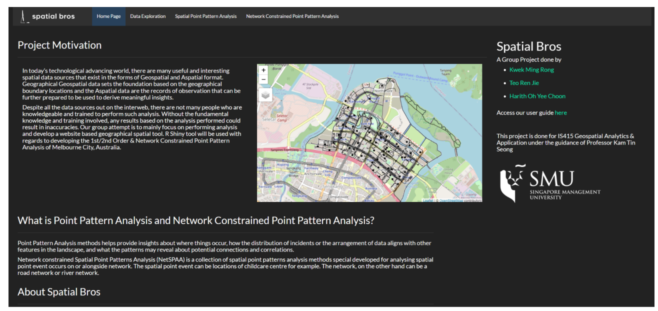

Fig 1: Home page of the application

The design of the application focuses on interactivity. The features include filtering, zooming in and out, displaying details and hovering over. We have adopted a dark color color for the background and light toned colors for the maps to enhance the perceived contrast between the data and the background, making it easier to distinguish different features and patterns. This is important for the maps where the user needs to be able to distinguish between different types of data points or regions. Moreover, it reduces eye strain and improves readability in low-light environments, as it reduces the amount of bright light emitted by the screen.

The user can access several tabs, including Home Page, Data Exploration, Spatial Point Pattern Analysis, and Network Constrained Point Pattern Analysis. The Home Page provides a concise overview of our project, while the Data Exploration tab showcases mapping of spatial and network points for exploratory data analysis. In the Spatial Point Pattern Analysis tab, users can find the outcomes of the 1st and 2nd order spatial point pattern models we conducted. Lastly, the Network Constrained Point Pattern Analysis tab offers results from the Network Kernel Density Estimation and Statistical Functions.

3.2 **Data Cleaning & Filtering**

For our analytical mapping, we intend to study the spatial points of the data that lies within the blue lines as shown below which is defined as the borders of the city of Melbourne, Australia. However, it was found that some parts of the localities fall outside of the city of Melbourne which could cause inaccuracies to our analysis.

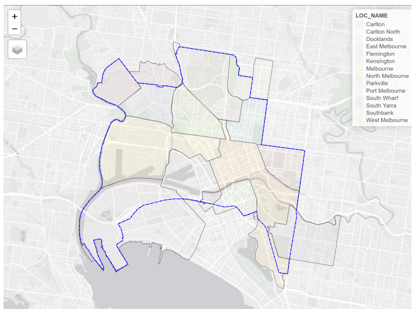

Fig 2: A polygon map showing the city limits of Melbourne and its localities

One way to overcome this issue is to constraint each locality to the city limits of the City of Melbourne, Through R programming, we use **st_intersection** with the spatial points that are within the City of Melbourne.

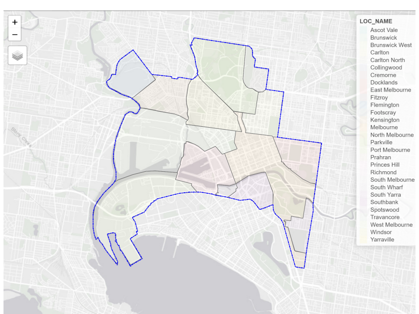

Fig 3: A polygon map showing the city limits of Melbourne and its localities with constraints

Moreover, for each spatial point on the map, we intend to keep the unique geometry points and those that fall under the latest year for each category respectively.

3.3 **ANALYTICAL MAPPING**

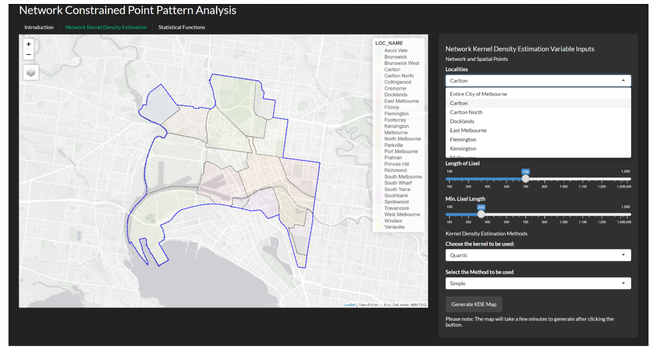

Fig 4: A polygon map with filters

To standardize the user interface design for our application, the map will be placed on the left while all the related functions are placed on the right side of the page. These are some of the main features for the functions implemented in the application for users to filter and interact:

Filtering by localities enables users to focus on the districts that they are looking for in the city of Melbourne. To choose the districts, there is a drop-down list for the user to pick the district. For example, a user may be interested in exploring only the childcare centers located in Carlton. By filtering the map view to show only Carlton, the user can quickly and easily explore the childcare centers in that area.

