ENVIRONMENTAL RESEARCH
INFRASTRUCTURE AND SUSTAINABILITY

PURPOSE-LED PUBLISHING™

Part of the ENVIRONMENTAL RESEARCH SERIES

PAPER • OPEN ACCESS

Tracking green space along streets of world cities

To cite this article: Giacomo Falchetta and Ahmed T Hammad 2025 *Environ. Res.: Infrastruct. Sustain.* **5** 025011

IOP Publishing Environ. Res.: Infrastruct. Sustain. 5 (2025) 025011 https://doi.org/10.1088/2634-4505/add9c4

ENVIRONMENTAL RESEARCH
INFRASTRUCTURE AND SUSTAINABILITY

CrossMark

OPEN ACCESS

RECEIVED
19 March 2025

REVISED
14 April 2025

ACCEPTED FOR PUBLICATION
16 May 2025

PUBLISHED
27 May 2025

PAPER

Tracking green space along streets of world cities

Giacomo Falchetta[1,2,*] **and Ahmed T Hammad**[3]

[1] International Institute for Applied Systems Analysis, Laxenburg, Austria
[2] Centro Euro-Mediterraneo sui Cambiamenti Climatici—RFF-CMCC European Institute for Economics and the Environment, Venice, Italy
[3] Decatab PTE LTD, Singapore, Singapore
[*] Author to whom any correspondence should be addressed.

E-mail: falchetta@iiasa.ac.at

Keywords: street green space, green view index, sustainable cities, environmental justice, remote sensing, machine learning

Supplementary material for this article is available online

Abstract

Street green space (SGS) - the presence of vegetation along streets of cities—is a key piece of urban infrastructure. SGS provides a broad range of functions, such as mitigating the urban heat island effect, reducing the impact of extreme precipitation events, and supporting human and animal well-being. Here we introduce an approach to estimate SGS based on the statistical modeling of a street-based indicator of canopy coverage (the green view index, GVI) with multispectral satellite observations and ancillary spatially granular data. Based on our trained and cross-validated non-parametric model, we conduct spatial sampling and prediction in 190 large cities distributed across twenty regions and estimate local to continental GVI trends between 2016–2023. Jointly considering such global pool of cities, we find evidence of a trend of GVI decrease of 0.3%–0.5% per year ($p < 0.01$). Yet, both the direction and magnitude of trends show high heterogeneity across and within regions and cities, which we explore, along with stark inequalities in SGS availability within each city. Our analysis provides an updated estimate of the GVI as a measure of SGS across a global pool of cities and an open-source, validated approach to assess its future changes and support the design of policies for sustainable cities.

1. Introduction

While appreciated already in ancient times [1], urban vegetation is being considered ever more crucial in the context of global urbanization and growing anthropogenic impact of human and natural systems [2–4], also in the ambition of the Sustainable Development Goal 11 of 'making cities inclusive, safe, resilient and sustainable' (https://sdgs.un.org/goals/goal11). Urban vegetation provides a broad range of ecosystem services [5]. For instance, urban vegetation is known to be an important factor that affects the effect of urban heat islands (e.g. through shade [6] and evaporative cooling [7, 8]) and therefore the active air cooling energy requirements in buildings [8–14]. Other important services include the flood control function [15, 16], carbon sequestration [17, 18] and biodiversity preservation [19]. Urban vegetation is also a source of intangible benefits for human health, psychological well-being [20–23] and happiness [24, 25] and it is positively associated with urban real estate prices [26]. Overall, because of this broad range of functions, the World Health Organization [27] recommends a minimum of 9 m^2 of urban vegetation-covered area per person with an ideal value of 50 m^2.

Street green space (SGS)-the presence of vegetation along public streets of cities—is a major contributor to urban green space (UGS). The key role of SGS and its canopy coverage is related to its general proximity to buildings and people [28] and its public nature [29]. Research demonstrated that it is mainly the density of SGS within the city core that drives the perception of greenness of a city, rather than large peripheral parks [30–32].

Previous studies have analyzed urban vegetation at different levels of geographical coverage [33–36]. For instance, at the global and continental scales some large-scale assessments based have been carried out based

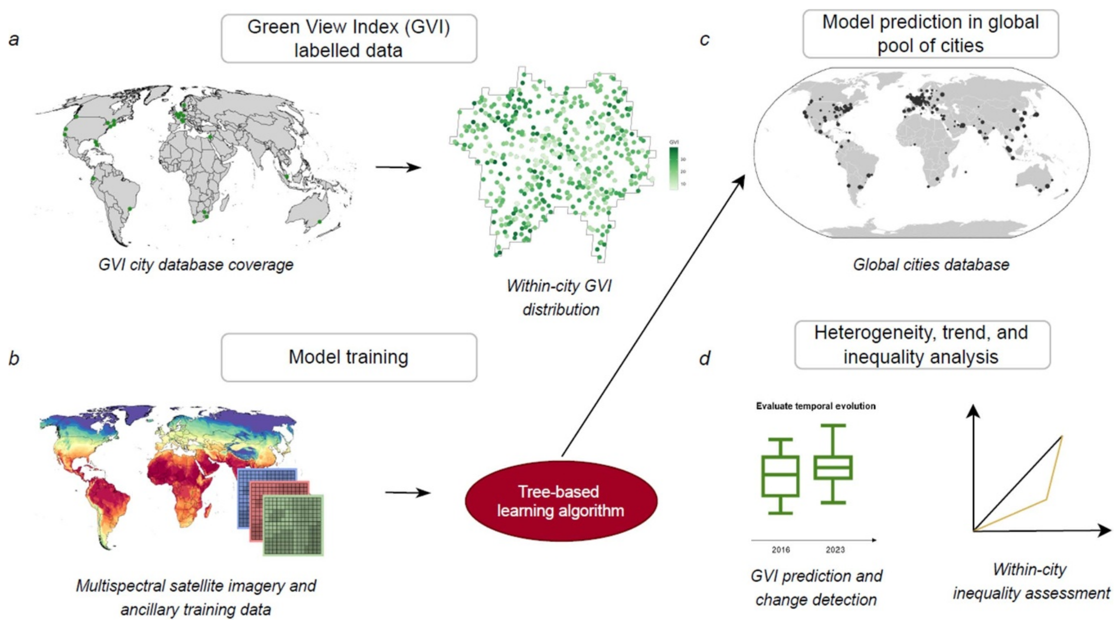

Figure 1. Illustrative representation of the green view index (GVI) assessment and monitoring analysis carried out in this study. (a) Graphical representation of the green view index (GVI) of street green space (SGS) city database from Seiferling *et al* [52] and within-city distribution of labeled GVI data; (b) and (c) machine learning (ML) model training and prediction across a pool of major world cities; (d) change detection and within-city inequality analysis.

on pre-classified land cover data [37–40], while other studies have derived satellite data-based estimates of different green space indicators [36, 41–49]. However, these estimates largely lack street-level validation and do not capture the specific features of vegetation along public streets, as opposed to other types of vegetation (such as private gardens and parks).

With specific regard to SGS, previous studies have developed approaches its quantification using street-based imagery across individual cities as well as regional or global pools of cities [50–58]. A related strand of research focused on quantifying the role of SGS relevance for within-city distribution (and environmental justice issues) [8, 59–62], as well as in terms of its temperature cooling potential [63, 64].

Yet, there is little evidence of an approach for estimating the spatial variability of SGS within and across cities with street-level relevance and for assessing its evolution over time based on globally available and frequently updated data. We attribute this gap to the challenge of upscaling existing localized street imagery-based assessment to a large-scale analysis. In a context where the release of public and spatially granular street-level imagery is limited and unwieldy, such goal requires resorting to granular, globally available datasets from earth observation to develop modeling approaches for emulating such street-based indicators [65, 66].

