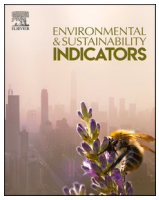
Environmental and Sustainability Indicators 27 (2025) 100864

Contents lists available at ScienceDirect

Environmental and Sustainability Indicators

journal homepage: www.sciencedirect.com/journal/environmental-and-sustainability-indicators



Propagating uncertainty in urban tree trait measurements to estimate socioeconomic inequities in ecosystem service accessibility: A machine learning and simulation framework

Jaime Francisco Pereña-Ortiz, Ángel Enrique Salvo-Tierra, Pablo Cozano-Pérez, Ángel Ruiz-Valero *

Department of Botany and Plant Physiology, Faculty of Sciences, Universidad de Málaga, 29010, Málaga, Spain

ARTICLE INFO

Keywords:
Bayesian
Ecosystem-services
Environmental-justice
Machine-learning
Urban-forestry

ABSTRACT

Achieving Sustainable Development Goal 11 requires addressing inequities in access to ecosystem services provided by urban trees to ensure a fair distribution of environmental benefits across socioeconomic groups. Ecosystem services estimates based on urban forest inventories often face challenges related to missing data or traits recorded in numerical ranges. Measurement-level uncertainty, if unaccounted for, can lead to overly optimistic estimates and limit their utility for urban planning. This study presents a methodological framework for propagating uncertainty in ecosystem services estimation with i-Tree. The approach integrates machine learning models to impute missing data and employs copulas and Monte Carlo simulations to assess the impact of trait uncertainty on provision. These simulations are further incorporated into Bayesian Hierarchical Models to evaluate how trait uncertainty influences estimates of inequities in ecosystem services accessibility across socioeconomic groups at the census tract level. Results indicate substantial uncertainty in tree-level estimates, which decreases with increasing spatial scale due to the central limit theorem. Model findings reveal inequities in benefits accessibility across social groups, even after accounting for census tract tree density, unobserved covariates and spatial autocorrelation via random effects, and tree trait uncertainty. Higher-income areas with greater income inequality and lower proportions of minority ethnic populations tend to have greater access to ecosystem services. This study provides a replicable methodology applicable to urban tree inventories with similar data limitations and offers urban planners an analytical tool for designing targeted tree-planting strategies that promote equitable ecosystem services provision distribution and foster healthier, more resilient urban communities.

1. Introduction

Currently, 57.7 % of the global population resides in urban areas, a figure projected to increase to 67.9 % by 2050 (UN, 2025). Urban environments are particularly vulnerable to the effects environmental hazards due to high population densities and the unique atmospheric processes that define urban climates (Oke et al., 2017). The impacts of extreme temperatures in cities have already led to severe public health consequences. The three most lethal extreme heat events of the 21st century alone have led to nearly 200,000 fatalities across Europe (Robine et al., 2008; Barriopedro et al., 2011; Ballester et al., 2023). Future projections of urban warming present an increasingly alarming scenario, driven by the intensification of Urban Heat Islands (UHI) (Chapman et al., 2017; Rosenzweig et al., 2018), the rising intensity, frequency, duration, and spatial extent of heatwaves (Lorenzo et al., 2021; Domeisen et al., 2023), and the interaction effect of UHI-heatwaves (Founda and Santamouris, 2017).

Environmental hazards in urban areas extend beyond extreme heat, with air pollution and flooding also posing significant public health risks (Lelieveld et al., 2015; Hu et al., 2018). In 2021, an estimated 8.08 million (6.71 million – 9.48 million) deaths were attributed to air pollution, with nearly 97 % linked to particulate matter exposure (Murray et al., 2020). In Europe, despite a 33.7 % reduction in $PM_{2.5}$ concentrations between 1990 and 2019, this pollutant remained

* Corresponding author.
E-mail addresses: jperena@uma.es (J.F. Pereña-Ortiz), salvo@uma.es (Á.E. Salvo-Tierra), pablocozano@uma.es (P. Cozano-Pérez), arvalero@uma.es (Á. Ruiz-Valero).

https://doi.org/10.1016/j.indic.2025.100864
Received 2 April 2025; Received in revised form 9 August 2025; Accepted 17 August 2025
Available online 20 August 2025

responsible for 368,006 deaths in 2019 (Juginović et al., 2021). Regarding flood hazards, the number of people affected by flooding has steadily increased over the past four decades (Hu et al., 2018). Currently, an estimated 1.81 billion people, 23 % of the global population, are directly exposed to flooding events with a 100-year return period (Rentschler et al., 2022). Moreover, flood-related disasters are projected to become more severe, prolonged, and frequent due to the increasing occurrence of extreme precipitation events and rising sea levels (Tellman et al., 2021).

These challenges highlight the urgent need for effective adaptation and mitigation strategies to address the environmental issues affecting urban areas. The scientific literature has explored various approaches to enhance urban livability and resilience (Yu et al., 2020). Among these, urban tree planting has increasingly been recognized as a nature-based solution for mitigating multiple environmental challenges due to its ability to provide essential ecosystem services (ES) (Pataki et al., 2021). Ensuring the equitable distribution of urban tree canopy coverage across socioeconomic groups is essential to guaranteeing fair access to its benefits throughout the city (Schwarz et al., 2015). While there is a well-established correlation between total ES provision and the overall amount of green cover, the generation and supply of these services are largely influenced by tree-specific traits (Miedema-Brown and Anand, 2022; Liang and Huang, 2023). Furthermore, studies on environmental justice yield different results and conclusions depending on whether ES provision is directly assessed or if tree canopy coverage is used as a proxy, highlighting the need for a direct evaluation of ES (Riley and Gardiner, 2020).

The implementation of Municipal Urban Tree Master Plans, alongside the development of Urban Tree Inventories (UTIs), is increasingly gaining traction in Spanish cities as part of broader efforts to align with the Sustainable Development Goals (SDGs) (UN, 2023). UTIs serve as detailed censuses of municipally managed urban tree populations, systematically recording essential attributes such as species composition, structural traits, and geographic coordinates for each individual tree. However, it is common for municipalities to have incomplete inventories due to missing data or the use of ranges rather than numerical values for certain structural traits (Baró et al., 2019). These characteristics introduce significant uncertainty into the ES provided by urban trees when using mechanistic models such as i-Tree (Nowak, 2024). If this uncertainty is not properly accounted for in subsequent analyses, it can lead to overconfident results (Gelman et al., 2013).

