Earth Observation and Geomatics Engineering 7(1) (2023) 45–55

EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

Identifying critical locations in urban transportation networks using network kernel density estimation (Net-KDE): a spatial analysis

Marjan Ghanbari [1], Mohammad Karimi[1]*, Christophe Claramunt[2], Claire Lagesse [3]

[1] Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran
[2] Naval Academy Research Institute Lanveoc-Poulmic, BP 600, 29240 Brest Naval, France
[3] Franche-Comté University, CNRS, ThéMA, F-25000 Besançon, France
Article history: Received: 2023-03-17, Accepted: 2023-05-01, Published: 2023-06-14

ABSTRACT

In the realm of urban transportation management, identifying critical locations within transportation networks is paramount for efficient urban planning, crisis management, and infrastructure development. This study investigates the effectiveness of the Network Kernel Density Estimation (Net-KDE) method in identifying these critical locations. The Net-KDE method is chosen for its inherent capacity to consider spatial patterns and path length between sample and estimation points, making it well-suited for capturing the complexities of urban transportation networks. The approach is experimented with a series of control maps applied to Tehran city. Overall, Net-KDE provides valuable and agreeable outputs that show its potential. The findings underscore the crucial role of scenario customization in enhancing the method's accuracy. Remarkably, featuring adaptive bandwidth and strategic sampling emerges as the most effective in identifying critical locations. This research has a broader impact on crisis responders, urban developers, and city planners in addition to improving our knowledge of urban network dynamics.

KEYWORDS

Network Kernel Density Estimation

Urban Transportation Network

Critical Locations

Sampling Strategy

Spatial Bandwidth

Urban Planning

1. Introduction

The foundation of everyday life in our globalized society is provided by vital infrastructure systems for energy, telecommunication, and transportation. Transportation networks in urban areas need to be resilient because any disruption in one of these systems could have far-reaching effects (Hosseini & Pishvaee, 2022). Disruptions, whether intentional or caused by natural disasters, can have a big effect on everyday life, especially in our fast-paced urban settings (Hosseini & Pishvaee, 2021). Thus, it is crucial to pinpoint the exact locations of the transport network's hubs during disruptive events (Martinez-Pastor, 2018; Martinez-Pastor et al., 2022).

The structural aspects of transport networks have been the primary focus of previous research on the identification of critical points (Bertolini & Dijst, 2003; Camagni & Salone, 1993). Various indicators that specifically address the physical layout of road networks from a structural perspective have been proposed and examined in some studies (Duan & Lu, 2013; Freeman, 2002; Porta et al., 2006). On the other hand, researchers have emphasized the crucial connection between the design of the transport network and urban functionality, emphasizing how important these connections are in identifying critical locations. Interestingly, in this context, the travel distribution pattern has become a useful indicator (Zhou et al., 2015). Numerous approaches currently in use in the literature require gathering a great deal of data and measuring several indicators in order to identify and assess the results (Martinez-Pastor et al., 2022). These techniques are frequently used for entire urban transport systems,

* Corresponding author

E-mail addresses: mkarimi@kntu.ac.ir

which places a significant computational and time burden, especially when working with dynamic networks. In an attempt to identify functionally critical locations, Zhou et al. (2015) employed the kernel density estimation method for spatial point pattern analysis. Taking into consideration both structural characteristics and the amount of traffic flowing through these intersections during the sampling process, they employed a restricted set of sample points collected from intersections. Their approach was based on the Ordinary KDE method, which is intended for two-dimensional spaces, even though their study yielded insightful information. Recognizing limitations in the ordinary KDE algorithm within a one-dimensional road network, two key issues materialize: 1) Estimation points in a two-dimensional space may lack specific road network alignment, affecting the accordance of identified critical locations, a concern addressed in this study. 2) The use of Euclidean distance in point-to-point calculations, as in the ordinary KDE algorithm, is questioned, advocating for the adoption of network distance (path length) for more practical results.

To overcome these limitations, we implement a modified substitute identified as Network Kernel Density Estimation (Net-KDE) (Xie & Yan, 2008). Net-KDE places estimation points precisely on the road network, often as lixel center points. It utilizes network distance calculations instead of Euclidean distance, following Xie and Yan (2008), ensuring accurate alignment with the road network. This approach, unlike typical studies, extends beyond proximity-based density calculations, enhancing spatial accuracy in critical location identification.

Net-KDE stands out in its ability to provide accurate spatial results by uniquely considering the specific characteristics of transportation networks. This accuracy is underpinned by the algorithm's incorporation of the network's structural features, such as road structures, in the density estimation process. Unlike traditional methods, Net-KDE ensures a more meticulous representation of the transport system, resulting in spatial outcomes that closely align with the actual one-dimensional network (McSwiggan et al., 2017).

For emergency response and the planning of urban transport networks, this precision is especially useful. One major advantage of Net-KDE is its efficiency, which lowers the computational burden when compared to conventional techniques. Because it doesn't require a lot of data or various indicators, it's vital to managing dynamic, constantly shifting transportation networks. We streamline the analysis procedure by utilizing Net-KDE for the analysis of spatial point patterns. By removing the need for complicated additional factors and simplifying the process of identifying critical locations, this method enhances accessibility and user-friendliness. In the end, Net-KDE is an acceptable alternative that can be used in actual urban transportation environments.

Within the urban context of Tehran, the Net-KDE methodology was used in a real-world case study. In Tehran's transport network, the critical locations were identified and then compared to reference thematic city maps. The term 'reference city maps' signifies thematic maps of Tehran, including master plans and grade separation data from transportation studies. These maps act as controls for comparing critical locations identified by Net-KDE. To validate the accuracy of Net-KDE in this complex urban setting, spatial point pattern similarity and proximity were evaluated. The knowledge obtained from this empirical approach prepares the way for improving urban management, especially with regard to the dependability and safety of the transportation network.

The remainder of the paper is structured as follows. The next section briefly reviews related work while section 3 develops the principles behind the NET- KDE methodology. The study area is presented in section 4 while the experimental results are reported in sections 5 and 6. Finally, section 7 discusses the findings and section 8 draws conclusions and outlines further work.

