It is not always greener on the other side: Greenery perception across demographics and personalities in multiple cities

Matias Quintana[a], Fangqi Liu[b], Jussi Torkko[c], Youlong Gu[a,d], Xiucheng Liang[d], Yujun Hou[d], Koichi Ito[d], Yihan Zhu[d], Mahmoud Abdelrahman[d], Tuuli Toivonen[c], Yi Lu[b], Filip Biljecki[d,e,*]

[a]*Future Cities Lab Global, Singapore-ETH Centre, CREATE campus, #06-01 CREATE Tower, Singapore, 138602, Singapore*
[b]*Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, , Hong Kong SAR*
[c]*Digital Geography Lab, Department of Geosciences and Geography, University of Helsinki, Helsinki, , Finland*
[d]*Department of Architecture, National University of Singapore, 4 Architecture Drive, Singapore, 117566, Singapore*
[e]*Department of Real Estate, National University of Singapore, 15 Kent Ridge Dr, Singapore, 119245, Singapore*

Abstract

Quantifying and assessing urban greenery is consequential for planning and development, reflecting the everlasting importance of green spaces for multiple climate and well-being dimensions of cities. Evaluation can be broadly grouped into objective (e.g., measuring the amount of greenery) and subjective (e.g., polling the perception of people) approaches, which may differ – what people see and feel about how green a place is might not match the measurements of the actual amount of vegetation. In this work, we advance the state of the art by measuring such differences and explaining them through human, geographic, and spatial dimensions. The experiments rely on contextual information extracted from street view imagery and a comprehensive urban visual perception survey collected from 1,000 people across five countries with their extensive demographic and personality information. We analyze the discrepancies between objective measures (e.g., Green View Index (GVI)) and subjective scores (e.g., pairwise ratings), examining

---

*Corresponding author

whether they can be explained by a variety of human and visual factors such as age group and spatial variation of greenery in the scene. The findings reveal that such discrepancies are comparable around the world and that demographics and personality do not play a significant role in perception. Further, while perceived and measured greenery correlate consistently across geographies (both where people and where imagery are from), where people live plays a significant role in explaining perceptual differences, with these two, as the top among seven, features that influences perceived greenery the most. This location influence suggests that cultural, environmental, and experiential factors substantially shape how individuals observe greenery in cities. We also found that the spatial arrangement of greenery in the sight, rather than its proximity to the person, influences perception. Our study provides a new understanding of the deep relationships between objective and subjective street-level greenery assessments, contributing to a more human-centric design of green urban environments.

---

1. Introduction

Quantifying and assessing greenery is crucial for planning and development, and has been the subject of research for decades. Particularly, green spaces are favored due to their benefits for climate mitigation (Fujiwara et al., 2024a,b), physical (Mao et al., 2025; Gao et al.; Lu, 2019; Lu et al., 2018), and mental (Jiang et al., 2014; Wang and Yao, 2025) well-being, and economic development (Wu et al., 2022). Traditionally, researchers have assessed greenery with data collected from aerial or satellite platforms (Li et al., 2015a), with Street View Imagery (SVI) becoming an increasingly popular urban data source for similar purposes (Kang et al., 2020; Biljecki and Ito, 2023). These images have enabled researchers and practitioners to gather objective (Ye et al., 2019; He et al., 2025; Li et al., 2015a; Ki and Lee, 2021a; Wu et al., 2022; Li et al., 2024; Zhang and Zeng, 2024; Mao et al., 2025; Merdymshaeva et al., 2025; Zhao et al., 2025), and subjective (Yang et al., 2009; Quintana et al., 2025; Lefosse et al., 2025) greenery measurements from the human-centric perspective.

One widely used objective measurement to assess eye level greenery is the Green View Index (GVI) (Yang et al., 2009; Ye et al., 2019; He et al., 2025; Li et al., 2015a; Ki and Lee, 2021a; Wu et al., 2022; Li et al., 2024; Zhang and Zeng, 2024; Mao et al., 2025; Merdymshaeva et al., 2025; Zhao et al., 2025),

which complements remote sensing, specifically orthophotography-based metrics such as the Normalized Difference Vegetation Index (NDVI) (Helbich et al., 2021; Huang et al., 2025; He et al., 2025). When computed from SVI, GVI extends traditional greenery variables such as the number of street trees and parks areas as it can capture broader greenery, including also other aspects such as lawns and green walls (Ki and Lee, 2021b). Besides objective measurements such as GVI, SVI has been handy to measure subjective dimensions (e.g., through surveys where people rate their perception of greenery). Evaluating both of these has been important because people's impressions of how green a place feels can differ significantly from technical measurements such as vegetation coverage (Torkko et al., 2023). This discord means that a place could be measured to be highly green but still feel less green to people, or vice versa. These differences stem from the measurement mechanism; biological vision focusing in spatial forms, elements, and colours, compared to ratio of vegetation pixels Cao et al. (2025); and are also driven by many factors that are inherent to human perception, such as personal preference, cultural background, and the visibility, density, and type of greenery (Liu et al., 2024).

Research comparing subjective and objective greenery has been conducted using both in-situ and image-based approaches, though their findings are not always directly comparable. In in-situ work, SVI-derived GVI has been shown to underestimate participants' field-based impressions of greenery (Torkko et al., 2023), potentially reflecting the broader sensory and contextual cues present during onsite experience. Image-based studies have typically relied on perceptual ratings or greenery manual selection directly from SVI. For example, Ye et al. (2019) asked respondents to judge greenery levels and Long and Liu (2017); Suppakittpaisarn et al. (2022) quantified perceived greenery through respondents' greenery manual segmentation. Both reported positive associations with objective metrics, but their analyses were limited to single-city or single-country contexts and did not examine the subjective–objective relationship further. Beyond greenery, Lefosse et al. (2025) analyzed the subjective perception of biophilia on imagery across biomes and demographics, providing global insights on the benefits of eye contact with nature-based elements. These location and demographic differences in subjective measurements are crucial for providing targeted context-aware urban solutions (Qi et al., 2022).

However, efforts to explore the relationships between objective and subjective metrics across locations and demographics, and to comprehensively explore other image-derived features, remain scarce. The few that combined GVI with greenery perception do so at a single city, limiting understanding of whether their

3

relationship differ across geographies and populations (Torkko et al., 2023; He et al., 2025). The rare multi-city efforts across demographics focus on broader topics such as biophilia (Lefosse et al., 2025) or focus solely on NDVI and GVI with no subjective measurements (Huang et al., 2025). Our work addresses these gaps by exploring the relationship between objective and subjective metrics across multiple locations and diverse participants who have volunteered to divulge their thorough personal information for this research.

In this work, we utilize the comprehensive Street Perception Evaluation Considering Socioeconomics (SPECS) dataset, which is a large-scale global survey we designed and collected to support multiple lines of urban perception research (Quintana et al., 2025), to analyze and understand the relationship between commonly used objective and subjective methods used to evaluate greenery. Among other information, the effort collected perception ratings on diverse but controlled imagery (i.e., carefully and systematically curated and selected) from 1,000 people from around the world, together with a series of information about the participants and images. First, we performed correlation and bias quantification analyses to thoroughly understand the relationship between objective and subjective greenery quantification at the street level. Then, we looked deeper into the images by analyzing aspects such as the arrangement of greenery within the image and greenery proximity to the viewer. Finally, we investigated which of these variables explains greenery perception the most by training a tree-based model using image-derived features and location information to predict greenery perception scores. Figure 1 shows an overview of this work. Through this comprehensive set of analyses of a multi-city dataset, we address three key questions that have not been investigated before at this scale and depth: a) How do subjective greenery evaluations relate to objective (GVI, vegetation arrangement, and proximity to the viewer) measurements? b) Which demographic factor most strongly influences greenery perception? c) How does the spatial arrangement of greenery affect subjective ratings?

2. Literature review

2.1. Subjective human perception of greenery

Perceived greenery is shaped by both individual characteristics and broader geographic contexts. Studies have shown that demographic groups can experience and evaluate nature-based elements differently. For instance, Lefosse et al. (2025) found that women preferred more nature-based elements than men, with additional variation attributable to city of residence. Beyond stated preferences,

other work has examined physiological and psychological responses to greenery. Jiang et al. (2014) demonstrated that tree-cover density yields diminishing stress-reduction benefits for men compared to women, while Lottrup et al. (2013) reported similar gender-differentiated relationships between workplace greenery and stress. Age-related differences have also been observed, with green space immediately outside the home providing stronger developmental benefits for young girls than for boys (Taylor et al., 2002). Together, these findings highlight that perceptions of and responses to greenery are not uniform, but embedded within demographic, cultural, and place-based contexts.

Studies has employed a range of approaches to measure these perceptions. Most studies evaluate greenery qualitative on-site (Falfán et al., 2018; Leslie et al., 2010; Sugiyama et al., 2008) or under controlled laboratory conditions (Aoki, 1991), often using street-level images or SVI (Yang et al., 2009; Quintana et al., 2025; Lefosse et al., 2025; Ogawa et al., 2024). In-situ assessments range from providing ratings from 0-100% (Falfán et al., 2018) to a Likert scale of different perceived green elements (Leslie et al., 2010; Sugiyama et al., 2008). Image-based evaluations range from manual green segmentation (Yang et al., 2009; Long and Liu, 2017) to tracking physiological reactions to 3D videos (Jiang et al., 2014), to pair-wise comparison tasks (Quintana et al., 2025; Lefosse et al., 2025; Ogawa et al., 2024). Compared with field-based assessments, image-based methods and online surveys are more time- and cost-effective (Kang et al., 2020) and allow large-scale collection of perceptual data (Gu et al., 2025; Salminen et al., 2025).

Although subjective evaluations are susceptible to human-induced biases (Gupta et al., 2012), they capture qualitative dimensions of greenery that quantitative approaches may overlook (Leslie et al., 2010). Importantly, they help reveal demographic- and place-specific preferences, providing valuable insights for inclusive, context-aware urban design practices and strategies (Qi et al., 2022).

2.2. Objective machine-derived perception of greenery

Numerous indices and methods have been developed for computing the amount of greenery, such as remote sensing-based NDVI (He et al., 2025; Helbich et al., 2021; Huang et al., 2025) and point-cloud based methods (MacFaden et al., 2012; Yu et al., 2016; Labib et al., 2021; Zięba-Kulawik et al., 2021). However, methods relying on SVI, such as GVI (Ye et al., 2019; He et al., 2025; Li et al., 2015a; Ki and Lee, 2021a; Wu et al., 2022; Li et al., 2024; Zhang and Zeng, 2024; Mao et al., 2025; Merdymshaeva et al., 2025; Zhao et al., 2025; Lu, 2019; Lu et al., 2018) are predominantly used as an objective or computational approach to assess street and eye level greenery. This method consists of calculating the visibility of greenery

from a specific position (Yang et al., 2009) by computationally extracting the ratio of vegetation elements' pixels over the entire image via color thresholding (Li et al., 2015b,a; Lu, 2019; Lu et al., 2018) or semantic segmentation (Aikoh et al., 2023; Hou et al., 2024).

Unlike remote sensing imagery, street-level imagery better captures what people experience (Ye et al., 2019), allowing GVI and other objective measurements to capture information that otherwise cannot be gathered from top-down perspectives (Gaw et al., 2022; Li et al., 2024; Helbich et al., 2021). Additionally, it is straightforward to acquire and not very sensitive to quality issues, as even single (non-panoramic) imagery is sufficient (Biljecki et al., 2023). This data also offers a wider global coverage and availability from proprietary (e.g., Google Street View or Baidu) and crowdsourced services (e.g., Mapillary or KartaView) (Biljecki and Ito, 2023; Hou et al., 2024).

GVI has been found to be a relevant measure and predictor in the built environment — it was used to examine the relationship between street greenery and socioeconomic profiles (Li et al., 2024, 2015a), examine its relationship with overall physical activity (Mao et al., 2025), measure greenery cooling effect for heat stress (Merdymshaeva et al., 2025), and used as labeled data for GVI prediction models based on multi-source remote-sensing data (Ma et al., 2025a). Despite this adoption and ease of use, GVI faces challenges.

In relation to NDVI, studies have reported systematic differences when using GVI, although the direction and magnitude of these differences are context dependent. While Huang et al. (2025) observed that GVI tended to yield lower greenery estimates than NDVI across their sampled cities, Ye et al. (2019) demonstrated that discrepancies between the two metrics are not consistently unidirectional but instead exhibit regional patterns, with GVI occasionally exceeding NDVI in certain urban contexts. These variations are partly attributable to methodological differences between street-level and satellite-based imagery, including seasonal effects (Zhao et al., 2025), spatial bias (Zhang and Zeng, 2024), and sampling strategies (Huang et al., 2025). In an attempt to overcome its shortcomings, researchers have developed more specialized versions of GVI to capture the full 360° view (Li et al., 2015b) or the view from multi-story buildings using floor-specific adjustment (Yu et al., 2016).

2.3. Objective meets subjective

Although each type of measurement has its benefits and both aim to assess greenery quantitatively, fundamentally, they measure different things: objective measurements quantify the visible greenery numerically and consistently, while

subjective measurements assess greenery based on people's perceptions, encapsulating both perceived quantity and quality. Thus, recent attempts to understand the relationship between objective and subjective measurements have resulted in approaches combining both types of measurements to obtain a complete picture of urban greenery. He et al. (2025) combined satellite imagery and SVI-based perceived safety to identify the relationship between greenery, through NDVI, and safety perception scores. Using SVI, approaches range from biological-inspired measurements (i.e., landscape imagery, spatial forms, landscape elements, and color features) for subjective perception scores prediction (Cao et al., 2025), color-emotion analysis for measuring the positive impact of blue and green spaces (Chen et al., 2025), to quantifying biophilia based on image-based elements (objective measures) and pair-wise comparisons (subjective measures) (Lefosse et al., 2025). Outside of greenery, urban design subjective preferences, and street physical features, both obtained from SVI, complement each other when used to explain housing prices (Qiu et al., 2022). These approaches highlight opportunities where both types of measurements complement each other or serve to understand the other. Nevertheless, these comparisons and hybrid methods, despite being done at a granular level within a city, have been constrained to single-city studies in places like Houston (USA) (He et al., 2025), Tianjin (China) (Chen et al., 2025), Helsinki (Finland) (Torkko et al., 2023), and Stockholm (Sweden) (Merdymshaeva et al., 2025), overlooking one of the main limitations of subjective-based measurements: its location- and people-specific nature. Current multi-city approaches either look solely at NDVI and GVI with no subjective measurements (Huang et al., 2025) or focus on broader topics such as biophilia (Lefosse et al., 2025).

In this paper, we address these gaps by conducting a multi-city, multi-population experiment to evaluate the relationships and potential systematic differences between objective greenery measures (GVI, vegetation arrangement, and proximity to the viewer) and subjective greenery perceptions. We further examine how demographic and personality characteristics – derived from our large global survey – help explain these patterns.