Kernel estimation methods are statistical methods used to estimate the probability density function of a random variable. By allowing users to filter by kernel estimation methods, they can explore the spatial patterns and densities of different types of locations on the map.

For each analysis, you can choose from three network options: Road Network, Pedestrian Network, and Tram Network under the drop-down list. Users can also select their areas of interest, such as Childcare Centres, Business Establishments, Drinking Fountains, Landmarks, and Public Toilets with the drop-drop list. .

3.4 **1st/2nd Order Spatial Point Pattern Analysis**

Another main function of the application is the kernel density estimation under the Spatial Point Pattern Analysis. The main features that are included in this tab are the kernel density estimation methods. Under the methods, it includes drop-down lists for bandwidth, kernel, method and the confidence level of Clarks Evans Test.

The bandwidth drop-down list includes fixed and adaptive bandwidth. Fixed bandwidth is often used when the data is evenly distributed and the user has prior knowledge about the appropriate bandwidth value. The fixed bandwidth method is very sensitive to highly skewed distribution of spatial point patterns over geographical units (eg. urban vs rural). For the fixed bandwidth, the KDE map will adopt a green-yellow color scheme where green being the more dense while yellow indicates less dense which is generated by the User Interface.

On the other hand, adaptive bandwidth is useful when the data is unevenly distributed, and the user does not have prior knowledge of the appropriate bandwidth value. In this case, an adaptive bandwidth can adjust to the local density of data points, resulting in a more accurate and robust estimation. For the adaptive bandwidth, the KDE map will adopt a red-yellow color scheme where red being the more dense while yellow indicates less dense as generated by the User Interface.

3.5 Network Constrained Point Pattern Analysis

One of the main functions of our application is the Network Constrained Point Pattern Analysis Density Estimation which is under its own tab. This is to allow users to study the spatial distribution of events that occur on or near a network, such as road networks or pedestrian networks. By understanding the spatial patterns of events occurring on or near a network, researchers and decision-makers can make informed decisions about resource allocation, service provision, and infrastructure planning. There are two types of analyses available: Network Kernel Density Estimation and G & K Function Analysis.

- **Network Kernel Density Estimation**

Under the **Network Kernel Density Estimation** tab, these are the variable inputs that are included.

Lixel, short for Local Index of Spatial Autocorrelation, is the distinct feature for this tab which is a method for identifying spatial clusters in data. The lixel filter includes the length of lixel and the minimum lixel length: This includes blue sliders for the users to adjust the lixel parameters.

- **Length of Lixel**: refers to how much length should the intensity of the lixel be reflected on the map from 100 to 1,500.
- **Min. Lixel Length**: to help predetermine the distance for splitting to avoid overlapping from 100 to 1,500.

The higher the length of the lixel and the minimum pixel length, the more obvious the results will be on the map as the pixels cover more distance. As generated by the UI, we adopt the red-yellow/ green-yellow color scheme where red/green shows the most dense area while yellow shows the least dense for each bandwidth respectively to make the results more visible,

This will allow users to quickly identify areas where there are high concentrations of a particular type of location, such as a business establishment or childcare center. This can be useful for identifying areas of the city that may be of particular interest to the user, such as trendy districts.

4. DEMONSTRATION - USE CASE

4.1 Data Exploration

- **Network**

By using the map with the WorldGrayCanvas background, the user can easily view the **road network, pedestrian network, and tram network** in a clear and detailed manner. The road network can help the user navigate the city and plan driving routes, while the pedestrian network provides information on walkways and pedestrian-friendly areas. The tram network, which is a vital part of the public transportation system, can be viewed to plan trips or to explore the city's public transportation options. With the WorldGrayCanvas background, the user can easily switch between various layers to obtain the desired level of detail and view the different types of networks simultaneously. This feature can be particularly useful for individuals who are unfamiliar with the city and need to navigate around efficiently.