2. The Related Work

Critical location identification in transportation networks has traditionally been based on graph theory, which frequently extends structural metrics such as node degree, clustering coefficient, cut vertex, and centrality measures to each network element (Demšar et al., 2008; Freeman, 2002; Fulkerson & Harding, 1977; Psaltoglou & Calle, 2018). Alternatively, studies have examined the consequences when certain network links are eliminated (Jenelius, 2009; Rodríguez-Núñez & García-Palomares, 2014). In the context of large and dynamic urban transportation networks, these methods are particularly laborious to calculate, even though they produce insightful results. To predict congestion and its effect on traffic performance, data-driven methods have been used, such as traffic flow prediction (Corley & David, 1982; Nguyen et al., 2016). However, in order to create accurate models, these methods require large databases that are updated often.

Zhou et al. (2015) presented a spatial point pattern analysis as an alternative method for locating critical locations throughout the road network. They restricted the subject matter of the analysis to a restricted amount of sample points by selecting components of the urban transport network according to functional and structural metrics. Then, using the Kernel Density Estimation (KDE) technique, critical locations were determined. This strategy benefited from the advantages of previous methods while addressing their shortcomings. The application of the KDE method as a spatial point pattern analysis produced an important shift. For point events in geographic space, KDE computes a continuous probability density surface

(Baddeley et al., 2015; Bernardo et al., 2015; Gatrell et al., 1996; Zhang, 2022). Its uses extend to a wide range of fields, including disaster preparedness (Krisp et al., 2005), crime analysis (Kalinic & Krisp, 2018; Levine, 2017), road accident assessment (Anderson, 2009), and wildlife-vehicle collision analysis (Borrajo et al., 2021). The problem, though, is that critical locations in urban transport networks are inevitably associated with road networks, indicating the necessity for a spatial analysis that considers the road network layout. A variety of events, organized as points along a network of lines to create distinct spatial patterns, are utilized in many applications. These events include traffic accidents, bicycle incidents, vehicle thefts, street crimes, roadside trees, and invasive species (Baddeley et al., 2021). These locations in such networks are inherently tied to road networks, necessitating a spatial analysis that fully considers the road network layout. The inherent limitation of KDE in capturing the intricacies of road networks prompts the need for more tailored approaches, to enhance the accuracy in location identification.

Spatial statistics and Geospatial Information Systems (GIS) communities have been concentrating more and further on studying point patterns throughout linear networks over the years (Baddeley et al., 2000; Briz-Redón et al., 2019; McSwiggan et al., 2020; Okabe & Yamada, 2001). The Net-KDE approach (McSwiggan et al., 2017; Okabe et al., 2009; Xie & Yan, 2008) has been proposed as a solution to these limitations, enabling effective kernel estimation on one-dimensional networks. The aforementioned method has been implemented in diverse settings, such as identifying areas of high risk in transportation networks (Briz-Redón et al., 2019) and calculating the relative risk in spatial point patterns (McSwiggan et al., 2020). In the area of critical location identification within transportation road networks, the adoption of the Net-KDE approach represents a noteworthy advancement.

This particular challenge can be effectively addressed by it because of its effectiveness, accuracy, and flexibility in one-dimensional networks.

With Net-KDE's benefits, we can more effectively manage urban areas and ensure the reliability and safety of these vital transport networks by utilizing Net-KDE's potential for accurate network-based identification of critical locations.

3. The Proposed Method

This study's methodology consists of multiple discrete steps. First, a graph representing the urban transportation network is constructed, with road segments acting as edges and intersections serving as nodes. These graph nodes serve as the basis for the strategic placement of sample points. The network is then divided into a continuous grid, and estimation points are chosen from the centers of these cell types. The path length between the sample points and the estimation points needs to be determined, in addition to the spatial bandwidth. After implementing these fundamental procedures in place, estimation points based on the network kernel function are used to perform spatial density estimation, subsequently, the methodology's last component is to identify critical locations of the urban transport network based on the estimated densities. The workflow of the study is depicted in Figure 1.

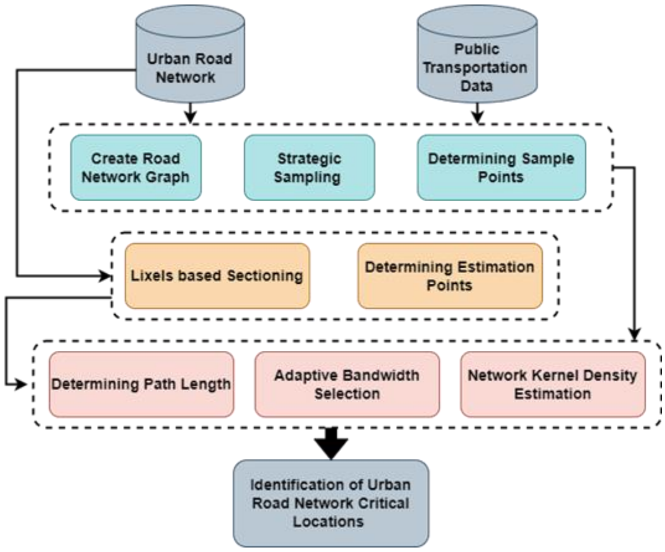


Figure 1. Research Workflow

3.1 Construction of the Transportation Network Graph

Graph theory principles serve as the foundation for the creation of the transportation network graph (Gould, 2012).

A road network is represented in this paper representation as $G = \{V, E\}$, where the set of nodes, or intersections within the network, is represented by $V = \{v\_1, v\_2, ..., v\_n\}$, and the set of edges, or road segments connecting these nodes, is represented by $E = \{e\_1, e\_2, ..., e\_m\}$.

It is important to emphasize that sample points are obtained by choosing samples from the graph's nodes. In the selection of sample points, a systematic approach was employed by filtering distinct intersections in Tehran, with priority given to those with high public transportation volume. In order to estimate the probability density function of a random variable, these sample points function as observed data points.

3.2 Definition of Lixels[1] and Network Construction

Every segment in the urban transport network serves as a link between two nearby road intersections in the reference network to form a segment-based graph. Each linear lixel unit, which has a specified network length, is created by further dividing these segments. The creation of evenly distributed points throughout the network, which are

[1] A fundamental linear unit corresponds to a cell in a 2D raster grid.

necessary for density estimation, is made considerably by this division (Xie & Yan, 2008).

The central points of each lixel, denoted by numbered grey circles in Figure 2, are regarded as estimation points. An estimation point is a spot where the probability density function is calculated by combining the contributions of kernel functions placed at each sample point.