Here we aim at providing a global analysis of recent trends of the green view index (GVI) [52], a widely used metric of UGS at the street level [50, 67]. We develop a machine learning (eXtreme Gradient Boosted decision trees) modeling approach using 10-meter resolution multispectral satellite imagery data, climate records, and ancillary datasets to estimate ground-truth GVI measurements obtained from street-based georeferenced imagery from 23 cites in 14 countries around the world, mapped in figure 1(A). Based on the trained and validated model, we estimate local to continental trends of GVI over the 2016–2023 period over 190 large cities (visualized in figure 1(C)) distributed in twenty world macro-regions. Our analysis provides an updated estimate of GVI of SGS across world cities and its recent evolution, together with an open-source, validated approach to assessing its changes in near real-time.

The remainder of the paper is structured as follows: section 2 presents and describes the data used for model training and prediction, as well as the data processing steps conducted; it then elaborates on the statistical approach and model validation techniques implemented, as well as on the prediction stage and change analysis over time and space. Section 3 presents the results of the analysis, and it is divided into two main parts: first, the results of the model prediction and change analysis over time are illustrated, then within-city inequality over space and in relation to the urban population is assessed. Section 4 concludes the paper by summarizing the main findings, their implications, as well as highlighting the key limitations of the paper and paving the way for future work in the domain, as well as discussing potential uses of the open-source methodology and data produced in the paper.

Figure 2. Illustrative examples of different levels of SGS and its correspondence to GVI values in four selected cities. The three selected locations in each city represent the location of the first, second, and third quartiles of the within-city distribution of the GVI of SGS. Captions above each street photograph report the corresponding GVI level. Street View image data (c) 2024 Google.

2. Materials and methods

2.1. Data

2.1.1. Labeled GVI data

Labeled GVI data are sourced from Seiferling *et al* [52], available at http://senseable.mit.edu/treepedia. These data express the percentage fraction of canopy coverage in a given location as perceived on the street level, and thus are a useful description of SGS. The data are obtained based on street-level Google Street View (GSV) imagery. Each point comes with latitude and longitude attributes, GVI, and date of measurement. To put the GVI indicator of SGS into perspective and visually associate it with perceived street-level greenness, figure 2 provides illustrative examples of GVI levels in the training data in relation to the original street photography upon which they are based on different cities located in different geographical and climatic zones, Amsterdam (Europe, oceanic climate); Miami (North America, tropical monsoon climate); Singapore (Asia, tropical rain-forest climate); and Cape Town (Africa, Mediterranean climate). Each photograph is based on the location where the first (25th), second (50th), and third (75th) quartiles of the local distribution of GVI in each city lie.

The GVI labeled data covers 23 cities across 14 countries (figure SI-1 maps them and shows the estimated GVI distribution, while table SI-14 provides count of labeled points by city). Figure SI-3 illustrates the distribution of the years of street imagery acquisition and thus the reference date of GVI estimation, highlighting how about two-thirds of the snapshots on which the GVI estimation is based are derived from street imagery from years 2015-2016. This informs the use of multispectral satellite data and historical climate and land use records from years 2015 and 2016 in the model training phase.

In addition, figure SI-4 shows that the vast majority of training data points are based on imagery taken in periods of the year when in each city deciduous trees are growing green leaves, thus rendering the training data suitable to capture the presence of canopy coverage. Finally, tables SI-16 and SI-17 show the distribution of training points across Köppen-Geiger macroclimate zones. To ensure consistency, we filter the residual GVI labeled data to remove observations based on street photographs taken in periods of the year and locations where canopy coverage is likely to be absent or strongly reduced, such as for the case of broadleaved trees and other seasonal vegetation. Specifically, we remove observations which—based on the within-city GVI distribution—are taken in periods of the year where the GVI is more than two standard deviations lower than the mean (of all periods of the year, within each city).

2.1.2. Urban areas definition

Comparing GVI values and trends across cities around the world requires a standard definition of city boundaries. To define urban areas we use vector data from the EC-JRC Urban Centers Database [68]—namely the GHS Urban Centre Database UCDB R2019A (https://ghsl.jrc.ec.europa.eu/ucdb2018visual.php). Urban Centres are defined in a consistent way (by specific cut-off values on resident population and built-up surface share in a 1-km uniform global grid) across geographical locations and over time, applying the 'Global Definition of Cities and Settlements' developed by the European Union to the Global Human Settlement Layer Built-up (GHS-BUILT) areas and Population (GHS-POP) grids. Finally, national boundaries are based on the GADM v4.1 database [69].

2.1.3. Multispectral satellite data

To predict GVI values we adopt an approach based on multispectral satellite data. We use Sentinel 2 data due to their open nature, high spatial resolution (10 – 60 m, depending on the band), and global coverage. In particular, as seen from table SI-1, we select bands having a 10-meter resolution covering wavelengths between 443.9 nm and 833 nm. The data are available from 2016 to the current date and are planned to be released in near real-time in future years, thus currently allowing both for a 7 year change assessment and for tracking of future changes.

Note that whilst labeled GVI data include the date of acquisition of the original street-based imagery, we do not merge satellite and ancillary data by the month of acquisition, as this would impair the possibility of making extrapolative predictions, which is the key purpose of our model and analysis. On the other hand, for each variable, we include monthly values for all 12 months of the year and we let the model determine the most important features to predict GVI values in different areas of the world.

2.1.4. Ancillary training data

As an additional data source to the statistical model, we consider the ERA5 Monthly Averaged by Hour of Day climate reanalysis data product from the European Centre for Medium-Range Weather Forecasts (ECMWF). We calculate the monthly median value for each year of interest considering five variables: temperature of air at 2 m above the surface; temperature of the soil in layer 1 (0–7 cm); surface solar radiation; accumulated liquid and frozen water falling to the Earth's surface; and amount of evaporation from vegetation transpiration. In addition, we also use the Dynamic World land cover [70] data (near real-time global 10 m resolution) as additional covariates in the model, providing probabilities of each pixel being classified as one of eight dominant land cover types. As a third ancillary spatial covariate, we consider gridded population Global Human Settlement Layer 2020 [71] to capture the density of population within the urban areas. To account for variations in the outcome variable between cities and geographical areas driven by different levels of economic development, we include GDP per capita in our set of explanatory variables (https://data.worldbank.org/).

2.1.5. Data processing and extraction

To extract data we use Google Earth Engine. In particular, we extract monthly median values within a 10 m buffer around each GVI point coordinates for each year between 2016 and 2023. Note that for both periods, the month-wise 2 year median is calculated to smooth the potential impact of an anomalous year on the assessment. The resulting dataset is a 12 month multi-band raster file.

The data is then processed to extract imagery onto 10 m radius GVI buffer polygons using the R scientific programming environment [72]. Latitude and longitude are transformed into polar coordinates to enhance the spatial representation of the data points. As additional features, we calculate the spatial median for the percentage coverage of trees, bare ground, and grass within the ten nearest neighbor points. We also determine the median population distribution for these neighboring points.

Missing data account for 2.6% of the whole training dataset and 1.9% after removing Canada and Australia, which are excluded from the model training data due to the small number of data points with non-missing data available. Note that in the prediction stage following model training, cities from Canada and Australia are nonetheless included in the pool of cities where GVI prediction is carried out. The final dataset compiled for model training comprises 1295 999 observations and 104 variables. Table 1 shows descriptive statistics for the variables included as features in the model training, where the monthly data for the satellite bands and for the climate variables are presented as yearly averages.

Table 1. Descriptive statistics of the training dataset.

