IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2023.0322000

Smart Cities and Access to Nature: A Framework for Evaluating Green Recreation Space Accessibility

JENNY MARTINEZ[1], and SACHIT MAHAJAN[2]
[1]Decision Sciences Group, Universidad Nacional de Colombia, Medellín, Colombia (e-mail: jrriosm@unal.edu.co)
[2]Computational Social Science (COSS), ETH Zurich, 8092 Zurich, Switzerland (e-mail: sachit.mahajan@gess.ethz.ch)

Corresponding author: Sachit Mahajan (e-mail: sachit.mahajan@gess.ethz.ch).

This work was supported through the project "CoCi: Co-Evolving City Life", which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No. 833168.

ABSTRACT As our world becomes increasingly urbanized, smart cities are leading the way in using technology to create more efficient, connected, and sustainable environments. However, amidst all the talk of connectivity and smartness, it's crucial not to lose sight of one of the most basic human needs: access to nature in cities. This research describes a novel open-source framework for investigating the availability and accessibility of green recreation spaces using open-source data and statistical analytic approaches. The framework includes a comprehensive set of tools for data extraction, processing, analysis, and visualization, thereby enabling reproducible geospatial research. We test our framework on five international cities: Medellin, Milan, Chicago, Singapore, and Mumbai. Through geospatial analysis and statistical modeling of data sourced from OpenStreetMaps, we explore and comprehend the characteristics and distribution of spatial accessibility related to green recreation spaces in five cities. We find significant clustering of green recreation spaces in all these cities, indicating that a majority of such spaces are located in close proximity to each other within small areas. Our findings also shed light on the potential implications of unequal distribution of green recreation spaces for the health and well-being of city residents and highlight the need for policies and initiatives that promote equitable access to green recreation spaces in smart cities.

INDEX TERMS Accessibility, Cities, Geospatial analysis, Recreation Spaces, Open-source

I. INTRODUCTION

Urbanization is one of the defining trends of the twenty-first century, with cities now housing more than half of the world's population [1]. While cities provide numerous advantages, such as economic opportunities and cultural diversity, they also pose significant challenges, such as traffic congestion, environmental pollution, and social inequality [2]–[6]. By harnessing technology to build more sustainable, efficient, and livable urban settings, smart cities have emerged as a possible answer to these difficulties [7]. However, the traditional emphasis on data-centric techniques and connectivity has frequently overlooked the relevance of nature and accessibility in influencing the urban quality of life. As cities continue to grow and become more densely populated, it is important to consider the role that green recreation spaces play in promoting community well-being [8], [9]. Green recreation spaces can provide a variety of benefits to city residents, including improved mental and physical health, social cohesion, and a sense of community pride [10], [11]. There is a growing body of evidence that demonstrates the positive impact of green spaces on health. For example, a recent study found that living near green space is associated with a lower risk of mortality [12], while another found that exposure to green space can lead to reductions in stress and anxiety [13]. In addition, green recreation spaces can also promote physical activity, which is an important factor in maintaining good health [14]. As we delve deeper into the development of smart cities, it becomes increasingly apparent that the significance of access to nature, such as green recreation spaces has been undervalued. The conventional understanding of what constitutes a smart city has largely centered on the deployment of cutting-edge technologies such as Artificial Intelligence (AI), the Internet of Things, etc. to optimize the efficacy and efficiency of government services [15], [16]. While this approach has yielded positive results in many parts of the world, there remains room for a paradigm shift from a data-driven to a data-informed approach. By leveraging technological advancements and available data, this approach could prioritize promoting community well-being and access to nature, instead of solely prioritizing government efficiency and productivity. As a result, there is a growing need for a more comprehensive strategy that integrates technology, community, and nature to develop sustainable, healthy, and inclusive urban environments that promote livability, sustainability, and resilience. This transition is critical for addressing urbanization-related issues such as limited access to green spaces and the detrimental impact on community well-being.

In this study, we propose an open-source[1], reproducible, and extendable framework that can be used to investigate the distribution and accessibility of green recreation spaces (as well as other amenities) in cities using open-source data. As compared to the previous works in this area that look at the accessibility of infrastructure like green spaces through the lens of socio-economic factors, we follow a more quantitative approach that looks into the spatial distribution of green recreation spaces by performing statistical analysis to understand their spatial arrangement in the cities followed by a network analysis to understand the accessibility of those spaces. We evaluate the framework on five global cities, namely Medellin, Milan, Chicago, Singapore, and Mumbai, in order to identify potential disparities in green recreation space distribution and access. Our findings suggest that despite the differences in geography, level of development, and planning strategies across these cities, green recreation spaces tend to cluster in certain areas, which impacts their accessibility to citizens. By analyzing the accessibility of green spaces, our approach offers a systematic and quantitative means for city planners to identify areas that require improvement in order to promote greater access for all members of society. In this way, our research contributes to the development of a more equitable and sustainable urban environment.

The rest of the paper is organized as follows: In Section 2, we discuss the related work. In Section 3, we describe the methods used, followed by the presentation of results in Section 4. Section 5 contains the discussion and conclusions of this study.

II. BACKGROUND

In this section, we briefly go through the related work in the area of accessibility analysis of city amenities like green recreation spaces. We also discuss the commonly used terms as well as methods that are widely used in the literature to discuss the accessibility of social infrastructure in cities as well as the use of digital technology and open-source data for decision-making in smart cities.

Accessibility is a multifaceted concept that encompasses several dimensions, including physical, and socio-economic aspects. In the context of urban planning, accessibility refers to the ease with which people can reach and use essential services and amenities, such as green recreation spaces [17]. These spaces, which include parks, gardens, and other natural environments for recreation, play an important role in promoting physical and mental well-being, fostering social cohesion, and enhancing the overall quality of life for urban residents [18], [19]. Many studies have investigated the accessibility of urban green spaces [20], [21] as well as other amenities in cities [22], employing various methods and metrics. Some researchers have used distance-based measures, such as the proximity of residential areas to the nearest green space [23], while others have used more sophisticated approaches, such as the two-step floating catchment area method, which accounts for both supply and demand factors [24]. Additionally, researchers have explored the role of socioeconomic factors in shaping accessibility patterns, revealing that lower-income and minority communities often face greater barriers to accessing green spaces [25], [26]. There has also been a significant increase in the use of open-source digital technologies as well as openly available data to map cities [27]–[29]. Several recent works have utilized data sources like OpenStreetMap (OSM) [22], [30], remote-sensing data [31], [32] as well as sensor-based data to understand how the resources and infrastructure is distributed in cities and how it has changed over the time [33].

