Ecological Indicators 163 (2024) 112108

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

greenR: An open-source framework for quantifying urban greenness

Sachit Mahajan

Computational Social Science, ETH Zurich, Stampfenbachstrasse 48, Zurich 8092, Switzerland

ARTICLE INFO

Keywords:
Urban greenness
Open source
Urban analytics
Cities
Street network

ABSTRACT

How do we quantify the levels of greenness within urban street networks? Numerous attempts to quantify this factor have been made through survey methodologies, remote sensing data, and street view imagery. The results are promising, but are often limited by scalability constraints, including limited data availability, high requirements of computational power, and validation challenges. This study introduces a comprehensive framework for urban greenness assessment, leveraging open-source data to overcome these limitations. Central to this framework is the development of the greenR R package and an accompanying Shiny app, designed to compute, visualize, and analyze a novel proximity-based green index for individual street segments. Beyond this, the framework innovatively extends its analytical capability by integrating accessibility analysis, Green View Index quantification as well as introducing the Green Space Similarity Index, a novel metric that evaluates and compares the characteristics of urban green spaces across different regions. This extension enriches the proposed framework, providing not just a measure of proximity to green spaces, but also insights into their spatial connectivity and distribution. The efficacy of greenR is demonstrated by studying urban greenness patterns across several cities, highlighting the potential impact of such an open-source framework for citizens, urban planners, and policy-makers. This study not only advances our methodological approach to quantifying urban greenness but also provides practical tools and metrics that can inform sustainable urban planning and policy decisions.

1. Introduction

Quantifying and assessing urban greenness is a complicated scientific challenge. As cities continue to expand, the conversion of natural landscapes into “concrete jungles” has led to a decline in greenery in cities, negatively impacting the quality of urban environments (Zhou and Wang, 2011; Fuller and Gaston, 2009; Cui et al., 2020). The scarcity of green spaces not only diminishes the aesthetic appeal of cities, but also poses significant environmental, social, and health-related concerns. From an environmental perspective, green spaces serve a pivotal role in counteracting the detrimental effects of urbanization (Chen et al., 2022; Wolch et al., 2014; Martinez and Mahajan, 2023). They function as carbon sinks, absorbing atmospheric pollutants and curtailing greenhouse gas emissions (Strohbach et al., 2012). Furthermore, they contribute to local climate regulation by providing shade, attenuating the urban heat island effect, and fostering air circulation (Arghavani et al., 2020). Green spaces also help with rainwater filtration, alleviating the strain on urban drainage systems and reducing flooding risk. On a social level, green spaces have been associated with enhanced physical and mental health (Wan et al., 2022). They offer venues for social interaction, community engagement, and foster a sense of belonging between inhabitants. As hubs for recreational activities, cultural events, and leisure, they enrich the social fabric of communities (Pincetl and Gearin, 2005; Martinez and Mahajan, 2023). At the street level, the extent of urban greenness can have significant implications. Furthermore, visible greenery can encourage outdoor recreational activities, thereby promoting physical well-being (Zhang and Dong, 2018; Villeneuve et al., 2018). Furthermore, urban greenness has been linked to a reduction in negative emotions (Helbich et al., 2019).

Urban greenness can be conceptualized in various ways. Traditional techniques of conceptualizing urban greenness have mostly focused on monitoring and quantifying green areas (Yang et al., 2009; Cadenasso et al., 2007). Other approaches to understanding urban greenness, on the other hand, take into consideration elements such as tree canopy cover (Ordóñez et al., 2023), vegetation density, and green infrastructure (Kumar et al., 2019). While these methods provide valuable insights into the extent and distribution of urban green spaces, they often overlook the critical aspects of proximity and accessibility, which significantly affect the usability and benefits of these green spaces to urban residents. Recognizing this gap, this study proposed a new “green index” defined as a quantified measure of the presence and influence of green spaces and trees along urban street segments. This metric is

E-mail address: sachit.mahajan@gess.ethz.ch.

https://doi.org/10.1016/j.ecolind.2024.112108
Received 26 November 2023; Received in revised form 25 March 2024; Accepted 2 May 2024
Available online 13 May 2024
1470-160X/

derived from a composite assessment of the distribution, and accessibility of green areas. It is not merely a measure of green area coverage but rather a composite assessment considering both the distribution and proximity-based influence of greenery relative to the street network.

With technological advancements and easy data access, there has been a steady growth in research initiatives that have attempted to quantify urban greenness. Let us discuss these approaches.

Traditional Approaches to Green Space Assessment: The challenge of quantifying and evaluating urban greenness has been addressed through several methodologies. Conventional techniques often rely on field surveys (Kothencz and Blaschke, 2017; Leslie et al., 2010) and the visual interpretation of high-resolution satellite imagery (Skokanová et al., 2020; Huerta et al., 2021). Although these methods yield valuable data, they are obstructed by several constraints that impede their scalability, cost-effectiveness, and ability to generate comprehensive green index maps.

Field surveys, a long-standing method for gathering data on vegetation cover and composition within urban areas, are labor-intensive and time-consuming (Gupta et al., 2012). These surveys also suffer from limitations in spatial coverage as well as self-reporting bias (Lu et al., 2019), making it difficult to obtain an accurate and complete picture of green spaces across an entire city (Tzoulas et al., 2007). Furthermore, field surveys are expensive and can pose logistic challenges, especially when dealing with large urban areas.

High-resolution satellite imagery has proven to be a valuable tool for mapping and monitoring urban areas (Nouri et al., 2020; Ludwig et al., 2021; Senanayake et al., 2013; Mahajan and Martinez, 2021). However, the acquisition and processing costs associated with these data also limit widespread use, particularly in resource-constrained settings (Grün, 2000). Moreover, the temporal resolution of public satellite imagery is often limited (Wang et al., 2021), making it challenging to capture dynamic changes in green spaces over time. The coarse spatial resolution of some satellite imagery also hampers the accurate identification and delineation of small-scale green features within urban environments (Rioux et al., 2019; Gr et-Regamey et al., 2014).

Emerging Approaches: The limitations of traditional methods have spurred researchers to investigate alternative data sources and methodologies. For instance, several works have proposed the use of remote sensing data, such as airborne LiDAR (Light Detection and Ranging), to capture detailed information on urban vegetation structure and density (Zhang and Shao, 2021; Pyszny et al., 2020). While LiDAR data can provide three-dimensional information about vegetation height and canopy structure, enabling more accurate quantification of green spaces, its high cost and limited availability restrict its widespread use in large-scale urban assessments.

In recent years, Google Street View and Mapillary have also emerged as valuable tools for urban greenness assessment (Li et al., 2015; Lu, 2018; Biljecki and Ito, 2021; Li, 2021). These platforms provide street-level imagery, which can be used to calculate the Green View Index (GVI), a measure of the proportion of visible greenery in urban environments (Zhang and Hu, 2022; Toikka et al., 2020; Cai et al., 2018). While these methods offer a different way of capturing urban greenness, they present challenges as well. The processing of large volumes of street-level images is computationally intensive and requires sophisticated image analysis techniques. Additionally, the coverage of Google Street View and Mapillary is not uniform globally, with lower coverage in developing countries and rural areas (Biljecki and Ito, 2021). Furthermore, the data is not publicly accessible and, therefore, lacking independent quality assessment and improvement.

To overcome the limitations of traditional data sources and the challenges associated with street-level imagery, some studies have started to leverage crowdsourced data and volunteered geographic information (VGI) platforms (Cui et al., 2021; McCall et al., 2015). For example, OpenStreetMap (OSM) (Haklay and Weber, 2008), a prominent VGI platform, allows citizen mappers to contribute spatial data, including information on urban features. OSM has been utilized to map various aspects of urban environments (Yap et al., 2023; Mahajan et al., 2019; Texier et al., 2018), including green spaces (Ye et al., 2019; Martinez and Mahajan, 2023; Liao et al., 2021). So far, however, the use of OSM data for calculating a comprehensive green index is underexplored. There are very limited studies focusing on its integration and validation for green space mapping in urban areas (Ludwig et al., 2021).

Given this context, the greenR framework emerges as an innovative solution, addressing the gaps in current methodologies. Table 1 compares widely used urban greenness indices, including the Normalized Difference Vegetation Index (NDVI) (Gascon et al., 2016), Tree Canopy Cover (TCC) (Ordóñez et al., 2023), Impervious Surface Cover (ISC) (Lane et al., 2017), Urban Neighborhood Green Index (UNGI) (Gupta et al., 2012) and GVI, with the contributions of greenR, highlighting its unique approach in integrating proximity, accessibility, and scalability. This comparison underscores the innovative aspects of greenR in addressing the critical dimensions of urban greenness assessment that are often overlooked in other indices. By focusing on a comprehensive, and accessible approach, greenR not only complements existing indices but significantly advances the field by offering a more nuanced understanding of the spatial distribution and accessibility of urban green spaces.

The primary aim of this research is to overcome the limitations of existing methodologies for quantifying urban greenness as well as create an integrated framework that provides tools to perform a thorough and detailed analysis of urban greenness. This objective is realized through the introduction of a novel, open-source framework that harnesses OSM data to estimate the green index for each edge within a city's network. This approach offers a detailed and scalable view of urban greenness, incorporating not only the quantification of green spaces but also their spatial distribution and accessibility.

The contributions of this work are twofold:

1. *Methodological Contribution*: This research introduces a novel approach for evaluating urban greenness at a granular level, leveraging the collective intelligence of citizen mappers. Through the application of OSM data, this framework addresses the scalability challenges, often observed in conventional methods, including limited data availability and substantial computational power needs. Testing and validation of the proposed methodology has been conducted with reference datasets to ensure the reliability and precision of the generated green indices. The introduction of a new metric for analyzing the inter and intra-city spatial distribution and connectivity of green spaces marks a significant advancement, providing nuanced insights into urban ecological landscapes.
2. *Technological and Social Contribution*: The development of greenR, an R package, facilitates the quantification, analysis, and visualization

Table 1
Comparison of Urban Greenness Indices with greenR Contributions.