Statistic	Mean	Median	St. Dev.	Min	Max
Green view index (GVI)	19.11	17.55	10.67	1.00	91.07
GDP per capita (USD)	55 345.82	63 768.20	20 261.28	11 720.64	116 486.50
Population density (pop km^{-2})	124.51	71.75	168.47	0.00	8344.41
Trees (%)	5.59	3.70	6.82	1.63	77.81
Bare (%)	4.37	3.69	2.69	1.86	60.60
Grass (%)	3.85	2.97	3.16	1.76	58.95
Water (%)	4.53	3.98	2.91	1.59	74.40
Shrub and scrub (%)	4.17	3.45	2.21	1.46	52.98
Flooded vegetation (%)	3.49	3.35	1.05	1.63	58.26
Built (%)	65.62	72.08	16.36	2.05	79.90
Snow and ice (%)	3.34	3.12	1.04	1.76	46.37
Crops (%)	4.52	3.27	5.04	1.68	71.17
B2	2685.80	2771.29	705.17	896.08	5407.75
B3	2456.04	2545.25	641.34	758.25	5417.58
B4	2510.91	2590.96	675.33	552.67	5907.21
B8	3039.24	3049.33	707.61	672.79	5823.17
2 m air temperature (°C)	16.21	17.30	5.02	10.28	27.24
Surface pressure (Pa)	99 159.76	100 928.00	4066.17	72 378.70	101 964.60
Total precipitation (m)	0.06	0.06	0.05	0.01	0.42

2.2. Statistical methods and model validation

We apply the extreme gradient boosted decision trees algorithm (XGBoost) [73, 74] as implemented in the H2O R library [75] to capture the underlying non-linear relations between the GVI and the set of predictors derived from satellite measurements. Huber loss function is used to handle outliers, and 10-fold spatial cross-validation (SCV) is adopted to avoid an overconfident assessment of model predictive power in the testing phase and to account for the inherited spatial dependence in our data points [76, 77][4]. We include pixels classified as water-covered as a spatial sampling variable to account for the presence of water bodies, as vegetation near water bodies often differs significantly from vegetation farther away due to variations in microclimate, soil moisture, and nutrient availability. As a final stage, after model tuning, we retrain the algorithm on the entire dataset before the prediction phase[5].

Table 2 shows different evaluation metrics on the training, testing and retrained model over the full dataset. Metrics included are the R-squared (R^2), the root mean squared percentage error (RMSPE) and the mean absolute percentage error (MAPE). Figure SI-5 visually represents the R^2 results for the three datasets. We estimate 10-fold SCV values R^2 of 75% and 68% in the training and testing set, respectively, resulting in a full sample model R^2 of 75%. Comparing the three error metrics indicates that our model is able to capture a substantial portion of the variation in GVI (R^2 of 75%) with an average deviation from the actual values of approximately 27.1% as measured by the MAPE[6]. The relatively high RMSPE suggests that there are outlier cases where the model performs poorly. In order to investigate the error distribution of the model, we carry out country-level and macroclimate zone error analysis. Figure SI-6 reports country-level accuracy results, with differences in model prediction skill likely driven by heterogeneous urban structures and different cloud coverage frequencies. Table SI-12 shows the results of error analysis for each city in the training sample. Finally, table SI-10 shows how the model performs best in the continental and dry climate zone with the lowest RMSPE and MAPE values. The MAPE values do not vary significantly across different climatic classes. On the other hand, a more pronounced misalignment with the actual values is found in the temperate and tropical climate zones when considering the RMSPE.

2.3. Prediction and statistical testing of change over time

We sample points along streets located within urban boundaries with a constant density across cities, proportional to the cumulative urban street network length. Street network data is obtained by querying Open Street Map (OSM), as exemplified in figure SI-4). The sampling of points along roads is first random to ensure independence and then refined via Latin hypercube sampling (LHS). While new sample points are generated in random sampling without considering the previously generated sample points, LHS ensures

[4] The use of SCV instead of standard cross-validation is still a topic of debate in the literature. The interested reader can also refer to the work of [78] for a different perspective.

[5] Additional packages used in this work include **data.Table** [79],**ggplot2**[80],**stargazer** [81].

[6] For comparison, we also trained a deep neuronal network model. However, this resulted in a lower level of accuracy compared with XGBoost (R^2 of 0.64 on both train and test).

Table 2. Model training and testing benchmarks.

Dataset	R-Squared	RMSPE	MAPE
Train	0.754	0.472	0.268
Test	0.684	0.609	0.31
All (post-tuning)	0.749	0.483	0.271

there is only one sample in each row and each column of the two-dimensional (latitude and longitude) spatial data. Table SI-15 provides count of sampled points in each city where model predictions are generated. Figure 2 provides a schematic visual representation of sampled points along streets, while tables SI-2–SI-9 provides descriptive statistics for the sampled prediction points for each year between 2016 and 2023, and figures SI-25–SI-33 demonstrate the consistency of the random sampling procedure with the spatial distribution of the training data.

To predict GVI values in cities out of the labeled data sample or for different time steps, we use the trained model and multispectral satellite imagery on the sampled points within each city for each of the years of interest. To appraise the statistical change in GVI levels, we estimate regression on the full global sample and on regional and city-level subsets of sampled points to test for the existence of time trends at different scales, quantified as β in the following equation:

$$\log\left(GVI_{ict}\right) = \alpha + \beta \text{year}_t + \mu_i + \theta_c + \varepsilon_{ict} \tag{1}$$

where, i is each sample point, c is each city, t is each year in the sample and μ and θ are vector of sample points and city-level fixed effects.

We evaluate the statistical significance of the β coefficient by calculating a $t - statistic$ and the related $p - value$ for the difference in means across years in the sample. As seen from table SI-20, we also test additional specifications with inclusion or removal of fixed effects and where the *year* variable is treated as a set of binary variables instead of a linear continuous time trend variable.

2.4. Within city inequality in population exposure to SGS assessment

To appraise within-city inequalities in exposure to GVI, we calculate the resident population in the surrounding (with a 250 m radius buffer) of each sampled point (based on the GHS-POP (2022 revision) [82] and the 2016–2023 mean predicted GVI at each point. This procedure allows to characterize the city population by GVI exposure at the within-city level and thus to characterize heterogeneity and inequalities, which are visualized through cumulative exposure curves.

3. Results and discussion

3.1. Mapping and tracking the evolution of SGS in world cities

Based on sampling points along streets in 190 large cities distributed across twenty global macro-regions, we estimate a median GVI of 15.5 in 2016–2023, with a trend of decrease of 0.3%–0.5% per year (0.02–0.03 GVI points/yr., depending on the regression model specification to estimate the time trend). Table SI-20 reports the results of the regression models estimated on the repeated predictions in the sampled points along streets within each city to evaluate the time trend at a global scale.

Mapping the results (figure 3) reveals that during the eight-year period investigated, the average GVI levels in Oceanic, European, South East Asian and Eastern and Southern African cities show significantly higher values than those in continental Asia, North and Western Africa, and American continents. Looking at the direction and magnitude of GVI evolution over 2016–2023 reveals a high level of heterogeneity both between and within regions (figure 4(A)). For instance, during the 2016–2023 eight-year period, we estimate median GVI declines by 1.7%/yr. in cities of Asia and 2.6% in urban areas in Oceania, while we find evidence of growth rates of similar magnitude (about 1%/yr.) in European and North American cities. Conversely, in African and Latin American cities we only find evidence of a small trend of change.

As an additional experiment, we compare the correlation of the estimated values of GVI with the yearly average NDVI over time and space (see table SI-21). This exercise is useful for evaluating the agreement between an estimated street-based indicator (the GVI) and an index directly calculated from multispectral remote sensing data (the NDVI) in the large, globally-relevant pool of locations and cities covered by our study. Previous studies [12, 83] have already shown that the two metrics are not substitutes, even if correlated. Accordingly, our experiment shows a significant correlation between the two measures across the years covered by the analysis and over space when aggregated at the city and country levels.