- **Spatial Points**

Users have the ability to observe the geographical distribution of points of interest within the City of Melbourne, displayed on WorldGrayCanvas as a background. They can **zoom in or out** to explore specific details about each location or choose to focus on **particular themes/sub-categories**. Furthermore, users can filter the area of interest according to their preference. In addition, a **tabular representation of point data** for respective points of interest is available for users to view.

4.2 Spatial Point Pattern Analysis

- **1st Order**

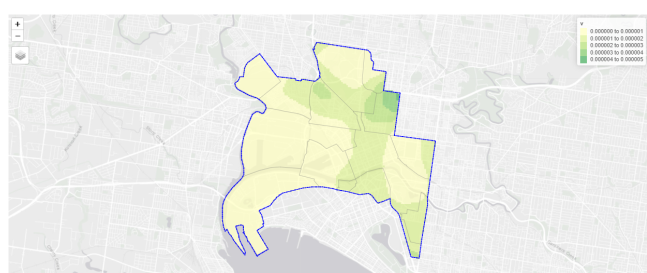

Fig 4: A Kernel Density Estimation map (KDE) showing child care centers within the entire city of Melbourne

If the filters are Localities: Entire City of Melbourne, Location of interest: Child Care centers, Bandwidth: Fixed, Kernel: Quartic, Method: Simple, Confidence: 95%:

There will be a green-yellow KDE map showing the patterns of the **childcare centers** within the **entire city of Melbourne** whereby the North-east side of the city will be the most dense (dark green).

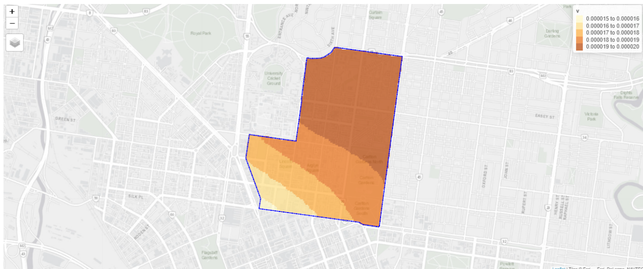

Fig 5: A Kernel Density Estimation map showing drinking fountains within Carlton

If the filters are Localities: Carlton, Location of interest: Drinking Fountains, Bandwidth: Adaptive, Confidence: 99%:

There will be a red-yellow choropleth map showing the patterns of the **drinking fountains within Carlton** whereby the North-east and the North side of the city will be the most dense (red) while the opposite side is least dense(yellow). Also, we can observe that there is a wave-like pattern from south to the North side of Carlton.

- **2nd Order**

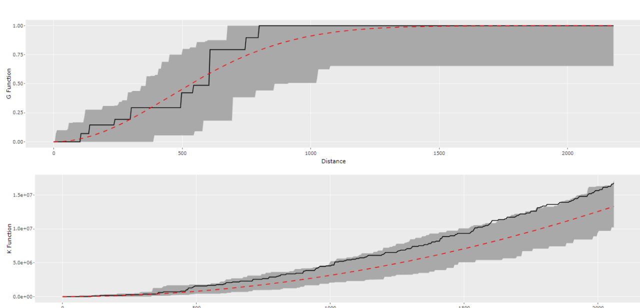

Fig 6: G and K Function for child care centers within the entire city of Melbourne

If the filters are Localities: Entire City of Melbourne, Location of interest: Child Care centers, Confidence: 95%:

There will be **2 gray line charts** showing the patterns of the **childcare centers** for every 1 meter within the entire city of Melbourne for both **G-function and K-function**.