3. Methods

3.1. Dataset

We performed a survey that comprises pairwise comparison ratings of street view imagery from five different cities (80 images per city, a total of 400) contributed by 1,000 participants from 5 countries (200 participants from each country), for which we collected detailed demographic and personality information.
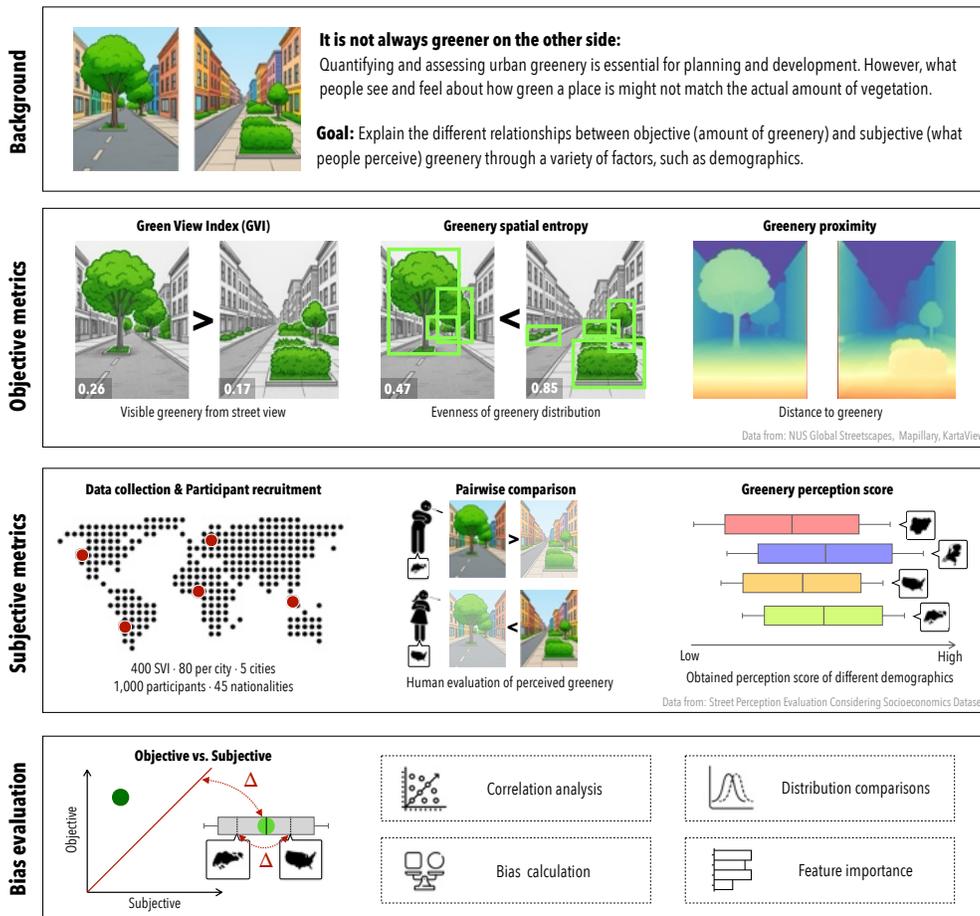
Figure 1: Workflow and methodology of this study: dataset identification, data preprocessing, and scores comparisons. We used the Street Perception Evaluation Considering Socioeconomic (SPECS) dataset due to its diverse samples and inclusion of a *greenery* perceptual indicator (Quintana et al., 2025). We calculated green perception scores based on ratings for all 400 images, obtained the Green View Index (GVI) using the panoptic semantic segmentation models recommended in (Hou et al., 2024), and computed greenery spatial arrangement metrics based on the segmented greenery. We then compared these subjective, i.e., greenery perception, and objective, i.e., GVI, measurements in terms of their correlation and bias. Finally, as a last comparison, we looked into the vegetation spatial arrangement within images among the least and most green-rated images.

The cities where the images come from are balanced worldwide: Santiago (Chile), Amsterdam (Netherlands), Abuja (Nigeria), Singapore (Singapore), and San Francisco (USA), and the imagery was initially obtained from the NUS Global Streetscapes dataset (Hou et al., 2024). The survey participants came from the same five countries. As we mentioned in (Quintana et al., 2025), where we introduced this dataset (called SPECS – Street Perception Evaluation Considering Socioeconomics), we chose these cities to following a three-tier filtering process: (1) continental representation to ensure global coverage, (2) availability in established datasets (MIT PP2 Dubey et al. (2016) and NUS Global Streetscapes (Hou et al., 2024)) to enable comparative analyses, and (3) consideration of cultural, urban form, and socioeconomic diversity within these constraints.

These criteria were essential for evaluating our core hypothesis that urban perception — across many perceptual indicators including *green* — varies across demographic groups and across cities with different spatial and cultural characteristics. The five cities therefore intentionally span diverse built forms, socioeconomic profiles, and cultural settings (Figure 1; Supplementary Information in Quintana et al. (2025)). This diversity is a necessary condition for testing perceptual robustness across heterogeneous environments.

From these cities, we sample images following the image download process from Hou et al. (2024) and download all available street view images within a 2.4 km^2 square in the city center (with the exception of Abuja, for which we manually selected a similarly sized but more developed area). We then filtered images based on available metadata (e.g., clear weather and image taken during the day); visual metrics such as visual complexity; and image composition ratios of roads, vegetation, cars, buildings, and sky) Quintana et al. (2025). None of the final 400 selected images showed snow or ground cover attributable to weather conditions. By keeping the quality and properties of imagery consistent, we minimise the potential bias of various aspects of imagery on perception.

The information we collected from 1,000 participants from the five countries include gender, age, nationality, country and city of residence, length of stay in the said city, annual income level, number of household members, race and ethnicity, and Big Five Inventory personalities (*extraversion*, *agreeableness*, *conscientiousness*, *neuroticism*, and *openness*). The dataset contains the participants' pairwise comparison ratings for the commonly used six indicators (*safe*, *lively*, *wealthy*, *beautiful*, *boring*, and *depressing*) proposed in the MIT Place Pulse 2.0 dataset (PP2) (Dubey et al., 2016), which have been widely and consistently adopted in the literature Cui et al. (2023); Kang et al. (2023); Meir and Oron-Gilad (2020); Hidayati et al. (2020); Ogawa et al. (2024); Liang et al. (2024). Using them also

facilitates comparative analysis with previous and future studies. In addition, we also included four new indicators of *live nearby*, *walk*, *cycle*, and *green* which are better aligned with livability, active mobility, and sustainability study goals. To support this research, we introduced this last mentioned indicator and was prompted to users in the form of "Which place looks **greener**?". A snippet of the dataset is shown in Table 1 and the final dataset contained 50,000 pairwise comparisons ratings across all participants and perceptual indicators (i.e., 1,000 participants $\times$ 50 pairwise comparisons per survey on which participants rated imagery from all five cities).

We calculated machine-derived metrics from the images, such as the amount of greenery, greenery arrangement, and greenery proximity, and computed the *greenery* perception scores from the pairwise comparison ratings based on the participants' location.

	Demographic data (multiple choice)				Big Five personality score (Bounded in [1, 7])			Pairwise comparisons (multiple choice: left, right, equal)			
Participant ID	Gender	Age group	Country of residence	...	Extraversion	...	Openness	Indicator	Left image ID	Right image ID	Choice
146	Female	40-49	Nigeria	...	3	...	2.5	safe	307	99	right
146	Female	40-49	Nigeria	...	3	...	2.5	green	88	284	equal
		⋮			⋮				⋮		
31	Male	21-29	Singapore	...	2.5	...	2.5	safe	397	148	left
31	Male	21-29	Singapore	...	2.5	...	2.5	green	20	340	right
		⋮			⋮				⋮		

Table 1: Snippet of the SPECS dataset containing its three main sections: participants' demographic data, Big Five personality scores, and pairwise comparisons ratings. Each participant completed an onboarding survey that included the demographic and personality questions (10 questions each) and 50 pairwise comparisons ratings (5 questions per indicator) of the 400 randomized images. The complete list of questions and dataset details are available in Quintana et al. (2025).

3.2. Objective measurements

3.2.1. Amount of greenery: Green View Index

GVI is the visibility of greenery from a specific position (Yang et al., 2009) and can be calculated as shown in the equation below (Li et al., 2015a).

$$GVI = \frac{\text{Number of Green pixels}}{\text{Number of Total pixels}} \quad (1)$$

With the index ranging from 0 to 1, when a lot of greenery is visible from a specific position, the index value is high. To compute GVI, instance-based image analysis is first used to segment green vegetation. This segmentation is commonly done with spectral information, i.e., using the different pixel values from the RGB bands (Li et al., 2015b,a), or with semantic segmentation machine learning models (Aikoh et al., 2023; Hou et al., 2024). We compared both approaches and, similar to (Torkko et al., 2023), chose the semantic segmentation approach as it is less sensitive to small variations in pixel values, such as lighting conditions or dimly/brightly lit greenery (Pietikainen et al., 1996; Batlle et al., 2000). Figure A.8 shows an example of a pair of images from the SPECS dataset with their respective GVI based on spectral and semantic segmentation (middle and right column, respectively). We ran the panoptic semantic segmentation, i.e., instance and semantic segmentation, inferences using the ZenSVI Python package (Ito et al., 2025a), which uses a model trained on the Mapillary Vistas version 1.2 dataset (Neuhold et al., 2017) as the backend. We computed the GVI by including areas segmented as both 'vegetation' and 'terrain' categories. Following (Torkko et al., 2025) and their GVI-based methodological recommendations, we retained 'terrain' segments to explicitly capture natural elements within the GVI calculation. To ensure consistency, we manually validated 'terrain' segments by examining the segmented street view imagery and including only those areas that represented grass surfaces or vegetated patches.

3.2.2. Greenery distribution: Spatial entropy

We quantified the distribution of vegetation within the image. While existing work in urban studies has used visual complexity as a measure of entropy, e.g., whether the image is composed predominantly of one element or multiple ones (Mayer et al., 2023; Yap et al., 2023), this computation does not take into account the spread of the elements. Calculating the entire image entropy would provide an overall vegetation proportion without accounting for spatial arrangements. For example, an urban scene with vegetation evenly distributed across the

image and an urban scene with vegetation concentrated on one side would have the same entropy values. Thus, as we are particularly interested in how the identified vegetation (e.g., trees, bushes, grass patches) is spatially distributed in an image, we opted for computing the spatial entropy as a measure of the spatial heterogeneity, or randomness, of vegetation within an image.

For a given sliding window W of size $s \times s$, the spatial entropy is defined using the Shannon entropy formula (Equation 2) where p is the proportion of vegetation pixels in the window W (Equation 3) and $1 - p$ represents the proportion of non-vegetation pixels. To obtain the overall spatial entropy for the image, we first generate a vegetation mask and then compute the mean entropy across all valid windows (Equation 4), where N is the total number of sliding windows across the image and $H(W_i)$ is the entropy computed for the i-th window. Spatial entropy is bounded in $[0, 1]$, and a high value indicates a more mixed and uncertain distribution, while lower entropy suggests more uniformity within the image.

$$H(W) = -p \log_2 p - (1 - p) \log_2(1 - p) \tag{2}$$

$$p = \frac{\sum W}{s^2} \tag{3}$$

$$H_{\text{mean}} = \frac{1}{N} \sum_{i=1}^{N} H(W_i) \tag{4}$$

Due to the different image sizes, a characteristic of crowdsourced SVI owing to heterogeneous equipment used (Hou et al., 2024), we opted for a window size based on the images' dimensions, i.e., relative window size, instead of a fixed dimension. We performed a sensitivity analysis (Figure A.9) on window size fractions ranging from 0.10 to 1.00 with 0.05 increments and found the maximum average value at a window size fraction of 0.45. This peak entropy window size represents the scale at which the image exhibits the most spatial variation between vegetation and non-vegetation. For the dataset, this selected window size fraction resulted in squared sliding windows from 480×480 to 921×921 pixels. Figure 2 shows an example of five images from the dataset arranged based on their respective aggregated *green* perception Q scores (y-axis) and GVI values (x-axis) (Figure 2a, and spatial entropy values (Figure 2b).

3.2.3. Greenery proximity: Distance to the viewer

In addition to quantifying the spatial entropy of vegetation within the image, we also measured its proximity to the viewer. This metric is crucial because iden-
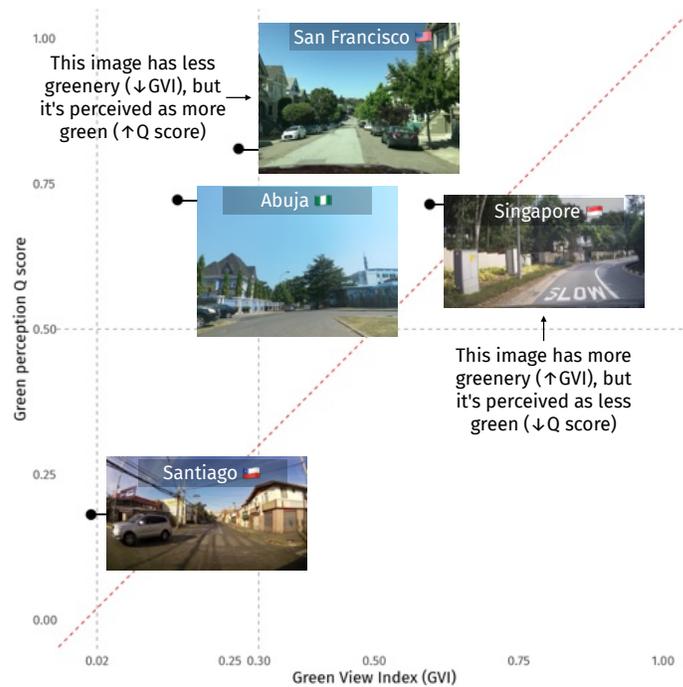
(a) Discrepancy between subjective and objective notions of urban greenery. A red dashed line shows the identity function and vertical lines are shown as thresholds for imagery with some vegetation (GVI = 0.3), very little vegetation (GVI = 0.02), and perception Q score of 0.5.
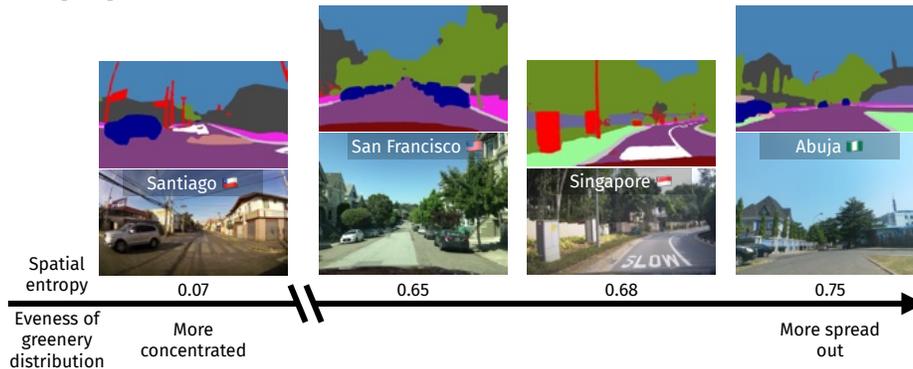
(b) Higher values of spatial entropy indicate vegetation distributed throughout the urban scene, and lower values indicate vegetation concentrated in specific locations. At comparable spatial entropy values (between 0.68 and 0.65), the SVI from Singapore has a higher GVI than the one from San Francisco (0.60 and 0.27, respectively).