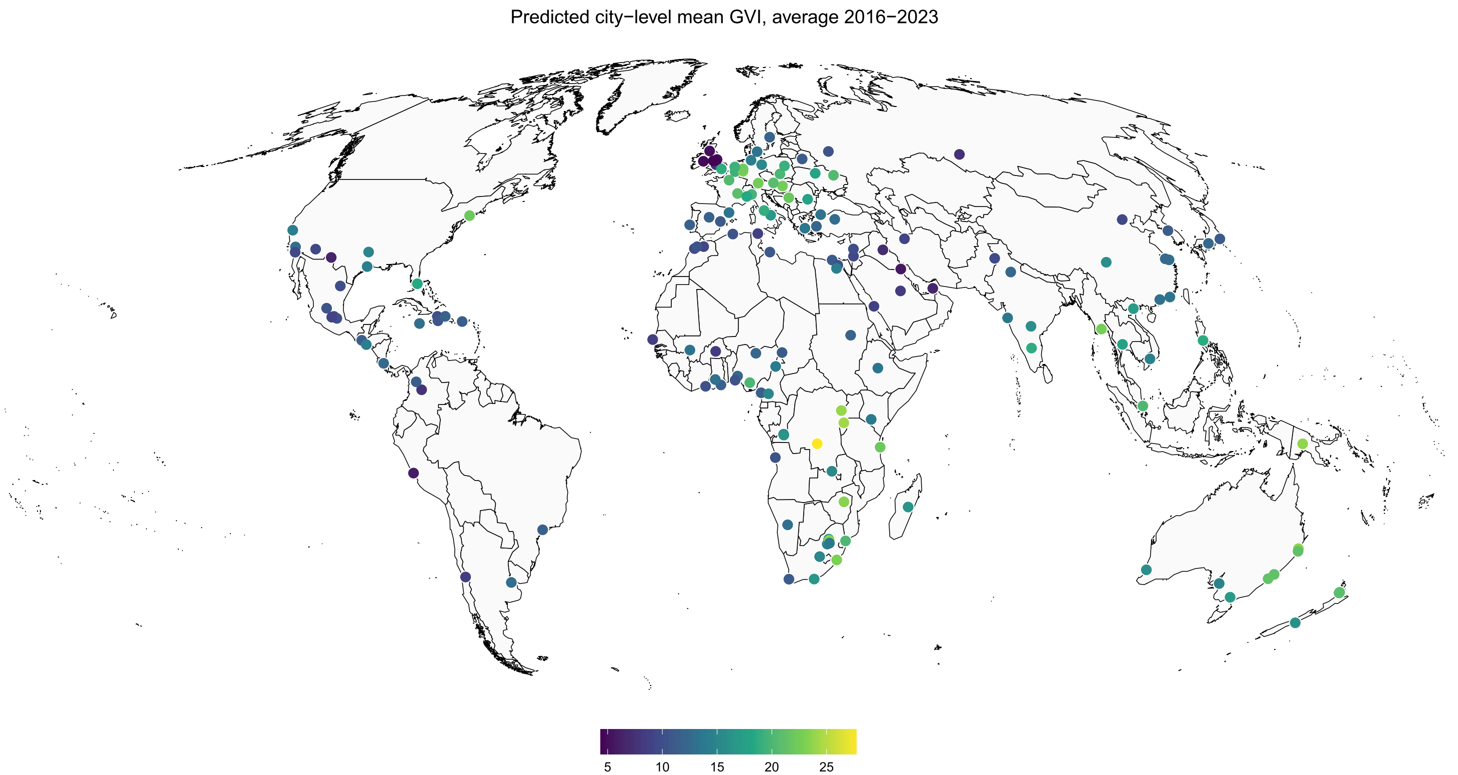

Figure 3. Map of the estimated green view index (GVI, a metric estimating local street canopy coverage in percentage terms) in 190 major world cities distributed in 20 regions. Each dot represents the geographical coordinates of the centroid of a city and its color coding describes the average value across the sampled points along the streets of the city in years 2016-2023. National boundaries are based on the GADM v4.1 database [69].

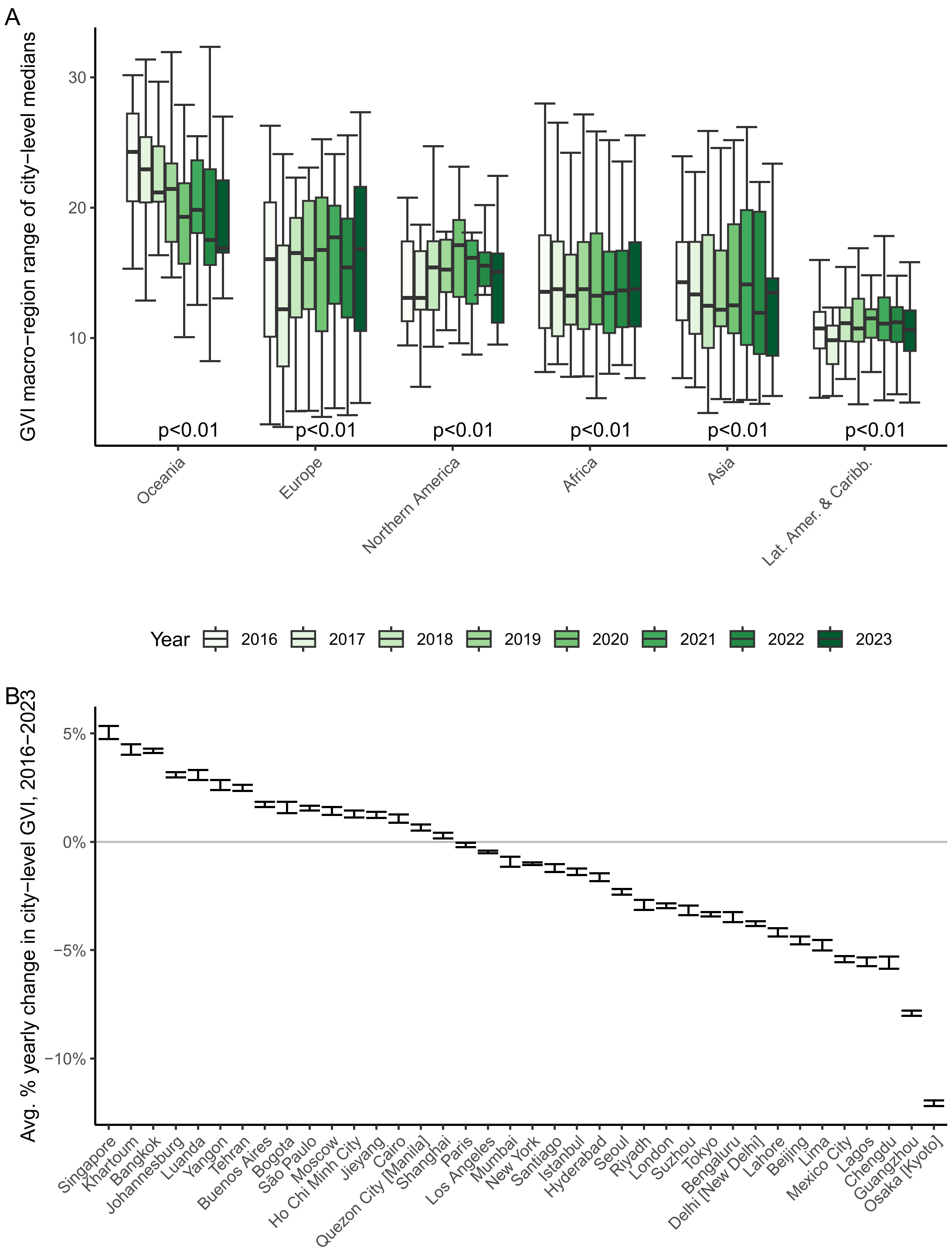

Figure 4. Distribution of the estimated green view index (GVI) in randomly sampled points along streets within (selected) cities boundaries in 2016 and 2023. (A) Regional distribution of city-level median GVI (based on sampled points in all cities analyzed in this paper) and *p-values* of regression coefficient of region-level mean GVI change linear trend; (B) city-level estimated average percentage change per year of GVI based on sampled points in the 50 largest cities analyzed, by population. In both panels, *p-values* above each box represent the probability of rejecting the null-hypothesis of no trend of change in the local GVI between 2016 and 2023.