Index	Data Source	Limitations	greenR Contribution
NDVI	Remote Sensing	Sensitive to soil moisture, shadows, doesn't measure accessibility	Complements NDVI by integrating accessibility and proximity-based analysis
TCC	Aerial, Remote Sensing	Ignores under-canopy, non-tree vegetation	Includes all vegetation types
ISC	Remote Sensing	Indirect greenness, doesn't capture all green spaces	Direct quantification with context, considers all green spaces
UNGI	GIS, Remote Sensing	Limited granularity due to grid-based approach	Enhances granularity by quantifying green index at street-level
GVI	Street-level Imagery	Limited coverage	Complements with scalable OSM-based assessments, enriching pedestrian perspective

of green indices. The open-source nature of this framework increases its accessibility to a diverse audience, such as researchers, urban planners, and policy-makers. The framework extends its utility with tools for accessibility analysis and the segmentation-based GVI, offering efficient ways to assess and visualize urban greenery. This tool enhances the understanding of greenness distribution and characteristics within cities, also providing a platform for additional research and application in urban studies.

The paper is structured as follows: The greenR methodology and the accompanying R package are detailed in Section 2. Section 3 outlines the results of the greenR framework. Section 4 provides a comprehensive discussion, and the paper concludes with Section 5.

2. Methods

This study adopts a network-based approach, leveraging OSM data to measure the "green index" of all the edges within the city's network. This index, ranging from 0 to 1, represents the relative greenness of each section, factoring in proximity to green spaces and tree density. The comprehensive research framework includes steps for data acquisition, processing, analysis and visualization, all of which are implemented in R. The emphasis of this methodology is on reproducibility and open-source innovation, fostering a collaborative and transparent research environment. This approach serves not only to enhance the validity and reliability of the data. It also serves to create a conducive environment, where both experts and non-experts can meaningfully contribute (Mahajan et al., 2022), a key aspect for addressing our societal challenges.

The greenR [1] framework streamlines this process, utilizing the greenR R package to encapsulate essential features. A principal challenge in developing greenR was the integration of its diverse functionalities into a cohesive toolkit. Given the multifaceted nature of urban greenness—encompassing aspects such as proximity, density, and accessibility—creating a suite of tools that not only addresses these dimensions comprehensively but also operates seamlessly was a complex undertaking. This challenge was met by leveraging a modular design, where each component of the framework was developed to interact harmoniously, ensuring that the toolkit could offer a holistic analysis of urban greenness while remaining user-friendly and accessible to a broad audience.

The greenR package provides a simplified and user-friendly approach for conducting green index analysis, facilitating a seamless process from data acquisition to analysis and visualization. Its integrated suite of functions makes it easier for researchers as well as for other interested stakeholders to quantify and evaluate urban greenness without the need for extensive coding. Bundling these functions within an R package has several advantages. First, it enhances reproducibility by ensuring consistent application of the methodology. Second, the package simplifies the green index calculation process, by abstracting away code that users do not need to deal with. This makes it accessible to a wider audience. The package is offered as an open-source resource to encourage collaboration, sharing of knowledge, and further development within the research community.

The following paragraphs will explore the design and features of the greenR package, along with a detailed overview of its primary functions.

2.1. Data acquisition

A key part of the methodology is the `get_osm_data()` function, which plays an important role in acquiring geospatial data from OSM. This function is designed to systematically gather essential data, setting the stage for all subsequent analysis stages. Utilizing a bounding box (bbox) to define the geographical boundaries, it is adaptable to various locations, making it versatile for diverse urban studies.

It downloads OSM data for the specified spatial area with regard to three key environmental features: highways, green areas, and trees. "Highways" refer to all roads and paths within the defined bbox, whereas "Green Areas" refers to a variety of land-use categories such as forest, vineyard, plant nursery, orchard, greenfield, recreation ground, allotments, meadow, village green, flowerbed, grass, farmland, garden, dog park, nature reserve, and park. These areas can be visualized using the `visualize_green_spaces()` function (Fig. 1(a)). The final category, "Trees" contains all individual tree entities listed in the OSM database inside the bbox. The function returns a list containing the data for these three features. The downloaded data then undergoes a transformation process, forming a graph network representation, $G(V, E)$, where V represents the set of nodes, and E signifies the set of edges. In this city network, nodes (V) symbolize intersections, and edges (E) denote the roads connecting these nodes.

2.2. Clustering urban green spaces

The `green_space_clustering()` function employs K-means clustering to categorize urban green spaces based on their respective area size. It produces a precise depiction of the distribution of green spaces (Fig. 1(b)), which is critical for urban planning and environmental studies. By inputting the desired number of clusters, users can segment green spaces into distinct groups, reflecting the variety in size and potentially highlighting spatial disparities in green space distribution.

K-means clustering (Hartigan and Wong, 1979) is an iterative algorithm that minimizes the within-cluster sum of squares (WCSS), the total squared distance between each point and the centroid of its cluster. The mathematical formulation for this algorithm is given by:

$$\text{WCSS} = \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 \tag{1}$$

where x represents an observation in d-dimensional space, S_i denotes the set of observations in the i^{th} cluster, and μ_i is the centroid of S_i.

In the context of `green_space_clustering()`, the function initially transforms the green spaces into an equal-area projection to calculate the areas accurately. Post-transformation, the K-means algorithm is applied to these areas, clustering the green spaces into user-defined categories. The function then visualizes these clusters on a Leaflet map with layer control for different basemap tiles, enhancing the interpretability of urban green space distribution. This clustering approach serves as an analytical tool, providing insights into the ecological structure of urban landscapes and informing sustainable urban development strategies.

2.3. Green space similarity index

The Green Space Similarity Index (GSSI) is a proposed as a new composite metric to evaluate and compare the characteristics of urban green spaces across different regions. This index intricately blends the quantification of green space areas with a measure of their spatial connectivity, thereby providing a comprehensive tool for urban green space analysis.

The methodology involves transforming spatial data into an equal-area projection, ensuring accurate and consistent area measurements. For each region i, the total area of green spaces, denoted as A_i, is computed:

$$A_i = \sum_{k=1}^{n_i} \text{area}(G_{ik}) \tag{2}$$

Here, G_{ik} represents the k-th green space in region i, and n_i is the total

[1] https://github.com/sachit27/greenR

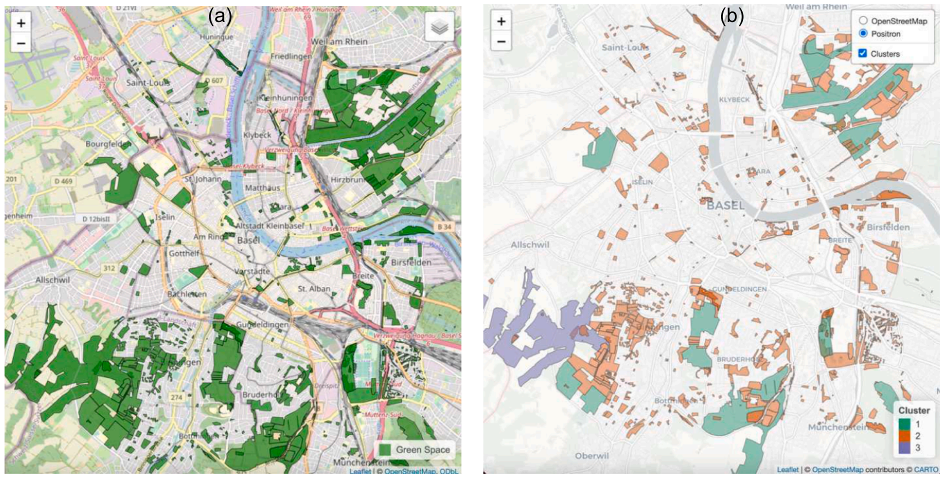

Fig. 1. Comparative Visualization of Urban Green Spaces. (a) a spatial analysis of distribution. (b) clustering based on area size.

number of green spaces within that region.

Beyond mere area calculations, the GSSI also incorporates spatial connectivity. The Average Nearest Neighbor Distance (ANND), D_i, is calculated using spatial statistics techniques. This distance represents the average shortest distance to the nearest green space for each green space within the same region, thus capturing the aspect of green space distribution and accessibility:

$$D_i = \frac{1}{n_i}\sum_{k=1}^{n_i}\min_{l\neq k}(d(G_{ik}, G_{il})) \tag{3}$$

The GSSI for a region i, denoted as GSSI_i, is then calculated by combining area-based metrics with the proximity-based connectivity measures. The function normalizes these values against the highest scoring city in the dataset, allowing for relative comparisons across different urban settings:

$$\text{GSSI}_i = \text{Normalized Value}\left(\frac{1}{\text{Variation in } A_i}\cdot\frac{1}{D_i}\right) \tag{4}$$

The function `gssi()` facilitates the calculation of GSSI within the greenR framework. While the GSSI offers valuable insights into urban green space characteristics, it inherently focuses on the size and spatial distribution of these spaces. The index can further be integrated with other qualitative aspects such as ecological value, type of green space, or user accessibility. Such expansions of the GSSI could provide a more holistic view of urban green infrastructure.

2.4. Accessibility analysis

The function `accessibility_greenspace()` facilitates the visualization of green spaces that are accessible within a predetermined walking duration from a given point of origin (Fig. 2). The function accepts geographical coordinates and a maximum walking time as inputs to generate a map that graphically delineates green areas that are within the specified walking time. The core mechanism of the function is the generation of isochrones, which are contours linking points that share an equal travel time from a central location (O'Sullivan et al., 2000). The isochrone generation leverages the Open Source Routing Machine (OSRM), which computes the most efficient paths through a network, considering the walking pace and the spatial configuration of the urban landscape (Luxen and Vetter, 2011).