The study of uncertainty in i-Tree has only recently gained attention (see Table 1 for a summary of the main contributions to date). Uncertainty in model estimations can be attributed to four main sources: (1) model misspecification and uncertainty in model parameter estimates, (2) sampling uncertainty, and (3) uncertainty in the values of independent variables. Previous research primarily focused on investigating the relationships between input and output variables in i-Tree Eco and identifying the most influential parameters for estimating urban forest structure and functions (Pace et al., 2018; Lin et al., 2020). While i-Tree Eco has proven valuable for urban forest planning and management (Lin et al., 2019), direct implementation of uncertainty quantification methods within i-Tree Eco has been limited exclusively to sampling uncertainty (Nowak et al., 2008). The quantification of the other types of uncertainty has been addressed only in isolated studies (Table 1), and despite some of these showing improved model performance, their results have yet to be fully incorporated into the i-Tree model framework.

With respect to model misspecification, this can be understood as the difficulty in accurately modeling the mechanistic processes underlying ES provision. This challenge has led to the development of various models with different structures (Lin et al., 2019), and the literature highlights the variability observed when comparing different model types (Aguaron and McPherson, 2012; Boukili et al., 2017), suggesting that uncertainty due to misspecification and model selection is a significant source of overall model uncertainty. In other fields of research, discrepancies in predictions across models have led to the development

Table 1
Summary of i-Tree Eco research on uncertainty quantification by type.

Uncertainty type	Studies	Main contribution	Key Findings
Integrated - Total	Lin et al. (2021)	First comprehensive Monte Carlo framework assessing input, sampling, and model uncertainties across multiple ecosystem services	Sampling uncertainty >90 % for most services; input/model <3 %; BVOC exception with 18 % sampling
Sampling	Martin et al. (2013)	Quantified inadequacy of standard 200-plot protocol against complete census data	Standard 200-plot protocol yields ±11–24 % error; requires 258–870 plots for ±10 % accuracy
	Lin et al. (2021)	Bootstrap simulation analysis demonstrating sampling uncertainty dominance	Sampling uncertainty accounted for >90 % of total uncertainty for leaf area, biomass, and carbon services
Input	Szkop (2020)	Quantified influence of air pollution monitoring station selection on model outputs	High variability between station types (CV: 0.24–0.57) vs. low within-type variability (CV: 0.05–0.23)
	Westfall et al. (2021)	Monte Carlo simulation framework assessing measurement variability impact using inter-observer replication	Tree measurement variability contributed minimally to ecosystem service uncertainty (standard error increase ≤0.4 %)
	Lin et al. (2021)	Monte Carlo simulations with independent normal distributions to propagate tree traits measurement error	Negligible effects of input uncertainty on total uncertainty (<3 %)
Model	McPherson et al. (2016) Timilsina et al. (2017)	Alternative leaf area estimation approaches using species-specific and DBH-based models vs. i-Tree defaults	Locally developed models outperformed default i-Tree Eco equations
	Lin et al. (2020)	Species-specific allometric equations for carbon storage estimation vs. generic forest-based equations	Species-specific equations improve carbon storage accuracy but limited by equation availability and applicability
	Pace et al. (2021)	Direct field validation of PM2.5 deposition using eddy covariance measurements	Species-specific deposition velocity parameterizations reduced PM2.5 estimation discrepancy from 6 to 20 × to 2.2-7.2 × vs. field data
	Morani et al. (2014)	Direct field validation of ozone fluxes using eddy covariance in Mediterranean climate	Climate-specific stomatal conductance parameterization (m = 3 to 6.49) reduced ozone overestimation from 94 % to 0.68–16.95 %
	Kofel et al. (2024)	Comparison of species-specific versus genus-level parameterization for BVOC estimation accuracy	Species-level approaches showed 3-4 × differences in pollutant removal estimates vs. genus-level i-Tree parameterization
	Manzini et al. (2023)	integrates species-specific parameters to enhance pollutant removal estimates	Climate-adapted species selection framework identifies high-efficiency species under varying

(continued on next page)

Table 1 (*continued*)

Uncertainty type	Studies	Main contribution	Key Findings
		beyond generalized i-Tree Eco outputs.	environmental conditions
	Coville et al. (2022)	Direct field validation of i-Tree Hydro validated using empirical using paired-catchment experiments of tree removal runoff data.	Calibration improved accuracy substantially; default settings overestimated runoff and underestimated benefits.
	Lin et al. (2021)	Application of variance-covariance matrices and Student's t-distributions to generate random coefficient sets from the allometric equations	Negligible effects of model uncertainty on total uncertainty (<3 %)

of integrated modeling frameworks, where a combination of estimates from different models is produced. Fletcher et al. (2019) summarized a broad spectrum of such techniques, ranging from simple data pooling to more robust methods such as ensemble or stacking models, as well as frameworks based on joint likelihood or shared components. In the urban forestry field, however, these integrated approaches are not commonly used. Model selection is further complicated by the use of either urban-specific parameters or equations, which remain relatively limited, or forest-derived equations adjusted for open-grown urban trees.

i-Tree Eco currently estimates leaf area using crown-based allometric equations developed from Chicago Park tree data (Nowak, 1996). Alternative approaches include species-specific equations (McPherson et al., 2016) and DBH-based models (Timilsina et al., 2017). Peper and McPherson (2003) found Nowak's method slightly overestimates leaf area in northern California, and Timilsina et al. (2017) demonstrated that locally developed models outperformed default i-Tree Eco equations for 74 urban trees across 5 species in Wisconsin. However, these comparisons are limited to few species and single sites. While locally developed allometric relationships can be superior, they require sufficiently intensive and representative datasets (Van Breugel et al., 2011) and substantially increase analysis costs.

Carbon storage and carbon sequestration in i-Tree Eco rely on forest-based allometric equations with biomass correction factors for open-grown environments (Nowak et al., 2013a). While these approaches are popular in urban forestry, they may yield conservative biomass estimates (Aguaron and McPherson, 2012). Lin et al. (2020) observed that employing species-specific allometric equations could substantially improve the accuracy of carbon estimations. However, urban-specific allometric relationships remain scarce and highly location-dependent, with existing equations typically developed for street trees (McPherson et al., 2016), thus limiting their widespread application.

For pollutant removal, the i-Tree Eco model applies constant deposition velocities and resuspension rates across all tree species, leading to substantial uncertainty when applied in diverse urban forests (Nowak et al., 2013b). Moreover, the model's assumption of complete particulate matter removal following rainfall events that exceed canopy storage capacity fails to capture species-specific retention dynamics (Xu et al., 2020). Field validation indicates that rainfall removes only 51–70 % of accumulated particles (Xu et al., 2017), with retention varying significantly depending on leaf surface traits. Pace et al. (2021), using eddy covariance in a Mediterranean urban forest for direct quantification of deposition fluxes at the ecosystem scale, revealed temporal mismatches between modeled and observed peak deposition rates. Canopy-scale measurements indicate that while i-Tree Eco simulations capture general deposition patterns, they respond weakly to atmospheric and meteorological variability. The standard model underestimates cumulative PM2.5 values, by 6–20 times compared to field measurements, whereas species-specific parameterizations reduce this gap to between 2.2 and 7.2 times. Authors suggest to incorporate species-specific leaf traits, such as deposition velocity, particle retention capacity, and wash-off rates to enhance model performance by avoiding generalized assumptions, such as uniform resuspension behavior across tree types (Hirabayashi et al., 2012). Pace and Grote (2020) also provide a detailed review of the factors that should be considered in pollutant removal models within i-Tree Eco to reduce the uncertainty associated with model specification.