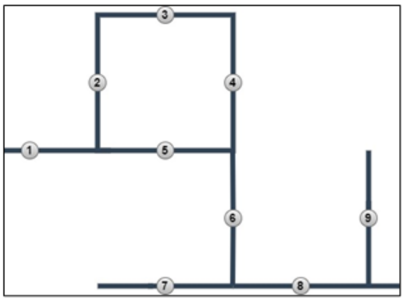


Figure 2. Determining estimation points based on lixels network construction

The adoption of lixels in the analysis of urban transport networks offers specific advantages over traditional network vertices. Unlike traditional vertices, lixels break down the network into smaller, interconnected cellular units, providing a more refined representation. This approach enables granular and detailed analysis of critical locations, capturing slight spatial patterns within the transport network. Lixels enhance spatial resolution by acting as cellular units, leading to increased precision in identifying critical locations. Additionally, they enhance the representation of the network structure by considering edges and adopting a cellular and point-based network structure, especially in areas with diverse road layouts. Overall, lixel-based analysis allows for a more comprehensive capture of both the density and spatial distribution of critical locations, providing a subtle insight of the transportation network (McSwiggan et al., 2017; Xie & Yan, 2008).

3.3 Network Kernel Density Estimation

In the Network Kernel Density Estimation process, estimation points undergo systematic processing based on a predefined spatial bandwidth. Utilizing a network kernel function, the spatial density assigned to each estimation point in relation to the sample points is measured. The proximity of data points within the network's spatial configuration determines how sample points affect the density estimation of estimation points.

The core equations (1-2) express this connection (McSwiggan et al., 2017):

$$\hat{f}_s(x_i) = \frac{1}{n}\sum_{j=1}^{n} K_{h_{s_i}}(x_i|x_j)\ \hat{f}_s(x_i) \qquad (1)$$

$$\hat{f}_{s_{(x_i)}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{\pi} a(\pi) \times \frac{1}{h_{s_i}\sqrt{2\pi}} e^{-0\cdot5\left(\frac{l(\pi)^2}{h_{s_i}}\right)} \qquad (2)$$

Where, K represents the kernel function, x_i is the estimation point, x_j is a sample point, h_{s_i} is the adaptive bandwidth (Abramson, 1982), and n is the number of sample points. The kernel function calculation is based on Equation 3-5:

$$K_{h_{s_i}}(x_i|x_j) = \sum_{\pi} a(\pi)\varphi_{h_{s_i}}(l(\pi)) \qquad (3)$$

$$a(\pi) = \prod_{v_i \epsilon \pi}\left(\frac{2}{\deg(v_i)} - \delta_i\right) \qquad (4)$$

$$h_{s_i} = \sigma_0\left(\frac{\left(\frac{\tilde{f}_{s_{(x_i)}}}{n}\right)^{-0\cdot5}}{\prod_{i=1} a_i^{\frac{1}{n}}}\right) \qquad (5)$$

In Equation 3. $\pi = (v_1.v_2.\dots.v_p)$, where $v_1.v_2\dots.v_p$ are vertices, and $\varphi_{h_{s_i}}(l(\pi))$ is the probability density for a total displacement $l(\pi)$, the summation accounts for different possible paths in the network, and $a(\pi)$ is a combinatorial weight for the path and is calculated based on Equation 4, in which v_i refers to a node in the path π, and δ_i is defined for 1<i<$P-1$, by $\delta_i = 1$, if the path goes backward at the node v_i, while $\delta_1 = \delta_{p-1} = 0$. Equation 5 incorporates key variables: σ_0 representing the global bandwidth; pre estimated value $\widetilde{f_{s}}_{(x_i)}$, calculated using a rule-of-thumb bandwidth $h_s = (3n)^{-(\frac{1}{5})}\bar{s}$ based on (Scott, 2015), where $\bar{s}$ signifies the root sum of squares of the sample standard deviation. Additionally, a_i is introduced, representing the initial bandwidth computed as $a_i = \sqrt{n/\tilde{f}_{s(x_i)}}$, with n denoting the sample size.

By capturing local density variations, the spatial density makes sure that the estimation points are influenced by nearby sample points on the network. See (McSwiggan et al., 2017) for a more thorough explanation and derivations of these equations.

4. Study Area and Data

The study area incorporates Tehran, the capital city of Iran, which has a population of about 10 million people (Figure 3). Natural disasters like earthquakes and floods are prevalent in Tehran, and dealing with the city's overcrowded transit system can make crisis management more difficult. Consequently, locating critical locations on the road network is the main goal of this investigation.

Figure 3. The study area, Tehran

The data utilized in this study, including Tehran Municipality's public bus transport data for the year 2020–2021 and road network datasets, were acquired in GIS format. The scale of the data is at the city level, with a specific emphasis on the urban transport network of Tehran.

In this study, Tehran's bus transport data was employed due to its significance in representing urban traffic patterns, a crucial factor in the sampling process, density estimation using Net-KDE, and identifying critical locations within the urban transport network. While the data may not cover the entire road network, it serves as a resource to gain an understanding of urban traffic dynamics and contribute to the comprehensive analysis of critical locations. The use of bus transit data aligns with the available resources for studying urban movement in Tehran and helps in understanding the complexities of the transportation network.

5. Implementation and Results

This section investigates the data's quantitative analysis and provides significant insights from this research, emphasizing the significance of the Net-KDE methodology's identification of critical locations in the urban transport network.

5.1 Sampling Process and Selection of Estimation Points

Strategically chosen samples and estimation points play a crucial role in representing Tehran's urban transport system within the comprehensive network. The selection criteria for intersections were based on distinct characteristics, with priority given to those with more than four-way connections and high-traffic-volume public transport intersections. This ensured a focused and organized sampling process.

In order to ensure a focused representation, high-traffic intersections are prioritized in sampling to capture important nodes in Tehran's urban transport system. Insights into the dynamics and performance, particularly with regarding public transportation operations and urban mobility, can be obtained from this strategic selection of significant locations for thorough network analysis.

Figure 4 (a, b) displays heat maps illustrating the distribution of sample points in both random and strategic sampling manners. The heat maps were created in ArcMap, using GIS software. The strategic sampling map illustrates clustered patterns by emphasizing prioritized locations with high public transportation traffic, while the random sampling map shows the distribution of sample points. The present analysis contributes to the comprehension of spatial distribution across different sampling scenarios by providing insights into the potential of strategic sampling to capture concentrated areas.