The mathematical underpinnings of isochrones are rooted in graph theory. The urban structure is modeled as a graph $G = (V, E)$, where V denotes the vertices representing intersections and E signifies the edges, reflecting the walkways or streets connecting these vertices. Given a

Fig. 2. Isochrone map depicting the accessibility of green spaces from a specified location within a 10-min walking radius in Zurich. The nested colored contours represent incremental time intervals from the center.

vertex v_0 symbolizing the user's starting location and a time budget T, the isochrone for v_0 at time T is defined by the subgraph $G_T = (V_T, E_T)$, where:

- $V_T \subseteq V$ includes all vertices reachable from v_0 with a path $P = (v_0, v_1, \ldots, v_k)$ such that the total traversal time $\sum_{(v_i, v_{i+1})\in P} w(v_i, v_{i+1}) \leqslant T$.
- $E_T \subseteq E$ comprises all edges that connect vertices in V_T.

The function $f : V \rightarrow \mathbb{R}$ calculates the minimum time required to reach any vertex v from v_0, which can be found using algorithms like Dijkstra's. The isochrone is then the level set of f for the value T:

$$Iso_T(v_0) = \{v \in V | f(v) \leqslant T\} \tag{5}$$

This isochrone $Iso_T(v_0)$ demarcates the boundary within which all points are reachable from v_0 within the time T. The geometric shape of the isochrone is contingent upon the graph's layout and the distribution of weights across the edges, which, in this scenario, are determined by pedestrian speeds and road conditions.

2.5. Green Index Calculation

Algorithm 1.

Green Index Calculation

Result: Green Index for each road segment
Input: Road network, Green areas, Trees
Output: Green Index for each road segment
Method:
Step 1: Download data for highways, green areas, and trees from OpenStreetMap
Step 2: Transform all data to the same coordinate reference system
Step 3: Define distance decay functions for green areas and trees
Distance Decay Function for Green Areas:
foreach *edge in the road network* **do**
  foreach *green area* **do**
    Calculate the distance from the edge to the green area
    Apply the decay function: $\exp\left(-\frac{\text{distance}}{D}\right)$, where D is a constant representing a typical decay range
    Limit the decay function value between 0 and 1
  end
  Sum all decay function values for the edge
end
Distance Decay Function for Trees:
foreach *edge in the road network* **do**
  foreach *tree* **do**
    Calculate the distance from the edge to the tree
  end
  Find the minimum distance to any tree
  Apply the decay function: $\exp\left(-\frac{\text{min distance}}{D}\right)$, where D is a constant representing a typical decay range
  Limit the decay function value between 0 and 1
end
Step 4: Initialize parallel processing
foreach *edge in the road network* **do**
  Step 5: Calculate the green index for green areas by applying the distance decay function to each green area and summing the results
  Step 6: Calculate the green index for trees by applying the distance decay function to the trees
  Step 7: Calculate the final green index as the average of the green area and tree indices, normalize it to a range of [0,1]
end

The `calculate_green_index()` function, is the core of the green index calculation. It takes as input the OSM data, a Coordinate Reference System (CRS) code, and parameter `D` for the distance decay functions. The algorithm extracts the highways, green areas, and trees data from the input list and transforms the data into the given CRS. It then defines distance decay functions for green areas and trees. For each edge in the highways data, the function calculates the green index using the decay functions and returns a data frame with the green index for each edge. The green index calculation is performed as per Algorithm 1.

2.5.1. Distance decay functions for green areas and trees

Distance decay is a geographic principle that illustrates the impact of increasing distance on spatial interactions. This phenomenon posits that the relationship between two places diminishes as the physical distance between them grows (Chen, 2015; O'Leary, 2011). Research has shown that the benefits of environmental features, such as green areas decrease with an increase in the distance (Hogendorf et al., 2020). When studying distance decay, researchers often rely on different decay functions, including the Gaussian, power, logistic, and exponential functions (Du et al., 2018). Each function has unique characteristics that make it suitable for specific situations. In this study, it was determined that the exponential decay function was the best option for capturing the impact of green spaces and trees on road segments because of its ability to accurately reflect the diminishing influence of greenery with increasing distance. The exponential decay function, characterized by its rapid decrease at shorter distances, effectively mirrors the real-world scenario where the perceived benefits of green spaces, such as aesthetic value and air quality improvement, decrease sharply as one moves away from them. This function has been widely used in ecosystem related research (Zluwa and Pitha, 2021; Rui et al., 2018; Kozak et al., 2011) and is particularly well-suited for urban environments where space is limited and the effects of greenery are most pronounced within a small radius.

Consider each edge $e \in E$ of the road network. For green areas, the distance d_{ga} to each green area $g \in G$ is computed, followed by an application of the decay function:

$$F_G(d_{ga}) = \exp(-d_{ga}/D) \tag{6}$$

where D controls the rate of decay. The exponential function ensures a fast decay of the influence with increasing distance. The value of this decay function is bounded between 0 and 1, i.e., $0 \leqslant F_G(d_{ga}) \leqslant 1$. The sum of all decay function values for edge e is computed as follows:

$$D_e^{green} = \sum_{g \in G} F_G(d_{ga}) \tag{7}$$

Similarly, for trees, the distance d_{te} to each tree $t \in T$ is computed and the minimum distance d_{te}^{min} to any tree is calculated. The decay function for trees is then applied to this minimum distance:

$$F_T(d_{te}^{min}) = \exp\left(-\left(d_{te}^{min}\big/D\right)\right) \tag{8}$$

This decay function is also bounded between 0 and 1, i.e., $0 \leqslant F_T(d_{te}^{min}) \leqslant 1$.

2.5.2. Final green index calculation

The final green index $I_{G,e}$ for each edge e is calculated as the average of the green area and tree indices. Subsequently, it is normalized to a value between 0 and 1 by subtracting the minimum green index value and dividing by the range of green index values:

$$I_{G,e} = \frac{I_{G,e}^{green} + I_{G,e}^{tree}}{2} \tag{9}$$

$$I_{G,e} = \frac{I_{G,e} - \min(I_{G,e})}{\max(I_{G,e}) - \min(I_{G,e})} \tag{10}$$

Herein, $I_{G,e}^{green}$ is the sum of the decay function values for green areas and $I_{G,e}^{tree}$ is the decay function value for trees. This normalization transforms the green index into a relative measure that falls between 0 and 1, making it comparable across different edges in the road network. An index of 0 implies a low degree of greenness, while an index of 1 implies a high degree of greenness.

This mathematical formulation ensures that the green index captures the influence of green areas and trees, with a higher weight given to features closer to each edge of the road network. The use of the distance decay functions ensures that the influence of green areas and trees decreases with increasing distance, reflecting the diminishing visual and environmental impact of green features with distance.

2.5.3. Green index analysis and visualization

The three key functions for green index analysis and visualization are `calculate_percentage()`, `plot_green_index()`, and `save_as_leaflet()`.

The function `calculate_percentage()` is focused on the analysis of the underlying data. It groups the edges by their respective green index and calculates the percentage of edges for each green index. This analysis sheds light on the overall distribution of greenness across the city's network, enabling quantification of the prevalence of various levels of greenness on the city's roads.

Data visualization is another critical aspect of the greenR package. The function `plot_green_index()`, creates a static as well as dynamic, color-coded map of the green index across the city's network. Each edge is plotted and color-coded based on its green index, providing an intuitive, easy-to-understand depiction of the city's spatial distribution of greenery (Fig. 3). This graphical output aids in identifying potential patterns or trends in the data, such as clusters of high or low green index areas, that may not be readily apparent from the numerical data alone.

Complementing the static visualizations, the `save_as_leaflet()` function generates an interactive geospatial visualization in the form of a Leaflet map. This map, saved as an HTML file, provides an interactive user experience, facilitating dynamic exploration of the data. Each edge in the network is displayed and color-coded based on its green index. The use of Leaflet maps enhances understanding of the data's spatial aspect, offering a more immersive and detailed view of the city's greenness distribution in the city's road network. This functionality brings the data to life, further aiding the identification of patterns and a deeper spatial analysis of the city greenness.

2.5.4. Data export and sharing

The final essential component of the greenR package is encapsulated by the `save_json()` function, which is responsible for data export. This function ensures the output – the calculated green index – is appropriately formatted and saved for further use, facilitating seamless communication and sharing of results.

The `save_json()` function operates by taking the output data frame, containing the green index values for each edge, and converting it into a GeoJSON file. This format is widely used for encoding a variety of geographic data structures. The resulting GeoJSON file retains the geographical properties of the data and can be readily employed in a broad range of GIS applications, as well as web-based data visualization libraries. This ensures that the results can be effectively integrated into different systems and easily shared.

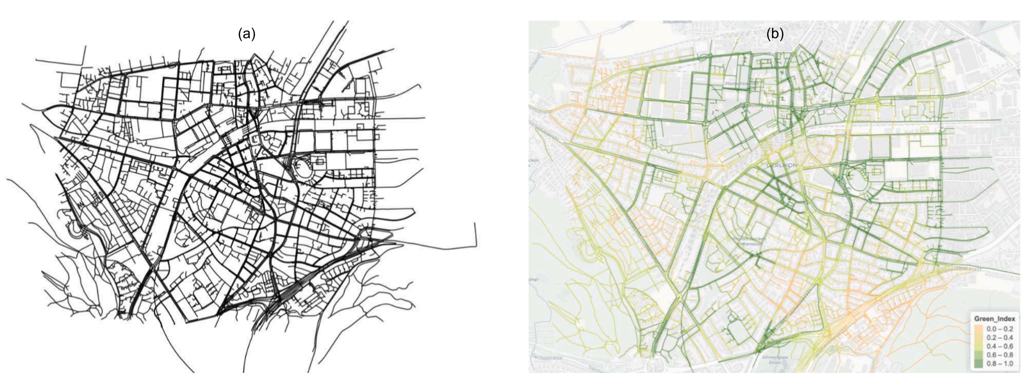

Fig. 3. Green index analysis. (a) displays the intricate street network of Oerlikon neighbourhood in Zurich, highlighting the roadway patterns and intersections. (b) superimposes the Green Index metric upon the same street network, ranging from 0 (low green index) to 1 (high green index).