Despite the growing body of literature on green space accessibility, its integration into the smart city discourse remains limited. Smart city initiatives often prioritize technological solutions, such as sensor networks, data analytics, and intelligent transportation systems, while overlooking the importance of equitable access to green spaces [34], [35]. The existing research gaps highlight the critical need for a comprehensive understanding of accessibility within the context of smart cities, as well as innovative techniques to enhance the availability and distribution of green recreation spaces. Furthermore, it is critical to use digital infrastructure to develop user-friendly frameworks and tools that allow decision-makers as well as other stakeholders to make informed decisions about urban planning initiatives. If done the right way, smart cities have the ability to seamlessly integrate technology, community, and nature, and to create equitable access to resources while increasing communal well-being.

III. METHODOLOGY

In this section, we will discuss in detail the methodology that is used in the work. Figure 1 provides an overview of the framework that is used for understanding the spatial distribution of green recreation spaces in a city as well as analyzing the accessibility of those recreation spaces. The framework comprises three integrated workflows. The first workflow is used to extract the street network of a location using OSM API. The OSM data is accessed using the "osmnx" package in Python [36]. This package provides easy-to-use functions for retrieving, processing, and manipulating OSM data. This workflow takes the name of the city as well as the network type to extract the street network. In this work, we used the "walk" network type to extract the street network suitable for pedestrians. The extracted street network is then converted

[1]https://github.com/sachit27/Accessibility-Analysis

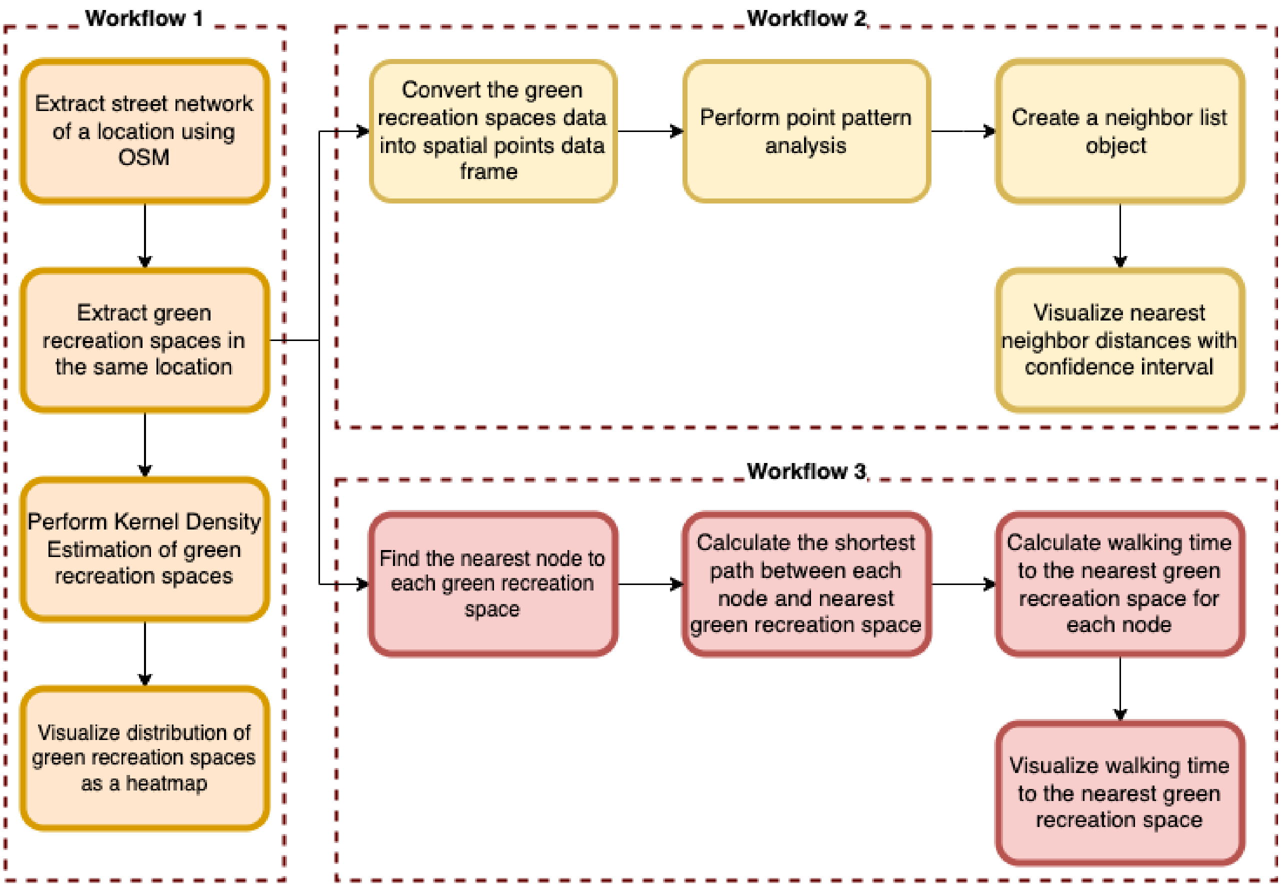

FIGURE 1. **A schematic representation of the proposed framework consisting of three workflows.**

into a GeoDataFrame that contains the nodes and edges of the street network. The second step in this workflow entails retrieving the location's green recreational spaces. We then extract the polygons of green recreation spaces in the city. We specify the tags to narrow down the green spaces with the tags "leisure" and "landuse" that are connected with parks and recreation areas.