4.2 **Network Constrained Point Pattern Analysis**

- **Network Kernel Density Estimation**

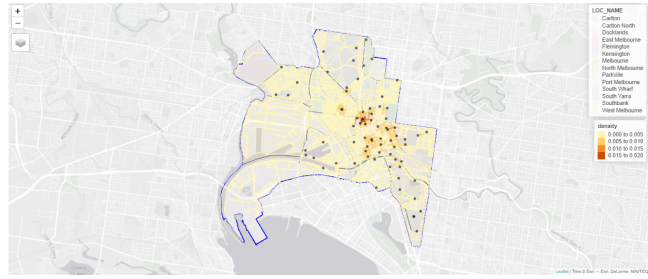

Fig 7: Line map between public toilets and road network within the entire city

If the filters are Localities: Entire City of Melbourne, Location of interest: Public Toilets, Types of Network: Road Network, Kernel: Tricube and Method: Simple, Length of lixel: 700, Minimum Lixel length: 350:

It will generate a line map with red-yellow lines displaying the patterns of the **road network while accounting for the accessibility of the public toilets** nearby. The more red the lines are, the more accessible the public toilets.

- **Statistical Functions**

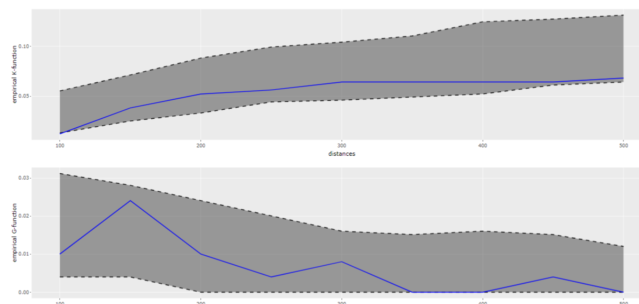

Fig 8: G and K Function for Drinking Fountains within Docklands

If the filters are Localities: Docklands, Location of interest: Drinking Fountains, Types of Network: Road Network, Start: 100, End: 500, Number of Simulations: 50, Aggregate Value: 50:

There will be 2 gray line charts showing the patterns of the **Drinking Fountains** for every 1 meter within Docklands for both e**mpirical G-function and empirical K-function** up to 500 meters.

5. DISCUSSION

5.1 Usability of Application

The application is designed to maximize interactivity by incorporating various features such as map zooming and filter selection to cater to user preferences. With reference to Jakob Nielsen's 10 heuristics for user interface design, this application has successfully met the criteria. (World Leaders in Research-Based User Experience, 2021).

- **Consistency & Standards**

The design of this application adheres to consistency and standards. Filters for all of the tabs are similar, allowing users to display maps based on the localities and locations of interest that users intend to study.

- **Aesthetic and minimalist design**

The aesthetic and minimalist design ensures that only necessary filters and maps are displayed on each tab, providing users with useful insights.

- **Help and documentation**

To aid users, these are some of the following features that we have added:

- There is an Introduction for every function tab to briefly explain the spatial model and its usage.
- The user guide is provided in the Home Page tab.
- A set of instructions and explanations below every function and an example to help users understand the functionalities.

5.2 Optimizing placement of water drinking fountains

One practical example that our application can apply on is the study of the spatial points of the water drinking fountains around the city.

By analyzing the spatial patterns of existing water drinking fountains, city officials can optimize their placement. For example, they can place fountains in areas where there is high pedestrian traffic, such as near tourist attractions or busy streets. This can help to ensure that the fountains are easily accessible and well-utilized.

By ensuring that there are sufficient water drinking fountains around Melbourne City, officials can also help to improve public health outcomes. Access to clean and safe drinking water can help to reduce the risk of dehydration and related health problems, particularly during hot summer months.

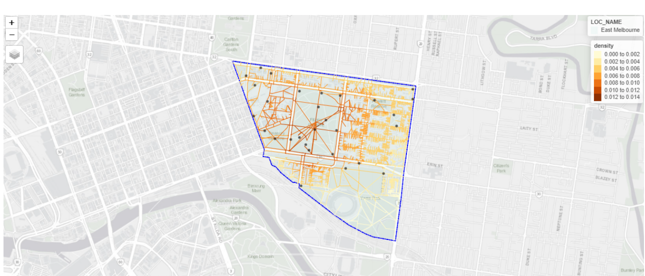

Fig 9: Network Kernel Density Estimation for Drinking Fountains in East Melbourne

For instance, we can first set the Localities: East Melbourne , Location of interest: Drinking Fountains , Types of Network: Pedestrian Network, Kernel: Quartic and Method: Simple, Length of lixel: 700, Minimum lixel length: 350:

From Figure 9, we can see that the middle of the district is the most dense in terms of the walking accessibility of drinking fountains. Therefore, city officials can try to look into other parts of East Melbourne to ensure that the public have easy access to safe drinking water as they travel in the pedestrian network.