5.2. Critical Locations

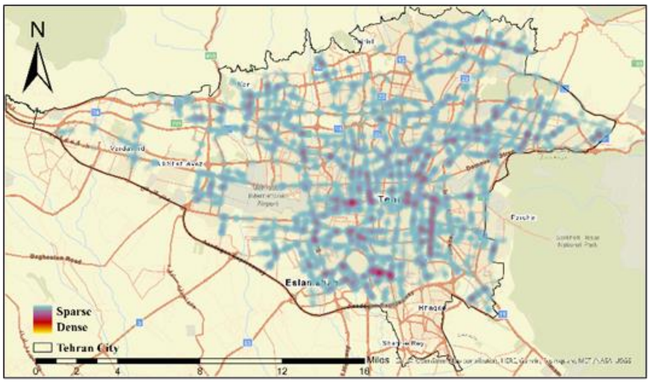
To evaluate and contrast the effectiveness of the approaches employed in this study, three distinct situations have been developed. These scenarios' primary objective is to evaluate the Net-KDE method's effectiveness with regard to various sampling strategies and selecting between fixed and adaptive spatial bandwidth.This study critically evaluates the Net-KDE method's efficacy in addressing urban challenges. The impact of bandwidth on smoothness and sensitivity to variations in the data is examined, resulting in a crucial parameter in kernel density estimation. The scenarios provide insights into how Net-KDE adapts depending on the bandwidth settings, ensuring a detailed assessment of its efficacy in capturing local variations despite overfitting.

The Net-KDE method is used in the first scenario, but its bandwidth is fixed. The Net-KDE method is used in the next two scenarios with both fixed and adaptive spatial bandwidth selection, along with strategic sampling. This investigation makes it possible to evaluate this approach's effectiveness in solving the given urban issue. A comprehensive overview of the procedure by which each scenario was implemented is provided in the subsequent section.

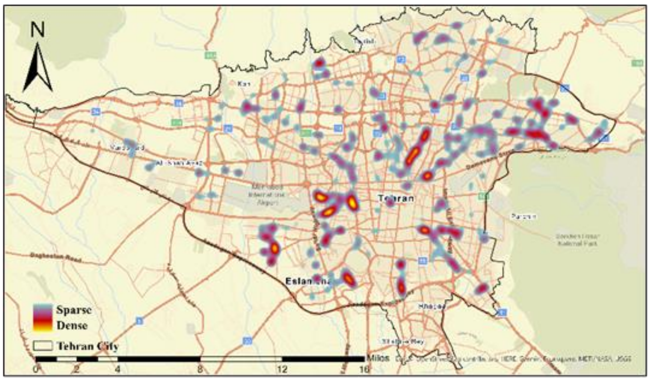
5.2.1 The First Scenario

In the first scenario, the Net-KDE method (Equation 2) was utilized for the identification of critical locations. The sampling approach included the random selection of 5000 sample points, to ensure a comprehensive coverage of the urban transport network. The nearly uniform distribution of these sample points across the study area is illustrated in Figure 4a.

To conduct density estimation effectively, a fixed bandwidth was opted for in this scenario. The selection of a fixed bandwidth was motivated by the relatively uniform distribution of sample points, as depicted in Figure 4a. Utilizing a fixed bandwidth (h) aimed to streamline computational processes and eliminate the potential computational overhead associated with an adaptive bandwidth.



(a)



(b)

Figure 4. Heat map of Samples Spatial Distribution (a) Random Sampling, (b) Strategic Sampling

The calculation of the suitable fixed bandwidth (h) involved employing Equation 6 (Scott, 2015).

$$h = (3n)^{-\left(\frac{1}{5}\right)} \bar{s} \qquad (6)$$

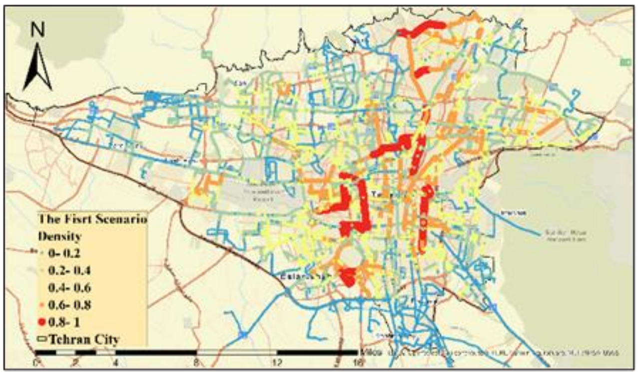
Here, $\bar{s}$ represents the root sum of squares of the sample standard deviations. The resulting fixed bandwidth was determined to be *1.3 km*. This choice aimed to strike a balance between capturing local variations in density and avoiding overfitting, thereby contributing to the effectiveness of the Net-KDE method in identifying critical locations within the urban transport network.

5.2.2 The Second Scenario

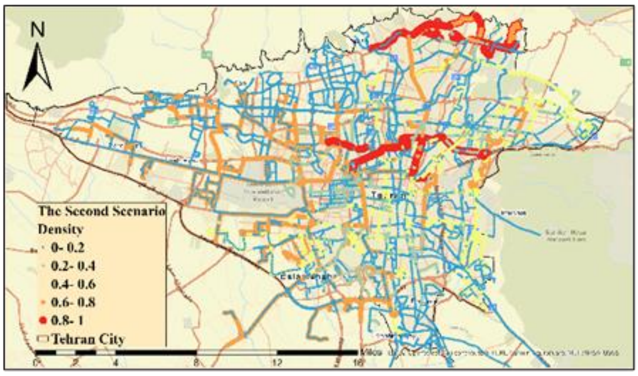
In the second scenario, a strategic sampling approach was implemented, leading to a reduced sample size of 1,411 points. This deliberate sampling strategy aims to focus on specific areas of interest within the urban transport network. As depicted in Figure 4b, the sample points display varied densities, deviating from the uniform distribution observed in the first scenario. To facilitate density estimation effectively, a fixed bandwidth is determined using Equation 6, resulting in a calculated bandwidth of 2.9 km. This fixed bandwidth is applied to streamline computational processes and eliminate potential computational overhead associated with an adaptive bandwidth, similar to the rationale employed in the first scenario.