Morani et al. (2014) validated i-Tree ozone flux estimates against 74 days of eddy covariance measurements in a Mediterranean forest near Rome. They found that the standard i-Tree model overestimated ozone uptake by 94.18 %. Sensitivity analyses showed that the stomatal conductance coefficient (m) was the key driver. Adjusting m to 3 reduced overestimation to 16.95 % in warm periods, and using $m = 6.49$ yielded only 0.68 % overestimation under cooler, well-watered conditions (>15 % soil moisture). They found climate-specific parameterization critical for accurate ozone flux and ES estimates in water-limited environments.

Similar limitations have been identified for Biogenic Volatile Organic Compounds (BVOC) estimations, where i-Tree's genus-level parameterization may lack sufficient refinement to identify species that disproportionately impact urban air quality. Kofel et al. (2024) compared species-specific BVOC calculations with i-Tree Eco estimates, revealing substantial discrepancies, with differences ranging from three-to fourfold between methods. The species-level approach likely overestimated removal by using annual mean conditions versus i-Tree Eco's hourly temporal variability, while species-specific emission factors captured greater within-genus variation than genus-level parameterization. Despite producing different absolute values, both approaches yielded results within the same order of magnitude, suggesting genus-level parameterization provides reasonable estimates but species-level refinement could improve accuracy for identifying key species with disproportionate air quality impacts, supporting enhanced modeling frameworks like FlorTree (Manzini et al., 2023).

Despite the widespread use of i-Tree models to estimate the hydrologic benefits of urban trees, empirical field validation at the sewershed scale remains critically limited (Kuehler et al., 2017; Coville et al., 2022). i-Tree Eco, an adaptation of i-Tree Hydro, using national average values and assumptions, such as full infiltration on permeable surfaces and total runoff on impermeable ones, can estimate avoided runoff by comparing tree-covered and treeless scenarios. In contrast, i-Tree Hydro applies a more detailed, process-based framework, incorporating dynamic water table tracking, advanced infiltration modeling, and distinctions between connected and disconnected impervious surfaces. Both models, however, rely on the same algorithms for plant-specific hydrological processes (Hirabayashi, 2013, 2015; USDA Forest Service, 2020). No empirical validation studies have been found for i-Tree Eco, while for i-Tree Hydro, validation beyond the original publications describing the model and its updates (Wang et al., 2008; Yang et al., 2011), which reported satisfactory results, is limited to a single study. Coville et al. (2022), using empirical data from Selbig et al. (2022), who measured runoff before and after the removal of 31 street trees (2990 m^2 of canopy) along a 400-m street segment, showed that calibrated i-Tree Hydro produced accurate results. The model estimated avoided runoff within 4 % of observed values, predicting 6120 L/tree and 63.5 L/m^2, compared to field measurements of 6380 L/tree and 66 L/m^2. However, uncalibrated model outputs predicted nearly twice the total runoff and only 34 % of the avoided runoff, highlighting discrepancies between default parameter settings and field-calibrated performance.

Beyond studies comparing alternative parameterizations of i-Tree or its validation against empirical data, which if used to inform future integrated prediction models or model selection, could be considered a form of model specification uncertainty, sampling uncertainty has received more attention than input uncertainty since the early

development of i-Tree Eco. The standard i-Tree Eco sampling protocol recommends a total of 200 plots; however, for example, Martin et al. (2013), comparing against a full urban tree census, estimated that between 258 and 870 plots with at least one tree present would be required to achieve an allowable error of ±10 %, depending on the ES being assessed. More recently, Lin et al. (2021) found that for leaf area, biomass, carbon storage, and carbon sequestration sampling uncertainty accounted for over 90 % of the total uncertainty.

Input uncertainty in i-Tree Eco can be disaggregated into two components: (1) uncertainty associated with meteorological and pollutant concentration data. Current i-Tree Eco restricts air pollution modeling to single monitoring station data inputs, despite significant spatial heterogeneity in urban pollutant concentrations. Szkop (2020) quantified the influence of air pollution data source selection on model outputs using data from different monitoring station types, revealing substantial variation in i-Tree Eco pollutant removal estimates based on input data source. Coefficient of variation analysis showed low variability among urban background stations but high variability between station types, demonstrating that single-station approaches substantially underestimate spatial variability in ES provision across heterogeneous urban pollution environments. (2) Uncertainty associated with measured tree traits. Current urban forest inventory applications largely overlook the systematic quantification of how measurement variability affects ES estimates, despite broad recognition that tree attribute measurements are inherently uncertain (Roman et al., 2017; Bancks et al., 2018). Common practices for handling incomplete or range-based observations often involve discarding data points or imputing missing values using genus-level averages or the midpoint of reported ranges (e.g., Baró et al., 2019; Kofel et al., 2024). Westfall et al. (2021) applied a Monte Carlo simulation framework to assess the influence of measurement variability on ES estimates by systematically perturbing the original measurements to replicate inter-observer variability. Each perturbed dataset was then processed through the i-Tree Eco software. The results showed that variability in tree measurements contributed minimally to overall uncertainty in ES estimates, with standard error increases of approximately 0.4 % or less across most services.

Lin et al. (2021) represent the most significant advancement to date in quantifying uncertainty in i-Tree Eco estimates. They employed a comprehensive Monte Carlo uncertainty analysis framework to assess input, sampling, and model structure uncertainties. Input uncertainty in tree measurements was quantified by modeling measurement error based on the USDA Forest Service's Forest Inventory and Analysis (FIA) field guide standards. These standards define measurement tolerance (MT) as the acceptable measurement range, and measurement quality objective (MQO) as the minimum percentage of measurements required to fall within that range. Trait-specific MT and MQO values were used for standard deviation calculation to construct independent normal distributions for each tree trait, serving as probability distributions in Monte Carlo simulations to propagate measurement error. Sampling uncertainty was evaluated through bootstrap resampling of plot data with replacement. Model uncertainty was incorporated by applying variance-covariance matrices and Student's t-distributions to generate random coefficient samples from the allometric equations used. Total uncertainty was then calculated under the assumption of independence among the three sources using variance summation. Their results showed that sampling uncertainty dominated total uncertainty, contributing more than 90 % for most ES, while input and model uncertainties remained negligible (<3 %). However, for BVOC emissions, sampling uncertainty averaged 18 %, suggesting that all three sources contribute meaningfully to the overall uncertainty in BVOC estimates.