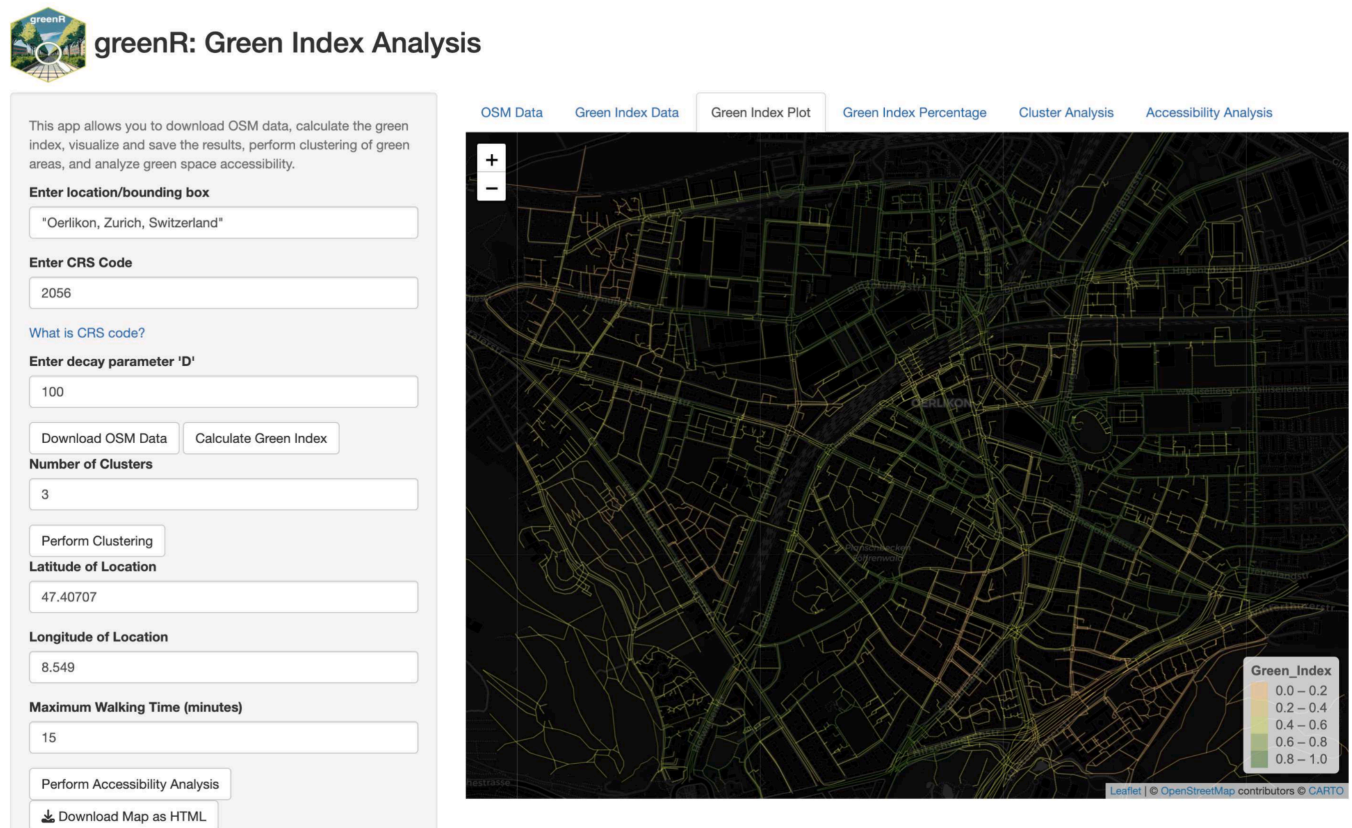

Fig. 4. Overview of greenR: Its core is an R package filled with intuitive high-level functions, including a function to launch the RShiny application. The greenR RShiny App provides an interactive platform for users to execute all functions without the necessity to write own code.

2.6. Green view index

The greenR framework extends its functionality with the integration of the GVI, a method that quantifies the proportion of visible greenery from the pedestrian's perspective (Zhang and Hu, 2022; Toikka et al., 2020). This extension enriches the package's ability to not only measure proximity to green spaces but also capture the experiential qualities such as canopy coverage, aesthetic appeal, and the shade comfort, thus offering a holistic view of urban greenery.

The GVI calculation (using the `calculate_and_visualize_GVI()` function) presented here utilizes advanced image processing techniques, with a primary focus on superpixel segmentation. This process is pivotal in partitioning a digital image into multiple segments, known as superpixels, thereby simplifying the image's representation for more effective analysis (Zhou, 2015; Stutz et al., 2018). The method employed for superpixel segmentation is detailed in the the R package library (Lampros, 2022).

The calculation of the GVI is based on a precise identification of green pixels within these superpixels. For each pixel $p_{i,j}$ in the segmented image I, the function uses specific criteria to determine its greenness, primarily based on RGB color thresholds. These criteria are: a green component G that is greater than both the red R and blue B components, and also greater than a certain threshold (in this case, $G > 0.2$). This method of classification aligns with the use of an indicator function $\mathbf{1}_{\text{green}}(p_{i,j})$ in the mathematical formula:

$$\text{GVI} = \frac{\sum_{i,j} \mathbf{1}_{\text{green}}(p_{i,j})}{\text{TotalPixels}} \tag{11}$$

where $\mathbf{1}_{\text{green}}(p_{i,j})$ returns 1 if the pixel $p_{i,j}$ is classified as green according to the defined RGB thresholds. Additionally, the function incorporates the calculation of the Excess Greenery (ExG) for each pixel, which is utilized in the visualization of green pixels. This ExG value is determined by the formula $\text{ExG} = (G - R) + (G - B)$, and is used to enhance the contrast in the visualization between green areas and other parts of the image.

The GVI, thus calculated, represents the ratio of green pixels to the total number of pixels in the image, providing a quantitative measure of visible greenery from a pedestrian's viewpoint. This measure extends the capabilities of the greenR framework, allowing for a more nuanced understanding of the distribution and quality of urban green spaces.

3. Results

3.1. Green index quantification

greenR is designed as a comprehensive and easy-to-use toolkit for greenness quantification. Fig. 4 shows the overview of greenR. It includes an R package and an accompanying R Shiny app for the acquisition, analysis, and visualization of green index data. The greenR package consists of several key integrated functions (as discussed in the Methods section), including the one to run the Shiny app.

The greenR package was thoroughly evaluated across four distinct cities: Zurich, Medellin, Taipei, and Toronto. Zurich, a well-balanced European city, blends modern urban spaces with historical zones. Medellin, located in the hilly terrain of Latin America, offers one to study a topographical challenge. Taipei, an Asian megacity, represents high-density urbanization. Finally, Toronto, a North American city, mixing green spaces and urban development, showcases the need for green space management in a multi-cultural setting. Fig. 5 shows the green index maps generated for all these cities, where different color codes correspond to different levels of green index. This visual representation facilitates a spatial exploration of city areas, enabling comparative assessments of their relative greenness.