While it is beyond the scope of this paper to address the mechanisms underlying the observed trend of GVI decrease in the pool of global cities analyzed—as well as the regionally and city-level heterogeneous trends—it should be noted that the detected SGS canopy coverage change via the GVI indicator may depend on a broad range of interacting factors. These might include global urbanization patterns determining land use change and deforestation [84], in particular in the Global South, but also urban afforestation policies [85]; climate change impacts on vegetation health and canopy growth [86] and the already documented

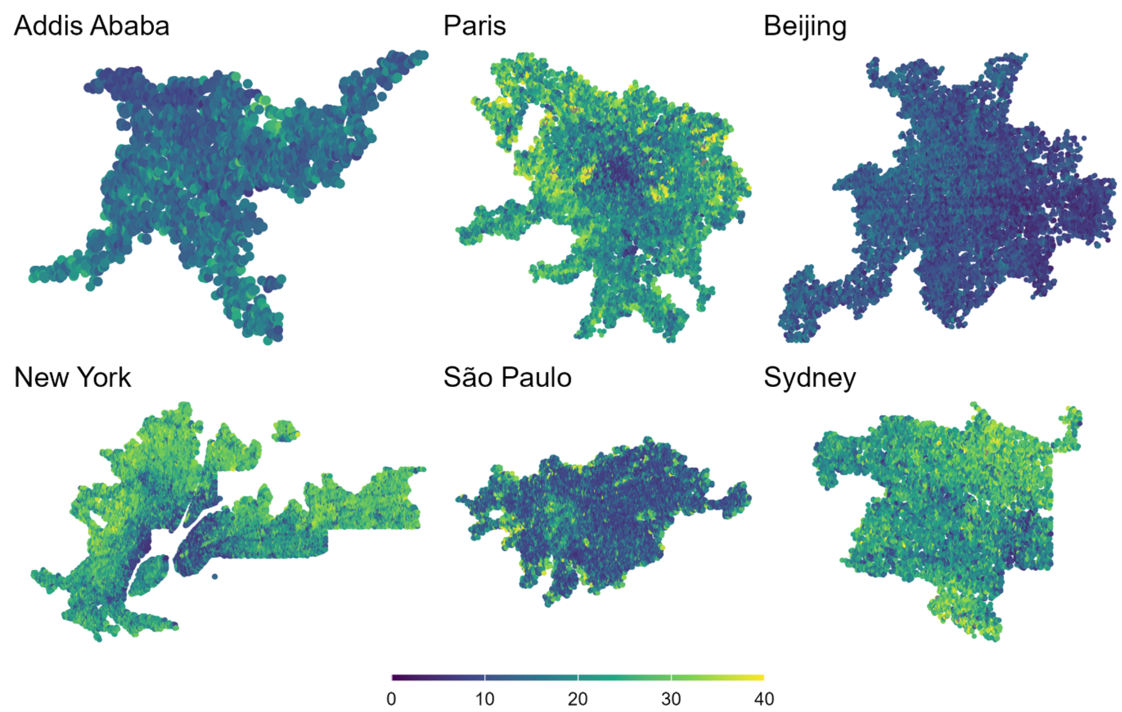

Figure 5. Map of the estimated green view index (GVI) for the illustrative example of six selected cities. Each dot in each city represents the geographical coordinates of a sampling point in a city and its color coding describes the average local estimated GVI value in the 2016–2023 period.

ongoing CO2 fertilization [87]; or even weather variability during the eight observation years (2016–2023), potentially affecting the observed variations.

Zooming into each city covered by our global analysis (see table SI-18 for a comprehensive summary of all cities covered), considering the ten most populated urban areas in each of six global macro-regions, we estimate the cities with the highest city-level median GVI (in 2016–2023) to be the following: Dar es Salaam (Tanzania) in Africa; Brisbane (Australia) in Oceania; Dortmund (Germany) in Europe; Kolkata (India), in Asia; Belo Horizonte (Brazil) in Latin America; and New York (USA) in North America. On the other hand, the large cities with the lowest values in each region are Cairo (Egypt) in Africa; Tokyo (Japan) in Asia; Saint Petersburg (Russia) in Europe; Lima (Peru) in Latin America; Phoenix (USA) in North America; and Adelaide (Australia) in Oceania.

Moving to the analysis of SGS change within each city over the eight-year period covered by our analysis, figure 4(B) shows (for a subset of highly populated urban areas) that we find evidence of statistically significant (p-$value < 0.01$) changes in the vast majority of city-level median GVI levels between 2016 and 2023. The five cities with the strongest, statistically significant mean GVI increase trend are Helsinki (Finland), Kyiv (Ukraine), Leeds (United Kingdom), Houston (US) and Ciudad Juárez (Mexico). Among others, additional large cities witnessing a GVI growth rate >5%/yr. include Windhoek (Namibia), Kolkata (India), Rio de Janeiro (Brazil), and Singapore. Conversely, the top five cities by mean GVI decreasing trend are Osaka (Japan), Guangzhou (China), Auckland (New Zealand), Douala (Cameroon) and Port Moresby (Papua New Guinea).

3.2. Measuring within-city inequalities in SGS

While appraising across-cities and regions trends is crucial to provide a global picture of change dynamics, SGS can also be highly unequally distributed within each city in terms of neighborhoods and accessibility of the residential population [38, 88–92]. This issue is widely discussed in the literature in relation to within-city gentrification and segregation dynamics and, more broadly, to urban environmental justice [93]. Figure 5 provides a representation of the within-city variation in the GVI for selected cities across the six continents covered by our study (see figures SI-7–SI-22 for similar maps of sampled points in additional cities).

Based on the within-city variations in each city covered by our analysis, we carry out an assessment of the inequality in the distribution of GVI (i.e. the estimated availability of SGS) with respect to the population living in proximity of each sampling point. Curves of the cumulative GVI exposure of the residential population (figure 6) demonstrate that substantial heterogeneity across cities within the same region is observed: in Europe, for instance, the GVI is much more equally distributed in relation to the local resident

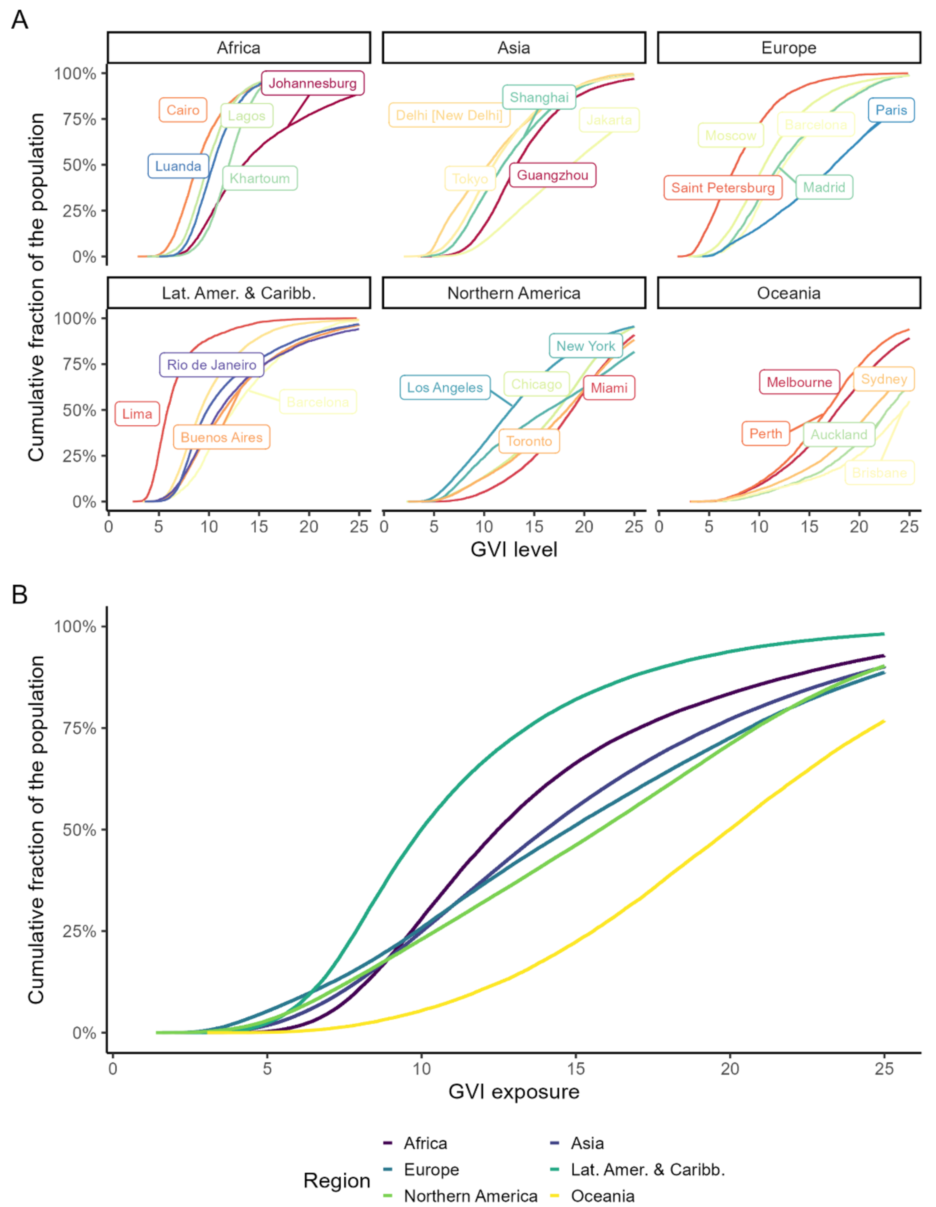

Figure 6. Residential population exposure inequality to the green view index (GVI). (A) city-level curves, faceted by region; (B) Region-level curves (based on city-level curves aggregation).