The second workflow uses the extracted green recreation spaces data to first visualize the distribution of the green recreation spaces as a density heatmap. The density estimation is performed using a kernel density estimation (KDE) algorithm [37]. KDE calculates the density at each point as a weighted sum of nearby points. This is followed by a point pattern analysis approach to understand the spatial distribution of green recreation spaces. Point pattern analysis is a statistical method used to study the spatial arrangement of points in two-dimensional space. This approach is widely used in geospatial analysis, ecology, and epidemiology to analyze the spatial patterns of different variables [38], [39]. In our case, we are interested in understanding how green recreation spaces are distributed in a region of interest. To do this, we create a point pattern object from the coordinates of the green recreation spaces within the specified window. The window represents the extent of the region of interest. We use the K function [40] to quantify the spatial distribution of points in a point pattern. It is defined as the expected number of points within a certain distance of any point in the pattern, normalized by the density of the points. K function can be represented as:

$$K(d) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\frac{1}{A}\mathbb{I}(\|x_i - x_j\| \leq d) \tag{1}$$

where $K(d)$ is the value of the K-function at distance d, n is the total number of points in the study region, x_i and x_j are the locations of the i-th and j-th points and A is the area of the study region. The function $\mathbb{I}$ returns the value 1 if the distance between points i and j is less than or equal to d, and 0 otherwise. The envelope test is a method of hypothesis testing that is commonly used to determine if the pattern of points is significantly different from a random spatial distribution. It involves generating a large number of simulated point patterns and calculating the K function for each simulation. The results are then used to create an envelope of confidence intervals around the observed K function values. Following the point pattern analysis, we use the nearest neighbor approach [41] to quantify the degree of clustering or dispersion of points in a point pattern. This approach calculates the distances between each point (in this case, the green recreation spaces) and its nearest neighbors. By examining the distances between the points, we can determine if the distribution is random, clustered or dispersed.

To perform the nearest neighbor analysis, we first create a neighbor list object and set the distance range between 0 and 1000 meters to identify the neighboring points within this range. The formula for the nearest neighbor distance (d) is given below:

$$d = \min(dist(x_i, x_j)) \tag{2}$$

where d is the nearest neighbor distance, x_i is the location of the i^{th} point in the pattern, and x_j is the location of the nearest neighbor to the i^{th} point. The G function is used to quantify the degree of clustering or dispersion in a point pattern [42]. It is defined as the cumulative distribution function of the nearest neighbor distances. The formula for the G function is given below:

$$G(d) = P(d(x_i) \leq d) \tag{3}$$

where $G(d)$ is the value of the G function at distance d, $d(x_i)$ is the nearest neighbor distance of point i, and $P(d(x_i) \leq d)$ is the probability that the nearest neighbor distance is less than or equal to d. The expected value of the G function for a random spatial distribution is given by the following equation:

$$G(r) = 1 - \exp\left(-\frac{\pi r^2}{\lambda}\right) \tag{4}$$

where $G(r)$ is the expected value of the G function at distance r, λ is the intensity of the point pattern (i.e., the number of points per unit area), and $\exp\left(-\frac{\pi r^2}{\lambda}\right)$ is the Poisson probability of finding a point within a distance r of a randomly chosen point in a Poisson point process with intensity λ. The ratio of the observed G function to the expected G function is called the K function, which is the same as the K function used in point pattern analysis. We also use an envelope approach to create a confidence interval around the nearest neighbor histogram. The confidence interval is calculated by simulating 1000 point patterns and calculating the nearest neighbor distances for each simulated pattern.

The point pattern analysis and nearest neighbor approach are appropriate methods to understand the distribution of green recreation spaces in cities. These methods allow us to determine if these spaces are randomly distributed or if they are clustered or dispersed. This information can help urban planners and policymakers make informed decisions about the location and distribution of green spaces in the city. Additionally, this approach can be applied to other cities to compare the distribution of green spaces and identify areas that may require additional green spaces.

The final workflow deals with understanding the accessibility of green recreation spaces in cities by calculating the shortest walking distance from the nearest node to the points of interest (POI). The POI here is the location data of the green recreation space. To perform this, we first calculate the nearest node to each POI using the nearest neighbor search algorithm. We then select the nearest node that is part of the pedestrian network. This is followed by the calculation of the shortest walking distance from the nearest node to the POI using the Dijkstra algorithm. The Dijkstra algorithm finds the shortest path between nodes in a graph, in this case, the pedestrian network. This is represented mathematically as:

City	Number of green recreation spaces
Medellin	816
Milan	790
Chicago	870
Singapore	432
Mumbai	963

TABLE 1. The number of green recreation spaces in cities analyzed in this article.

$$D(i,j) = \min_{v \in V} dist(i, v) + D(v, j) \tag{5}$$

where $D(i,j)$ is the shortest path between node i and node j, v is the set of all nodes in the graph, $dist(i, v)$ is the distance between node i and node v, and $D(v, j)$ is the shortest path between node v and node j.

After calculating the shortest walking distance from each POI to the nearest node on the pedestrian network, we aggregate this information to understand the accessibility of green recreation spaces in the city. It is important to consider here that our analysis did not take into consideration the total number of green recreation spaces, but rather the proximity of the nearest green recreation space to each node. This technique is based on the convention that accessibility measures are often linked to the number of amenities at specific places, such as POI numbers or other POI features [43].

IV. RESULTS

In this section, we will discuss the results obtained after testing the proposed framework in five global cities: Medellin, Milan, Chicago, Singapore, and Mumbai. We chose these five global cities for their different geographic locations, cultural settings, and varying levels of urban growth in order to thoroughly examine the performance of our approach. Medellin, a South American city noted for its modern urban planning projects, is included to provide unique insights into how green recreation spaces are dispersed and utilized in a developing urban setting. Milan, on the other hand, is a European city known for its commitment to sustainability, serving as a model for green space distribution in a densely populated metropolitan setting. As a major North American city, Chicago provides insight into green space accessibility in a densely populated and culturally diverse metropolis. Singapore, a Southeast Asian city-state famed for its rigorous urban planning, exemplifies how a dense urban setting may provide universal access to green recreational spaces. Lastly, Mumbai, one of the most populous cities in the world, presents a unique case study on the challenges of providing equitable access to green spaces in a densely populated, rapidly growing urban