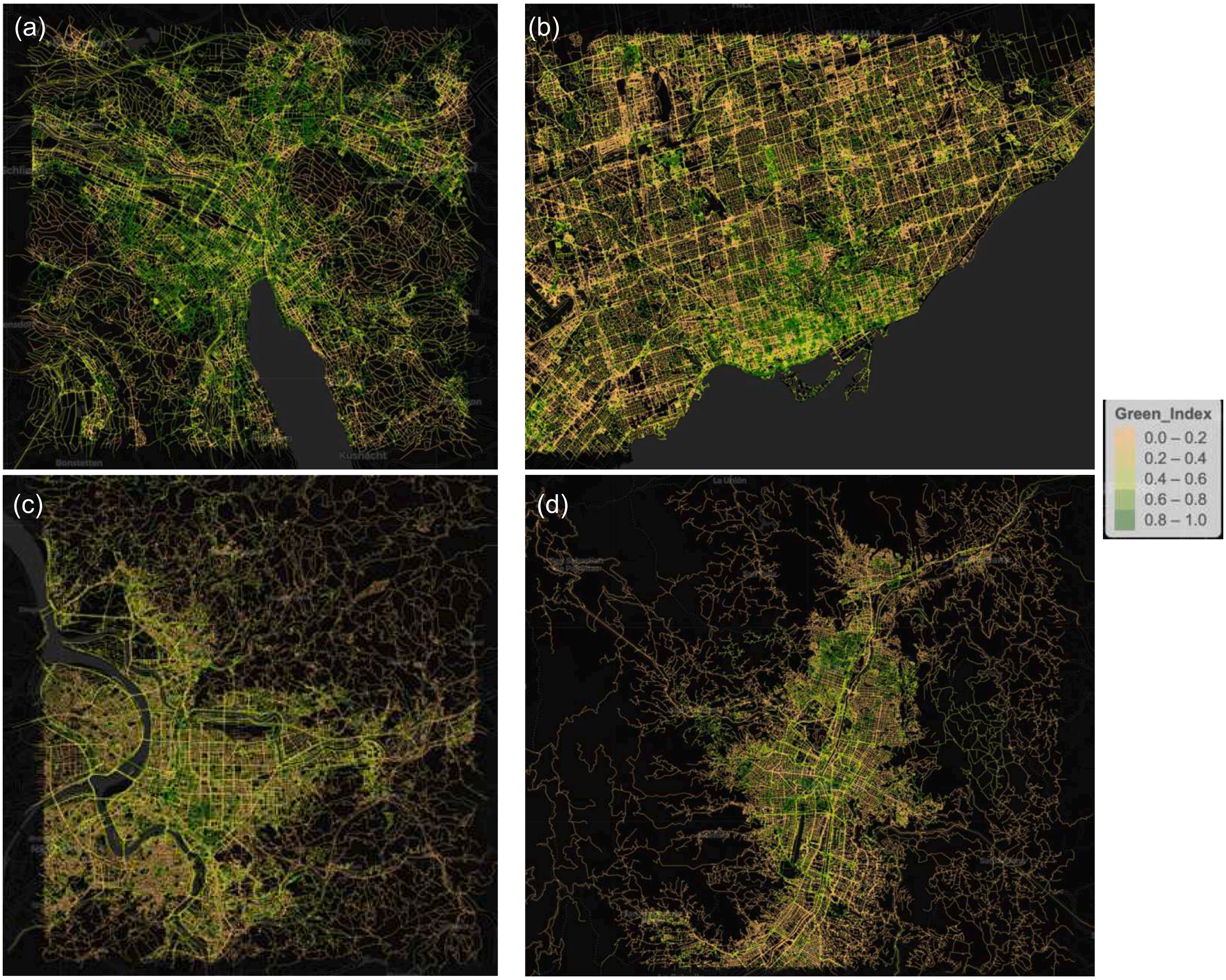

Fig. 5. Visual representations of the green index for (a) Zurich, (b) Toronto, (c) Taipei, and (d) Medellin.

Table 2
Summary of green index outcomes for four cities, where the figures within brackets represent the number of edges. The listed percentages correspond to the proportion of edges with the green index scores shown in the left column.

	City			
green index	Zurich	Medellin	Taipei	Toronto
<0.4	32% (23125)	72.4% (32837)	61.9% (45656)	60.4% (108938)
0.4–0.7	43.3% (31339)	23.2% (10532)	34.1% (25146)	28% (50508)
>0.7	24.7% (17888)	4.4% (1987)	4% (3003)	11.6% (20843)

The summarized outcomes in Table 2 reveal the green index distribution across four diverse cities. The green index is strategically classified into three ranges to facilitate comprehension and comparison: below 0.4 indicating low greenness, between 0.4 and 0.7 denoting moderate greenness, and above 0.7 signifying high greenness. This categorization allows for a nuanced evaluation of urban greenery. Medellin and Taipei exhibit a significantly larger proportion of street segments with a green index below 0.4, reflective of urban areas with lesser greenery. Zurich displays a balanced distribution, with a significant proportion of segments in the moderate range, while Toronto, despite a high count of segments in the low greenness range, also showcases numerous segments in the high greenness range, indicative of plentiful greener spaces.

These varied green index profiles underscore the greenR framework's effectiveness in discerning the extent of urban greenness at street-level. The visual representations enable spatial insights, allowing the identification of areas with varying degrees of greenery within the urban matrix.

It's worth noting that there are instances in the visualizations where certain forested regions exhibit a lower green index. This phenomenon typically arises in areas not fully mapped in OSM, highlighting not a shortfall of the method, but an opportunity to enrich the dataset. The greenR framework, in this sense, serves as a catalyst for community engagement in data curation, inviting enhancements to the mapping of green spaces. This collaborative potential not only bolsters the accuracy

Table 3
Comparative performance metrics of the greenR framework against the Helsinki and Cardiff datasets.

	Accuracy	Precision	Recall	F1 Score
Helsinki Dataset	0.71	0.71	0.71	0.70
Cardiff Dataset	0.90	0.88	0.90	0.89

of green indices but also fosters a richer understanding of urban green infrastructure.

3.1.1. Comparative analysis

The evaluation of computational models for urban greenness, a measure inherently subjective and dynamic, requires thoughtful consideration, particularly in the absence of standardized ground truth data. Traditional validation approaches are often constrained by the availability and quality of such data. Recognizing these limitations, the study presented herein opts for a comparative analysis rather than a direct validation, utilizing available datasets as a reference to examine the performance of the greenR framework.

Acknowledging the absence of ground truth data for the cities of Zurich, Toronto, Taipei, and Medellin, this study leverages datasets from Cardiff and Helsinki that offer an alternative measure of urban greenness through the GVI, calculated using Google Street View imagery. It is crucial to note that the GVI methodology—predicated on the analysis of visible greenery from a pedestrian's perspective—differs fundamentally from the greenR's reliance on OSM data, which captures greenness based on horizontal spatial data. In contrast, the GVI assesses the vertical dimension (Z dimension) of greenness, such as the visibility of tree canopies and vegetation. Consequently, the comparison between these methodologies illustrates not a validation but a methodological juxtaposition that highlights how different dimensions of greenness analysis may correlate.

The first dataset draws from the GVI for Helsinki (Toikka et al., 2020), and constitutes 3000 randomly selected observations. The second dataset comprises 3000 random observations, sourced from street greenery percentage data for Cardiff, acquired from the Office for National Statistics, United Kingdom (Office for National Statistics, 2018). To achieve uniformity with the scaling system utilized within the greenR framework, the Green Index values within the validation datasets undergo rescaling to range between 0 and 1. This ensures consistency in scale for comparison and further analysis.

The evaluation metrics used to assess the performance of the greenR framework against these datasets include accuracy, precision, recall, and the F1 score. Each metric provides insight into different aspects of the framework's performance:

- *Accuracy*: Reflects the overall correctness of the model's predictions.
- *Precision*: Measures the model's ability to identify truly green segments as green.
- *Recall*: Indicates how well the model captures all relevant green segments.
- *F1 Score*: Balances precision and recall in a single metric, useful when seeking a measure that takes both false positives and false negatives into account.

The data for Helsinki, utilized in this study, is from 2020, whereas the Cardiff data is from 2018. This temporal discrepancy is significant, especially considering the dynamic nature of urban green spaces which can undergo notable changes in short periods. The greenR framework's results, as presented in Table 3, are based on the most recent Open OSM data available for both cities. These results show a promising potential of the greenR framework to approximate greenness in urban environments, with accuracy levels of 0.71 for Helsinki and 0.90 for Cardiff. However, there are also cases where the predicted green index values do not align with the GVI categorization. This divergence underscores the inherent differences in the methodologies and the dimensions of greenness they represent, highlighting the complementary nature of these tools rather than suggesting direct equivalence.

3.2. Green Space Similarity Index (GSSI) Analysis

In this study, we utilized the `gssi()` function to conduct a comparative analysis of urban green spaces across a diverse set of cities. This analysis focused on evaluating the size variability and connectivity of green spaces within each city. Specifically, we included Zurich, Mumbai, New Delhi, Toronto, Mexico City, Medellin, Taipei, Manchester, San Francisco, and Milan, encompassing a broad spectrum of urban landscapes and green space configurations for comparative purposes. The resulting GSSI values for each city were derived based on these criteria (Fig. 6).

The GSSI values obtained are as follows: Manchester (1.00), Zurich (0.61), Medellin (0.51), Milan (0.42), San Francisco (0.25), Taipei (0.19), New Delhi (0.37), Toronto (0.13), Mumbai (0.14), and Mexico City (0.03). It is crucial to note that these values represent a relative comparison among the cities in our dataset. In this analysis, Manchester, with the highest GSSI score, served as a reference point against which the other cities were scaled. Consequently, a city's GSSI value indicates its performance in terms of green space size variability and connectivity relative to the best-performing city in the set.

The use of a relative comparison approach, such as the GSSI, offers several advantages. Primarily, it allows for benchmarking cities against a top performer, providing clear targets for urban development and environmental planning. This method also aids in prioritizing efforts by highlighting cities that might require more attention in improving their green spaces. Furthermore, a relative comparison sheds light on how cities fare against one another in specific urban and environmental

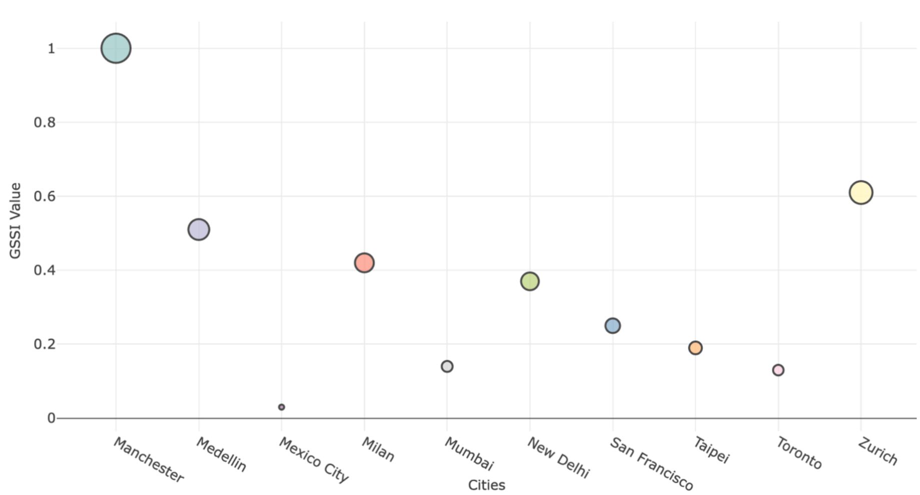

Fig. 6. Bubble chart representing the GSSI values for various cities. Each bubble's size corresponds to the GSSI value, providing a visual comparison of green space size variability and connectivity across the cities. Manchester, with a GSSI value of 1.00, serves as the benchmark, depicted by the largest bubble.