5.2.3 The Third Scenario

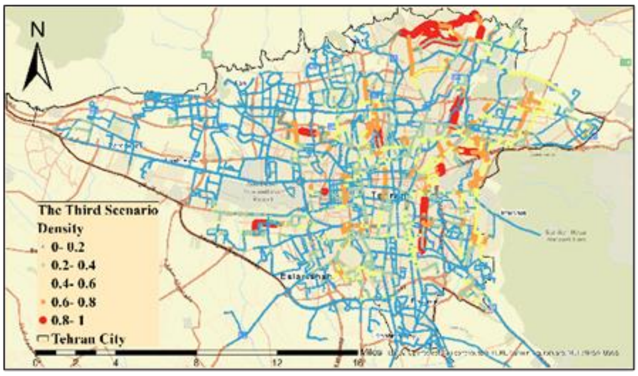
In the third scenario, due to the non-uniform distribution resulting from strategic sampling, a fixed bandwidth is deemed suboptimal. Unlike the first two scenarios, the third scenario adopts an adaptive bandwidth approach. The adaptive bandwidth is determined using Equation 5, yielding values ranging from 0.81 km to 3.6 km. This variation in bandwidth aims to dynamically adjust to the spatial characteristics of the sample points, providing a more nuanced density estimation. The third scenario, akin to the second, seeks to capture the complexity of the urban transport network. The outcomes of these three scenarios are presented in Figure 5.



(a)



(b)

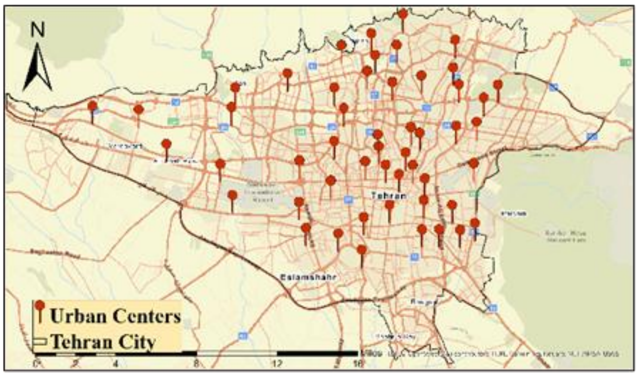


(c)

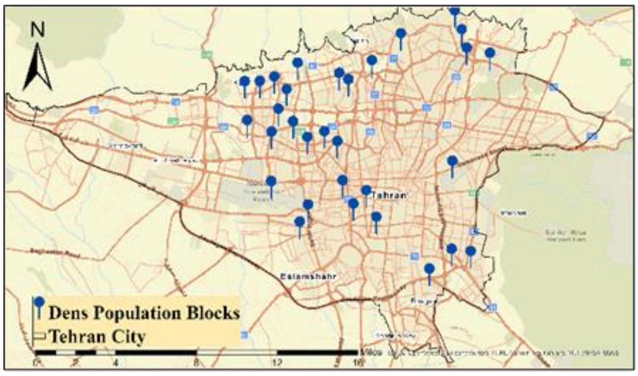
Figure 5. Primary Critical locations; (a) the first scenario, (b) the second scenario, (c) the third scenario

6. Evaluation

Performance evaluation is crucial to ensuring the method's reliability. In order to provide comparable ground truth data, our assessment includes the utilization of additional data from transportation and traffic research in addition to multiple thematic maps of Tehran, including three master plan reference maps. Several aspects of the urban environment are depicted on these reference maps:

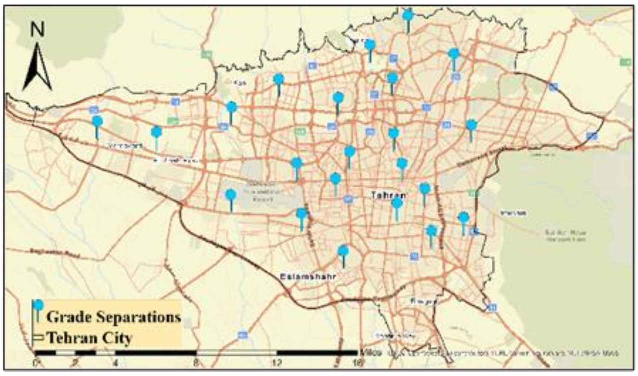


(a)



(b)

(c)



(d)

Figure 6. Control Maps (a) Urban center, (b) Dense population, (c) Mixed zones, (d) Grade separations

The first reference map (Figure 6(a)) focuses on urban centers, including places of commerce, recreation, culture, and landmarks.

Blocks with over 3,000 inhabitants are depicted in the second reference map (Figure 6(b)).

Mixed-use areas are marked on the third map (Figure 6(c)).

The fourth map shows grade separation locations with significant movement flows and a variety of land uses based on transportation reference data (Figure 6(d)).

These reference maps operate as controls for determining the critical locations determined by the various scenarios.

In order to establish a correspondence between the scenarios and the mixed zones and grade separation control maps, we have employed the K-Means clustering approach.

The optimal number of clusters is determined using the Calinski-Harabasz (CH) index (Caliński & Harabasz, 1974), which measures an object's similarity to its cluster compared to others, based on distance. A higher CH index indicates denser and more independent clusters. As shown in Figure 7, we determine the number of clusters for these two control maps as 21 and 28, respectively. Ultimately, these cluster centers are considered control locations.

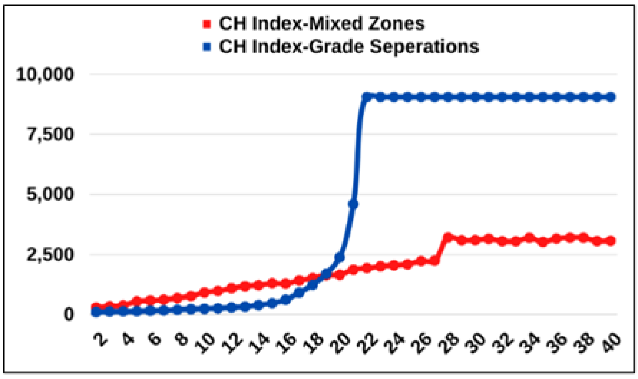
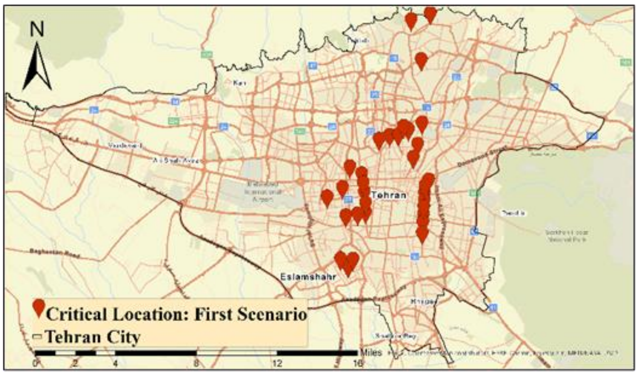