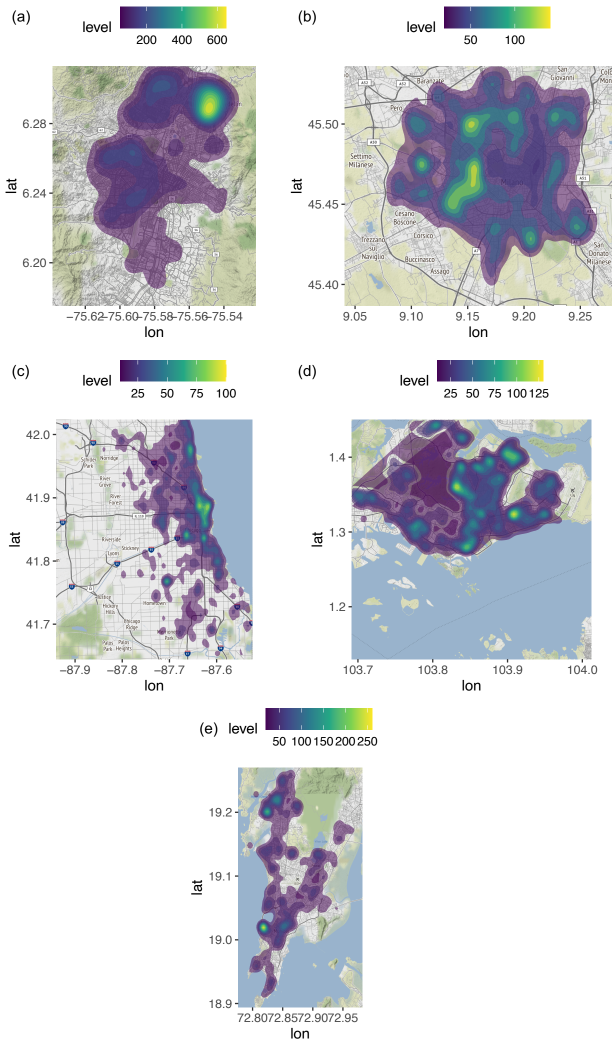

FIGURE 2. Density maps showing the distribution of green recreation spaces in (a) Medellin, (b) Milan, (c) Chicago, (d) Singapore, and (e) Mumbai. The legend "level" refers to the density level of the green recreation spaces.

area. By selecting these diverse cities, we aimed to capture a broad range of urban landscapes and examine the distribution and accessibility of green recreation spaces across different contexts. Table 1 shows the cities and the number of green recreation spaces. This is based on the data till December 2022. Figure 2 shows the density of green recreation spaces based on their spatial distribution in different cities. The regions are divided into bins and the density of green recreation spaces is calculated in each bin. The intensity or concentration of green recreation spaces in various parts of the map is represented by these density maps. The maps show kernel density surfaces, with colored values representing a condensed representation of the spatial variance in the density of green recreation spaces across the study areas. Figure 2 shows that the distribution of green recreation places is not uniform across cities. The heat maps reveal distinct patterns in different cities. Medellin and Mumbai show notable clusters of green spaces in specific areas within the city. In comparison to the other cities, Milan has a less concentrated distribution of green spaces. In Chicago, there is a large clustering of green recreation places around the city center, whereas, in Singapore, significant clustering is evident in the city's eastern part. While the heat maps provide a visual understanding of the distribution of green recreation spaces, further statistical analysis is required to comprehensively understand the distribution patterns. Therefore, our next step involves conducting a point pattern analysis.

Following up on the previous analysis, we performed a point pattern analysis to get a more detailed picture of the distribution of green recreation spaces in the selected cities. Our primary focus was on investigating distribution patterns using the nearest-neighbor approach, as shown in Figure 3. To do this, we created histograms of the actual data's distances to the nearest neighbor, as well as an envelope of expected values acquired through simulation. This involved creating 1000 random point patterns with a Poisson distribution, the intensity of which was determined by the density function produced from the original data. Then, for each bar in the histogram, we generated the 95% confidence interval and superimposed it on the original histogram. This methodology draws inspiration from the work of Bevan et. al in [41]. The primary goal here is to establish if the observed distribution of nearest-neighbor distances matches our expectations or is completely random. This could be done using the Clark and Evans test [44] as it has been used in the past, but we used a more robust alternative. We generate random point patterns through Monte Carlo simulation [45] to simulate a comparable number of random points in the study areas. These simulated point patterns are used to establish an expected distribution and create an envelope of expected values for the nearest-neighbor distances. By comparing the observed distribution of nearest-neighbor distances to the simulated distribution, we can determine whether the observed pattern deviates significantly from randomness. There are three ways to analyze the relationship between the bars and the envelope (Figure 3).

1) If the bars are constantly within the envelope, it indicates that the observed distribution of nearest neighbor distances is consistent with a random pattern. In this scenario, the green recreation spaces are dispersed in the manner that would be expected if they were randomly placed within the study area.
2) If the bars continuously surpass the upper bounds of the envelope, this suggests a clustering pattern. In this case, it suggests that the green recreation spaces tend to be closer to each other than would be expected by chance, indicating a non-random spatial organization with significant clustering.
3) If the bars continually fall below the envelope's lower bounds, this indicates a dispersed pattern. In this scenario, this indicates that the green recreation spaces are