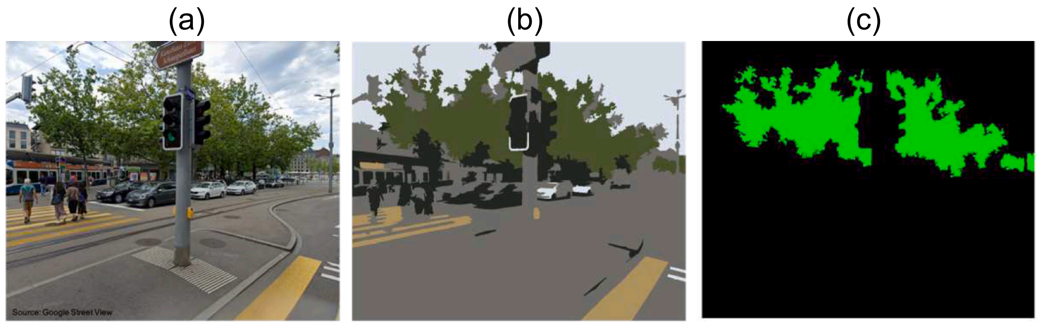

Fig. 7. Stages of GVI Calculation: (a) Street view, (b) Segmentation, and (c) Green pixel identification.

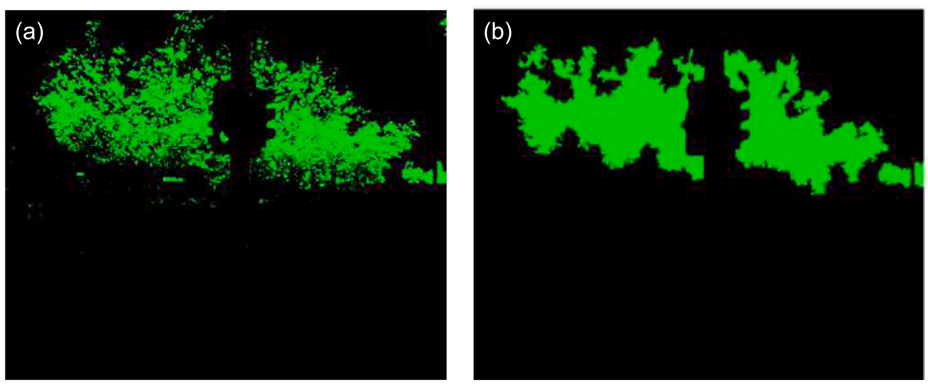

Fig. 8. Comparison between the GVI results obtained after using the HSV approach and the segmentation-based approach.

aspects, which can be instrumental in comparative urban studies. However, it is important to interpret these results within the context of this study, recognizing that they reflect the green space characteristics relative to the cities included in our dataset and not absolute standards of green space quality.

3.3. Green view index calculation

The GVI calculation, as implemented in the greenR framework, involves a multi-stage process that enhances the accuracy and relevance of green space identification in urban imagery. Fig. 7 delineates these stages, showcasing the progression from the original image to the final GVI output.

- Original Image: The process begins with the original image, which serves as the raw data input. This image captures a typical urban scene, providing the visual context for green space analysis.
- Superpixel Segmentation: The second stage involves superpixel segmentation. In this phase, the image is partitioned into superpixels, each representing a coherent group of pixels. This segmentation is crucial for reducing noise and enhancing the contextual understanding of green spaces in the image. Superpixels group similar pixels together, making it easier to differentiate between vegetative and non-vegetative elements.
- Green Pixel Identification: Following segmentation, the next step is to identify green pixels within each superpixel. This identification is based on predefined color thresholds that distinguish green vegetation from other elements. The process benefits from the prior segmentation, as it allows for a more nuanced approach to identifying greenery, accounting for variations in shade and texture that might be present within urban green spaces.

The final stage involves calculating the GVI and visualizing the identified green areas. The GVI (0.15 in Fig. 7(c)) is computed as the ratio of green pixels to the total number of pixels in the image, providing a quantitative measure of the visible greenery. The visualization highlights these green areas, offering a clear and interpretable representation of the green spaces within the urban landscape.

3.3.1. Comparison with standard GVI calculation method

In the conventional approach to GVI calculation, green pixels are identified using predefined color thresholds in the hue, saturation and value (HSV) color space (Yang et al., 2015; Cheng et al., 2017; Liu and An, 2019). While this method is straightforward and computationally less intensive, it often lacks precision. The conventional method tends to overgeneralize, potentially misidentifying non-vegetative elements as green space due to similar color hues. Such inaccuracies can lead to a skewed representation of urban greenery, particularly in complex urban landscapes where various elements may share color characteristics with vegetation.

For example, Fig. 8 shows a comparison between the GVI results obtained after using the HSV approach as well as the approach proposed within the greenR framework. As we can see in Fig. 8(a), the HSV approach provides a general indication of green areas but may include non-vegetative elements or miss finer details due to the simplistic nature of the pixel-by-pixel analysis. We can see how the green traffic light is also identified as a green pixel leading to misidentification of green space. On the other hand, in segmentation-based GVI calculation (Fig. 8 (b)), the identification of green areas appears more refined. This method benefits from the grouping of pixels into meaningful segments, allowing for a more context-aware and accurate identification of greenery. The segmentation helps in distinguishing true vegetative elements from other objects with similar colors.

4. Discussion

Benefits and Added Value of the greenR Framework: In addition to addressing the technical challenges of quantifying urban greenness, the proposed approach offers significant social benefits and added value. By generating comprehensive green index maps, using open-source tools and citizen-contributed data, we can enhance the understanding of urban greenness and provide valuable information and services to both city planners and residents.

One notable benefit of the greenR approach is the ability to identify alternative and more appealing routes within cities. Green index maps can serve as a basis to offer improved walking, cycling, or recreational routes that prioritize green spaces. By incorporating greenness as a factor in route planning, this approach can contribute to promoting active and healthy lifestyles among urban residents. People can choose routes that not only connect destinations efficiently, but also offer a pleasant and rejuvenating experience by maximizing exposure to green spaces. This will also have a positive impact on public health, physical activity, and reduced stress levels (Lanki et al., 2017; Zhang and Tan, 2019).

Furthermore, the greenR framework can help cities identify areas that lack sufficient green spaces. This is useful to prioritize interventions improving greenness in those locations. By analyzing the green index at the level of individual edges in the city network, we can identify corridors or specific routes with low greenness values. This information can guide urban planners and policymakers in implementing targeted interventions, such as planting trees, establishing pocket parks, or creating green buffers along these routes. Enhancing the greenness of these areas can not only improve the aesthetic appeal of the city, but also provide numerous benefits, including mitigation of extreme temperatures, air quality improvement, and noise reduction (Zhang et al., 2014; Lacasta et al., 2018). By conducting a thorough analysis of the green index for each edge in the city network, we can accurately identify corridors and routes that are severely lacking in greenery. This critical information should be utilized by urban planners and policymakers to take targeted and immediate action, such as planting trees, creating green buffers, or establishing pocket parks along these routes to significantly enhance the greenness value and improve the overall health and well-being of the community.

Democratizing Access to Green Index Information: In recent years, there has been a consistent emphasis on how research outputs and digital tools might be used to assist people and other stakeholders in making informed decisions (Chen et al., 2018; Münster et al., 2017; Mahajan et al., 2018). The greenR framework can empower citizens by providing them with valuable information about the greenness of their neighborhoods and the city as a whole. Accessible green index maps can raise awareness among residents about the availability and distribution of green spaces, fostering a sense of ownership and promoting engagement in environmental stewardship. Citizens can utilize this information to advocate the preservation of existing green spaces, the establishment of new parks, or the improvement of green spaces in their communities. Such citizen-led initiatives have the potential to drive positive change and create more sustainable and livable cities (Buijs et al., 2016).

Moreover, the availability of open-source tools and methodologies can democratize access to green index information. By providing the framework as an open-source resource, it becomes possible for other researchers, practitioners, and city governments to apply and adapt this approach to their respective contexts. This promotes knowledge sharing, collaboration, and the collective advancement of green space assessment methodologies. The open-source nature of this framework also encourages transparency and reproducibility in research, allowing for independent validation and refinement of the methodology by the wider scientific community. In summary, while traditional methods for quantifying green spaces have their merits, they also face limitations in terms of scalability, costs, and comprehensive mapping. The proposed approach, utilizing OSM data and an open-source framework, represents a promising alternative.

Embracing the Dynamic Nature of OSM Data: Challenges and Opportunities: The utilization of OSM data, as in the greenR framework, is not without limitations, although these do not reduce the value of the tool. Rather, they provide avenues for further refinement and enhancement. One of the key challenges is the dynamic nature of OSM data. As a crowd-sourced mapping project, the quality and completeness of the data can vary greatly, depending on the contributors' level of engagement, their geographic location, and the temporal frame of data collection (Zheng and Zheng, 2014). This dynamic aspect of OSM can introduce a degree of uncertainty into the green index calculations (Herfort et al., 2023). Furthermore, OSM data may have limited representation of private properties due to privacy concerns, and issues with semantic tags and annotations (Vargas-Munoz et al., 2020). These challenges can impact the accuracy of analysis based on OSM data. Nevertheless, this dynamic nature also creates opportunities, such as the integration of multimodal data, increased citizen participation and urban data democratization.