Figure 2: Four images from the SPECS dataset (Quintana et al., 2025) arranged based on their (a) *green* perception Q scores (y-axis) and their Green View Index (GVI) values (x-axis), and their (b) spatial entropy.

tical GVI values can arise from vastly different spatial configurations, potentially leading to distinct subjective perceptions. As noted in classic Gestalt psychology, the principle of proximity is fundamental to visual perceptual grouping, determining how discrete elements are organized into a coherent whole (Wagemans et al., 2012). For instance, a scene dominated by large canopy structures in the distance may yield the same GVI as one with smaller vegetation in the immediate foreground. However, purely pixel-based metrics fail to distinguish between these scenarios. By incorporating the distance variable, we aim to disentangle these spatial attributes, allowing us to examine whether human perception is driven more by the immediate visual impact of nearby vegetation or the extensive coverage of distant landscapes.

To estimate the distance of vegetation, we utilized the Depth Anything V2 model (Yang et al., 2024). This advanced monocular depth estimation algorithm has already demonstrated strong performance in urban domains such as street-view analysis and greenery perception (Ito et al., 2025b; Ma et al., 2025b). Trained on a combination of high-quality synthetic images and over sixty million pseudo-labeled real images, the model delivers high accuracy and robust generalization, making it well-suited for analyzing street-view scenes without requiring ground-truth depth annotations.

We calculated vegetation distances using the same masks employed for GVI calculation, i.e., masks containing pixels classified as vegetation and green terrain from the panoptic semantic segmentation inference. These masks were overlaid on the depth maps to isolate vegetation pixels. For each image, we then derived two proximity indicators: *absolute depth* (the mean distance of vegetation pixels to the observer) and *relative depth* (the depth of vegetation compared to other scene elements). Together, these measures capture whether greenery is predominantly near or far from the viewer and extend the GVI by incorporating a proximity dimension into the assessment.

3.3. Subjective measurements: Perception ratings

The numerical perception scores were computed using the Strength of Schedule (SOS) method and are referred to as Q scores, ranging from [0, 10] and reflecting the magnitude of the perception. The dataset also provides the Q scores based on demographics, i.e., Q scores calculated using ratings from only a specific demographic group and based on the image and participants' locations. Another widely used numerical scoring method for pairwise comparisons in urban visual perception studies is the Trueskill ranking system (Kang et al., 2023; Qiu et al.,

2025). Both scores, Q scores and Trueskill scores (Herbrich et al., 2006), inherently depend on the ratings considered for their calculations. Thus, rather than being an absolute measure of an image's perceptual dimensions, these scores are only comparative; they are only valid within the rating pool in which they were calculated.

We compared the relative scores for the *green* perceptual indicator and found mostly similar distributions. Figure A.10 shows a Quantile-Quantile plot of Trueskill scores and Q scores based on location-pair (i.e., country-city pair) groups based on the images' and participants' country and city, respectively. While there are some discrepancies around lower scores (e.g., values above the diagonal line in Figure A.10, Chile-Abuja location pair), we opted for Q scores as we used them in the original SPECS dataset (Quintana et al., 2025) and recommended in recent relevant work (Gu et al., 2025). We also followed the SPECS dataset preprocessing criteria of only retaining images with at least four pairwise comparisons in each grouping context, as this was the reported average comparison in PP2 (Dubey et al., 2016). Finally, we normalized the computed Q scores into the $[0, 1]$ range to match the range of GVI.

3.4. Analyzing the measurements relationships

To systematically evaluate the agreement between subjective *green* perception Q scores and objective GVI values, we employed a multi-step statistical approach that accounts for the different measurement scales and distributional properties of these metrics. First, we performed a Pearson's correlation analysis, following previous objective-subjective greenery studies (Suppakittpaisarn et al., 2022; Torkko et al., 2023), to quantify the linear relationship between perception and measurement. However, correlation alone does not reveal systematic biases or the magnitude of disagreement between metrics. Therefore, following the analysis in (Suppakittpaisarn et al., 2022), we conducted a Bland-Altman analysis, which visualizes the arithmetic difference, i.e., perception Q score - GVI, against the average of both metrics for each image. This approach identifies whether disagreements are consistent across the measurement range or vary systematically with the magnitude of greenness. Given that GVI represents an absolute metric while perception Q scores are relative measures, we normalized GVI values to ensure comparable scales. Then, to quantitatively confirm our visual assessments, we applied the non-parametric Wilcoxon signed-rank test. We chose this test because the perception Q scores exhibited non-normal distributions across different participants and image groups. The test compares the distribution of arithmetic differences, i.e., perception Q score - normalized GVI, against zero, testing the

null hypothesis H_θ : the median difference equals zero. Rejection of this hypothesis indicates that one metric is consistently higher or lower than the other, revealing systematic measurement bias.

We also investigated whether the spatial characteristics of greenery within images influence green perception ratings. Specifically, we examined greenery distributions using spatial entropy measures and proximity analysis to test whether perceived greenness relates not only to the amount of greenery but also to its spatial arrangement. Images were divided into two groups based on their *green* perception Q scores: those within the 25th ($\leq$Q1, perceived as less green) and those within the 75th ($\geq$Q3, perceived as more green) among their respective image-participant location pairs. We applied the Mann-Whitney U test, a non-parametric alternative that does not assume normality, to compare spatial arrangements metrics between these groups. This analysis tests whether urban scenes perceived as highly green (*green* perception Q score $\geq$Q3) exhibit systematically different greenery spatial patterns – such as more dispersed vegetation (higher spatial entropy) or different peripheral distribution – compared to scenes perceived as less green (*green* Q score $\leq$Q1). Understanding these spatial-perceptual relationships could inform urban design strategies that optimize not just the quantity but also the arrangement of green elements to maximize positive greenery perception.

Finally, we trained a Random Forest model to predict *green* perception Q scores using both participant characteristics, image-derived features, and perception Q scores for the remaining indicators. The input features included participants' place of residence (whose pairwise comparisons generated the Q scores); along with image-based measures: Sky View Index, Green View Index, spatial entropy, and Shannon entropy; and the perception Q scores, calculated based on participants' place of residence, for the remaining nine indicators: *safe*, *lively*, *wealthy*, *beautiful*, *boring*, *depressing*, *live nearby*, *walk*, and *cycle*.

Features such as greenery proximity and image location were excluded due to their limited relevance to the our modeling objectives. Greenery proximity distributions could not be meaningfully reduced to single scalar metrics without substantial information loss, unlike the spatial entropy measures, which effectively capture spatial patterns in single values. We also excluded the image location (i.e., the city where the SVI was captured) because it functions largely as a proxy identifier; including it would risk data leakage by enabling the model to exploit city-level differences rather than the street-level visual features that this analysis aims to isolate.

Our modeling goal is not to compare greenery levels across cities but to examine how different participant groups (by country) respond to the same set of

urban scenes based on their visual characteristics. Random Forest was selected for its interpretability and suitability for tabular data. To ensure robust and geographically balanced evaluation, we applied an 80-20 train-test split stratified by image-participant location combinations.

After computing the perception Q scores, based on the participants' locations and a combined location 'All', for all 10 indicators on each image we retained scores with at least four pairwise comparisons, matching exiting thresholds (Quintana et al., 2025) and other urban visual datasets (Dubey et al., 2016) The final dataset resulted in 1,903 images or around 79% of the maximum 4,800 images (i.e., 400 images × (5 cities + 1 combined location)). Of these, over 297 images were rated by participants from single locations (i.e., cities), with 1,522 images used for training and 381 for testing. Model hyperparameters were optimized using 5-fold cross-validation on the training set. Detailed hyperparameter settings and training procedures are provided in Table A.3.

Although multicollinearity might be expected among image-derived features given their shared visual and spatial characteristics, its impact in Random Forest is minimised. Because of its algorithm design, Random Forests are relatively robust to multicollinearity as each tree is train on random subset of features, reducing the chances of correlated features being used together. Nevertheless, we further assess the model's performance with conditional permutation feature importance to account for potential multicollinearity. Unlike standard permutation importance where a feature is dropped independently to all others, conditional permutation feature importance (Strobl et al., 2008) permutes a feature while preserving its correlation structure with other features. This permutation analysis prevents inflated or deflated importance metrics and produces unbiased importance estimates in a multicollinearity setting.

4. Results

The results show that perceived greenery and GVI are moderately related but not interchangeable: correlations are often medium to moderate across locations, yet perception scores tend to exceed GVI values. Beyond quantity, the spatial arrangement and image-derived elements such as Sky View Index emerge as important drivers of perception, underscoring that how greenery is arranged within the streetscape matters alongside how much of it is present.

18
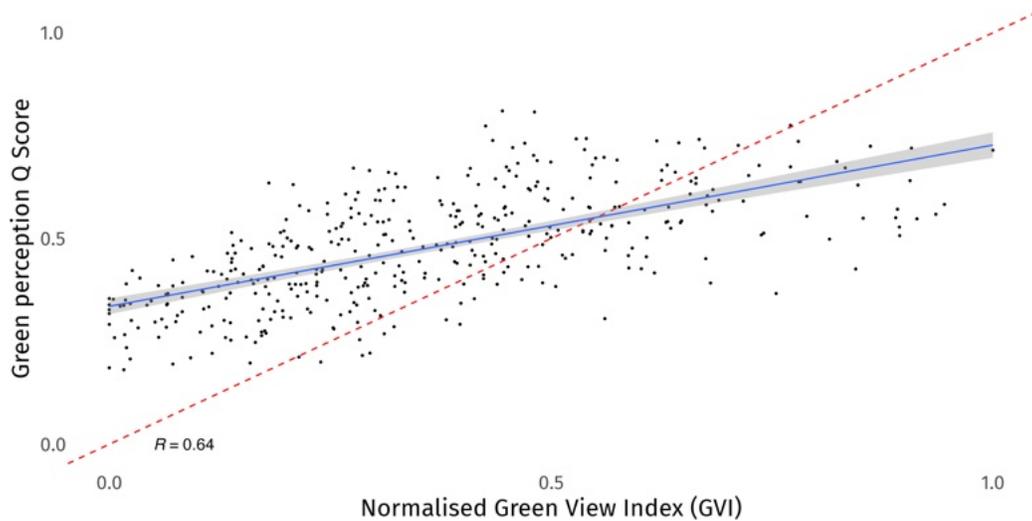
(a) Pearson correlation scatter plot between the *green* perception Q scores and normalised Green View Index (GVI) with the linear model in blue and an identity function as a red dashed line. The correlation is significant ($p < 0.5$) and all 400 images ($n = 400$) are used.

(b) Pearson correlation heatmap between the *green* perception Q scores and Green View Index (GVI) of images grouped by their location (x-axis) as well as all combined imagery (right-most column), rated by participants from different countries (y-axis). Most correlations are significant (bolded with a $*p < 0.05$) and with an R-value > 0.50. Each location pair (country-city pair) has at least 54 images with more than four pairwise comparisons ($n \geq 54$). The top row presents a summarised overview of the correlation of all participants rating images from particular locations and the value in the top right represent the correlation for all combinations in the dataset.

Figure 3: Pearson correlation between the *green* perception Q scores and Green View Index (GVI) using all images and responses (a) and grouped by the images' and participants location (b).

4.1. Correlation analysis

Figure 3 shows the overall correlations between *green* perception Q scores and normalised GVI (Figure 3a) and based on the images' and participants' location (Figure 3b). The analysis reveals moderate correlations overall and weak correlations within location pairs. The overall correlation is significant with an R-value of 0.64 and out of the 25 image-participant location pairs, all but one correlation is statistically significant, and in half of them (12), the R-value is ≥ 0.50. Most of the highest but still moderate correlations (R-value ranging from 0.43 to 0.65) are found on images from Abuja rated by participants across all countries, with the highest correlation (R-value of 0.70) on images from Amsterdam rated by participants from Chile. Weaker correlations (R value ranging from 0.19 to 0.40) are found on images from Singapore rated by participants across all countries. More specifically, the two weakest combinations are: participants from USA rating greenery from imagery in Singapore (0.19) and participants from Singapore rating greenery in images from there (0.28), while the two strongest pairs are participants from Chile rating greenery from imagery in Amsterdam (0.70) and participants from Singapore rating greenery from imagery in Abuja (0.65).

A more detailed Pearson's correlation analysis is presented as scatter plots in Figure B.11a and Figure B.11b. Unlike Figure 3b, in the latter figure, images are grouped solely based on either the images' location or the participants' location, not both. The analysis reveals relatively strong correlations between *green* perception Q scores and GVI. In Figure B.11b, correlations average R $\approx$ 0.62 for ratings given by all participants to imagery from each city (top row), and R $\approx$ 0.55 for ratings given by participants from each country to all imagery (bottom row).

Following Jiang et al. (2014), who found an inverse relationship between tree cover density and stress reduction for men at around 30% tree coverage, we examined regions with GVI < 0.3. in detail. We identified two distinct regions of interest. The green region captures imagery with moderate vegetation (GVI < 0.3) that is perceived as moderately green (*green* perception Q score > 0.5). The yellow region captures imagery with very little vegetation (GVI < 0.02, more than ten times lower than the green region GVI threshold) that is perceived as not particularly green (*green* perception Q score < 0.5).

For the green region, substantial variation exists across location pairs. In Figure B.11a, between 10% and 50% of images meet these criteria, with the lowest percentage for Santiago imagery rated by participants from Nigeria and the highest for Abuja imagery rated by participants from the Netherlands. When images are grouped by location (Figure B.11b, top row), the range is 9% to 41%, with Santiago showing the lowest and Abuja the highest percentages. However, when

grouped by participants' location (Figure B.11b, bottom row), the distribution is more consistent, with approximately one-fourth (24% to 27%) of images meeting these criteria across all five countries.

The yellow region shows different patterns. No imagery from Abuja or Singapore, and almost none from San Francisco, falls within this region regardless of participant location (Figure B.11a). In contrast, 13% to 17% of Amsterdam imagery consistently meets these criteria. When examining all participants' ratings (Figure B.11b, top row), the percentage varies dramatically by city location (0% in Singapore and Abuja, 19% in Amsterdam), while ratings by participants from specific countries show consistently low percentages (4% to 5%) across all cities (Figure B.11b, bottom row).