population in Paris than in Madrid, Barcelona, or Moscow. To explore such within-city inequalities, we also calculate Gini index values of GVI inequality weighted by the local population (see table SI-19) suggest that among the cities covered by our global assessment, those with the most unequal GVI accessibility are Cape Town (South Africa), Athens (Greece), Alexandria (Egypt), Nairobi (Kenya), and Rabat (Morocco), whilst the (relatively) most equal distribution is estimated in several Oceanian and Caribbean cities.

When aggregating the sampling point-level results for the full set of cities analyzed in the paper to the macro-regional level, Oceania stands out as the region with the most equal distribution of GVI, followed by Europe and North America. The most unequal regions are found to be Latin America and Africa. This result is in line with recent research, such as Li *et al* [8], demonstrating strong inequalities in green space quantity and quality between cities in the Global North and South as a consequence of both socio-economic and natural factors. Such environmental injustice manifests itself the most in Global South cities, where dynamics of 'green apartheid' [94] have been highlighted, in relation to both income [95] and social and racial factors, as well as urban center-edge gradients [96].

4. Conclusion

4.1. Summary of the study and implications

The approach developed to estimate and track the evolution of GVI—an estimate of canopy coverage along urban roads—by modeling street-level data with openly and globally available, frequently updated remotely sensed, geospatial, and climate reanalysis datasets has important implications. This is because SGS has great relevance for several dimensions of sustainable cities related to SDG 11. For instance, in relation to climate change mitigation and adaptation, greener cities bear greater potential to reduce the socio-economic impacts of extreme climate events [97], adverse heat-related impacts on human health [98, 99], and decrease thermal energy needs in buildings, such as cooling demand [8, 13]. Thus, it is key that national and local decision-makers and public planners seek to expand SGS and tackle the drivers of SGS decline where this is ongoing. Global initiatives such as the FAO's Tree Cities of the World (https://treecitiesoftheworld.org/) are encouraging cities to join a global effort for urban vegetation expansion.

The methodology and results of this study can be used by policymakers in the design and the evaluation of multi-scale policies (e.g. from national funding allocation to local council implementation actions) for greener, more sustainable, ecosystem services-rich and equitably designed cities to pursue environmental goals for climate change mitigation and adaptation and pollution reduction, as well as other social development goals. To achieve these aims, the output data should be combined with other key indicators of urban resource use and impacts to appraise the relevance and potential benefit of SGS expansion and densification. Such dimensions and data include local temperature records from weather stations, energy demand for thermal regulation in buildings, records of the impact of extreme weather events on people and urban infrastructure, climate-related adverse impacts on human health such as morbidity and mortality, as well as subjective well-being indicators. Moreover, environmental justice dimensions are increasingly pivotal in the quest for a just transition [100]. Thus, also within-region and within-city inequalities are worth careful equity-oriented planning to ensure that the benefits of urban vegetation are equitably accessible and infrastructure inequalities are reduced [101]. Indicators of SGS distribution should hence be analyzed in relation to socio-demographic group distribution inside and across cities to identify the context of environmental justice relative to income, gender, age, race and other dimensions.

4.2. Limitations and future work

Coming to the limitations of this study, although model training and validation show satisfactory accuracy levels (figure SI-5), the analysis is not without caveats and the nature of the training data has to be taken into account. First, the GVI metric used to assess SGS, describes the percentage fraction of canopy coverage as perceived from street-level 360-degree photography taken at a given geographical point. Thus, the coverage and quality of the street-based imagery underlying the labeled GVI data is likely mixed across and within the cities for which it is available and on which the ML statistical model is trained. Sources of bias associated with mixed coverage of sample data points of vegetation within cities and across the globe include the presence of water bodies, cloud cover, varying types of canopy and trees, as well as differences in land use and management practices. Water bodies can create microclimates that enhance vegetation health, potentially skewing results if not evenly distributed across the sample. Cloud cover can obscure satellite imagery, leading to incomplete or inaccurate data collection, particularly in regions with frequent cloud cover. The diversity in canopy types and tree species can influence the reflectance values captured by remote sensors, causing variability in vegetation indices. From a more general standpoint, the primary challenge we are undertaking involves predicting a street greenness indicator, which manifests in a three-dimensional space, using satellite data derived through extrapolation from images captured in a two-dimensional space. Moreover, the latter provides a more coarse-grained picture of SGS. While this makes reaching high accuracy challenging, it might, in turn, induce the model to yield heterogeneously accurate estimates across different geographical, urban, and social settings across and within the global pool of cities analyzed.

Irrespective of limitations, our analysis provides a globally relevant estimate of the GVI of SGS across world cities validated against street-level imagery in a range of highly diverse cities in terms of climate and socio-economic conditions, and it contributes to future research through an open-source, open-data, statistically validated approach to assess future changes in near real-time. The approach is suitable for scalability to additional cities and geographies and for low-cost and rapid tracking of future SGS change, as the statistical model can be updated at the high frequency of the release of multispectral satellite imagery and climate records. Future work can build on ever more available and high-coverage street imagery databases [102] to further assess and improve the understanding of the trends and determinants of urban landscape change.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.13886667.

Acknowledgment

Giacomo Falchetta gratefully acknowledges support from the IIASA 2024 Innovative and Bridging Grant *URGED*. The authors are thankful to participants to the EGU 2023 session ITS1.11/NP0.2 '*Digitalisation processes of complex systems: Urban Geosciences and Geo-Health*' and EGU 2024 session HS1.1.7 '*Looking for resilience at building scale: Nature-based Solutions to face water related and energetic challenges*' for their useful feedback.

Code availability

The replication code is made publicly available in a Github repository (https://github.com/giacfalk/urban_green_space_mapping_and_tracking). Correspondence and requests for materials should be addressed to Giacomo Falchetta (falchetta@iiasa.ac.at).

Author contributions statement

G F and A T H designed the research and carried out the formal analysis; G F generated the figures; all authors contributed to writing and editing the paper.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Giacomo Falchetta https://orcid.org/0000-0003-2607-2195
Ahmed T Hammad https://orcid.org/0000-0003-3327-2435