Another fascinating prospect lies in the realm of participatory co-creation workshops, a method that has already been tried and evaluated using digital tools for societal innovation and joint endeavors (Helbing et al., 2023; Jarke, 2019; Mahajan, 2022). Here, green index quantification can be integrated with human perception, bridging the gap between objective data and subjective experience. Such workshops can facilitate the sharing of local knowledge, allow for the recognition of uncharted green spaces, and bring potential areas for improvement to light. They can further serve as a platform for educating citizens about urban greenness, fostering awareness, and empowering collective urban development initiatives. Thus, the challenges posed by OSM data usage can be turned into opportunities for engagement, continuous model improvement, and more holistic urban greenness analysis. This would reinforce the value of greenR, emphasizing its role as a catalyst for informed, sustainable, and inclusive urban development. It also underscores the inherent advantages of an open-source, community-driven approach in addressing complex urban challenges.

5. Conclusion

In conclusion, this study introduces an innovative and scalable approach to quantify urban greenness, addressing significant challenges associated with the existing methods. Using the vast amount of data from OSM, this study proposed a reliable method for determining the green index of each street segment, providing a comprehensive overview of the level of green space in urban areas. Testing and comparison with different existing methodologies substantiate the robustness and reliability of this approach. The consequent development of greenR, a new R package, has been instrumental in enabling easy quantification, analysis, and visualization of the green index. An exploration of urban greenness patterns across various cities underscores the potential effectiveness and usability of the greenR framework. Beyond the green index, the introduction of the GSSI within the framework marks an advancement in urban green space analysis. The ability to evaluate and compare green spaces across regions based on area and spatial connectivity enriches the study's insights into urban ecology. Furthermore, the incorporation of accessibility analysis tools and the segmentation-based GVI expands the scope of the study, offering new perspectives on the accessibility and visual impact of urban greenery.

The potential implications of this study and the greenR tool are, in fact, far-reaching, potentially going beyond academic interest and having a real-world impact. The open-source framework has the potential to greatly contribute to informed decision-making, which can benefit people, urban planners, and policy-makers alike. As cities continue to grow and evolve, it is anticipated that the ability to quantify and analyze urban greenness at a granular level and at scale will increasingly be essential, e.g. to foster urban environments that are both sustainable and conducive to high-quality life.

Future work could explore potential enhancements to the framework. Additional methods could be explored to improve data accuracy, such as integrating OSM data with other datasets or using machine learning algorithms to predict missing or inaccurate tags. Furthermore, closer collaboration with the OSM community could foster better guidelines for tagging green spaces, reducing subjectivity and improving the consistency of data.

Code Availability

The greenR code can be found at https://github.com/sachit27/greenR.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The author acknowledges support 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. The author would also like to thank Mr. Javier Argota Sánchez-Vaquerizo, Mr. Alexander Eggerth, Dr. Damian Dailisan and Prof. Dirk Helbing for their insightful feedback during the development of greenR framework.

References

Arghavani, S., Malakooti, H., Bidokhti, A.A.A.A., 2020. Numerical assessment of the urban green space scenarios on urban heat island and thermal comfort level in tehran metropolis. J. Clean. Prod. 261, 121183.
Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and gis: A review. Landscape Urban Plan. 215, 104217.
Buijs, A.E., Mattijssen, T., van der Jagt, A.P., Ambrose-Oji, B., Andersson, E., Elands, B.H.M., Møller, M.S., 2016. Active citizenship for urban green infrastructure: fostering the diversity and dynamics of citizen contributions through mosaic governance. Curr. Opin. Environ. Sustain. 22, 1–6.
Cadenasso, M.L., Pickett, S.T., Schwarz, K., 2007. Spatial heterogeneity in urban ecosystems: reconceptualizing land cover and a framework for classification. Front. Ecol. Environ. 5, 80–88.
Cai, B.Y., Li, X., Seiferling, I., Ratti, C., 2018. Treepedia 2.0: applying deep learning for large-scale quantification of urban tree cover. In: 2018 IEEE International Congress on Big Data (BigData Congress), IEEE, pp. 49–56.
Chen, B., Wu, S., Song, Y., Webster, C., Xu, B., Gong, P., 2022. Contrasting inequality in human exposure to greenspace between cities of global north and global south. Nat. Commun. 13, 4636.
Chen, L.J., Ho, Y.H., Hsieh, H.H., Huang, S., Lee, H.C., Mahajan, S., 2018. Adf: An anomaly detection framework for large-scale pm2.5 sensing systems. IEEE Internet Things J. 5, 559–570.
Chen, Y., 2015. The distance-decay function of geographical gravity model: Power law or exponential law? Chaos, Solitons & Fractals 77, 174–189.
Cheng, L., Chu, S., Zong, W., Li, S., Wu, J., Li, M., 2017. Use of tencent street view imagery for visual perception of streets. ISPRS Int. J. Geo-Inform. 6, 265.
Cui, L., Wang, J., Sun, L., Lv, C., 2020. Construction and optimization of green space ecological networks in urban fringe areas: A case study with the urban fringe area of tongzhou district in beijing. J. Clean. Prod. 276, 124266.
Cui, N., Malleson, N., Houlden, V., Comber, A., 2021. Using vgi and social media data to understand urban green space: a narrative literature review. ISPRS Int. J. Geo-Inform. 10, 425.
Du, Q., Wu, C., Ye, X., Ren, F., Lin, Y., 2018. Evaluating the effects of landscape on housing prices in urban china. Tijdschrift voor economische en sociale geografie 109, 525–541.
Fuller, R.A., Gaston, K.J., 2009. The scaling of green space coverage in european cities. Biol. Lett. 5, 352–355.
Gascon, M., Cirach, M., Martínez, D., Dadvand, P., Valentín, A., Plasència, A., Nieuwenhuijsen, M.J., 2016. Normalized difference vegetation index (ndvi) as a marker of surrounding greenness in epidemiological studies: The case of barcelona city. Urban Forest. Urban Green. 19, 88–94.
Grêt-Regamey, A., Weibel, B., Bagstad, K.J., Ferrari, M., Geneletti, D., Klug, H., Schirpke, U., Tappeiner, U., 2014. On the effects of scale for ecosystem services mapping. PloS one 9, e112601.
Grün, A., 2000. Potential and limitations of highresolution satellite imagery. In: Proceedings of the 21st Asian Conference on Remote Sensing, Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry.
Gupta, K., Kumar, P., Pathan, S.K., Sharma, K.P., 2012. Urban neighborhood green index–a measure of green spaces in urban areas. Landsc. Urban Plan. 105, 325–335.
Haklay, M., Weber, P., 2008. Openstreetmap: User-generated street maps. IEEE Pervasive Comput. 7, 12–18.
Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: A k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics) 28, 100–108.
Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., Wang, R., 2019. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in beijing, china. Environ. Int. 126, 107–117.
Helbing, D., Mahajan, S., Fricker, R.H., Musso, A., Hausladen, C.I., Carissimo, C., Carpentras, D., Stockinger, E., Sanchez-Vaquerizo, J.A., Yang, J.C., et al., 2023. Democracy by design: Perspectives for digitally assisted, participatory upgrades of society. J. Comput. Sci. 71, 102061.
Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., Zipf, A., 2023. A spatio-temporal analysis investigating completeness and inequalities of global urban building data in openstreetmap. Nat. Commun. 14, 3985.
Hogendorf, M., Groeniger, J.O., Noordzij, J.M., Beenackers, M.A., van Lenthe, F.J., 2020. Longitudinal effects of urban green space on walking and cycling: A fixed effects analysis. Health & place 61, 102264.
Huerta, R.E., Yépez, F.D., Lozano-García, D.F., Guerra Cobian, V.H., Ferrino Fierro, A.L., de León Gómez, H., Cavazos Gonzalez, R.A., Vargas-Martínez, A., 2021. Mapping urban green spaces at the metropolitan level using very high resolution satellite imagery and deep learning techniques for semantic segmentation. Remote Sensing 13, 2031.
Jarke, J., 2019. Open government for all? co-creating digital public services for older adults through data walks. Online Inf. Rev. 43, 1003–1020.
Kothencz, G., Blaschke, T., 2017. Urban parks: Visitors' perceptions versus spatial indicators. Land use policy 64, 233–244.
Kozak, J., Lant, C., Shaikh, S., Wang, G., 2011. The geography of ecosystem service value: The case of the des plaines and cache river wetlands, Illinois. Appl. Geogr. 31, 303–311.
Kumar, P., Druckman, A., Gallagher, J., Gatersleben, B., Allison, S., Eisenman, T.S., Hoang, U., Hama, S., Tiwari, A., Sharma, A., et al., 2019. The nexus between air pollution, green infrastructure and human health. Environ. Int. 133, 105181.
Lacasta, A.M., Peñaranda, A., Cantalapiedra, I.R., 2018. Green streets for noise reduction.
Lampros, M., 2022. SuperpixelImageSegmentation: Image Segmentation using Superpixels, Affinity Propagation and Kmeans Clustering. URL: https://CRAN.R-project.org/package=SuperpixelImageSegmentation. r package version 1.0.5.
Lane, K.J., Stokes, E.C., Seto, K.C., Thanikachalam, S., Thanikachalam, M., Bell, M.L., 2017. Associations between greenness, impervious surface area, and nighttime lights on biomarkers of vascular aging in chennai, india. Environmental health perspectives 125, 087003.
Lanki, T., Siponen, T., Ojala, A., Korpela, K.M., Pennanen, A.S., Tiittanen, P., Tsunetsugu, Y., Kagawa, T., Tyrväinen, L., 2017. Acute effects of visits to urban green environments on cardiovascular physiology in women: A field experiment. Environ. Res. 159, 176–185.
Leslie, E., Sugiyama, T., Ierodiaconou, D., Kremer, P., 2010. Perceived and objectively measured greenness of neighbourhoods: Are they measuring the same thing? Landsc. Urban Plan. 95, 28–33.
Li, X., 2021. Examining the spatial distribution and temporal change of the green view index in new york city using google street view images and deep learning. Environ. Plan. B: Urban Anal. City Sci. 48, 2039–2054.
Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015. Assessing street-level urban greenery using google street view and a modified green view index. Urban Forest. Urban Green. 14, 675–685.
Liao, Y., Zhou, Q., Jing, X., 2021. A comparison of global and regional open datasets for urban greenspace mapping. Urban Forestry & Urban Greening.
Liu, H., An, H., 2019. Urban greening tree species classification based on hsv colour space of worldview-2. Journal of the Indian Society of Remote Sensing 47, 1959–1967.
Lu, Y., 2018. The association of urban greenness and walking behavior: Using google street view and deep learning techniques to estimate residents' exposure to urban greenness. International journal of environmental research and public health 15, 1576.
Lu, Y., Yang, Y., Sun, G., Gou, Z., 2019. Associations between overhead-view and eye-level urban greenness and cycling behaviors. Cities 88, 10–18.
Ludwig, C., Hecht, R., Lautenbach, S., Schorcht, M., Zipf, A., 2021. Mapping public urban green spaces based on openstreetmap and sentinel-2 imagery using belief functions. ISPRS International Journal of Geo-Information 10, 251.
Luxen, D., Vetter, C., 2011. Real-time routing with openstreetmap data, in: Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, New York, NY, USA. pp. 513–516. URL: http://doi.acm.org/10.1145/2093973.2094062, doi:10.1145/2093973.2094062.
Mahajan, S., 2022. Design and development of an open-source framework for citizen-centric environmental monitoring and data analysis. Scientific Reports 12.
Mahajan, S., Chung, M.K., Martinez, J., Olaya, Y., Helbing, D., Chen, L.J., 2022. Translating citizen-generated air quality data into evidence for shaping policy. Humanities and Social Sciences Communications 9, 1–18.