4.2. Bias evaluation

Besides the correlation between *green* perception Q scores and GVI values, and following the analysis based on images' and participants' locations, we developed a Bland-Altman plot or difference plot (Figure 4) to confirm the agreement between the two measurements. Using the Wilcoxon signed-rank test, we found the difference between *green* perception Q scores and normalized GVI values is predominantly different from zero ($p < 0.05$), suggesting the subjective and objective scores differ significantly across all location pairs, with a few noticeable exceptions. There are no significant measurement differences in imagery from Chile rated by participants from Santiago, nor in all imagery from Singapore across all participants ratings. This suggests a better alignment between participants respective *green* perception Q scores and these urban scenes' GVI. For the remaining image-participant location pairs, the direction of this difference is visualized in Figure 4 and results grouped instead by participants' and images' location separately in Figure B.12. These Bland-Altman plots show the vast majority of score differences, i.e., *greenery* perception Q score - normalized GVI, within two Standard Deviations (SD) of the mean (dashed orange line) and slightly more than have of them above the zero line (dotted gray line). This result suggests that, in most cases, *green* perception Q scores tends to be consistently higher than GVI values for the same urban scenes.

4.3. Greenery distribution

Figure 5 shows the greenery spatial entropy (a) and greenery depth (b) distributions with significant differences in more than half (16 out of 25) and almost all (20 out of 25) location pairs, respectively. On one hand, greenery spatial entropy
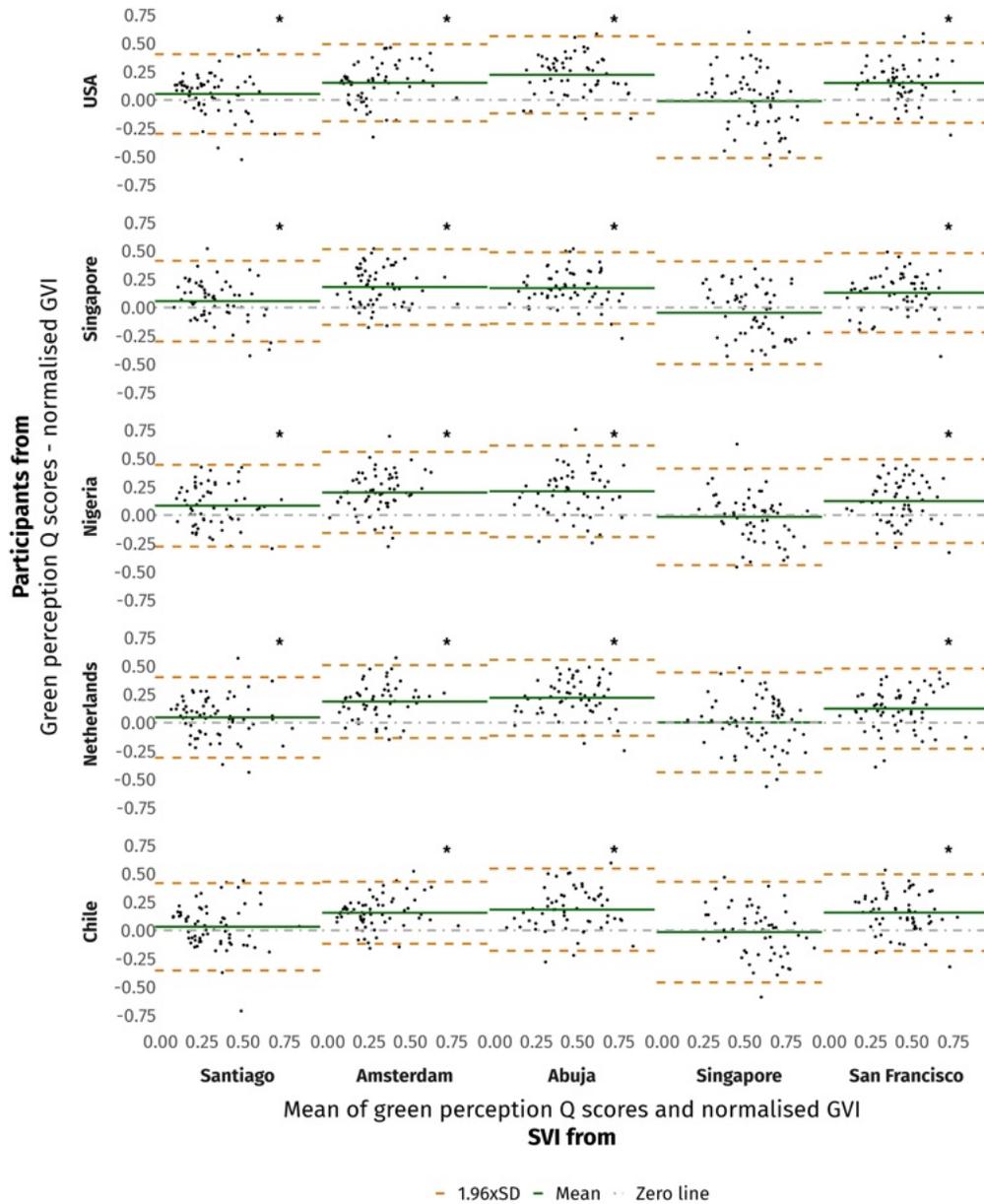
Figure 4: Bland-Altman plot comparing differences between *green* perception Q Scores and normalized Green View Index (GVI) values across different location pairs (country-city pairs). The plot shows the difference between measurements (y-axis) against their mean (x-axis) for all images and participants' location pairs, with green horizontal lines representing the mean difference, dashed orange lines indicating ±1.96 standard deviations, and a dotted gray line at zero. Each location pair has more than 54 images with more than four pairwise comparisons ($n \geq 54$). Location pairs with significant differences between *green* perception Q score and normalized GVI are shown (*$p < 0.05$).
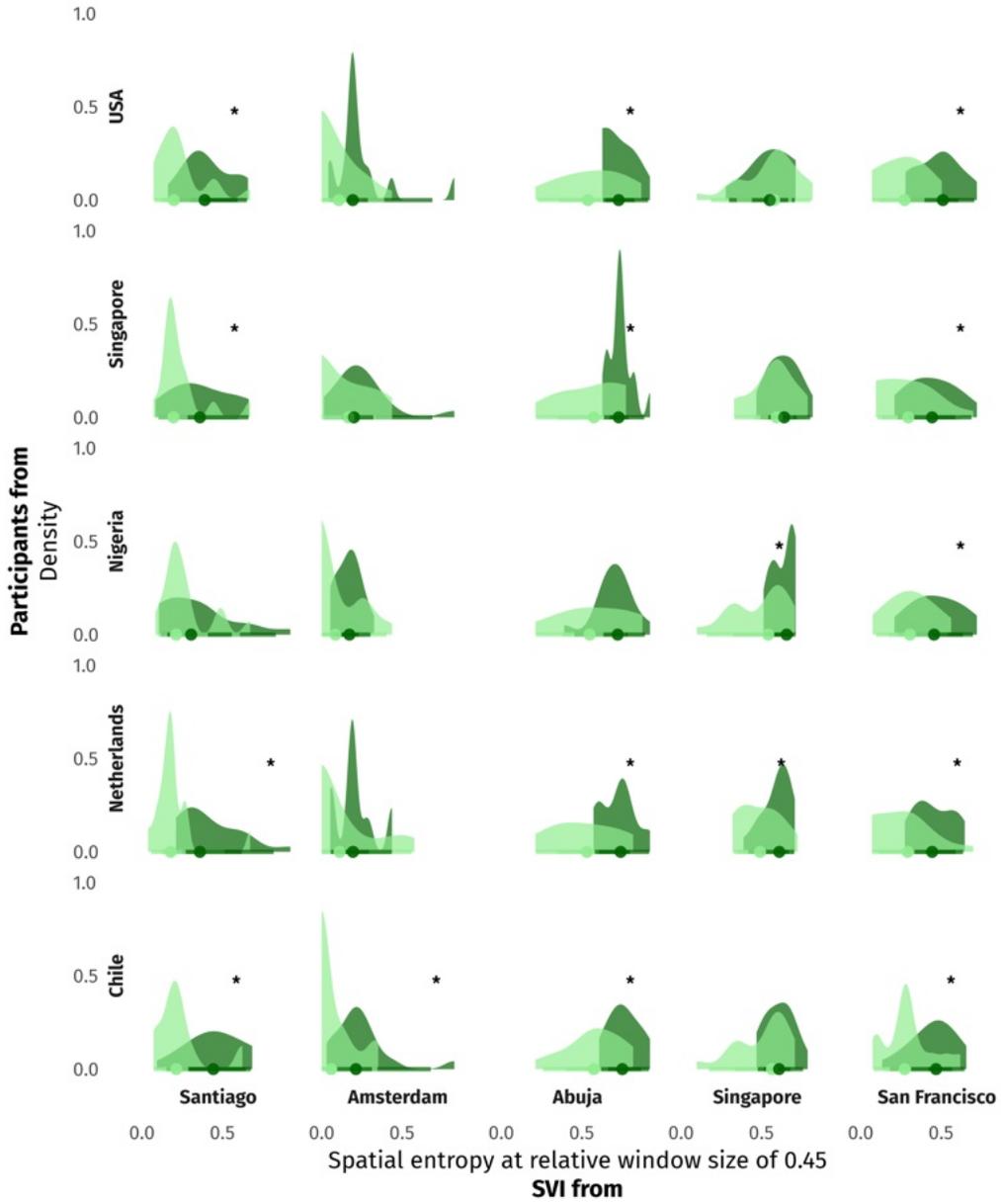
(a) Greenery spatial entropy distributions.

23
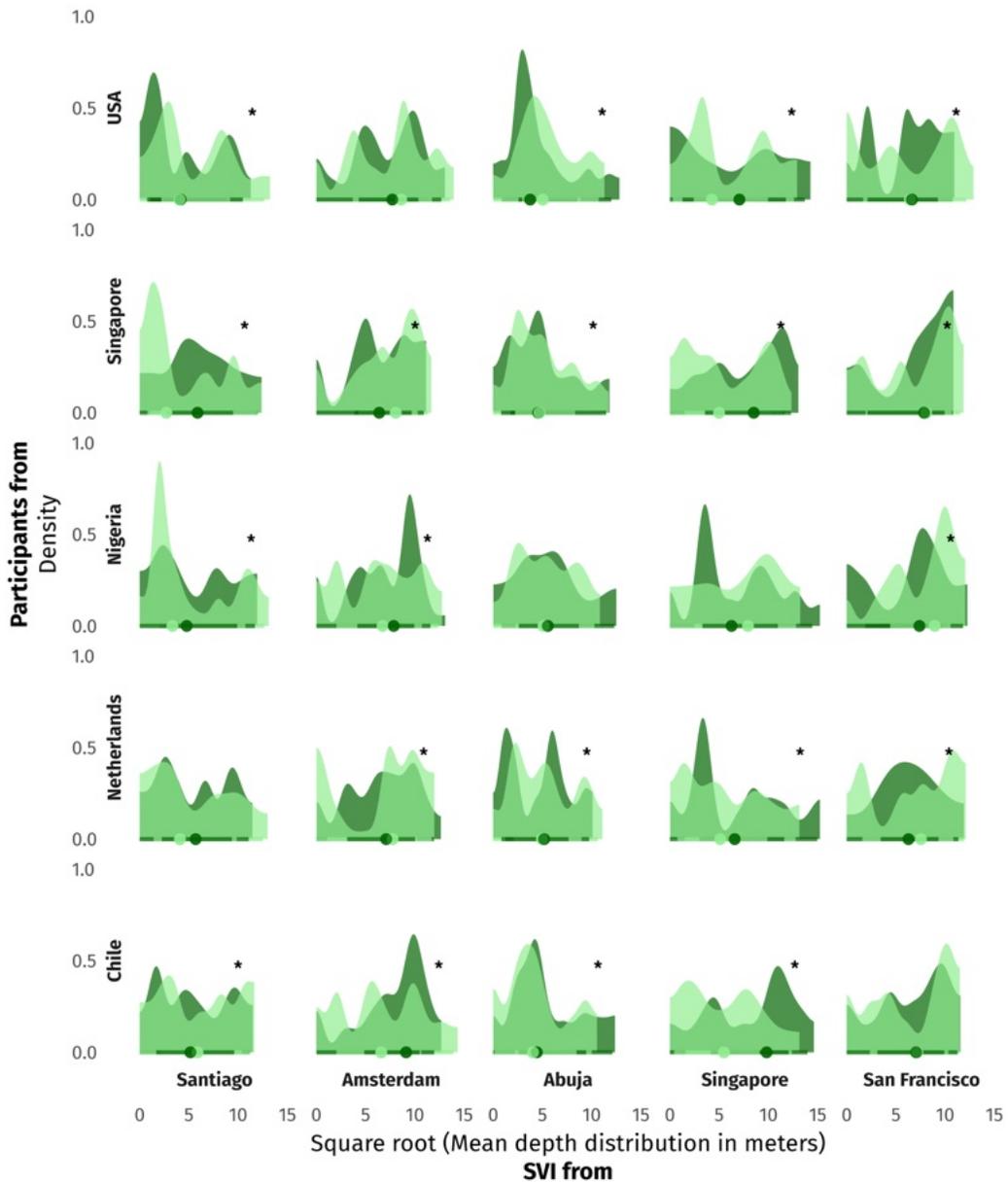
(b) Greenery absolute depth distributions (in square root for visualization purposes).

Figure 5: Greenery distributions and statistical comparison of street view imagery (SVI) within the 25th ($\leq$Q1 in lighter green) or within the 75th ($\geq$Q3 in darker green) percentile of their *green* perception Q scores. *Green* perception Q scores are calculated based on the images' and participants' location pairs (country-city pairs), with the median represented by a filled circle. Each location pair has more than 14 images with more than four pairwise comparisons ($n \geq 14$). Significance thresholds $*p < 0.05$.

values in images perceived as most green ($\geq$Q3) were significantly higher (Figure 5a) than in the images perceived as least green (darker green filled circle on the right side of the light green filled circle). This suggests that images perceived as very green typically have greenery distributed throughout the urban scene rather than concentrated in specific locations. On the other hand, greenery depth distributions showed wider spread and multiple peaks across all location pairs. Among the 20 location pairs with statistically significant differences in mean depth distances, eight showed higher mean distances in the least perceived *green* images ($\leq$Q1) while 12 showed higher mean distances in the most perceived *green* images ($\geq$Q3) (Figure 5b).

4.4. Greenery perception prediction

The best-performing Random Forest model was selected from 50 candidates with diverse hyperparameters (Table A.3). The model achieved low Mean Squared Error (MSE) and moderate predictive performance (i.e., 0.0191 MSE and 0.49 R^2, respectively) with comparable performance between 5-fold cross-validation and the test set, indicating good generalization without overfitting. To identify which features most strongly explain greenery perception, we conducted a conditional permutation feature importance analysis. This method measures the performance drop (increase in MSE) when individual feature values are randomly shuffled while preserving its correlation structure with other features, thereby quantifying each feature's contribution to model predictions while accounting for multicollinearity. Figure 6 presents the mean performance drop and standard deviation across multiple permutations for each feature. The results confirm that GVI is the strongest predictor of *green* perception Q scores. Additionally, image-derived features (spatial entropy, Shannon entropy, and Sky View Index) are among the most important features, followed by the perception Q scores for the *safe* perceptual indicator. Other positive perceptual indicators, such as *live nearby*, *beautiful*, *walk*, *cycle*, and *wealthy*, provide a marginal contribution while the remaining perceptual indicators and the participants' location have a negligible contribution to the prediction model. Complementing Figure 6, Table 2 shows an summary of the participants-based (demographics and personalities) and image-based (machine-derived metrics) features and their influence, or lack of it, in perceived greenery.