References

[1] Rostami R, Lamit H, Khoshnava S M, Rostami R and Fitry Rosley M S 2015 Sustainable cities and the contribution of historical urban green spaces: a case study of historical Persian gardens *Sustainability* **7** 13290–316
[2] 2024 Our cities, ourselves *Nat. Cities* **1** 1–1
[3] Bille R A, Jensen K E and Buitenwerf R 2023 Global patterns in urban green space are strongly linked to human development and population density *Urban Forest. Urban Green.* **86** 127980
[4] Russo A and Cirella G T 2018 Modern compact cities: how much greenery do we need? *Int. J. Environ. Res. Public Health* **15** 2180
[5] Derkzen M L, van Teeffelen A J and Verburg P H 2015 Quantifying urban ecosystem services based on high-resolution data of urban green space: an assessment for Rotterdam, the Netherlands *J. Appl. Ecol.* **52** 1020–32
[6] Turner V K, Middel A and Vanos J K 2023 Shade is an essential solution for hotter cities *Nature* **619** 694–7
[7] Zhan W *et al* 2024 Fraction-dependent variations in cooling efficiency of urban trees across global cities *ISPRS J. Photogramm. Remote Sens.* **216** 229–39
[8] Li Y *et al* 2024 Green spaces provide substantial but unequal urban cooling globally *Nat. Commun.* **15** 7108
[9] Wong N H, Tan C L, Kolokotsa D D and Takebayashi H 2021 Greenery as a mitigation and adaptation strategy to urban heat *Nat. Rev. Earth Environ.* **2** 166–81
[10] Seyam S 2019 The impact of greenery systems on building energy: systematic review *J. Build. Eng.* **26** 100887
[11] Massaro E *et al* 2023 Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes *Nat. Commun.* **14** 2903
[12] Aram F, García E H, Solgi E and Mansournia S 2019 Urban green space cooling effect in cities *Heliyon* **5** e01339
[13] Zhang B *et al* 2014 The cooling effect of urban green spaces as a contribution to energy-saving and emission-reduction: a case study in Beijing, China *Build. Environ.* **76** 37–43
[14] Du H *et al* 2017 Quantifying the cool island effects of urban green spaces using remote sensing data *Urban Forest. Urban Green.* **27** 24–31
[15] Staccione A, Essenfelder A H, Bagli S and Mysiak J 2024 Connected urban green spaces for pluvial flood risk reduction in the metropolitan area of Milan *Sustain. Cities Soc.* **104** 105288
[16] Kim H, Lee D-K and Sung S 2016 Effect of urban green spaces and flooded area type on flooding probability *Sustainability* **8** 134
[17] Sun Y, Xie S and Zhao S 2019 Valuing urban green spaces in mitigating climate change: a city-wide estimate of aboveground carbon stored in urban green spaces of China's capital *Glob. Change Biol.* **25** 1717–32
[18] Fryd O, Pauleit S and Bühler O 2012 The role of urban green space and trees in relation to climate change *CABI Rev.* 1–18

[19] Wooster E, Fleck R, Torpy F, Ramp D and Irga P 2022 Urban green roofs promote metropolitan biodiversity: a comparative case study *Build. Environ.* **207** 108458
[20] Reyes-Riveros R *et al* 2021 Linking public urban green spaces and human well-being: a systematic review *Urban Forest. Urban Green.* **61** 127105
[21] Bratman G N *et al* 2019 Nature and mental health: an ecosystem service perspective *Sci. Adv.* **5** 0903
[22] Tost H *et al* 2019 Neural correlates of individual differences in affective benefit of real-life urban green space exposure *Nat. Neurosci.* **22** 1389–93
[23] Hedblom M *et al* 2019 Reduction of physiological stress by urban green space in a multisensory virtual experiment *Sci. Rep.* **9** 10113
[24] Kwon O-H *et al* 2021 Urban green space and happiness in developed countries *EPJ Data Sci.* **10** 28
[25] Olszewska-Guizzo A, Sia A, Fogel A and Ho R 2022 Features of urban green spaces associated with positive emotions, mindfulness and relaxation *Sci. Rep.* **12** 20695
[26] Han L, Heblich S, Timmins C and Zylberberg Y 2024 Cool cities: the value of urban trees. *Technical Report* (National Bureau of Economic Research)
[27] WOrld Health Organization 2012 Health indicators of sustainable cities in the context of the rio+ 20 un conference on sustainable development WHO: Geneva, Switzerland
[28] Mullaney J, Lucke T and Trueman S J 2015 A review of benefits and challenges in growing street trees in paved urban environments *Landsc. Urban Plan.* **134** 157–66
[29] Coisnon T, Musson A, Pene S D and Rousselière D 2024 Disentangling public urban green space satisfaction: exploring individual and contextual factors across European cities *Cities* **152** 105154
[30] Leslie E, Sugiyama T, Ierodiaconou D and Kremer P 2010 Perceived and objectively measured greenness of neighbourhoods: are they measuring the same thing? *Landsc. Urban Plan.* **95** 28–33
[31] Shanahan D F *et al* 2017 Variation in experiences of nature across gradients of tree cover in compact and sprawling cities *Landsc. Urban Plan.* **157** 231–8
[32] Hoyle H, Jorgensen A and Hitchmough J D 2019 What determines how we see nature? Perceptions of naturalness in designed urban green spaces *People Nat.* **1** 167–80
[33] Farkas J Z, Hoyk E, de Morais M B and Csomós G 2023 A systematic review of urban green space research over the last 30 years: a bibliometric analysis *Heliyon* **9** e13406
[34] Kabisch N, Strohbach M, Haase D and Kronenberg J 2016 Urban green space availability in European cities *Ecol. Indicators* **70** 586–96
[35] Kabisch N and Haase D 2013 Green spaces of European cities revisited for 1990–2006 *Landsc. Urban Plan.* **110** 113–22
[36] Ju Y, Dronova I and Delclòs-Alió X 2022 A 10 m resolution urban green space map for major Latin American cities from sentinel-2 remote sensing images and OpenStreetMap *Sci. Data* **9** 586
[37] Zhou Q, Liao Y and Wang J 2022 Mapping global urban greenspace: an analysis based on open land-cover data *Urban Forest. Urban Green.* **74** 127638
[38] Chen B *et al* 2022 Contrasting inequality in human exposure to greenspace between cities of global North and global South *Nat. Commun.* **13** 4636
[39] Liao Y, Zhou Q and Jing X 2021 A comparison of global and regional open datasets for urban greenspace mapping *Urban Forest. Urban Green.* **62** 127132
[40] Duncan J M and Boruff B 2023 Monitoring spatial patterns of urban vegetation: a comparison of contemporary high-resolution datasets *Landsc. Urban Plan.* **233** 104671
[41] Wu S, Song Y, An J, Lin C and Chen B 2024 High-resolution greenspace dynamic data cube from sentinel-2 satellites over 1028 global major cities *Sci. Data* **11** 909
[42] Xu F, Yan J, Heremans S and Somers B 2022 Pan-European urban green space dynamics: a view from space between 1990 and 2015 *Landsc. Urban Plan.* **226** 104477
[43] Kopecká M, Szatmári D and Rosina K 2017 Analysis of urban green spaces based on sentinel-2a: case studies from Slovakia *Land* **6** 25
[44] Huang C *et al* 2021 Mapping the maximum extents of urban green spaces in 1039 cities using dense satellite images *Environ. Res. Lett.* **16** 064072
[45] Abutaleb K, Mudede M F, Nkongolo N and Newete S W 2021 Estimating urban greenness index using remote sensing data: a case study of an affluent vs poor suburbs in the city of Johannesburg *Egypt. J. Remote Sens. Space Sci.* **24** 343–51
[46] Reid C E, Kubzansky L D, Li J, Shmool J L and Clougherty J E 2018 It's not easy assessing greenness: a comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York city *Health Place* **54** 92–101
[47] Wang J, Liu W and Gou A 2022 Numerical characteristics and spatial distribution of panoramic street green view index based on segnet semantic segmentation in Savannah *Urban Forest. Urban Green.* **69** 127488
[48] Neyns R and Canters F 2022 Mapping of urban vegetation with high-resolution remote sensing: a review *Remote Sens.* **14** 1031
[49] Ludwig C, Hecht R, Lautenbach S, Schorcht M and Zipf A 2021 Mapping public urban green spaces based on openstreetmap and sentinel-2 imagery using belief functions *ISPRS Int. J. Geo-Inf.* **10** 251
[50] Li X *et al* 2015 Assessing street-level urban greenery using Google street view and a modified green view index *Urban Forest. Urban Green.* **14** 675–85
[51] Xia Y, Yabuki N and Fukuda T 2021 Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning *Urban Forest. Urban Green.* **59** 126995
[52] Seiferling I, Naik N, Ratti C and Proulx R 2017 Green streets- quantifying and mapping urban trees with street-level imagery and computer vision *Landsc. Urban Plan.* **165** 93–101
[53] Zhang T, Wang L, Hu Y, Zhang W and Liu Y 2024 Measuring urban green space exposure based on street view images and machine learning *Forests* **15** 655
[54] Tong M *et al* 2020 Evaluating street greenery by multiple indicators using street-level imagery and satellite images: a case study in Nanjing, China *Forests* **11** 1347
[55] Richards D R and Edwards P J 2017 Quantifying street tree regulating ecosystem services using Google street view *Ecol. Indicators* **77** 31–40
[56] Zhang L *et al* 2023 Decoding urban green spaces: deep learning and Google street view measure greening structures *Urban Forest. Urban Green.* **87** 128028