Building upon the reviewed literature, an integrated approach to address missing data and range-measured traits (henceforth referred to as traits uncertainty) remains underdeveloped. Furthermore, it is unclear how traits uncertainty propagates to ES estimates across spatial scales relevant to urban planning and when used as response variables in second-step regression analyses. This study develops a methodological framework to propagate tree traits uncertainty in ES estimation to socioeconomic inequalities in its accessibility, using the UTI of Málaga, Spain. The methodological approach integrates machine learning (ML) models, copula distributions, Monte Carlo simulations, and Bayesian hierarchical models (BHM). This integrated framework leverages the predictive capabilities of ML algorithms to estimate both the mean and standard error of missing functional trait values, rather than discarding incomplete observations or relying on averaging approaches. The incorporation of copula distributions enables the modeling of complex multivariate dependencies among tree traits while preserving their observed correlation structures within the UTI, thereby avoiding the limitations of univariate uncertainty simulations that ignore trait interdependencies. Finally, the combination of Monte Carlo simulations with BHMs provides a framework for propagating trait uncertainty through ES estimations and subsequently analyzing relationships with socioeconomic variables, enabling systematic evaluation of environmental justice patterns and quantification of how trait uncertainty affects the detection of inequitable access to urban forest benefits. This framework offers a replicable methodology for UTIs with similar data gaps and provides urban planners with a tool to guide equitable tree planting strategies under uncertainty in measured tree traits.

2. Material and methods

2.1. study area

Málaga has a population of 586,770 inhabitants, making it the sixth most populated city in Spain (NIS, 2023). It is a coastal city characterized by warm summers and mild winters, with an average annual temperature of 19.0 °C and an average annual total precipitation of 494.9 mm (NIS, 2022a). Málaga is divided into eleven municipal districts and 435 census tracts (Fig. 1), which constitute the smallest observational unit for which open-access socioeconomic data are available.

2.2. method

The traits uncertainty accounting framework and its influence on the assessment of inequities in accessibility to urban tree ES and exposure to biogenic organic volatile compounds (BVOCs) at census tract scale is summarized in Fig. 2.

2.2.1. urban tree inventory

UTIs only consider public managed trees, which means that trees on private properties, forested areas, unmanaged spaces, or vacant lands were excluded. For each tree, the following traits were recorded: species, diameter at breast height (DBH), trunk and total heights, crown height, crown diameter (measured in both N-S and E-W directions), and spatial coordinates. DBH and height were measured in diametric and altimetric classes generated by divisions of 10 cm and 5 m respectively.

2.2.2. machine learning (ML) imputation models

2.2.2.1. data pre-processing. Independent ML imputation models for crown diameter and trunk height up to the crown (trunk height) have been developed exclusively for trees. Palms were excluded from the imputation process since all their observations lack measurements for both variables (15,203 palm trees). Including them would lead to significant extrapolation issues, primarily due to the differing structural traits between palms and trees. Approximately 21 % of the trees (22,071) had missing data for the covariables, making them the target observations for imputation.

The following preprocessing steps were applied to the urban tree inventory to avoid including species with very low representation. (1) All observations belonging to the genera *Mespilus* and *Taxus* were removed as no individuals had recorded measurements for the variables

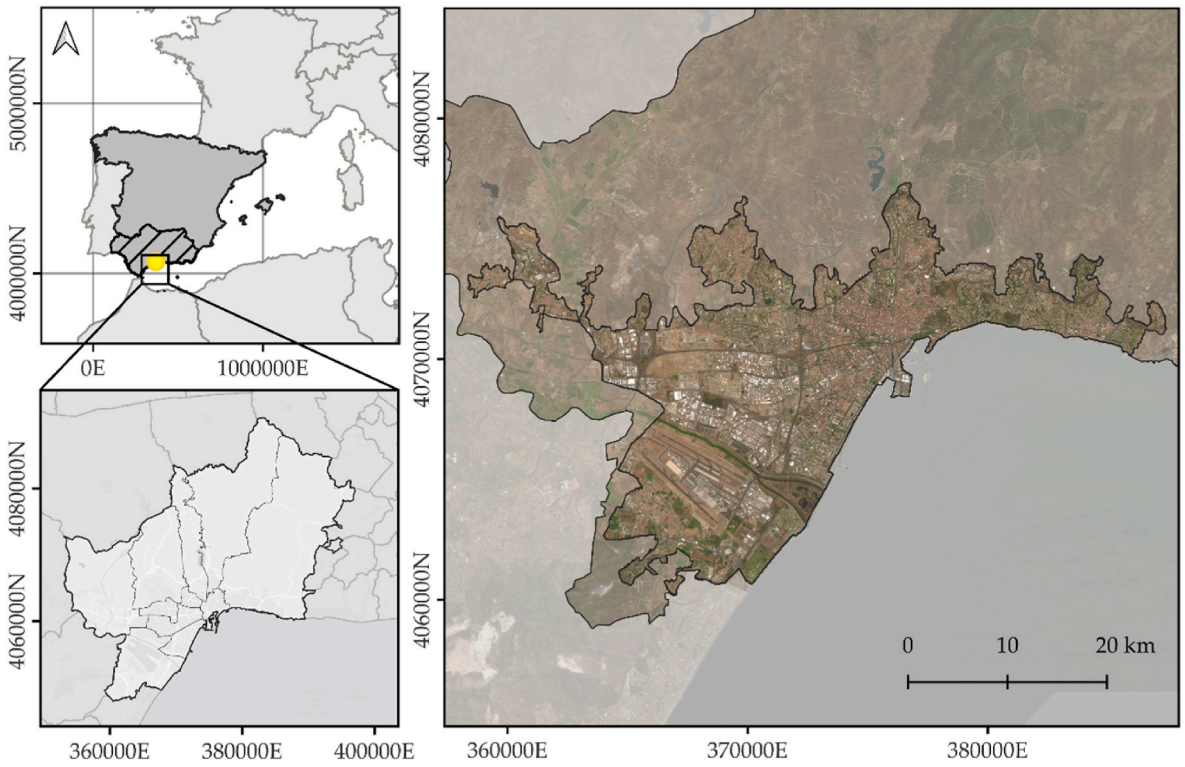


Fig. 1. Geographic Location of the Study Area focusing on the urban matrix. The background information includes data from the Spanish National Geographic Institute for the boundaries of Spain and its municipalities. The orthophotography used is from the 2022 National Plan for Aerial Orthophotography project (NIG, 2022). Map represented in coordinate reference system: EPSG 25830.

to be imputed. (2) Observations of *Albizia procera*, *Bauhinia candicans* and *Pinus sylvestris* were removed due to the absence of any recorded observation for the imputation variable. A total of 2576 trees were moved from the general training set to the prediction set due to missing crown diameter measurements. This step was necessary because the crown diameter was intended to be used as a predictor for trunk height up to the crown. Importantly this removal did not result in any species being composed exclusively of individuals with missing values thereby avoiding the need for additional preprocessing to mitigate extrapolation issues. Following these preprocessing steps the final dataset consisted of 62,387 observations for training the imputation models and 26,517 trees for which trunk height and crown diameter needed to be imputed.