Figure 7. The Calinski-Harabasz score for K-Means clustering of Mixed Zones map and Grade separation map

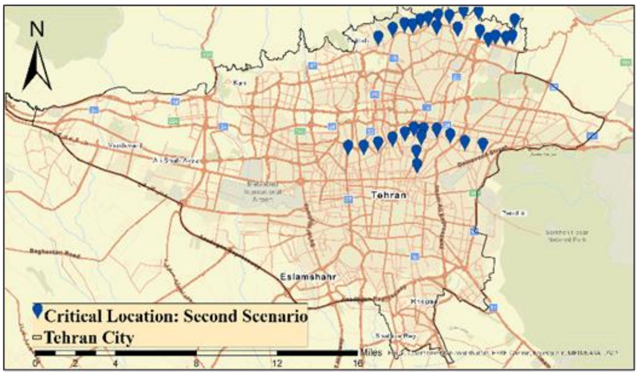
The scenarios produce different results, as shown in Figure 5, which makes cross-comparison a challenging task. In order to overcome this difficulty, we utilize the K-Means clustering technique, which divides the average number of all control map points into 32 clusters, thereby grouping critical points from various scenarios. These cluster centers are identified as critical locations since they indicate the average position of the cluster members. Figure 8 depicts the critical locations that each scenario's post-K-Means clustering revealed.

6.1. Spatial Similarity Evaluation

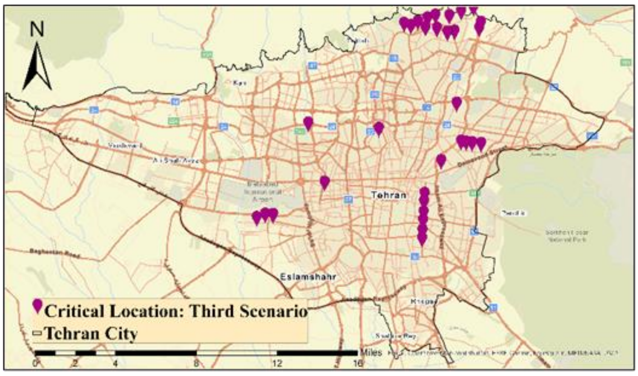
Spatial point pattern analysis is employed to assess the similarity between the scenarios and control maps by examining the spatial distribution of points (Vidanapathirana et al., 2022). The observed data in this study, represented by the critical locations identified in each scenario, are compared to the spatial patterns of control points on each control map. This comparison utilizes *Ripley's K-Function* (Equation7) and the *Average Nearest Neighbor (ANN) Ratio* (Equation8) (Clark & Evans, 1954).



(a)



(b)



(c)

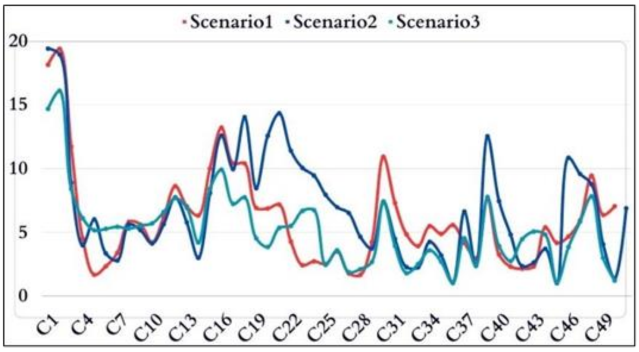
Figure 8. Final Critical Locations (a) first scenario, (b) second scenario, (c) third scenario

$$K(r) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} I(\|x_i - x_j\| \le r) \qquad (7)$$

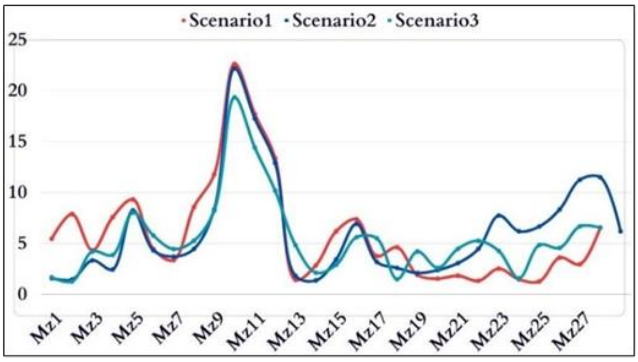
Where $K(r)$ is the *K-Function* at distance r, n is the number of sample points, $x_i.x_j$ are the coordinates of points *i, j,* and $I(\|x_i - x_j\| \le r)$ is an indicator function that equals 1 if the distance between points *i, j* is less than or equal to r, and 0 otherwise.

$$ANN\ Ratio = \frac{Observed\ Mean\ Distance}{Expected\ Mean\ Distance} \qquad (8)$$

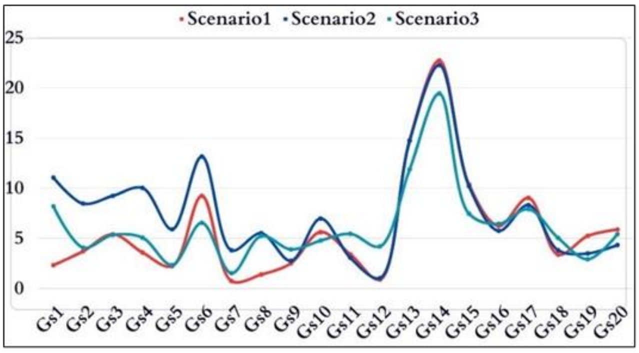
The observed mean distance is calculated based on the actual distances between points in the sample data, while the expected mean distance is determined based on hypothetical random distributions.



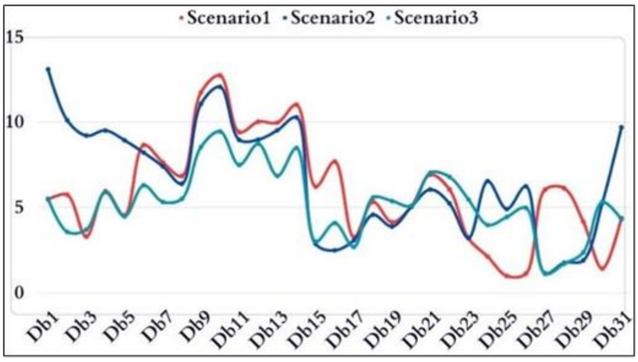
(a)



(b)



(c)



(d)

Figure 9. Average distance (Kilometers) of the control points (a) Urban Centers, (b) Mixed zones, (c) Grade separations, (d) Dens population blocks from the K nearest critical locations

Ripley K-Function's significantly higher values indicate spatial clustering, whereas its lower values represent spatial dispersion.