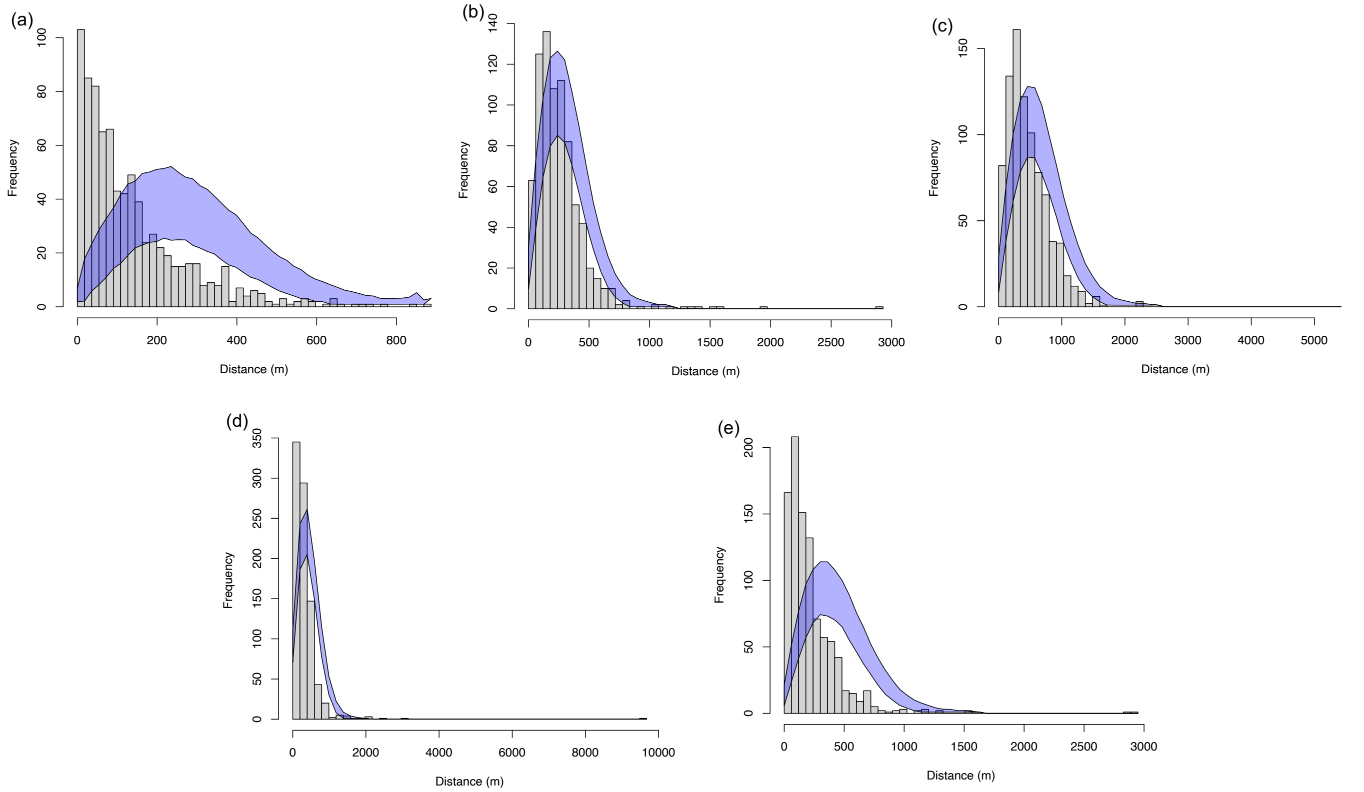

FIGURE 3. Histograms of observed nearest neighbor distances, with envelopes of expected frequencies for (a) Medellin, (b) Milan, (c) Chicago, (d) Singapore, and (e) Mumbai.

more equally spaced apart than would be expected by random chance, showing a non-random spatial organization with a dispersed distribution.

As observed in Figure 3, in the case of Medellin, variations from the expected nearest neighbor distances for a Poisson process with similar density are seen. The graph demonstrates a clustering pattern, with a larger concentration of green recreation spaces, than expected observed within 100 meters. Beyond around 200 meters, however, fewer spaces are concentrated, indicating a more dispersed distribution. For Milan, little clustering is observed within a distance of 150 meters, indicating a tendency for green recreation spaces to be located in closer proximity to each other within this range. Above this threshold, however, the observed distribution more closely resembles a random pattern, as indicated by the bars falling within the envelope. In the case of Chicago, a clustering pattern for green recreation spaces within a 300-meter distance is seen. After 500 meters, the distribution falls below the expected values, indicating a deviation from a random spatial layout. Similarly, for Singapore, clustering can be observed within approximately 400 meters, followed by a dispersed distribution of green recreation spaces beyond that point. Lastly, for Mumbai, a similar pattern emerges, with a strong clustering observed within 200 meters, followed by dispersed distribution of green recreation spaces. One takeaway from the analysis of all the cities is that a high frequency of green recreation spaces is clustered within an average distance of 200 meters (approximately). There are higher densities at shorter distances and there is a gradual decrease in density of green recreation spaces as the distance increases. There are some occasional fluctuations in counts that suggest the existence of distinct clusters or zones with varying densities.

While the nearest-neighbor analysis has provided valuable information about the distribution and clustering patterns of green recreation spaces within the study areas, understanding the proximity of these spaces alone is not enough in assessing their true accessibility. In the next step, we aimed to analyze the accessibility of green recreation spaces by calculating the shortest distance from the nearest node to the POI i.e. a green recreation space. The results of our analysis are presented in Figure 4, which shows heatmaps for each city that illustrate the walking time (in minutes) from all nodes in the pedestrian network to the nearest green recreation space. Additionally, Figure 5 presents a histogram for all five cities, showing the number of nodes, and walking time (in minutes) required to reach the nearest green recreation space. Our findings show that the average walking time to the nearest green space is 7.8 minutes for Medellin, 8.3 minutes for Milan, 6.7 minutes for Mumbai, 10 minutes for Chicago, and 14 minutes for Singapore. Despite the reported differences in walking time to the nearest green in these cities, it is critical to account for changes in node density and distribution within each

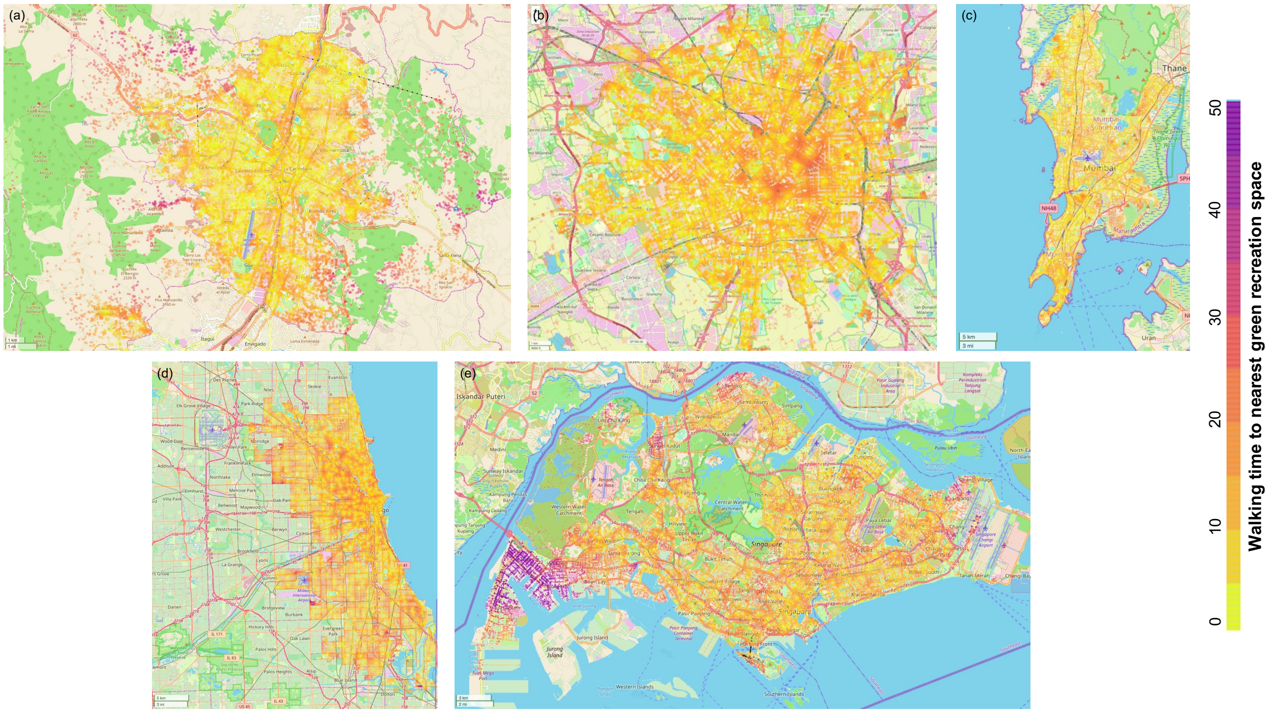

FIGURE 4. The heatmaps display the walking time required to reach the closest green recreation space for each of the five cities analyzed in this study, including (a) Medellin, (b) Milan, (c) Mumbai, (d) Chicago, and (e) Singapore.