Group	Category / Variable	Influence
People's profiles	**Demographics**	
	Gender	✗
	Age group	✗
	Annual household income	✗
	Education level	✗
	Race and ethnicity	✗
	City of residence	✔
	Personalities	
	Extraversion	✗
	Agreeableness	✔
	Conscientiousness	✗
	Neuroticism	✗
	Openness	✗
Image	**Machine-derived**	
	Green View Index (GVI)	✔
	Spatial entropy	✔
	Proximity to viewer	✗
	Sky View Index	✗

Table 2: Summary of demographics (gender, age group, annual household income, education level, race and ethnicity, and city of residence), personalities (extraversion, agreeableness, conscientiousness, neuroticism, and openness), image-derived features (Green View Index (GVI), spatial entropy, proximity to viewer, and Sky View Index), and perceptual indicators (*safe*, *lively*, *wealthy*, *beautiful*, *boring*, *depressing*, *live nearby*, *walk*, and *cycle*) and their influence in perceived greenery evaluation (in Q scores) in street view imagery. ✔ = influence, ✗ = no influence.
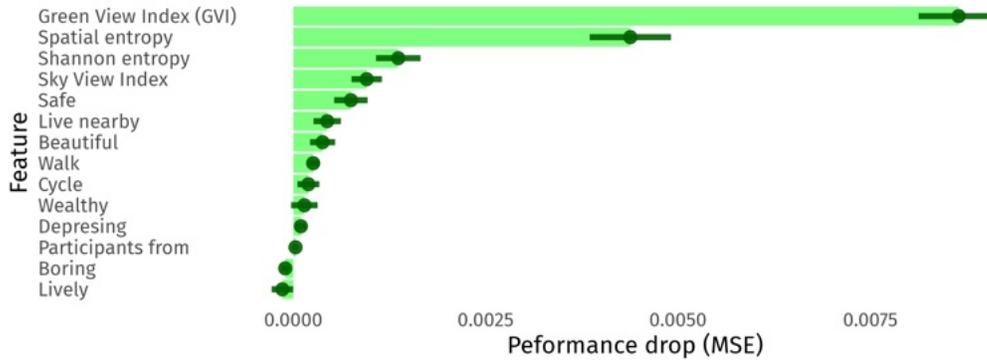
Figure 6: Conditional permutation feature importance analysis of the best performing Random Forest model. Mean and standard deviation of performance degradation (in Mean Squared Error (MSE), x-axis) when the values of a given feature (y-axis) are randomly shuffled while preserving its correlation structure with other features.

5. Discussion

5.1. How little is too little? Perceived versus measured greenery

As Figure 3b shows, perceived greenery and GVI values demonstrate strong agreement across nearly all images' and participants' location pairs. This strong agreement is underscored by GVI being the most influential feature when predicting *green* perception Q scores (Figure 6). However, participants often overestimated greenery compared to objective measurements (Figure 4), a finding that aligns with previous studies (Suppakittpaisarn et al., 2022; Torkko et al., 2023). Huang et al. (2025) found an underestimation of GVI when comparing to NDVI, attributing the discrepancy to differences in data collection methods, such as sparse sampling points to match satellite-imagery with street view imagery, and analyzable image aspects. In contrast, our study uses the same data source, i.e., SVI, for both subjective and objective measurements, suggesting that alternative sources of bias drive the observed differences. One potential explanation lies in the fundamental nature of our measurements: GVI provides absolute greenery values, while perception Q scores are inherently relative measures. As noted in our previous work (Quintana et al., 2025), relative scoring methods like Q scores and TrueSkill scores depend entirely on the specific rating pool used for their calculation. Consequently, these scores are only valid within their original rating context, i.e., within the set of images used.

Another possible, and complementary, explanation relates to the inherently nuanced nature of subjective evaluation. When assessing greenery, participants

may have incorporated contextual cues, such as spatial arrangement or perceived distance, beyond the visible amount of vegetation. This contextualization suggests that subjective greenery perception, unlike objective measures such as GVI, reflects a more holistic assessment informed by multiple contextual factors. Our correlation analysis aligns with this interpretation, showing systematic differences in how different populations perceive greenery. Imagery from Singapore consistently yielded the lowest correlations with GVI across all participant groups (Figure 3b), while participants from the USA exhibited the lowest correlations when rating imagery from any city. For example, the differences between normalized GVI and *green* perception Q scores for imagery from Singapore were not statistically significant, regardless of participants' location (Figure 4) even when all participants' responses were combined (Figure B.12, top row).

This finding suggests that both the urban context being evaluated and the cultural background of the evaluator influence the relationship between objective and subjective greenness measures. These patterns likely reflect location-specific perceptual baselines developed through extended urban exposure, as the location of the imagery is among the top three most influential features for predicting *green* perception Q scores (Figure 6). Given that the majority of participants have lived in their respective cities for more than five years (Quintana et al., 2025), their local urban characteristics may have become internalized reference points. For instance, participants from San Francisco and Santa Clara (USA) may not perceive dense urban scenes from Singapore as particularly green when compared to their urban greenery.

We noticed specific examples in three images – one each from Abuja, Singapore, and San Francisco – that were consistently rated as highly green (*green* perception Q score > 0.5) by all participants regardless of their residence, despite having low greenery coverage (GVI < 0.3) (Figure 2a, middle left image). These findings demonstrate that greenery perception can transcend measurable vegetation, likely incorporating other natural elements, built environment features, or contextual cues that contribute to an overall sense of 'greenery'. These results also align with research on biophilia – the positive emotional and sensory effects experienced when humans interact with nature. While Lefosse et al. (2025) found that people perceive nature differently depending on their geographic context, they confirmed that greenery remains the dominant factor in positive environmental perception. Our findings extend this understanding by demonstrating that perceived greenery operates as a more complex construct than objective vegetation metrics suggest, incorporating cultural, contextual, and possibly aesthetic factors that merit further investigation in urban planning and design applications.

28

5.2. Maybe it is not greener on the other side

Looking into spatial greenery placement, our analysis of greenery location reveals distinct patterns in how spatial arrangements influences greenery perception. Images perceived as most *green* (≥Q3) tend to feature greenery distributed throughout the scene rather than concentrated in specific areas (Figure 5a). In contrast, greenery proximity to the viewer showed a context-dependent relationship with *green* perception, demonstrating divergent patterns across location pairs (Figure 5b). For instance, observers (e.g., from the USA) viewing Abuja associated closer greenery with higher scores (aligning with the "proximity effect" described in Section 3.2.3), whereas perceptions of Singaporean imagery often favored greater depths, likely due to continuous green corridors that maintain visual impact even at a distance. To validate this finding, we conducted additional analysis comparing *green* perception scores between images with close greenery (median depth ≤Q1) and distant greenery (median depth ≥Q3) (Figure B.13). The perception rating distributions showed substantial overlap, indicating that greater distance does not inherently diminish perceived greenness. These findings nuance prior observations that distance has limited effect on perception Han (2021), suggesting instead that its role varies across different urban contexts.

Consistent with our *green* perception Q scores predictive modeling, spatial entropy emerged as the most influential factor after GVI, with other image-derived features ranking just below (Figure 6). Notably, greenery perception diverged from positive perceptual dimensions such as beauty or liveliness, which showed marginal and negligible predictive contribution, respectively. This disconnection may reflect the framing of the task: participants were not explicitly asked to rate how beautiful or lively the greenery in the urban scene is, but to rate the scene as a whole. We hypothesize that greenery judgments follow Gestalt principles, where vegetation is perceived holistically based on its overall spatial arrangements rather than its proximity to the viewer. This could explain why dispersed vegetation across the scene consistently received higher greenery ratings, even when located further from the camera.

5.3. Demographic and personality differences

We did not find significant differences in greenery perceptions across demographic groups (gender, age group, income, education level, and race and ethnicity) Quintana et al. (2025). This suggests that these demographic characteristics plays no role in perceiving greenery and that studies may be largely generalisable globally. Personality also has no influence, except in a few specific combinations,

e.g., there were differences between people with high and low agreeableness personality trait scores in participants from Singapore (Quintana et al., 2025). Moreover, despite our dataset being comprehensive (with respect to the state of the art), it still does not have enough responses per participant in nested demographics subgroups (e.g., responses from men in the 21-29 age group) to perform exhaustive demographic interaction analyses that are statistically significant. Our analysis revealed that imagery location and participants' geographic background (i.e., city of residence) significantly influence greenery perception, as the relationship between this perceptual indicator and objective greenery measurements varies by location. As such, ignoring residents' responses or using non-residents' responses for greenery subjective evaluation could lead to incorrect assessments.

Figure 7 illustrates differences in perceptions among participants from the same country when evaluating imagery from two different cities. Although both urban scenes contained very little greenery (GVI ≈ 0.04), participants disagreed on which was perceived as greener. Participants from the USA perceived the image from Abuja to be more *green*, while those from the Netherlands and Chile perceived the opposite, favoring the image from Amsterdam. In contrast, participants from Singapore and Nigeria considered both images comparable in perceived greenery. This finding suggests that, while subjective evaluation offers benefits such as providing a more holistic assessment through the lens of lived experiences (Chaney et al., 2024), perceived greenery should be considered with caution, as location-specific norms and culture contexts strongly influence such evaluations.

5.4. Recommendations

Based on our findings, we have the following recommendations when assessing street-level greenery:

GVI alone is not sufficient for evaluating greenery in human-centric applications. Our findings obtained from multiple, diverse cities affirm existing work highlighting that GVI values are not strongly correlated with perceived greenery. The former consistently underestimates the latter. In addition to the recommendations on GVI usage highlighted in (Torkko et al., 2023), researchers should be aware that relying solely on these measurements shifts the assessment away from a human-centric perspective that accounts for cultural, environmental, and experiential factors.

Subjective measurement of greenery is valuable but global models might miss location-specific responses. Our findings show that perceived greenery varies the most depending on the participants' location, which remains the only

Figure 7: Example of different *green* perception Q scores from two scenes from Abuja (left) and Amsterdam (right) between participants by their location. Both images have little greenery (GVI of 0.03 and 0.04, respectively) and are perceived as not too green (*green* perception Q score < 0.5). The individual street view image score, given by participants from a given country, is shown at the center of the figure next to the respective scene. The color gradient is calculated based on the delta of each score and the participants', grouped by location, average score, i.e., SVI Q score - average of scores of both images by participants from the same country. For example, participants from the US gave an average score of 0.27 to the images from Abuja and Amsterdam, which is 0.06 higher than the score they gave to Abuja and 0.06 lower than the score they gave to Amsterdam. The more similar the scores participants give to different scenes, the clearer the color gradient.

demographic factor that shapes perception of greenery. To ensure trustworthy results that reflect the perception of the local population, subjective assessment of greenery should be conducted by residents instead of global model predictions, i.e., prediction models trained on aggregated responses that are usually based on a geographically diverse pool of participants.

Arrangement over proximity in perceived greenery. Differences in perceived greenery were mainly found when imagery was grouped based on its greenery spatial arrangements rather than when imagery was grouped based on the vegetation distance from the viewer. Although this pattern may indicate that spatial arrangement influences subjective evaluations, our findings do not reflect participants' explicit preferences for particular configurations. Rather, they highlight a direction in which perceived greenery may be especially sensitive. Further work that directly examines preference for specific vegetation arrangements, as well as the type of greenery (Torkko et al., 2025; Korpilo et al., 2025), would help clarify which aspects of greenery most strongly shape subjective assessments.

5.5. Limitations and prospects for future work

Subjective perception Q scores and objective GVI measurements each have distinct data requirements that may limit their applicability across different urban contexts. Moreover, despite their widespread adoption, SVI is not a panacea and faces inherent limitations in data coverage and representativeness that could affect street-level greenery measurements (Fan et al., 2025). We inherit some of these limitations from our dataset (Quintana et al., 2025). Since the dataset was initially designed to evaluate ten perceptual indicators rather than focusing solely on greenery, the surveys were not optimized for demographic comparisons specific to greenery perception. Moreover, relative scores are commonly used to assess urban visual perception, making them susceptible to distribution shifts when calculated based on a different pool of participants. This means that while it is possible to find images with GVI values of 0 or 1, it is almost impossible to find images with a 0 or 1 *greenery* perceived Q score or Trueskill score. Additionally, while the image and city selection were stratified and conducted at a global scale, it did not explicitly focus on cities with different vegetation types. Although most participants (78%) had lived in their respective city for more than five years – suggesting familiarity with the city's broader environmental and urban-form context – intracity variation in greenery can be substantial. Participants may therefore reside in neighbourhoods that are significantly greener or less green than the citywide average, which could influence their perceptual baselines. To address this limitation, future work should examine greenery differences within countries and cities,

particularly those that are geographically large or environmentally diverse countries. The imagery was limited to specific weather conditions and times of day, which standardized the analyses but precluded any temporal or seasonal aspects of greenery evaluation. Because participants were not provided with a definition of the perceptual*green* indicator, their assessments likely reflected a holistic appraisal that integrated vegetation coverage alongside contextual cues such as colour tone, spatial arrangement, and visible sky. Future research should test the robustness of perceptual–objective relationships by analysing images at different seasonal conditions and climate contexts Torkko et al. (2025); Zhao et al. (2025).

Regarding the SVI selection within each city, our stratified sampling supports demographic representativeness within cities, but does not capture the geographical diversity needed to compare suburban, urban, or distinct land-use categories. Analyses on subjective and objective discrepancies across different urban contexts (e.g., urban vs. suburban areas or by land-use type) is a valuable direction for future research.

Within the images themselves, the vegetation color and saturation could impact GVI values. For example, trees without leaves in winter may still be classified as green after segmentation, but can differ significantly from a perceived score (e.g., Figure 1 in (Zhao et al., 2025)). Moreover, our GVI measurements were further constrained by the semantic segmentation model's performance and capabilities. Using models trained on different datasets could potentially produce varying segmentation results and thus different GVI values (Torkko et al., 2025). However, previous research across multiple datasets often reports a similar trend of GVI underestimating perceived greenery (Torkko et al., 2023; Leslie et al., 2010; Falfán et al., 2018), suggesting this limitation may be systematic rather than dataset-specific.

As highlighted by Torkko et al. (2025), future work should consider vegetation type diversity to analyze the differential impact of various forms of vegetation on perception. This vegetation type and spatial arrangements could be part of a modified GVI that could bridge the gap between objective and subjective greenery. Promising research directions include longitudinal analysis through time-series street view imagery to examine greenery changes in urban development (Liu et al., 2025), seasonal variation studies to capture temporal dynamics of vegetation perception (Zhao et al., 2025), or expand indices to for easy comparison between regions (Mahajan, 2024).