Feature engineering was conducted via mlr3pipelines v.0.7.1 (Binder et al., 2021) and the following predictors were considered: family, genus, species, average crown diameter, trunk height, average DBH, average height, pruning history, and crown shape. (1) Family, genus, and species were transformed using conditional target value impact encoding. (2) Average crown diameter was used as a predictor for the imputation model of trunk height, while trunk height up to the crown was used as a predictor for crown diameter. (3) For DBH and total tree height, the mean value of both recorded ranges was used as a predictor in the models. (4) The type of the most recent pruning performed on each tree was included in the models using one-hot encoding. The pruning types recorded in the UTI dataset are as follows: maintenance pruning, repeated tertiary pruning with lengths above 5 m, repeated tertiary pruning with lengths below 2 m, crown lifting, no pruning, topping, topiary pruning, and thin branches removal. (5) Crown morphology was also included using one-hot encoding. The inventory records the following crown shape categories: globular/spherical, natural, and pyramidal/conical. Each type can also be classified as full-crown or half-crown. For trees without a crown, not associated with senescence, the "no crown" category was used.

2.2.2.2. model implementation, tuning and validation. The development ML models were carried out using the mlr3 framework in its version 0.22.1 in R (Lang et al., 2019; Bischl et al., 2024). Random Forest (RF) model (Breiman, 2001) was selected as the learner. RF was chosen due to its demonstrated ability to perform well across various types of tabular datasets, i.e., data without spatial or temporal dimensions (Ziegler and König, 2014). Although RF is recognized for its strong predictive performance and robustness, even with default hyperparameter values (Probst et al., 2019), the impact of tuning is dataset dependent. Therefore, optimizing hyperparameters is essential to ensure the model achieves its best possible predictive performance and minimizes generalization error for unseen data (Bischl et al., 2023).

Nested Cross-Validation (nested-CV) was employed instead of classical cross-validation methods due to its ability to mitigate bias when used for both hyperparameter tuning and model evaluation (Bischl et al., 2023). Unlike classical cross-validation, nested-CV ensures that test data remain entirely separate from hyperparameter tuning (Schratz et al., 2019; Raschka, 2020; Bischl et al., 2023). As the outer resampling scheme, a Leave-Group-Out approach based on Feature Space Clustering (LGO-FSC), also known as "environmental blocking", was implemented (Roberts et al., 2017). K-prototypes clustering, implemented in the clustMixType v.0.4–2 R package (Szepannek, 2018), was applied to the predictor space to stratify it into ten groups. LGO-FSC was selected due to the interest in evaluating the predictive and generalization capabilities of ML models across trees with differing traits, to test if this approach could be considered for use in UTIs from other cities that may contain data combinations not observed by the models developed in the present study. A random 5-fold partition has been chosen as the inner resampling scheme. Root Mean Squared Error (RMSE) was selected as the minimization criterion in the nested-cv framework.

For hyperparameter tuning, Bayesian Model-Based Optimization (MBO) was employed using the mlr3mbo package v.0.2.8 (Schneider et al., 2025). The hyperparameter tuning space was selected based on the recommendations of Bischl et al. (2023). A stagnation criterion of no

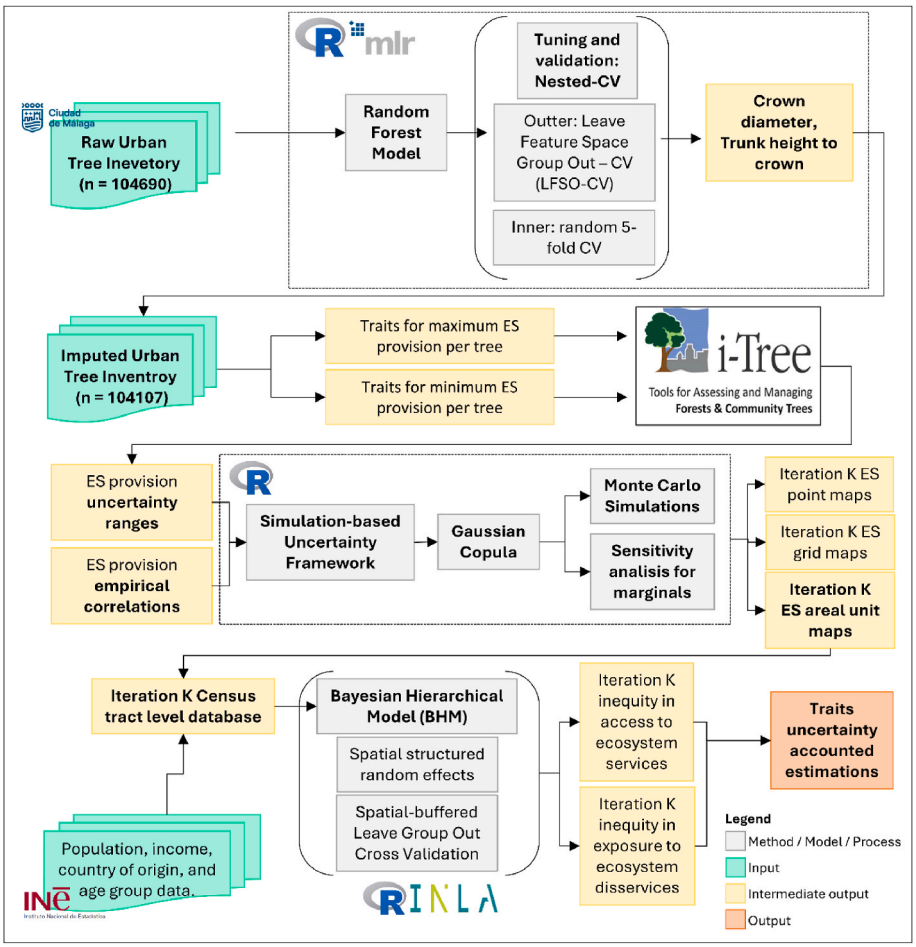


Fig. 2. Flowchart summarizing the methodology developed for assessing inequities in accessibility to urban tree ecosystem services and exposure to ecosystem disservices considering traits uncertainty.

improvement of at least 0.025 in the RMSE for the last 25 iterations was chosen as the termination criterion of the optimization process.

The hyperparameter configuration yielding the best performance during model evaluation for both RF models, was selected to fit the final model on the complete training set. RF models were then used to make predictions of the mean and standard error for their respective response variables.

2.2.3. *iTree-Eco*

To assess the ES provision and BCOV production of urban trees, the i-Tree Eco model, which is widely used in urban forest studies (e.g. McPherson et al., 2011; Nowak et al., 2014; Nyelele and Kroll, 2020), was used. For this study, the selected weather station is part of the AEMET (State Meteorological Agency) monitoring network and is situated at Málaga Airport (36° 39′ 58″ N, 4° 28′ 56″ O) at 8 km from the city center for the year 2015.