Table 1. The difference between the Ripley's K-Function and ANN indicators for each control map and scenario map

	Urban Centers		Mixed Zones		Grade Separations		Dens Population Blocks	
	K-Function	*ANN*	*K-Function*	*ANN*	*K-Function*	*ANN*	*K-Function*	*ANN*
Scenario1	2.018	1.7699	3.327	2.9061	2.1505	3.2302	0.2373	1.5298
Scenario2	5.1366	1.9749	6.445	3.111	5.2685	3.4352	2.881	1.2653
Scenario3	1.7466	1.6969	3.055	2.833	1.8785	3.1572	0.5093	0.9874

By taking into account the average distance between all nearest neighbors, the ANN Ratio determines the distance between each point and its nearest neighbor. When features are dispersed, their average distance is greater than what would be expected in a hypothetical random distribution, but when features are clustered, their average distance is less than that of the distribution.

These indicators are instrumental in our spatial point pattern analysis, allowing us to quantitatively assess the spatial distribution of critical locations in scenarios compared to control maps. They play a crucial role in discerning patterns of clustering or dispersion, providing a robust foundation for our evaluation of urban transport system dynamics.

Enhanced authenticity in a scenario, concerning the spatial structure of the city, is achieved when spatial patterns of critical locations become more readily comparable with a broader set of control maps. To accomplish this, the differences between the two indicators for each control map and scenario are estimated by calculating the respective indicators. According to Table 1, the spatial pattern similarity is indicated by the smallest differences, expressed in kilometers, between the indicators for two comparable maps.

The K-nearest neighbor method is applied in terms of spatial proximity. The total average distance for each control map and scenario is illustrated in Figure 9. In the scenario where the critical locations reach the minimum average distance to the control map, the control locations are spatially adjacent to each other. Using the rule of thumb method, $K = \sqrt{(N/2)}$ is the value of K, where N is the total number of critical locations in each case. K is therefore selected as 5.

Table 1 shows the variations for each control map and scenario map between the ANN and Ripley's K-Function indicators. Across all control maps, ***the Third Scenario*** has the highest spatial pattern similarity as measured by the K-Function indicator.

In this context, it's crucial to understand that lower values of the *K-Function* indicate higher spatial similarity. Therefore, when reviewing the *K-Function values* in Table 1, lower values represent a stronger resemblance in spatial patterns between control maps and scenario maps.

There are also notable similarities in the spatial patterns shown by ***the First Scenario***, especially for densely populated blocks. As a result of the decreased similarity in spatial patterns, ***the Second Scenario*** ranks lowest among all the control maps.

The total average distance in kilometers between the control points and the K closest critical locations is shown in Figure 8. Particularly when it comes to urban centers, mixed zones, and densely populated blocks, ***the Third Scenario*** performs better in terms of spatial proximity than the other scenarios. While ***the First Scenario*** performs well across a variety of control maps, ***the Second Scenario*** falls short in capturing the spatial patterns and proximity of the control maps.

In conclusion, the findings highlight how differently the three scenarios perform. While ***the First Scenario*** achieves a good degree of similarity, especially for dense population blocks, ***the Third Scenario*** exhibits the highest level of spatial pattern similarity to control maps. On the other hand, ***the Second Scenario*** performs poorly in these domains.

7. Discussion

The results obtained from the three different scenarios in our research highlight how crucial it is to use the appropriate strategy when applying the Net-KDE method to identify critical locations in urban transport networks.

The majority of sample points in ***the First Scenario***, which contains random sampling, are intersections in the urban network that have no particular significance. Decreased spatial similarity to the control map is the outcome of this random sampling strategy along with the Net-KDE method's dependence on the path length between sample points and estimation points as the main factor for density estimation. As random sampling tends to reduce the impact of actual critical network locations, it presents challenges. The scenario's inability to capture the true spatial patterns of critical locations is one instance of these challenges.

In order to improve the accuracy of the results, ***the Second Scenario*** introduces strategic sampling into the density estimation procedure. However, using a fixed bandwidth in this scenario has drawbacks that restrict its ability to adapt flexibly to various urban conditions. When compared to the control map, ***the Second Scenario*** struggles to achieve a sufficient level of spatial pattern similarity because of this constraint. Although fixed bandwidth makes things more straightforward, it might not be the best option when the underlying network is complex.

The Third Scenario, on the other hand, employs adaptive bandwidth and strategic sampling in a different manner. This approach outperforms other scenarios in terms of both closeness to control locations and similarity in spatial patterns, underscoring its advantages. Adaptive bandwidth provides density estimation more dynamic and responsive, enabling it to capture the spatial variations of critical locations. The strategic sampling approach additionally ensures that the sample points are selected strategically, emphasizing locations of significant importance within the urban network.

These results highlight how important scenario selection is to the accuracy and dependability of the Net-KDE approach. The effective implementation of the third Scenario, which is distinguished by its strategic sampling and adaptive bandwidth, emphasizes the significance of a methodological approach that is in line with the complexity of urban transportation networks. It is clear that not every situation can benefit from a global approach, which highlights the necessity of a customized strategy based on the unique features of the urban network under study.

8. Conclusion

In this study, we investigated the effectiveness of the Network Kernel Density Estimation (Net-KDE) method in identifying critical locations in urban transportation networks. Through the assessment of three different scenarios, we have demonstrated the significant influence that scenario selection possesses on method performance. The results highlight the vital need for scenario customization to align with the distinct features of urban transport networks. There might not be an appropriate response, especially when it pertains to intricate urban structures.

This research has broad consequences across various domains, including urban development, crisis management, and urban planning. It provides valuable insights for enhancing decision-making and resource allocation.

There are numerous directions that require further study as we approach the future. To improve the accuracy and computational efficiency of the Net-KDE approach, one area of interest is the development of more effective sampling techniques. Furthermore, to make the approach flexible for changing urban environments, we propose investigating sophisticated methods for incorporating real-time data. Lastly, possibilities of improving the accuracy of the approach arise from the optimization of the Net-KDE parameters and bandwidth selection procedures.