city network. For example, Chicago, with almost 16k nodes, and Singapore, with more than 10k nodes, have a more vast and complicated network structure than Medellin, which has only 3.5k nodes, and Mumbai which has only 4k nodes. The accessibility and closeness of green recreation areas may be influenced by differences in network complexity.

V. DISCUSSION AND CONCLUSION

The notion of smart cities is still developing, and many elements are taken into account while creating and implementing smart city programs. While access to nature and green spaces is crucial in urban development, it is not necessarily the major emphasis of smart city projects. There is a growing recognition of the importance of access to nature and green areas for human health and well-being, and this should be reflected in the development of smart city programs. Access to green recreation spaces has been shown to have numerous benefits for mental and physical health. Studies have shown that spending time in nature can reduce stress and improve overall well-being [46]. In addition, green space can help to improve air and water quality, reduce urban heat island effects [47], and provide habitat for urban wildlife [48]. Furthermore, green space can serve as an important public gathering space for communities. Parks, gardens, and other green spaces can provide a place for people to come together and socialize, promoting a sense of community and improving social cohesion.

While there have been a lot of ongoing discussions about nature in the cities, potential benefits as well as frameworks to create more inclusive cities [49], [50], there is still a lot to do when it comes to access to green recreation spaces in cities. The uneven spatial distribution of green spaces in cities can have a significant impact on the city's sustainability, environment, and quality of life for city residents [51]. One of the major issues is a lack of precise data and sufficient tools for understanding the distribution and accessibility of these areas. Despite considerable advances in recent years, utilizing this data for informed decision-making remains difficult due to limited technological infrastructure and complexity. Addressing these challenges requires a coordinated effort to create and implement tools and technology that enable data-informed decision-making, not just for the decision-makers but also for city residents. Open-source tools and technology have demonstrated significant promise in facilitating collective intelligence and participatory resilience [52], [53]. These tools can aid in democratizing information availability and encouraging public engagement and feedback in decision-making processes.

In this study, we proposed an open-source and extendable framework that gives statistical insights and visualizations of the distribution and accessibility of green recreation places in cities. The proposed framework was applied to five global cities: Medellin, Milan, Chicago, Singapore, and Mumbai, to analyze the distribution of green recreation spaces and assess their pedestrian accessibility. While the cities represented different geographical and cultural settings, we found that for all of them, most of the green recreation spaces were clustered in small areas resulting in uneven distribution. In terms of accessibility, we found that the walking time to the nearest green space was highest in Chicago and Singapore.

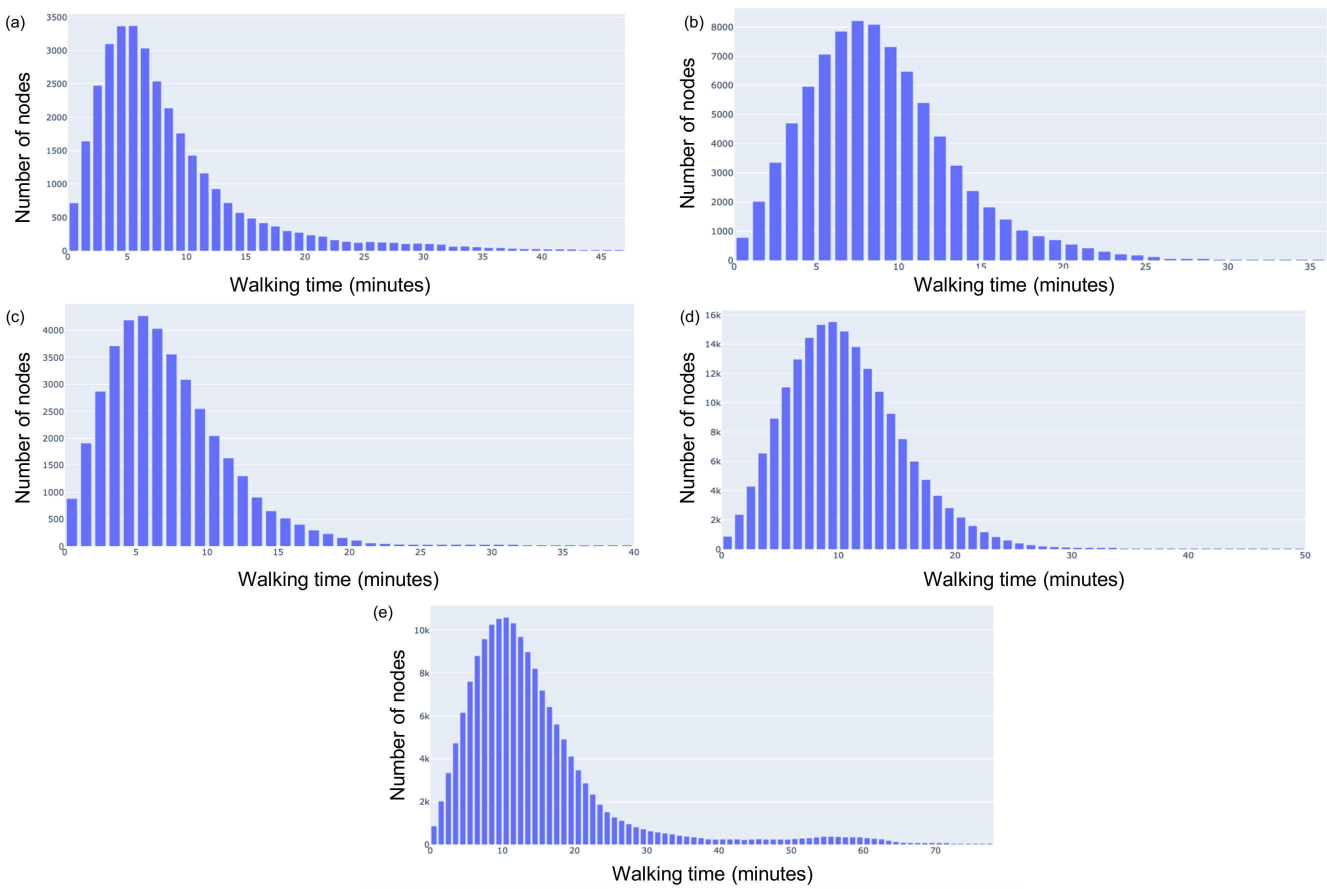

FIGURE 5. The histograms show the count of nodes against the walking time required to access the nearest green recreation spaces in: (a) Medellin, (b) Milan, (c) Mumbai, (d) Chicago, and (e) Singapore.