ES estimates reported from iTree and further analyzed were: (1) Yearly removal of carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), particulate matter less than 10 μm (PM_{10}), particulate matter less than 2.5 μm ($PM_{2.5}$), and sulfur dioxide (SO_2); (2) carbon storage; (3) annual gross carbon sequestration; (4) oxygen production; (5) annual emission of BVOCs (as isoprene and monoterpenes emissions); and, (6) annual avoided run-off. Further details of the i-Tree Eco methods are available at Nowak (2024). All the indicators were estimated at the individual tree level as point-support data. Data was rasterized to a 10 x 10-m grid and aggregated to a census tract support to allow statistical analyses with socio-demographic variables.

2.2.3.1. data set post-processing for iTree. Considering the measured trait in ranges and the predicted mean values and standard errors from the imputation models, the following datasets were generated to evaluate the range of uncertainty in ES provision. The selection of trait values combinations to determine the maximum and minimum provision estimates was based on the guidelines outlined by Lin et al. (2021), Kofel et al. (2024), and Nowak (2024).

For palm, only DBH, height, and species information were used in i-Tree, due to lack of data in the other traits. For tree species crown diameter, live crown height, base-of-crown height (i.e., trunk height to crown), and the percentage of missing crown were also included. Given the absence of measurements for the latter, a default assumption of 0 % missing crown was applied to all trees. However, for trees classified as having “half a crown” in their crown morphology field, a 50 % missing crown value was assigned.

(1) Dataset of minimum ES provision considering trait uncertainty. For palm trees, the lowest values within their DBH and height ranges were selected. For other tree species, the lowest values for DBH and height were also considered. Imputation consisted of mean plus one standard error. In cases where the imputed base-of-crown height exceeded total tree height (3729 trees, 4.19 %), the mean prediction was used instead. For crown diameter, the imputed value was set to the mean prediction minus one standard error. For trees recorded in the UTI without a measured crown, as well as those with an imputed crown diameter of zero or less, a minimum crown width of 0.5 m was assigned since i-Tree does not allow a crown diameter of zero.

(2) Dataset of maximum ES provision considering trait uncertainty. The highest values within the DBH and height ranges were used, along with the mean prediction for crown diameter plus one standard error and the mean prediction for base-of-crown height minus one standard error.

2.2.4. uncertainty simulation framework

2.2.4.1. copula and Monte Carlo simulations for ES provision uncertainty. A copula is a function that captures the dependence structure between random variables, independent of their marginal distributions. A d-dimensional copula is a cumulative distribution function (CDF) with uniform marginals, satisfying monotonicity, marginality, boundary conditions, and non-negativity constraints. Sklar (1959) states that any joint distribution can be decomposed into a copula and its marginal distributions. If the marginals are continuous, the copula is unique, making it a fundamental tool for modeling dependencies in multivariate distributions. Further details can be found in Nelsen (2005) & Schmidt (2007).

A Gaussian copula was adjusted, using copula package v.1.1–4 (Hofert et al., 2025), because the only prior knowledge available were the ranges of the marginal distributions and the correlations between ES. A copula was specified for each tree. The provision of every ES is considered as a random variable, which calculated ranges are selected as ranges of the marginal distributions. Correlations between marginals of every tree were selected depending on the number of observations. For trees belonging to species with more than 50 individuals, species-specific linear correlations are used. For trees that do not meet this condition but belong to genera with more than 10 observations, genus-specific linear correlations are assigned. Trees with fewer than 10 individuals in their respective genus are assigned family-specific correlations. Trees that do not meet any of these criteria are assigned correlations based on the entire UTI dataset.

To transpose the uncertainty in the ES estimates for each tree into the BHM for studying inequities, a Monte Carlo simulation process was implemented. A total of 1000 iterations were performed, during which a provision vector, representing the set of provision values for each ES for a tree, was randomly sampled independently for all 104,107 copulas. It represents plausible provision values based on the traits uncertainty. The results from each iteration are stored for subsequent aggregation at both the grid and census tract levels, with the latter serving as the observational unit for fitting the BHM in each iteration.

2.2.4.2. sensitivity analysis. Due to the lack of prior knowledge regarding the appropriate probability distribution for copula marginals, a sensitivity analysis was conducted to evaluate its impact on simulations and estimated inequities in ES provision. Various truncated distributions were selected using the crch package v.1.2-1 (Messner et al., 2016). The framework transitions from low to increasing certainty in mean provision values, with probability distributions in ascending order: uniform, truncated t-Student (high standard deviation), truncated normal (high standard deviation), truncated t-Student (low standard deviation), and truncated normal (low standard deviation). Truncated normal distributions had 100 degrees of freedom, while truncated t-Student had one. High and low standard deviations were set to one-quarter and one-eighth of the range, respectively.

2.2.5. Bayesian Hierarchical Model

i-Tree outputs were used as response variables in the modeling process after aggregation at the census tract level. Only microscale ES and disservices were considered in the modeling process, as their benefits are realized in their immediate surroundings, making them particularly relevant for urban distributional justice analyses. These include oxygen production, avoided runoff, BVOC emissions, and total pollutant removal. Transpiration, as a proxy for cooling capacity, was excluded due to the predominant role of tree canopy shading in temperature reduction (Zardo et al., 2017). Additionally, as shown in Figure A.1, transpiration exhibits a high correlation (0.968) with avoided runoff at the individual tree level, indicating that any estimated inequities in access to avoided runoff are directly applicable to tree transpiration as well. Carbon stocks and carbon sequestration were also excluded, as their benefits operate on a global rather than a local scale.

2.2.5.1. socio-economic covariables. Initially, the following socioeconomic variables were considered as covariates to assess inequalities in accessibility and exposure to ES: (1) population distribution by country of origin and age group (Annual Population Census, NIS, 2024); (2) average household income and Gini index (Household Income Distribution Atlas, NIS, 2022b); and, (3) unemployment rate and percentage of the population with higher education (Quarterly Economic Activity Indicators of the Population of Andalusia, ISCA, 2023).

The population data were further disaggregated into: (1) total population within each census tract, (2) percentage distribution by country of origin, and (3) percentage by age-group. Population data by country of origin were reclassified into five broad categories: Africans, Americans, Spaniards, Europeans, and Asians/Oceania. These regional population data were transformed into simplex covariates. Similarly, the population distribution by age group was categorized into three groups: under 29 years old, 30–64 years old, and over 65 years old, and subsequently converted into simplex covariates.