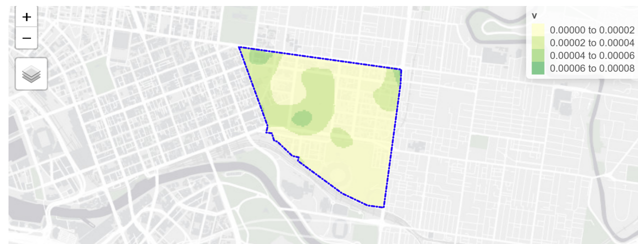

Fig 10: 1st Order Kernel Density Estimation for Drinking Fountains in East Melbourne

By generating the 1st Order Spatial Points Analysis Kernel Density Map in Figure 10, we can see a slightly more macro level view of the coverage of each drinking fountain; these are not network-constrained and give a general idea of which areas have a higher density of drinking fountains.

This is different from Network Constrained Kernel Density maps where pedestrian network in this case is important to the analysis as users might not be able to access a Drinking Fountain as easily as there is no pedestrian path present in a euclidean manner.

6. CONCLUSION

In conclusion, this study investigated the use of network-constrained analysis & first- and second-order spatial point pattern analysis approaches to look at the spatial patterns for the points of interests in Melbourne City, Australia. The findings demonstrate the understanding of spatial points with reference to the environment through 1st Order Spatial Point Analysis and between spatial point events through 2nd Order Spatial Point Analysis. Next, utilizing Network Constrained Spatial Point Analysis, we can tell the density of spatial points with reference to its surrounding networks, such as Drinking Fountains along Pedestrian paths. In this specific scenario, the use of Network Constrained SPatial Point Analysis is important as pedestrians can only walk on pedestrian paths and cannot cut across the built environment features, such as building walls or private land in a euclidean manner.

6. FUTURE WORK

The Spatial Bros application is an interactive tool which aims to allow citizens of the City of Melbourne and users in general to understand the use of Spatial Point Pattern Analysis (SPPA) and Network Constrained Point Pattern Analysis (NCPPA). Since visualizations are currently only offered as Spatial Points on the portal, it limits our data exploration and we hope that datasets are readily available and regularly updated for public use.

Our application tool currently further enhances the democratization of data by allowing the layman to understand the distribution of points of interest in the City of Melbourne using SPPA and NCPPA to see how the distributions differ. The tool will include instructions and explanations on the interpretation of distributions and statistical results to allow them to formulate their own findings. This could lead to heightened awareness and action from citizens, for example, to increase the amount of drinking fountains to increase the convenience of refilling bottles and reduce the use of single use disposable bottles.

In future works, we hope to be able to further enhance data democratization by allowing the import of any network and dataset in the world to understand data in their region. Side by side analysis could also be included to help understand the various urban typologies across cities and the differing distribution of Spatial Points across cities.

7. REFERENCES

- Glynn, L. (2023, January 16). *Yes! Melbourne has been ranked third in the top trending travel destinations for 2023*. Time Out Melbourne. Retrieved from https://www.timeout.com/melbourne/news/melbourne-ranked-third-in-the-top-trending-travel-destinations-for-2023-011623

- Rui, Y., Yang, Z., Qian, T., Khalid, S., Xia, N., & Wang, J. (2015). Network-constrained and category-based point pattern analysis for Suguo retail stores in Nanjing, China. *International Journal of Geographical Information Science*, *30*(2), 186–199. https://doi.org/10.1080/13658816.2015.1080829

- Špatenková, O., & Stein, A. (2010). Identifying factors of influence in the spatial distribution of domestic fires. *International Journal of Geographical Information Science*, *24*(6), 841–858. https://doi.org/10.1080/13658810903143634

- World Leaders in Research-Based User Experience. (April 24, 1994). 10 usability heuristics for user interface design. Nielsen Norman Group. Retrieved April 14, 2023 from https://www.nngroup.com/articles/ten-usability-heuristics/.