As the framework is open-source and based on open-source OSM data, it adds flexibility to the framework and makes it easy to test it for other amenities as well as locations. Furthermore, the framework's open-source structure allows collaboration and encourages community participation in the tool's development and enhancement. The flexibility for developers and researchers to add new functionality and features to the existing code base considerably saves the time and effort required to build equivalent tools from the ground up. Here it is also important to acknowledge the limitations of using OSM data. While OSM has shown to be a great resource for understanding urban landscapes, its reliance on volunteer contributions may result in data gaps or discrepancies, especially in areas with low levels of community engagement or technical expertise. Nonetheless, OSM remains a useful and effective data source for urban data analysis, providing a plethora of geospatial information that may be used to better understand the spatial distribution and accessibility of green recreation spaces as well as other amenities within cities.

ACKNOWLEDGEMENT

Jenny Martinez acknowledges the support from the Colombian Administrative Department of Science, Technology, and Innovation (Colciencias) National Doctoral Grant 727.

REFERENCES

[1] UN. 68% of the world population projected to live in urban areas by 2050, says un. 5 2018.
[2] Margit Mayer. The 'right to the city'in the context of shifting mottos of urban social movements. *City*, 13(2-3):362–374, 2009.
[3] Sharon L Harlan, Anthony J Brazel, Lela Prashad, William L Stefanov, and Larissa Larsen. Neighborhood microclimates and vulnerability to heat stress. *Social science & medicine*, 63(11):2847–2863, 2006.
[4] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3):1–55, 2014.
[5] Sachit Mahajan, Jennifer Gabrys, and Joanne Armitage. Airkit: a citizen-sensing toolkit for monitoring air quality. *Sensors*, 21(12):4044, 2021.
[6] Sachit Mahajan, Cyuan-Heng Luo, Dong-Yi Wu, and Ling-Jyh Chen. From do-it-yourself (diy) to do-it-together (dit): Reflections on designing a citizen-driven air quality monitoring framework in taiwan. *Sustainable Cities and Society*, 66:102628, 2021.
[7] Vito Albino, Umberto Berardi, and Rosa Maria Dangelico. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of urban technology*, 22(1):3–21, 2015.
[8] Bin Chen, Shengbiao Wu, Yimeng Song, Chris Webster, Bing Xu, and Peng Gong. Contrasting inequality in human exposure to greenspace between cities of global north and global south. *Nature Communications*, 13(1):4636, 2022.
[9] Jennifer R Wolch, Jason Byrne, and Joshua P Newell. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and urban planning*, 125:234–244, 2014.
[10] Calvin Wan, Geoffrey Qiping Shen, and Stella Choi. Underlying relationships between public urban green spaces and social cohesion: A systematic literature review. *City, Culture and Society*, 24:100383, 2021.

[11] Andrew CK Lee and Ravi Maheswaran. The health benefits of urban green spaces: a review of the evidence. *Journal of public health*, 33(2):212–222, 2011.

[12] Paul J Villeneuve, Michael Jerrett, Jason G Su, Richard T Burnett, Hong Chen, Amanda J Wheeler, and Mark S Goldberg. A cohort study relating urban green space with mortality in ontario, canada. *Environmental research*, 115:51–58, 2012.

[13] Shiyu Wan, David Rojas-Rueda, Jules Pretty, Charlotte Roscoe, Peter James, and John S. Ji. Greenspace and mortality in the u.k. biobank: Longitudinal cohort analysis of socio-economic, environmental, and biomarker pathways. *SSM - Population Health*, 19, 2022.

[14] Feng Li, Rusong Wang, Juergen Paulussen, and Xusheng Liu. Comprehensive concept planning of urban greening based on ecological principles: A case study in beijing, china. *Landscape and Urban Planning*, 72, 2005.

[15] NTU. Smarter cities for a better future. August 2022.

[16] Feng Li, Alberto Nucciarelli, Sinead Roden, and Gary Graham. How smart cities transform operations models: A new research agenda for operations management in the digital economy. *Production Planning and Control*, 27, 2016.

[17] Emily Talen and Luc Anselin. Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. *Environment and planning A*, 30(4):595–613, 1998.

[18] Jolanda Maas, Robert A Verheij, Peter P Groenewegen, Sjerp De Vries, and Peter Spreeuwenberg. Green space, urbanity, and health: how strong is the relation? *Journal of epidemiology & community health*, 60(7):587–592, 2006.

[19] Anna Chiesura. The role of urban parks for the sustainable city. *Landscape and urban planning*, 68(1):129–138, 2004.

[20] Dong Liu, Mei-Po Kwan, and Zihan Kan. Analysis of urban green space accessibility and distribution inequity in the city of chicago. *Urban Forestry & Urban Greening*, 59:127029, 2021.

[21] Jun Yang, Chuang Li, Yuechen Li, Jianhao Xi, Quansheng Ge, and Xueming Li. Urban green space, uneven development and accessibility: a case of dalian's xigang district. *Chinese Geographical Science*, 25:644–656, 2015.

[22] Leonardo Nicoletti, Mikhail Sirenko, and Trivik Verma. Disadvantaged communities have lower access to urban infrastructure. *Environment and Planning B: Urban Analytics and City Science*, 50(3):831–849, 2023.

[23] Heather E Wright Wendel, Rebecca K Zarger, and James R Mihelcic. Accessibility and usability: Green space preferences, perceptions, and barriers in a rapidly urbanizing city in latin america. *Landscape and urban planning*, 107(3):272–282, 2012.

[24] Wei Luo and Fahui Wang. Measures of spatial accessibility to health care in a gis environment: synthesis and a case study in the chicago region. *Environment and planning B: planning and design*, 30(6):865–884, 2003.