To address multicollinearity issues inherent to simplex covariates, the isometric log-ratio (ILR) transformation was applied using the compositions R package (v.2.0–8) (van den Boogaart et al., 2024). Pearson's correlation coefficient and the generalized variance inflation factor (GVIF) (car R package (v.3.0–12), Fox and Weisberg, 2019) were computed prior to model implementation. Pairs of variables with high correlation (Pearson's $r > 0.6$) or high GVIF (GVIF >5) were identified, and only one variable per pair was retained in the model. Due to a GVIF exceeding 6 and a correlation coefficient greater than 0.6 with income, the unemployment rate and the percentage of the population with higher education were excluded from the final model.

The final set of selected variables comprised the Gini index, log-transformed household income, log-transformed total population, ILR-transformed percentage of population by origin, and ILR-transformed percentage of population by age group and tree density which accounts for potential effects related to variations in tree cover availability across census tracts. All retained covariates exhibited a maximum GVIF of 2.42. To facilitate model fitting in R-INLA, all covariates were standardized (z-score transformation).

2.2.5.2. Model specification. BHM are widely used in spatial statistics due to their capacity to account for spatial autocorrelation by incorporating spatially structured priors on random effects. Properly accounting for spatial autocorrelation is essential, as neglecting these dependencies can lead to overly narrow credibility intervals and biased estimates of the relative importance of model coefficients (Banerjee et al., 2014). A Latent Gaussian Model, which can be considered a type of BHM with an additive structure for the linear predictor and observed data modeled through a likelihood function that depends only on the value of the linear predictor, was implemented using the Integrated Nested Laplace Approximation (INLA) framework and the INLA R package (v.22.12.16) (Rue et al., 2009).

Different model structures were evaluated as detailed in Section 2.2.4.3. Equation (1) describes the selected model structure

$$log(Y_{i,j})|\eta_{i,j}, \phi_j \sim Gaussian\ (\eta_{i,j}, \phi_j)$$

$$\eta_{i,j} = \alpha_j + \sum_{k=1}^{K} \beta_{k[j]} X_{ik} + \theta_{i,j}$$

$$\beta_{k[j]} \sim N(0, \delta_j)$$

$$\theta_{i,j} = \frac{1}{\sqrt{\tau_j}} \left(\sqrt{1-\psi_j}\, u_{i,j} + \sqrt{\psi_j}\, v_{i,j} \right) \quad (1)$$

$$v_{i,j} \mid v_{-i,j} \sim N\left(\frac{\sum_{h \in N_h} v_h}{n_i}, \frac{1}{n_i} \right)$$

$$u_{i,j} \sim N(0,1)$$

$$\alpha_j \sim N(0, 0.001)$$

$$\phi_j \sim \text{Gamma}(0.001, 0.001)$$

$$P(\tau_j) = \frac{\lambda}{2} \tau_j^{-\frac{3}{2}} exp\left(-\lambda \tau_j^{-\frac{1}{2}} \right), \tau > 0, \lambda > 0, \lambda = \frac{-ln\, a}{U}$$

$$P\left(\frac{1}{\sqrt{\tau_j}} > U_1 \right) = a_1, U_1 = 2, a_1 = 0.5$$

$$P\left(\frac{1}{\sqrt{\delta_j}} > U_2 \right) = a_2, U_2 = 1, a_2 = 0.2$$

$$P(\psi_j < U_3) = a_3, U_3 = 0.5, a_3 = 0.5$$

$Y_{i,j}$ represents the response variable (i.e., ES) j in census tract i. $Y_{i,j}$ was log-transformed to address its right-skewed distribution. Φ_j denotes the precision of the normal distribution of the data (likelihood). $\eta_{i,j}$ is the value of the linear predictor. α_j represents the model intercept. $\beta_{k[j]}$ represents the varying slope effects for the covariate k on ES j. A partial-pooling approach was applied, assuming that inequity patterns in access to ES should be similar across different services and originate from a common distribution. For the hierarchical priors of the effects, penalized complexity (PC) priors were employed (Fuglstad et al., 2019). The marginal standard deviation of the random effect (with zero mean) was set such that there is a 0.2 probability of exceeding a standard deviation of 1 for the varying slopes distribution for each covariate. $\vartheta_{i,j}$ represents the effect from a BYM2 model for observation i and ES j. BYM2 is an extension of the BYM model (Besag-York-Mollié Model) (Besag et al., 1991), proposed by Riebler et al. (2016), where a scaled variance approximately equal to one exists between the i.i.d. (independent and identically distributed) effects and the CAR (Conditional Autoregressive Model) components of the BYM2. In the CAR specification, h denotes the set of neighboring observations for i based on an adjacency matrix W, using the queen contiguity condition. The main advantage of the BYM2 model is its ability to separately capture the impact of spatial dependence and the effect of data variability. The random effects v_i, which constitute a structured effect via the CAR prior, and u_i, modeled as an i.i.d. effect, together form the implementation of BYM2. These effects are modeled through τ_j, which represents the marginal precision contribution from both random effects, and ψ_j, which indicates the fraction of the variance explained by the spatially structured random effect. Both parameters were specified using PC priors, assuming a 0.5 probability for the standard deviation of τ_j to exceed 2, and for ψ_j, a probability of 0.5 that the fraction of the total variance attributed to the spatial dependency structure exceeds 0.5. In contrast to the precision parameter τ_j, the resulting PC prior for ψ_j cannot be specified in closed form due to the dependence on the graph structure of the Gaussian Markov Random Field (Riebler et al., 2016).

2.2.5.3. Model selection and validation. Thirty models were tested to compare: (1) whether incorporating spatially structured random effects as prior information improves the model's predictive capacity and fit, and (2) whether incorporating multivariate structure to the spatial models leads to improvement. Additionally, three approaches were considered for estimating the covariate effects for each model structure: common fixed effects (complete-pooling), independent fixed effects (no-pooling), and random varying slopes across different ES (partial-pooling). M0 represents the model in which spatial effects are not included, it represents a linear regression model. M1-M6 are shared component effect models (SCM), where the spatially structured random effect estimated for avoided runoff is shared among the other ES (Held et al., 2005). M7-M18 are models that, in addition to the SCM, incorporate spatially structured random effects specific to each response variable. M19-M20 represent models with independent intrinsic MCAR (Multivariate CAR) and independent proper MCAR (Palmí-Perales et al., 2022) respectively. M21-M29 are models where a specific spatially structured effect is adjusted independently for each response variable. For further details on the model structures, please refer to Table A.1.

A Buffering-Leave Group Out Cross Validation (b-LGOCV) approach was implemented to evaluate model performance, as recent literature suggests that LGOCV is a more appropriate alternative to Leave-One-Out Cross Validation (LOOCV) when assessing the predictive performance of models incorporating structured random effects (Adin et al., 2024). LGOCV was applied using the recently developed framework by Liu and Rue (2022), which enables the calculation of cross-validation scores without the need to rerun the model for each resampling iteration. A 1500 m buffer was applied around each observation, based on the average estimated range for each response variable. The range was derived from an exponential semivariogram fitted to the residuals of a linear model, using the centroids of the census tracts. The semivariograms were computed using the gstat R package (v.2.0–7) (Pebesma, 2004). Observations within this buffer were used as the test set.