6. Conclusion

Our work advances research on human perception and urban greenery by examining the relationships and differences between perceived and measured greenery across multiple cities and explaining them through human-centric factors such as cultural, environmental, and experiential. Conducting a multi-city, demographically diverse survey, we identified a predominantly linear but moderate relationship between perceived and measured street-level greenery, confirming that there is a systematic discrepancy between how green a street feels to people and how green it is actually.

A key result of our study is that most demographic factors make almost no difference in how people perceive greenery, suggesting that survey results may be broadly generalisable and do not necessarily require a demographically balanced sample. However, we also show that place of residence – the only demographic factor that showed a measurable influence – plays a significant role in explaining perceptual variations. This influence suggests that cultural, environmental, and experiential contexts shape how individuals evaluate urban greenery, underscoring the importance of capturing residents' perceptions when designing such studies. Our findings also corroborate previous research showing that people's perception of greenery often overestimate objective measurements of greenery. Image analysis further revealed that residents notice the placement of greenery over mere proximity, suggesting that perception extends beyond simple vegetation presence or coverage. This confirms that greenery perception encompasses more than visible vegetation elements and involves nuanced qualitative assessments, not necessarily related to the aesthetics of the urban scene.

Further research and greenery-focused urban interventions should incorporate both objective and subjective measures, acknowledging their respective strengths and limitations. The finding that location trumps proximity suggests that investment in fewer, higher-quality green spaces may be more effective than simply maximizing green coverage, but further work is needed in understanding the type of greenery that influences people's perception the most.

CRediT authorship contribution statement

M.Q.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, **F.L.**: Methodology, Software, Data Curation, Writing - Review & Editing, **J.T.**: Methodology, Writing - Review &

Editing, **Y.G.** Methodology, Writing - Review & Editing, Visualization, **X.L.**: Writing - Review & Editing, Visualization, **Y.H.** : Data Curation, Writing - Review & Editing, **K.I.**: Methodology, Writing - Review & Editing, **Y.Z.**: Writing - Review & Editing, **M.A.**: Writing - Review & Editing, **T.T.**: Writing - Review & Editing, **Y.L.**: Methodology, Writing - Review & Editing, **F.B.**: Conceptualization, Methodology, Resources, Writing - Review & Editing, Visualization, Supervision, Funding acquisition.

Declaration of Competing Interesting

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 article.

Data availability

Survey responses and participants' demographic data (SPECS dataset) are released openly (https://github.com/matqr/specs) and the step-by-step process and code for all analyses are available in the public repository: https://github.com/matqr/greenery-perception.

Acknowledgements

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Claude for proofreading, assistance for code development, and the generation of urban scenes sketches for Figure 1. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Appendix A. Methods

| Street view image | Spectral segmentation | Semantic segmentation |

GVI = 0.1   GVI= 0.01

GVI = 0.24   GVI = 0.25

GVI = 0.23   GVI = 0.42
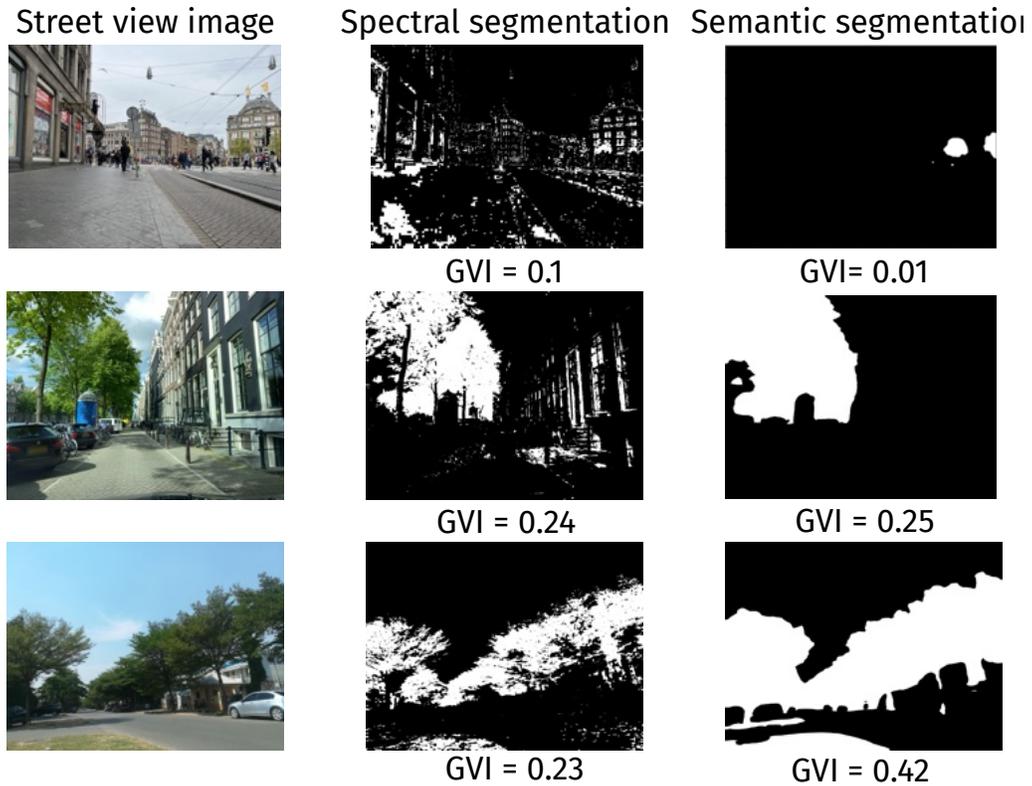
Figure A.8: Comparison of using spectral and semantic segmentation to mask out pixels in the street view image that represent vegetation. The Green View Index (GVI) is calculated following Equation 1 on each segmented image. Spectral segmentation is more susceptible to noise (middle column, top and bottom row) due to its dependency on pixel color-values and can lead to a 10× difference in GVI values compared to semantic segmentation (top row).
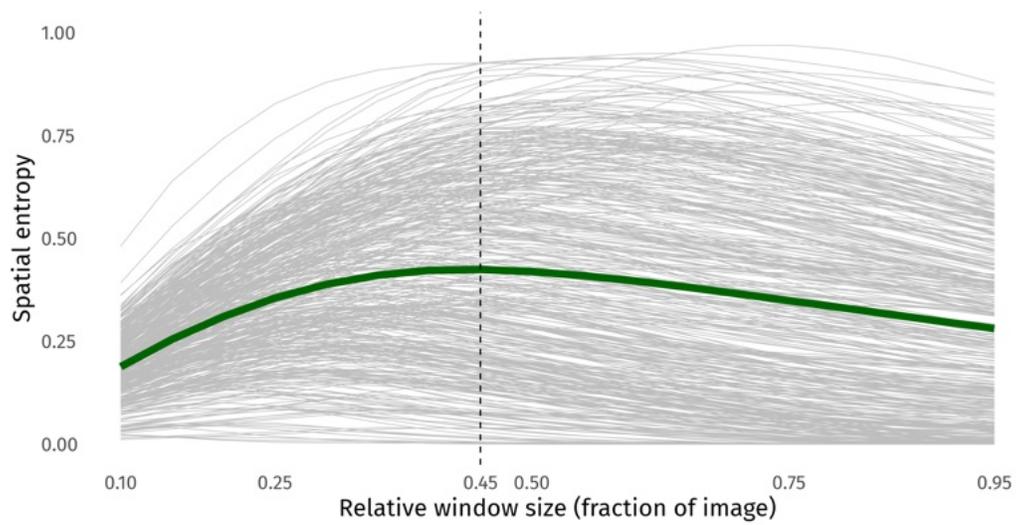
Figure A.9: Spatial entropy sensitivity analysis with different relative window sizes ([0.1, 1.0] in 0.05 increments) for all 400 images. The average line is highlighted in dark green and its maximum entropy is achieved at a relative window size of 0.45.
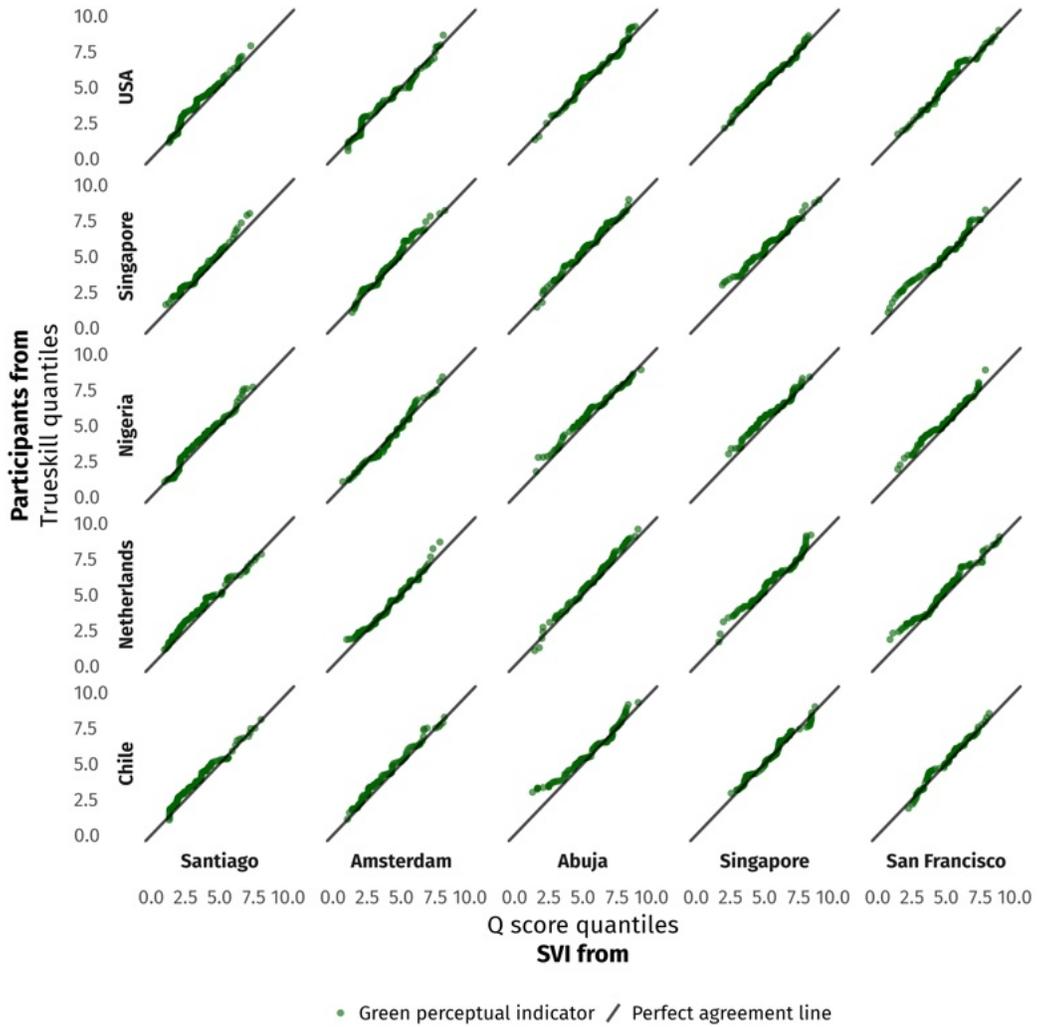
Figure A.10: Quantile-Quantile comparison between Q scores and Trueskill scores for the *green* perceptual indicators grouped by images' and participants' location pairs (country-city pairs). Each location pair has at least 54 images with more than four pairwise comparisons ($n \geq 54$).

(a) Hyperparameters	
n_estimators	100, 150, 200, ..., 500
max_depth	10, 20, 30, None
min_samples_split	2, 5, 10
min_samples_leaf	1, 2, 4
max_features	*sqrt*, *log2*, None

(b) Best performing model	
n_estimators	200
max_depth	None
min_samples_split	2
min_samples_leaf	2
max_features	None
$MSE_{cross-validation}$	0.0187
$MSE_{test\ set}$	0.0191
R^2	0.49

Table A.3: Overview of the Random Forest model training process and results. We used a 5-fold cross-validation for 50 candidates, based on a randomized combination of the available **(a)** hyperparameters, and reported the **(b)** best performing model based on Mean Squared Error (MSE).