In conclusion, our study demonstrates the adaptability of the Net-KDE method as an effective tool for analyzing urban transport networks. This research provides the possibility for a deeper comprehension of the spatial dynamics influencing critical locations within transport networks and facilitates more intelligent urban planning.

Acknowledgements

This work was supported by Tehran Research and Planning Center [137.1204088].

References

Abramson, I. S. (1982). On bandwidth variation in kernel estimates-a square root law. *The annals of Statistics*, 1217-1223.

Anderson, T. K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*, *41*(3), 359-364.

Baddeley, A., Nair, G., Rakshit, S., McSwiggan, G., & Davies, T. M. (2021). Analysing point patterns on networks—A review. *Spatial Statistics*, *42*, 100435.

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: methodology and applications with R*. CRC press.

Baddeley, A. J., Møller, J., & Waagepetersen, R. (2000). Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, *54*(3), 329-350.

Bernardo, C., Paleti, R., Hoklas, M., & Bhat, C. (2015). An empirical investigation into the time-use and activity patterns of dual-earner couples with and without young children. *Transportation Research Part A: Policy and Practice*, *76*, 71-91.

Bertolini, L., & Dijst, M. (2003). Mobility environments and network cities. *Journal of urban design*, *8*(1), 27-43.

Borrajo, M., Comas, C., Costafreda-Aumedes, S., & Mateu, J. (2021). Stochastic smoothing of point processes for wildlife-vehicle collisions on road networks. *Stochastic Environmental Research and Risk Assessment*, 1-15.

Briz-Redón, Á., Martínez-Ruiz, F., & Montes, F. (2019). Identification of differential risk hotspots for collision and vehicle type in a directed linear network. *Accident Analysis & Prevention*, *132*, 105278.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, *3*(1), 1-27.

Camagni, R. P., & Salone, C. (1993). Network urban structures in northern Italy: elements for a theoretical framework. *Urban studies*, *30*(6), 1053-1064.

Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, *35*(4), 445-453.

Corley, H. W., & David, Y. S. (1982). Most vital links and nodes in weighted networks. *Operations Research Letters*, *1*(4), 157-160.

Demšar, U., Špatenková, O., & Virrantaus, K. (2008). Identifying critical locations in a spatial network with graph theory. *Transactions in GIS*, *12*(1), 61-82.

Duan, Y., & Lu, F. (2013). Structural robustness of city road networks based on community. *Computers, Environment and Urban Systems*, *41*, 75-87.

Freeman, L. C. (2002). Centrality in social networks: Conceptual clarification. *Social network: critical concepts in sociology. Londres: Routledge*, *1*, 238-263.

Fulkerson, D. R., & Harding, G. C. (1977). Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming*, *13*(1), 116-118.

Gatrell, A. C., Bailey, T. C., Diggle, P. J., & Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British geographers*, 256-274.

Gould, R. (2012). *Graph theory*. Courier Corporation.

Hosseini, A., & Pishvaee, M. S. (2021). Extended computational formulations for tolerance-based sensitivity analysis of uncertain transportation networks. *Expert Systems with Applications*, *183*, 115252.

Hosseini, A., & Pishvaee, M. S. (2022). Capacity reliability under uncertainty in transportation networks: An optimization framework and stability assessment methodology. *Fuzzy Optimization and Decision Making*, *21*(3), 479-512.

Jenelius, E. (2009). Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*, *17*(3), 234-244.

Kalinic, M., & Krisp, J. M. (2018). Kernel density estimation (KDE) vs. hot-spot analysis–detecting criminal hot spots in the City of San Francisco. Proceeding of the 21?? Conference on Geo-Information Science,

Krisp, J. M., Virrantaus, K., & Jolma, A. (2005). Using explorative spatial analysis to improve fire and rescue services. In *Geo-information for disaster management* (pp. 1283-1296). Springer.

Levine, N. (2017). CrimeStat: A Spatial Statistical Program for the Analysis of Crime Incidents. In.

Martinez-Pastor, B. (2018). *Resilience of traffic networks to extreme weather events: analysis and assessment* Trinity College Dublin].

Martinez-Pastor, B., Nogal, M., O'Connor, A., & Teixeira, R. (2022). Identifying critical and vulnerable links: A new approach using the Fisher information matrix. *International Journal of Critical Infrastructure Protection*, *39*, 100570.

McSwiggan, G., Baddeley, A., & Nair, G. (2017). Kernel density estimation on a linear network. *Scandinavian Journal of Statistics*, *44*(2), 324-345.

McSwiggan, G., Baddeley, A., & Nair, G. (2020). Estimation of relative risk for events on a linear network. *Statistics and Computing*, *30*, 469-484.

Nguyen, H., Liu, W., & Chen, F. (2016). Discovering congestion propagation patterns in spatio-temporal traffic data. *IEEE Transactions on Big Data*, *3*(2), 169-180.

Okabe, A., Satoh, T., & Sugihara, K. (2009). A kernel density estimation method for networks, its computational method and a GIS-based tool. *International Journal of Geographical Information Science*, *23*(1), 7-32.

Okabe, A., & Yamada, I. (2001). The K-function method on a network and its computational implementation. *Geographical analysis*, *33*(3), 271-290.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and its Applications*, *369*(2), 853-866.

Psaltoglou, A., & Calle, E. (2018). Enhanced connectivity index–A new measure for identifying critical points in urban public transportation networks. *International Journal of Critical Infrastructure Protection*, *21*, 22-32.

Rodríguez-Núñez, E., & García-Palomares, J. C. (2014). Measuring the vulnerability of public transport networks. *Journal of Transport Geography*, *35*, 50-63.

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.

Vidanapathirana, N., Wang, Y., McLain, A. C., & Self, S. (2022). Cluster Detection Capabilities of the Average Nearest Neighbor Ratio and Ripley's K Function on Areal Data: an Empirical Assessment. *arXiv preprint arXiv:2204.10882*.

Xie, Z., & Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. *Computers, environment and urban systems*, *32*(5), 396-406.

Zhang, G. (2022). Detecting and visualizing observation hot-spots in massive volunteer-contributed geographic data across spatial scales using GPU-accelerated kernel density estimation. *ISPRS International Journal of Geo-Information*, *11*(1), 55.

Zhou, Y., Fang, Z., Thill, J.-C., Li, Q., & Li, Y. (2015). Functionally critical locations in an urban transportation network: Identification and space–time analysis using taxi trajectories. *Computers, Environment and Urban Systems*, *52*, 34-47.