[25] Jennifer Wolch, John P Wilson, and Jed Fehrenbach. Parks and park funding in los angeles: An equity-mapping analysis. *Urban geography*, 26(1):4–35, 2005.

[26] Alessandro Rigolon. A complex landscape of inequity in access to urban parks: A literature review. *Landscape and Urban Planning*, 153:160–169, 2016.

[27] Ling-Jyh Chen, Yao-Hua Ho, Hu-Cheng Lee, Hsuan-Cho Wu, Hao-Min Liu, Hsin-Hung Hsieh, Yu-Te Huang, and Shih-Chun Candice Lung. An open framework for participatory pm2. 5 monitoring in smart cities. *Ieee Access*, 5:14441–14454, 2017.

[28] P Hamel, AD Guerry, S Polasky, B Han, JA Douglass, M Hamann, B Janke, JJ Kuiper, Harold Levrel, H Liu, et al. Mapping the benefits of nature in cities with the invest software. *npj Urban Sustainability*, 1(1):25, 2021.

[29] Sachit Mahajan. Design and development of an open-source framework for citizen-centric environmental monitoring and data analysis. *Scientific Reports*, 12(1):14416, 2022.

[30] Timur Dogan, Yang Yang, Samitha Samaranayake, and Nikhil Saraf. Urbano: A tool to promote active mobility modeling and amenity analysis in urban design. *Technology| Architecture+ Design*, 4(1):92–105, 2020.

[31] Stephen Law, Brooks Paige, and Chris Russell. Take a look around: using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(5):1–19, 2019.

[32] Sachit Mahajan and Jenny Martinez. Water, water, but not everywhere: Analysis of shrinking water bodies using open access satellite data. *International Journal of Sustainable Development & World Ecology*, 28(4):326–338, 2021.

[33] Yu Liu, Fahui Wang, Yu Xiao, and Song Gao. Urban land uses and traffic 'source-sink areas': Evidence from gps-enabled taxi data in shanghai. *Landscape and Urban Planning*, 106(1):73–87, 2012.

[34] Rob Kitchin. The real-time city? big data and smart urbanism. *GeoJournal*, 79:1–14, 2014.

[35] Margarita Angelidou. The role of smart city characteristics in the plans of fifteen cities. *Journal of Urban Technology*, 24(4):3–28, 2017.

[36] Geoff Boeing. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

[37] Christopher C Weiss, Marnie Purciel, Michael Bader, James W Quinn, Gina Lovasi, Kathryn M Neckerman, and Andrew G Rundle. Reconsidering access: park facilities and neighborhood disamenities in new york city. *Journal of urban health*, 88:297–310, 2011.

[38] Anthony C Gatrell, Trevor C Bailey, Peter J Diggle, and Barry S Rowlingson. Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British geographers*, pages 256–274, 1996.

[39] Thorsten Wiegand and Kirk A Moloney. *Handbook of spatial point-pattern analysis in ecology*. CRC press, 2013.

[40] Peter Haase. Spatial pattern analysis in ecology based on ripley's k-function: Introduction and methods of edge correction. *Journal of vegetation science*, 6(4):575–582, 1995.

[41] Andrew Bevan, Eva Jobbová, Christophe Helmke, and Jaime J Awe. Directional layouts in central lowland maya settlement. *Journal of Archaeological Science*, 40(5):2373–2383, 2013.

[42] George LW Perry, Ben P Miller, and Neal J Enright. A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant ecology*, 187:59–82, 2006.

[43] Walter G Hansen. How accessibility shapes land use. *Journal of the American Institute of planners*, 25(2):73–76, 1959.

[44] Brian D Ripley. Tests of'randomness' for spatial point patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(3):368–374, 1979.

[45] Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

[46] Jenny J Roe, Catharine Ward Thompson, Peter A Aspinall, Mark J Brewer, Elizabeth I Duff, David Miller, Richard Mitchell, and Angela Clow. Green space and stress: evidence from cortisol measures in deprived urban communities. *International journal of environmental research and public health*, 10(9):4086–4103, 2013.

[47] Keith Van Ryswyk, Natasha Prince, Mona Ahmed, Erika Brisson, J David Miller, and Paul J Villeneuve. Does urban vegetation reduce temperature and air pollution concentrations? findings from an environmental monitoring study of the central experimental farm in ottawa, canada. *Atmospheric Environment*, 218:116886, 2019.

[48] Matthew B Haverland and Joseph A Veech. Examining the occurrence of mammal species in natural areas within a rapidly urbanizing region of texas, usa. *Landscape and Urban Planning*, 157:221–230, 2017.

[49] Qi Liu, Hidayat Ullah, Wanggen Wan, Zhangyou Peng, Li Hou, Tong Qu, and Saqib Ali Haidery. Analysis of green spaces by utilizing big data to support smart cities and environment: a case study about the city center of shanghai. *ISPRS International Journal of Geo-Information*, 9(6):360, 2020.

[50] Daniela D'Alessandro, Maddalena Buffoli, Lorenzo Capasso, GM Fara, Andrea Rebecchi, Stefano Capolongo, et al. Green areas and public health: improving wellbeing and physical activity in the urban context. *Epidemiologia e prevenzione*, 39(5):8–13, 2015.

[51] Margarita Triguero-Mas, Payam Dadvand, Marta Cirach, David Martínez, Antonia Medina, Anna Mompart, Xavier Basagaña, Regina Gražulevičienė, and Mark J Nieuwenhuijsen. Natural outdoor environments and mental and physical health: relationships and mechanisms. *Environment international*, 77:35–41, 2015.

[52] Thomas W Malone and Mark Klein. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization*, 2(3):15–26, 2007.

[53] Sachit Mahajan, Carina I Hausladen, Javier Argota Sánchez-Vaquerizo, Marcin Korecki, and Dirk Helbing. Participatory resilience: Surviving, recovering and improving together. *Sustainable Cities and Society*, page 103942, 2022.

JENNY MARTINEZ received an M.E. degree in systems engineering from the National University of Colombia in 2013. She is currently a Ph.D. candidate at the National University of Colombia. Her current research interests include data science applied to the environment and urban areas.

SACHIT MAHAJAN is a postdoctoral researcher at the Professorship of Computational Social Science at ETH Zurich. He has PhD in Social Networks and Human-Centered Computing, and his research interests lie at the intersection of Ubiquitous Systems, Data Science, and Citizen Science.