Appendix B. Results
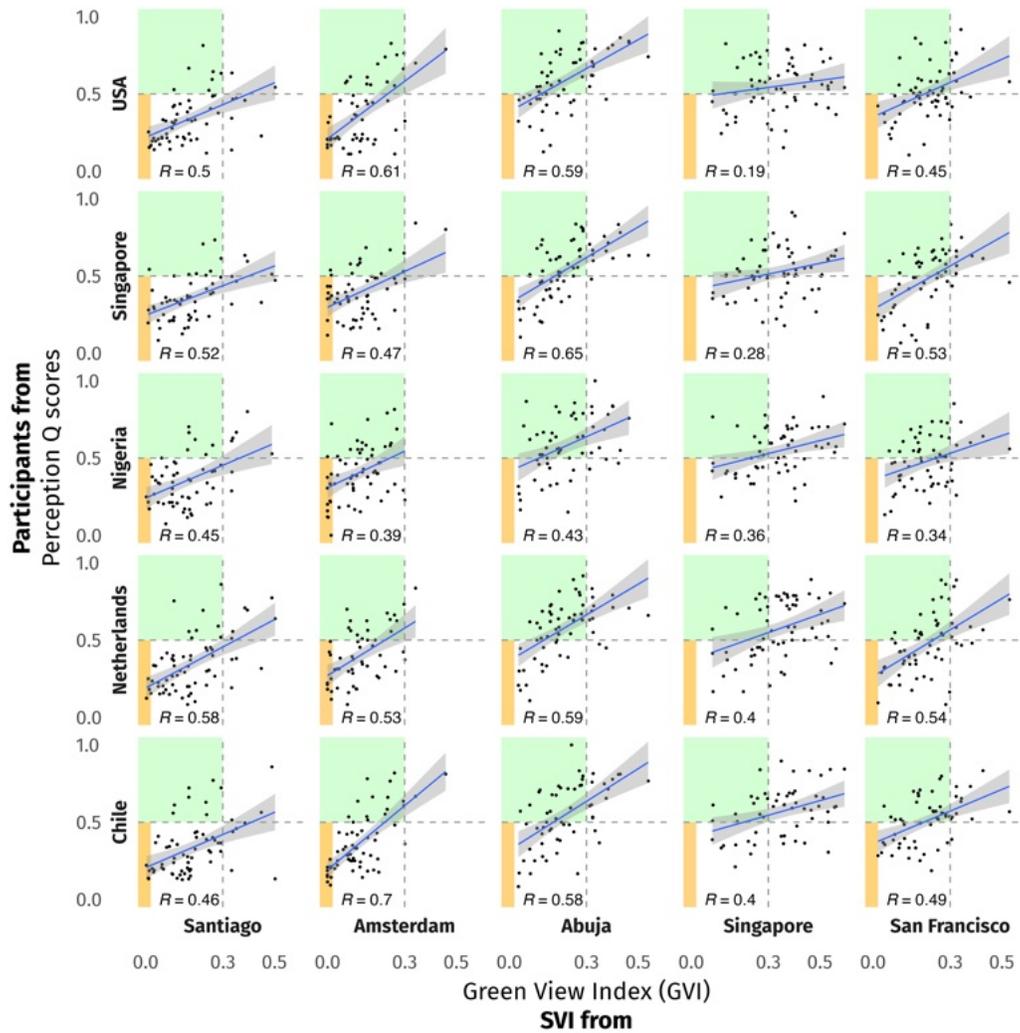
(a) *Green* perception Q scores are calculated based on the images' and participants' location pair. Each location pair has at least 54 images with more than four pairwise comparisons ($n \geq 54$).
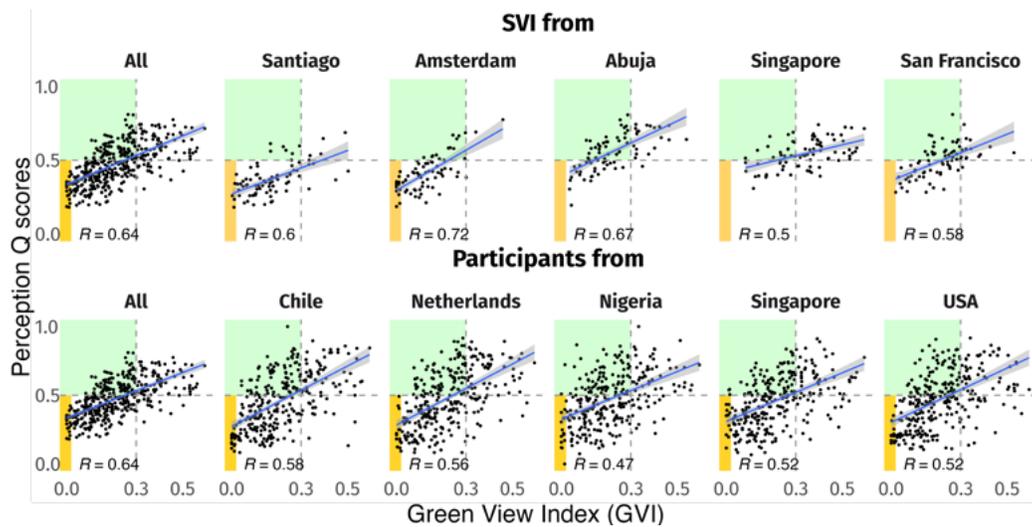
(b) *Green* perception Q scores are calculated based on the images' or participants' individual location. Each grouping has 80 images with more than four pairwise comparisons ($n = 80$) (top row, 'All' group $n = 400$) and at least 297 images with more than four pairwise comparisons ($n \geq 297$) (bottom row).

Figure B.11: Correlation between green perception Q scores and GVI values within images' and participants' location pairs (country-city pairs) (a) and images' or participants' individual location (b). Region in green captures the imagery with some vegetation (GVI< 0.3) perceived moderately green (*green* perception Q score > 0.5) and region in yellow captures the imagery with very little vegetation (GVI< 0.02) perceived as not so green (*green* perception Q score < 0.5).
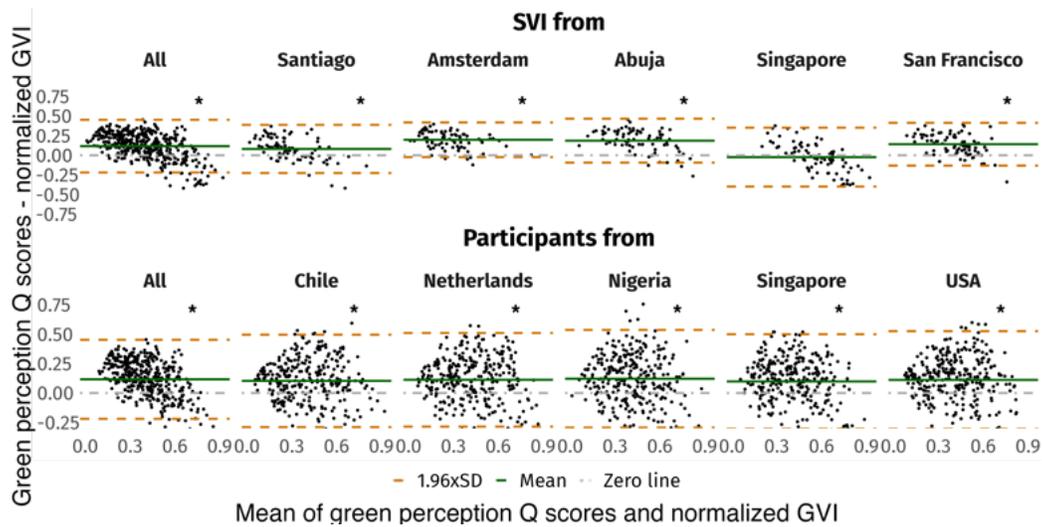
Figure B.12: Bland-Altman plot comparing differences between *green* perception Q Scores and normalized Green View Index (GVI) values across different location pairs (country-city pairs). The plot shows the difference between measurements (y-axis) against their mean (x-axis) grouped by images' and participants' location, with green horizontal lines representing the mean difference, dashed orange lines indicating ±1.96 standard deviations, and a dotted gray line at zero. Each grouping has 80 images with more than four pairwise comparisons ($n = 80$) (top row, 'All' group $n = 400$) and at least 297 images with more than four pairwise comparisons ($n \geq 297$) (bottom row). Plots with significant differences between green perception Q score and normalized GVI are shown (*$p < 0.05$).
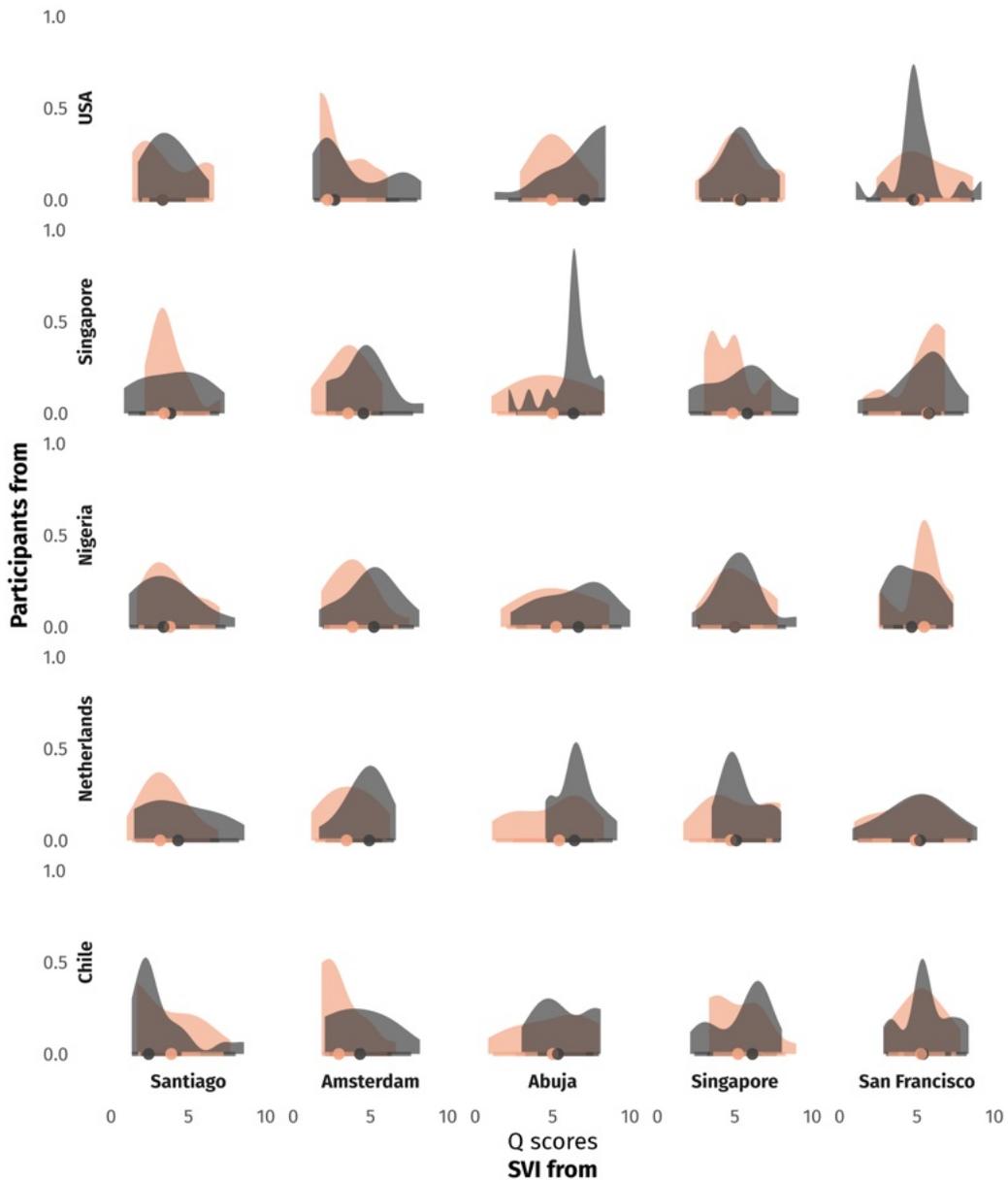
Figure B.13: *Green* perception Q score distributions of street view imagery (SVI) within the 25th ($\leq$Q1 in dark grey) or within the 75th ($\geq$Q3 in pink) percentile of their vegetation proximity in meters (median of depth values). Subjective Q scores are calculated based on the images' and participants' location pairs (country-city pairs), median represented by a filled circle. No significant differences between quantiles were found, and each location pair has more than 14 images with more than four pairwise comparisons ($n \geq 14$).

44

References

Aikoh, T., Homma, R., Abe, Y., 2023. Comparing conventional manual measurement of the green view index with modern automatic methods using google street view and semantic segmentation. Urban Forestry & Urban Greening 80, 127845. doi:10.1016/j.ufug.2023.127845.

Aoki, Y., 1991. Evaluation methods for landscapes with greenery. Landscape Research 16, 3–6. doi:10.1080/01426399108706344.

Batlle, J., Casals, A., Freixenet, J., Martí, J., 2000. A review on strategies for recognizing natural objects in colour images of outdoor scenes. Image and Vision Computing 18, 515–530. doi:10.1016/S0262-8856(99)00040-2.

Biljecki, F., Ito, K., 2023. Street view imagery in urban analytics and GIS: A review. Landscape and Urban Planning 215, 104217. doi:10.1016/j.landurbplan.2021.104217.

Biljecki, F., Zhao, T., Liang, X., Hou, Y., 2023. Sensitivity of measuring the urban form and greenery using street-level imagery: A comparative study of approaches and visual perspectives. International Journal of Applied Earth Observation and Geoinformation 122, 103385. doi:10.1016/j.jag.2023.103385.

Cao, Y., Yang, P., Xu, M., Li, M., Li, Y., Guo, R., 2025. A novel method of urban landscape perception based on biological vision process. Landscape and Urban Planning 254, 105246. doi:10.1016/j.landurbplan.2024.105246.

Chaney, R.A., Baer, A., Tovar, L.I., 2024. Gender-Based Heat Map Images of Campus Walking Settings: A Reflection of Lived Experience. Violence and Gender 11, 35–42. doi:10.1089/vio.2023.0027.

Chen, N., Wang, L., Xu, T., Wang, M., 2025. Perception of urban street visual color environment based on the CEP-KASS framework. Landscape and Urban Planning 259, 105359. doi:10.1016/j.landurbplan.2025.105359.

Cui, Q., Zhang, Y., Yang, G., Huang, Y., Chen, Y., 2023. Analysing gender differences in the perceived safety from street view imagery. International Journal of Applied Earth Observation and Geoinformation 124, 103537. doi:10.1016/j.jag.2023.103537.

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep Learning the City : Quantifying Urban Perception At A Global Scale. arXiv doi:10.48550/arxiv.1608.01769.

Falfán, I., Muñoz-Robles, C.A., Bonilla-Moheno, M., MacGregor-Fors, I., 2018. Can you really see 'green'? Assessing physical and self-reported measurements of urban greenery. Urban Forestry & Urban Greening 36, 13–21. doi:10.1016/j.ufug.2018.08.016.

Fan, Z., Feng, C.C., Biljecki, F., 2025. Coverage and bias of street view imagery in mapping the urban environment. Computers, Environment and Urban Systems 117, 102253. doi:10.1016/j.compenvurbsys.2025.102253.

Fujiwara, K., Ito, K., Ignatius, M., Biljecki, F., 2024a. A panorama-based technique to estimate sky view factor and solar irradiance considering transmittance of tree canopies. Building and Environment 266, 112071. doi:10.1016/j.buildenv.2024.112071.

Fujiwara, K., Khomiakov, M., Yap, W., Ignatius, M., Biljecki, F., 2024b. Microclimate Vision: Multimodal prediction of climatic parameters using street-level and satellite imagery. Sustainable Cities and Society 114, 105733. doi:10.1016/j.scs.2024.105733.

Gao, M., Cheng, X., Bao, Y., Zhou, X., . Beyond green: Unveiling the impact of urban park quality and greenery on children's physical activity. Environment and Planning B: Urban Analytics and City Science doi:10.1177/23998083241304258.

Gaw, L.Y., Chen, S., Chow, Y.S., Lee, K., Biljecki, F., 2022. Comparing street view imagery and aerial perspectives in the built environment. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-4/W3-2022, 49–56. doi:10.5194/isprs-annals-x-4-w3-2022-49-2022.

Gu, Y., Quintana, M., Liang, X., Ito, K., Yap, W., Biljecki, F., 2025. Designing effective image-based surveys for urban visual perception. Landscape and Urban Planning 260, 105368. doi:10.1016/j.landurbplan.2025.105368.

Gupta, K., Kumar, P., Pathan, S., Sharma, K., 2012. Urban Neighborhood Green Index – A measure of green spaces in urban areas. Landscape and Urban Planning 105, 325–335. doi:10.1016/j.landurbplan.2012.01.003.

Han, K.T., 2021. Effects of visible greenness, quantity and distance of indoor plants on human perceptions and physical parameters. Indoor and Built Environment 30, 1353–1372. doi:10.1177/1420326X20939595.

He, Q., Wu, L., Lee, C.S., Zhu, C., Bai, W., Guo, W., Ye, X., 2025. Greener the safer? Effects of urban green space on community safety and perception of safety using satellite and street view imagery data. Journal of Criminal Justice 97, 102372. doi:10.1016/j.jcrimjus.2025.102372.

Helbich, M., Poppe, R., Oberski, D., Emmichoven, M.Z.v., Schram, R., 2021. Can't see the wood for the trees? An assessment of street view- and satellite-derived greenness measures in relation to mental health. Landscape and Urban Planning 214, 104181. doi:10.1016/j.landurbplan.2021.104181.