Mahajan, S., Martinez, J., 2021. Water, water, but not everywhere: Analysis of shrinking water bodies using open access satellite data. International Journal of Sustainable Development & World Ecology 28, 326–338.
Mahajan, S., Tang, Y.S., Wu, D.Y., Tsai, T.C., Chen, L.J., 2019. Car: The clean air routing algorithm for path navigation with minimal pm2.5 exposure on the move. IEEE Access 7, 147373–147382.
Mahajan, S., Wu, W.L., Tsai, T.C., Chen, L.J., 2018. Design and implementation of iot-enabled personal air quality assistant on instant messenger. In: Proceedings of the 10th International Conference on Management of Digital EcoSystems.
Martinez, J., Mahajan, S., 2023. Smart cities and access to nature: A framework for evaluating green recreation space accessibility. IEEE. Access.
McCall, M.K., Martinez, J., Verplanke, J., 2015. Shifting boundaries of volunteered geographic information systems and modalities: Learning from pgis. ACME: An International Journal for Critical Geographies 14, 791–826.
Münster, S., Georgi, C., Heijne, K., Klamert, K., Noennig, J.R., Pump, M., Stelzle, B., van der Meer, H., 2017. How to involve inhabitants in urban design planning by using digital tools? an overview on a state of the art, key challenges and promising approaches. In: International Conference on Knowledge-Based Intelligent Information & Engineering Systems.
Nouri, H., Nagler, P., Chavoshi Borujeni, S., Barreto Munez, A., Alaghmand, S., Noori, B., Galindo, A., Didan, K., 2020. Effect of spatial resolution of satellite images on estimating the greenness and evapotranspiration of urban green spaces. Hydrol. Process. 34, 3183–3199.
Office for National Statistics, 2018. How green is your street? URL: https://www.ons.gov.uk/economy/environmentalaccounts/articles/howgreenisyourstreet/2018-11-19.
O'Leary, M., 2011. Modeling criminal distance decay. Cityscape 161–198.
Ordóñez, C., Labib, S., Chung, L., Conway, T.M., 2023. Satisfaction with urban trees associates with tree canopy cover and tree visibility around the home. npj Urban Sustainability 3, 37.
O'Sullivan, D., Morrison, A., Shearer, J., 2000. Using desktop gis for the investigation of accessibility by public transport: an isochrone approach. International Journal of Geographical Information Science 14, 85–104.
Pincetl, S., Gearin, E., 2005. The reinvention of public green space. Urban geography 26, 365–384.
Pyszny, K., Sojka, M., Wróżyński, R., 2020. Lidar based urban vegetation mapping as a basis of green infrastructure planning, in: E3S Web of Conferences, EDP Sciences. p. 02008.
Rioux, J.F., Cimon-Morin, J., Pellerin, S., Alard, D., Poulin, M., 2019. How land cover spatial resolution affects mapping of urban ecosystem service flows. Frontiers in Environmental Science 7, 93.
Rui, L., Buccolieri, R., Gao, Z., Ding, W., Shen, J., 2018. The impact of green space layouts on microclimate and air quality in residential districts of nanjing, china. Forests 9, 224.
Senanayake, I., Welivitiya, W., Nadeeka, P., 2013. Urban green spaces analysis for development planning in colombo, sri lanka, utilizing theos satellite imagery–a remote sensing and gis approach. Urban forestry & urban greening 12, 307–314.
Skokanová, H., González, I.L., Slach, T., 2020. Mapping green infrastructure elements based on available data, a case study of the czech republic. Journal of Landscape Ecology 13, 85–103.
Strohbach, M.W., Arnold, E., Haase, D., 2012. The carbon footprint of urban green space—a life cycle approach. Landscape and Urban Planning 104, 220–229.
Stutz, D., Hermans, A., Leibe, B., 2018. Superpixels: An evaluation of the state-of-the-art. Comput. Vis. Image Underst. 166, 1–27.
Texier, M.L., Schiel, K., Caruso, G., 2018. The provision of urban green space and its accessibility: Spatial data effects in brussels. PLoS ONE 13.
Toikka, A., Willberg, E., Mäkinen, V., Toivonen, T., Oksanen, J., 2020. The green view dataset for the capital of finland, helsinki. Data in brief 30, 105601.
Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., James, P., 2007. Promoting ecosystem and human health in urban areas using green infrastructure: A literature review. Landscape and urban planning 81, 167–178.
Vargas-Munoz, J.E., Srivastava, S., Tuia, D., Falcao, A.X., 2020. Openstreetmap: Challenges and opportunities in machine learning and remote sensing. IEEE Geoscience and Remote Sensing Magazine 9, 184–199.
Villeneuve, P.J., Ysseldyk, R.L., Root, A., Ambrose, S., DiMuzio, J., Kumar, N., Shehata, M., Xi, M., Seed, E., Li, X., et al., 2018. Comparing the normalized difference vegetation index with the google street view measure of vegetation to assess associations between greenness, walkability, recreational physical activity, and health in ottawa, canada. International journal of environmental research and public health 15, 1719.
Wan, S., Rojas-Rueda, D., Pretty, J., Roscoe, C., James, P., Ji, J.S., 2022. Greenspace and mortality in the u.k. biobank: Longitudinal cohort analysis of socio-economic, environmental, and biomarker pathways. SSM - Population. Health 19. https://doi.org/10.1016/j.ssmph.2022.101194.
Wang, T., Yin, Q., Lin, Z., Liu, T., Wu, S., Xiao, C., Li, M., An, W., 2021. Detecting and tracking moving tiny vehicles from satellite based on spatio-temporal constraints, in: International Conference on Computer Vision, Application, and Design (CVAD 2021), SPIE. pp. 287–291.
Wolch, J.R., Byrne, J., Newell, J.P., 2014. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. Landscape and urban planning 125, 234–244.
Yang, J., Zhao, L., Mcbride, J., Gong, P., 2009. Can you see green? assessing the visibility of urban forests in cities. Landscape and Urban Planning 91, 97–104.
Yang, W., Wang, S., Zhao, X., Zhang, J., Feng, J., 2015. Greenness identification based on hsv decision tree. Information Processing in Agriculture 2, 149–160.
Yap, W., Stouffs, R., Biljecki, F., 2023. Urbanity: automated modelling and analysis of multidimensional networks in cities. npj Urban Sustainability 3, 45.
Ye, Y., Richards, D.R., Lu, Y., Song, X., Zhuang, Y., Zeng, W., Zhong, T., 2019. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. Landscape and Urban Planning.
Zhang, B., Xie, G., xi Gao, J., Yang, Y., 2014. The cooling effect of urban green spaces as a contribution to energy-saving and emission-reduction: A case study in beijing, china. Build. Environ. 76, 37–43.
Zhang, J., Hu, A., 2022. Analyzing green view index and green view index best path using google street view and deep learning. Journal of Computational Design and Engineering 9, 2010–2023.
Zhang, L., Tan, P.Y., 2019. Associations between urban green spaces and health are dependent on the analytical scale and how urban green spaces are measured. International Journal of Environmental Research and Public Health 16.
Zhang, Y., Dong, R., 2018. Impacts of street-visible greenery on housing prices: Evidence from a hedonic price model and a massive street view image dataset in beijing. ISPRS International Journal of Geo-Information 7, 104.
Zhang, Y., Shao, Z., 2021. Assessing of urban vegetation biomass in combination with lidar and high-resolution remote sensing images. Int. J. Remote Sens. 42, 964–985.
Zheng, S., Zheng, J., 2014. Assessing the completeness and positional accuracy of openstreetmap in china. Thematic cartography for the society. Springer 171–189.
Zhou, B., 2015. Image segmentation using slic superpixels and affinity propagation clustering. Int. J. Sci. Res 4, 1525–1529.
Zhou, X., Wang, Y.C., 2011. Spatial–temporal dynamics of urban green space in response to rapid urbanization and greening policies. Landscape and urban planning 100, 268–277.
Zluwa, I., Pitha, U., 2021. The combination of building greenery and photovoltaic energy production—a discussion of challenges and opportunities in design. Sustainability 13, 1537.