[57] Zhang J and Hu A 2022 Analyzing green view index and green view index best path using Google street view and deep learning *J. Comput. Des. Eng.* **9** 2010–23

[58] Sun Y *et al* 2021 Using machine learning to examine street green space types at a high spatial resolution: application in Los Angeles county on socioeconomic disparities in exposure *Sci. Total Environ.* **787** 147653

[59] Rocha A D *et al* 2024 Unprivileged groups are less served by green cooling services in major European urban areas *Nat. Cities* **1** 424–35

[60] Yutian L, Running C, Bin C and Jiayu W 2024 Inclusive green environment for all? An investigation of spatial access equity of urban green space and associated socioeconomic drivers in china *Landsc. Urban Plan.* **241** 104926

[61] Wu J, Feng Z, Peng Y, Liu Q and He Q 2019 Neglected green street landscapes: a re-evaluation method of green justice *Urban Forest. Urban Green.* **41** 344–53

[62] Wu J, Peng Y, Liu P, Weng Y and Lin J 2022 Is the green inequality overestimated? Quality reevaluation of green space accessibility *Cities* **130** 103871

[63] Liu Z *et al* 2022 Nonlinear cooling effect of street green space morphology: evidence from a gradient boosting decision tree and explainable machine learning approach *Land* **11** 2220

[64] Hu Y *et al* 2024 Which street is hotter? Street morphology may hold clues-thermal environment mapping based on street view imagery *Build. Environ.* **262** 111838

[65] Larkin A, Gu X, Chen L and Hystad P 2021 Predicting perceptions of the built environment using gis, satellite and street view image approaches *Landsc. Urban Plan.* **216** 104257

[66] Chen G, Zhou Y, Voogt J A and Stokes E C 2024 Remote sensing of diverse urban environments: from the single city to multiple cities *Remote Sens. Environ.* **305** 114108

[67] Yang J, Zhao L, Mcbride J and Gong P 2009 Can you see green? Assessing the visibility of urban forests in cities *Landsc. Urban Plan.* **91** 97–104

[68] Florczyk A *et al* 2019 Description of the GHS urban centre database 2015 *Public Release* **1** 1–75

[69] Hijmans R, Garcia N and Wieczorek J 2022 Gadm: database of global administrative areas, version 4.1 *GADM Maps and Data*

[70] Brown C F *et al* 2022 Dynamic world, near real-time global 10 m land use land cover mapping *Sci. Data* **9** 1–17

[71] Pesaresi M *et al* 2016 Operating procedure for the production of the global human settlement layer from landsat data of the epochs 1975, 1990, 2000 and 2014 *Publ. Office Eur. Union* 1–62

[72] Team, R. C. 2020 *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing)

[73] Chen T and Guestrin C 2016 Xgboost: a scalable tree boosting system *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining KDD'16* (ACM) pp 785–94

[74] Friedman J H 2001 Greedy function approximation: a gradient boosting machine *Ann. Stat.* **29** 1189–232

[75] LeDell E and Poirier S 2020 H_2O automl: scalable automatic machine learning *Proc. AutoML Workshop at ICML* vol 2020

[76] Meyer H, Reudenbach C, Hengl T, Katurji M and Nauss T 2018 Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation *Environ. Model. Softw.* **101** 1–9

[77] Ploton P *et al* 2020 Spatial validation reveals poor predictive performance of large-scale ecological mapping models *Nat. Commun.* **11** 4540

[78] Wadoux A M-C, Heuvelink G B, De Bruin S and Brus D J 2021 Spatial cross-validation is not the right way to evaluate map accuracy *Ecol. Model.* **457** 109692

[79] Dowle M and Srinivasan A 2021 data.Table: extension of 'data.frame R package version 1.14.2

[80] Wickham H 2016 *Ggplot2: Elegant Graphics for Data Analysis* vol 174 (Springer)

[81] Jia W, Shaw K A and Martonosi M 2012 Stargazer: automated regression-based gpu design space exploration *2012 IEEE Int. Symp. on Performance Analysis of Systems & Software* (IEEE) pp 2–13

[82] Pesaresi M, Florczyk A, Schiavina M, Melchiorri M and Maffenini L 2019 GHS settlement grid, updated and refined regio model 2014 in application to GHS-built r2018a and GHS-pop r2019a, multitemporal (1975-1990-2000-2015), r2019a

[83] Li T *et al* 2021 Spatial relationship between green view index and normalized differential vegetation index within the sixth ring road of Beijing *Urban Forest. Urban Green.* **62** 127153

[84] Dijkstra L *et al* 2021 Applying the degree of urbanisation to the globe: a new harmonised definition reveals a different picture of global urbanisation *J. Urban Econ.* **125** 103312

[85] Yao N *et al* 2019 Beijing's 50 million new urban trees: strategic governance for large-scale urban afforestation *Urban Forest. Urban Green.* **44** 126392

[86] Esperon-Rodriguez M *et al* 2024 Mapping the climate risk to urban forests at city scale *Landsc. Urban Plan.* **248** 105090

[87] Zhu S *et al* 2023 Explainable machine learning confirms the global terrestrial CO_2 fertilisation effect from space *IEEE Geosci. Remote Sens. Lett.* **20** 2503705

[88] Leng S, Sun R, Yang X and Chen L 2023 Global inequities in population exposure to urban greenspaces increased amidst tree and nontree vegetation cover expansion *Commun. Earth Environ.* **4** 464

[89] Ekkel E D and de Vries S 2017 Nearby green space and human health: evaluating accessibility metrics *Landsc. Urban Plan.* **157** 214–20

[90] Wüstemann H, Kalisch D and Kolbe J 2017 Access to urban green space and environmental inequalities in Germany *Landsc. Urban Plan.* **164** 124–31

[91] You H 2016 Characterizing the inequalities in urban public green space provision in shenzhen, China *Habitat Int.* **56** 176–80

[92] Liu D, Kwan M-P and Kan Z 2021 Analysis of urban green space accessibility and distribution inequity in the city of Chicago *Urban Forest. Urban Green.* **59** 127029

[93] Anguelovski I *et al* 2022 Green gentrification in European and North American cities *Nat. Commun.* **13** 3816

[94] Venter Z S, Shackleton C M, Van Staden F, Selomane O and Masterson V A 2020 Green apartheid: Urban green infrastructure remains unequally distributed across income and race geographies in South Africa *Landsc. Urban Plan.* **203** 103889

[95] Chen Y, Yue W and La Rosa D 2020 Which communities have better accessibility to green space? An investigation into environmental inequality using big data *Landsc. Urban Plan.* **204** 103919

[96] Chen Y *et al* 2022 Inequalities of urban green space area and ecosystem services along urban center-edge gradients *Landsc. Urban Plan.* **217** 104266

[97] Bai T, Mayer A L, Shuster W D and Tian G 2018 The hydrologic role of urban green space in mitigating flooding (Luohe, China) *Sustainability* **10** 3584

[98] Choi H M *et al* 2022 Effect modification of greenness on the association between heat and mortality: a multi-city multi-country study *EBioMedicine* **84** 104251

[99] Zhang S *et al* 2023 Assessment of short-term heat effects on cardiovascular mortality and vulnerability factors using small area data in Europe *Environ. Int.* **179** 108154

[100] McCauley D and Heffron R 2018 Just transition: integrating climate, energy and environmental justice *Energy Policy* **119** 1–7

[101] Pandey B, Brelsford C and Seto K C 2022 Infrastructure inequality is a characteristic of urbanization *Proc. Natl Acad. Sci.* **119** e2119890119

[102] Hou Y *et al* 2024 Global streetscapes-a comprehensive dataset of 10 million street-level images across 688 cities for urban science and analytics *ISPRS J. Photogramm. Remote Sens.* **215** 216–38