Herbrich, R., Minka, T., Graepel, T., 2006. TrueSkill™: a Bayesian skill rating system, in: Proceedings of the 19th International Conference on Neural Information Processing Systems, Cambridge, MA, USA. pp. 569–576.

Hidayati, I., Tan, W., Yamu, C., 2020. How gender differences and perceptions of safety shape urban mobility in Southeast Asia. Transportation Research Part F: Traffic Psychology and Behaviour 73, 155–173. doi:10.1016/j.trf.2020.06.014.

Hou, Y., Quintana, M., Khomiakov, M., Yap, W., Ouyang, J., Ito, K., Wang, Z., Zhao, T., Biljecki, F., 2024. Global Streetscapes — A comprehensive dataset of 10 million street-level images across 688 cities for urban science and analytics. ISPRS Journal of Photogrammetry and Remote Sensing 215, 216–238. doi:10.1016/j.isprsjprs.2024.06.023.

Huang, Y., Sanatani, R.P., Liu, C., Kang, Y., Zhang, F., Liu, Y., Duarte, F., Ratti, C., 2025. No "true" greenery: Deciphering the bias of satellite and street view imagery in urban greenery measurement. Building and Environment 269, 112395. doi:10.1016/j.buildenv.2024.112395.

Ito, K., Zhu, Y., Abdelrahman, M., Liang, X., Fan, Z., Hou, Y., Zhao, T., Ma, R., Fujiwara, K., Ouyang, J., Quintana, M., Biljecki, F., 2025a. ZenSVI: An open-source software for the integrated acquisition, processing and analysis of street view imagery towards scalable urban science. Computers, Environment and Urban Systems 119, 102283. doi:10.1016/j.compenvurbsys.2025.102283.

Ito, K., Zhu, Y., Abdelrahman, M., Liang, X., Fan, Z., Hou, Y., Zhao, T., Ma, R., Fujiwara, K., Ouyang, J., Quintana, M., Biljecki, F., 2025b. Zensvi: An open-source software for the integrated acquisition, processing and analysis of street view imagery towards scalable urban science. Computers, Environment and Urban Systems 119, 102283. doi:10.1016/j.compenvurbsys.2025.102283.

Jiang, B., Chang, C.Y., Sullivan, W.C., 2014. A dose of nature: Tree cover, stress reduction, and gender differences. Landscape and Urban Planning 132, 26–36. doi:10.1016/j.landurbplan.2014.08.005.

Kang, Y., Abraham, J., Ceccato, V., Duarte, F., Gao, S., Ljungqvist, L., Zhang, F., Näsman, P., Ratti, C., 2023. Assessing differences in safety perceptions using GeoAI and survey across neighbourhoods in Stockholm, Sweden. Landscape and Urban Planning 236, 104768. doi:10.1016/j.landurbplan.2023.104768.

Kang, Y., Zhang, F., Gao, S., Lin, H., Liu, Y., 2020. A review of urban physical environment sensing using street view imagery in public health studies. Annals of GIS 26, 261–275. doi:10.1080/19475683.2020.1791954.

Ki, D., Lee, S., 2021a. Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. Landscape and Urban Planning 205, 103920. doi:10.1016/j.landurbplan.2020.103920.

Ki, D., Lee, S., 2021b. Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. Landscape and Urban Planning 205, 103920. doi:10.1016/j.landurbplan.2020.103920.

Korpilo, S., Willberg, E., Müürisepp, K., Klein, R., Taylor, R., Torkko, J., Hasanzadeh, K., Toivonen, T., 2025. Restoring nature, enhancing active mobility: The role of street greenery in the EU's 2024 restoration law. Ambio doi:10.1007/s13280-025-02178-w.

Labib, S., Huck, J.J., Lindley, S., 2021. Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. Science of The Total Environment 755, 143050. doi:10.1016/j.scitotenv.2020.143050.

Lefosse, D.C., Duarte, F., Sanatani, R.P., Kang, Y., van Timmeren, A., Ratti, C., 2025. Feeling Nature: Measuring perceptions of biophilia across global biomes using visual AI. npj Urban Sustainability 5, 4. doi:10.1038/s42949-025-00192-1.

Leslie, E., Sugiyama, T., Ierodiaconou, D., Kremer, P., 2010. Perceived and objectively measured greenness of neighbourhoods: Are they measuring the same thing? Landscape and Urban Planning 95, 28–33. doi:10.1016/j.landurbplan.2009.11.002.

Li, W., Sun, R., He, H., Yan, M., Chen, L., 2024. Perceptible landscape patterns reveal invisible socioeconomic profiles of cities. Science Bulletin 69, 3291–3302. doi:10.1016/j.scib.2024.06.022.

Li, X., Zhang, C., Li, W., Kuzovkina, Y.A., Weiner, D., 2015a. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. Urban Forestry & Urban Greening 14, 751–759. doi:10.1016/j.ufug.2015.07.006.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015b. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban Forestry & Urban Greening 14, 675–685. doi:10.1016/j.ufug.2015.06.006.

Liang, X., Chang, J.H., Gao, S., Zhao, T., Biljecki, F., 2024. Evaluating human perception of building exteriors using street view imagery. Building and Environment 263, 111875. doi:10.1016/j.buildenv.2024.111875.

Liu, D., Kwan, M.P., Yang, Z., Kan, Z., 2024. Comparing subjective and objective greenspace accessibility: Implications for real greenspace usage among adults. Urban Forestry & Urban Greening 96, 128335. doi:https://doi.org/10.1016/j.ufug.2024.128335.

Liu, Y., Wang, Z., Ren, S., Chen, R., Shen, Y., Biljecki, F., 2025. Physical urban change and its socio-environmental impact: Insights from street view imagery. Computers, Environment and Urban Systems 119, 102284. doi:10.1016/j.compenvurbsys.2025.102284.

Long, Y., Liu, L., 2017. How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. PLOS ONE 12, e0171110. doi:10.1371/journal.pone.0171110.

Lottrup, L., Grahn, P., Stigsdotter, U.K., 2013. Workplace greenery and perceived level of stress: Benefits of access to a green outdoor environment at the workplace. Landscape and Urban Planning 110, 5–11. doi:10.1016/j.landurbplan.2012.09.002.

Lu, Y., 2019. Using Google Street View to investigate the association between street greenery and physical activity. Landscape and Urban Planning 191, 103435. doi:10.1016/j.landurbplan.2018.08.029.

Lu, Y., Sarkar, C., Xiao, Y., 2018. The effect of street-level greenery on walking behavior: Evidence from Hong Kong. Social Science & Medicine 208, 41–49. doi:10.1016/j.socscimed.2018.05.022.

Ma, Y., Chen, P., Qin, Y., Yang, Z., Li, S., 2025a. From sky to ground: Monitoring visible street greenery via multisource remote sensing imagery with deep learning. Urban Forestry & Urban Greening 109, 128866. doi:10.1016/j.ufug.2025.128866.

Ma, Y., Chen, P., Qin, Y., Yang, Z., Li, S., 2025b. From sky to ground: Monitoring visible street greenery via multisource remote sensing imagery with deep learning. Urban Forestry & Urban Greening 109, 128866. doi:10.1016/j.ufug.2025.128866.

MacFaden, S.W., O'Neil-Dunne, J.P.M., Royar, A.R., Lu, J.W.T., Rundle, A.G., 2012. High-resolution tree canopy mapping for New York City using LIDAR and object-based image analysis. Journal of Applied Remote Sensing 6, 063567. doi:10.1117/1.JRS.6.063567.

Mahajan, S., 2024. greenR: An open-source framework for quantifying urban greenness. Ecological Indicators 163, 112108. doi:10.1016/j.ecolind.2024.112108.

Mao, Y., Xia, T., Hu, F., Hu, J., He, Y., Yan, J., Wang, L., Xu, H., Zhang, J., Chen, D., 2025. Greener is not always better: Exploring the non-linear relationships between three-dimensional green and gray spaces exposure and various physical activities. Building and Environment 272, 112654. doi:10.1016/j.buildenv.2025.112654.

Mayer, K., Haas, L., Huang, T., Bernabé-Moreno, J., Rajagopal, R., Fischer, M., 2023. Estimating building energy efficiency from street view imagery, aerial imagery, and land surface temperature data. Applied Energy 333, 120542. doi:10.1016/j.apenergy.2022.120542.

Meir, A., Oron-Gilad, T., 2020. Understanding complex traffic road scenes: The case of child-pedestrians' hazard perception. Journal of Safety Research 72, 111–126. doi:10.1016/j.jsr.2019.12.014.

Merdymshaeva, E., Mora, S., Machida, Y., Zhang, F., Duarte, F., Paul, S., Ratti, C., Mörtberg, U., 2025. Assessing the impact of greenery on urban heat using opportunistic drive-by sensing. Environment and Planning B: Urban Analytics and City Science doi:10.1177/23998083241311068.

Neuhold, G., Ollmann, T., Bulò, S.R., Kontschieder, P., 2017. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. 2017 IEEE International Conference on Computer Vision (ICCV) , 5000–5009doi:10.1109/iccv.2017.534.

Ogawa, Y., Oki, T., Zhao, C., Sekimoto, Y., Shimizu, C., 2024. Evaluating the subjective perceptions of streetscapes using street-view images. Landscape and Urban Planning 247, 105073. doi:10.1016/j.landurbplan.2024.105073.

Pietikainen, M., Nieminen, S., Marszalec, E., Ojala, T., 1996. Accurate color discrimination with classification based on feature distributions, in: Proceedings of 13th International Conference on Pattern Recognition, IEEE. pp. 833–838 vol.3. doi:10.1109/ICPR.1996.547285.

Qi, J., Ding, L., Lim, S., 2022. A decision-making framework to support urban heat mitigation by local governments. Resources, Conservation and Recycling 184, 106420. doi:10.1016/j.resconrec.2022.106420.

Qiu, W., Zhang, Z., Liu, X., Li, W., Li, X., Xu, X., Huang, X., 2022. Subjective or objective measures of street environment, which are more effective in explaining housing prices? Landscape and Urban Planning 221, 104358. doi:10.1016/j.landurbplan.2022.104358.

Qiu, Y., Wu, M., Huang, Q., Kang, Y., 2025. Do You Know Your Neighborhood? Integrating Street View Images and Multi-task Learning for Fine-Grained Multi-Class Neighborhood Wealthiness Perception Prediction. Cities 158, 105703. doi:10.1016/j.cities.2025.105703.

Quintana, M., Gu, Y., Liang, X., Hou, Y., Ito, K., Zhu, Y., Abdelrahman, M., Biljecki, F., 2025. Global urban visual perception varies across demographics and personalities. Nature Cities 2, 1092–1106. doi:10.1038/s44284-025-00330-x.

Salminen, E.A., Hausner, V.H., Murguzur, F.J.A., Engen, S., 2025. Optimizing recruitment in an online environmental PPGIS—is it worth the time and costs? International Journal of Geographical Information Science ahead-of-print, 1–23. doi:10.1080/13658816.2024.2427267.

Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 307. doi:10.1186/1471-2105-9-307.

Sugiyama, T., Leslie, E., Giles-Corti, B., Owen, N., 2008. Associations of neighbourhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships? Journal of Epidemiology & Community Health 62, e9–e9. doi:10.1136/jech.2007.064287.

Suppakittpaisarn, P., Lu, Y., Jiang, B., Slavenas, M., 2022. How do computers see landscapes? comparisons of eye-level greenery assessments between computer and human perceptions. Landscape and Urban Planning 227, 104547. doi:10.1016/j.landurbplan.2022.104547.

Taylor, A.F., Kuo, F.E., Sullivan, W.C., 2002. Views of nature and self-discipled: Evidence from inner city children. Journal of Environmental Psychology 22, 49–63. URL: https://linkinghub.elsevier.com/retrieve/pii/S0272494401902415, doi:10.1006/jevp.2001.0241.

Torkko, J., Malekzadeh, M., Willberg, E., Toivonen, T., 2025. Same greenery, different Green View Index: Inconsistent greenery assessment in semantic segmentation and street view imagery. Urban Forestry & Urban Greening 112, 128969. doi:10.1016/j.ufug.2025.128969.

Torkko, J., Poom, A., Willberg, E., Toivonen, T., 2023. How to best map greenery from a human perspective? Comparing computational measurements with human perception. Frontiers in Sustainable Cities 5, 1160995. doi:10.3389/frsc.2023.1160995.

Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., von der Heydt, R., 2012. A century of gestalt psychology in visual perception:

I. Perceptual grouping and figure-ground organization. Psychological Bulletin 138, 1172–1217. doi:10.1037/a0029333.

Wang, R., Yao, Y., 2025. Exploring the pathways linking visual green space to depression in older adults in Shanghai, China: Using street view data. Aging & Mental Health 29, 78–86. doi:10.1080/13607863.2024.2363370, arXiv:38940438.

Wu, C., Du, Y., Li, S., Liu, P., Ye, X., 2022. Does visual contact with green space impact housing prices? An integrated approach of machine learning and hedonic modeling based on the perception of green space. Land Use Policy 115, 106048. doi:10.1016/j.landusepol.2022.106048.

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. doi:10.1016/j.landurbplan.2008.12.004.

Yang, L., Kang, B., Huang, Z., Xu, X., Feng, J., Zhao, H., 2024. Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data. doi:10.48550/arXiv.2401.10891, arXiv:2401.10891.

Yap, W., Stouffs, R., Biljecki, F., 2023. Urbanity: automated modelling and analysis of multidimensional networks in cities. npj Urban Sustainability 3, 45. doi:10.1038/s42949-023-00125-w.

Ye, Y., Richards, D., 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 191, 103434. doi:10.1016/j.landurbplan.2018.08.028.

Yu, S., Yu, B., Song, W., Wu, B., Zhou, J., Huang, Y., Wu, J., Zhao, F., Mao, W., 2016. View-based greenery: A three-dimensional assessment of city buildings' green visibility using Floor Green View Index. Landscape and Urban Planning 152, 13–26. doi:10.1016/j.landurbplan.2016.04.004.

Zhang, W., Zeng, H., 2024. Spatial differentiation characteristics and influencing factors of the green view index in urban areas based on street view images: A case study of Futian District, Shenzhen, China. Urban Forestry & Urban Greening 93, 128219. doi:10.1016/j.ufug.2024.128219.

Zhao, T., Liang, X., Biljecki, F., Tu, W., Cao, J., Li, X., Yi, S., 2025. Quantifying seasonal bias in street view imagery for urban form assessment: A global analysis of 40 cities. Computers, Environment and Urban Systems 120, 102302. doi:10.1016/j.compenvurbsys.2025.102302.

Zięba-Kulawik, K., Skoczylas, K., Wężyk, P., Teller, J., Mustafa, A., Omrani, H., 2021. Monitoring of urban forests using 3D spatial indices based on LiDAR point clouds and voxel approach. Urban Forestry & Urban Greening 65, 127324. doi:10.1016/j.ufug.2021.127324.