The model performance was formally evaluated using the LGOCV-based log-score function (LGOCV score) (Equation (2)). Additional metrics were calculated and are presented in Tables A.2 and A.3.

$$LGOCV\ score = \frac{1}{n} \sum_{i=1}^{n} \pi\left(Y_i = y_i \middle| \boldsymbol{y}_{g_i} \right) \quad (2)$$

where, for each observation i, its Posterior Predictive Density (PPD, π) is computed according to the group structure g defined in the resampling scheme of b-LGOCV for that observation.

2.2.5.4. Uncertainty propagation in inequities estimates. Each Monte Carlo iteration of its respective sensitivity analysis case was used to fit the model. In the Bayesian framework, every variable estimated in the model, whether observable or latent, is treated as a random variable with an available posterior probability distribution (Gelman et al., 2013). Thus, in each iteration, a sample of 1000 observations was drawn from the posterior distributions of the covariate effects. By concatenating the vectors containing these posterior samples for a given variable k, a new distribution is generated, thereby accounting for the uncertainty in the estimation of ES and propagating this uncertainty into the patterns of socioeconomic inequity in access to ES.

3. Results

3.1. urban tree inventory

The urban forest of Málaga consists of 104,690 trees, of which 14.6 % (15,252) are palms and 85.4 % (89,438) are trees. Further details on the distribution by district can be found in Table A.4. A total of 195 different genera have been identified in Málaga, with the 44 most abundant genera accounting for 95 % of the urban forest composition. The five most abundant genera are *Citrus* (12,343, 11.8 %), *Ficus* (9,786, 9.35 %),

Washingtonia (9,619, 9.19 %), *Brachychiton* (7,758, 7.41 %), and *Jacaranda* (6,943, 6.63 %). The specific richness is 417. Half of Málaga's urban forest is represented by the 11 most abundant species, while the 104 most abundant species make up 95 % of its composition. The 10 most represented species are: *Citrus aurantium* (11,754, 11.2 %), *Washingtonia robusta* (8,286, 7.91 %), *Jacaranda mimosifolia* (6,941, 6.63 %), *Tipuana tipu* (5,892, 5.63 %), *Ficus microcarpa* (3,854, 3.68 %), *Brachychiton populneus* (3,267, 3.12 %), *Melia azedarach* (2,775, 2.65 %), *Cupressus sempervirens* (2,759, 2.64 %), *Brachychiton acerifolius* (2,248, 2.15 %), and *Ficus benjamina* (2,174, 2.08 %).

3.2. Urban trees ecosystem services

Results of the provision of ES by the whole urban forest are provided in terms of average supply with uncertainty ranges defined by minimum and maximum provision. The urban forest of Málaga stores 17,769 Tn C [13,437–22,191], sequesters 1285 Tn C/y [975–1604], reduces runoff by 0.011 hm^3/year [0.009–0.014], emits BVOCs totaling 15.38 Tn/y [11.98–18.83], produces oxygen at an average rate of 3429 Tn/y (T O_2/year) [2602–4278], and removes pollutants at an annual rate of 38.72 Tn/y [30.59–47.00]. Additional queries regarding the provision for each ES by district can be found in Table A.5, and for each census tract through Figs. A.2 – A.7.

The spatial distribution of ES provision and BVOC production is represented in Figures A.8–A.13. Significant variability per square meter is observed across the study area, reflecting the wide diversity in species composition and structural characteristics of the urban tree canopy within 100 m^2 areas. On average, urban tree cover is estimated to store 2.5 kg C/m^2 ± 3.5 [0.08, 12.35], with an annual carbon sequestration rate of 0.18 kg C/m^2 ± 0.2 [0.01, 0.69], avoiding 1.56 L/m^2 ± 2.29 [0.02, 7.46] of runoff, producing 2.17 g/m^2 ± 8.14 [0, 11.75] of BVOCs, generating 0.48 kg O_2/m^2 ± 0.52 [0.03, 1.85], and removing 5.45 g/m^2 ± 7.98 [0.08, 26.06] of air pollutants.

The uncertainty associated with the simulation process decreases as the spatial scale of analysis increases (Table 2). Shifting from examining variations in ES provision linked to trait uncertainty at the individual tree level or rasterized data to the census tract scale results in an approximate 20 % reduction in the mean coefficient of variation, with an even greater decrease observed for carbon storage. At the finer spatial scales, carbon storage exhibits the highest uncertainty across simulations. However, at the census tract scale, BVOC production becomes the most uncertain ES.

3.3. ML imputation models performance

The RF model selected for imputing canopy diameter had the following hyperparameters: minimum node size = 17, proportion of variables selected per split (mtry) = 0.245, number of trees = 675, and sample fraction = 0.81. This model achieved a median RMSE of 2.14 ± 3.04 m and a median R^2 of 0.299 ± 0.109 across the 10 test sets. Similarly, the RF model for imputing trunk height had a minimum node size of 23, mtry = 0.405, number of trees = 615, and a sample fraction of 0.4, yielding a median RMSE of 1.08 ± 0.41 m and a median R^2 of 0.265 ± 0.107. Notably, if a robust validation approach such as nested-CV were not employed and instead a random 10 fold CV was used, the models would have exhibited inflated performance metrics. Specifically, the crown diameter model would have achieved an R^2 of 0.72 ± 0.018 and an RMSE of 2.00 ± 0.13 m, while the trunk height model would have yielded an R^2 of 0.32 ± 0.014 and an RMSE of 1.05 ± 0.03 m, highlighting the risk of overestimation when using less rigorous validation methods.

Table 2
Distributions of the coefficients of variation for different ecosystem services across three spatial scales: individual trees (point support), 100 m^2 plots associated with rasterization (pixel support), and census tract scale (areal unit support). Results are presented as the mean ± standard deviation, along with the 0.025 and 0.975 quantile values derived from the simulations.

Ecosystem Service	Point support	Pixel support	Aerial support
Carbon storage	28.3 ± 14.9 [0.07, 57.9]	24.9 ± 14.4 [0.1, 57]	2.76 ± 3.45 [0.55, 12.99]
Carbon sequestration	19.9 ± 12.6 [0, 49.5]	17.7 ± 11.8 [0, 45.3]	2.34 ± 2.79 [0.56, 9.52]
Avoided runoff	17.9 ± 15.4 [0.9, 55.4]	16.3 ± 14.4 [0.8, 53.8]	2.33 ± 2.67 [0.51, 8.62]
BVOC production	19.1 ± 15.1 [0.6, 56.1]	17.3 ± 14.2 [0.6, 54.4]	3.34 ± 3.58 [0.74, 15.98]
Oxygen production	19.8 ± 12.6 [0, 49.5]	17.6 ± 11.8 [0, 45.3]	2.34 ± 2.8 [0.56, 9.43]
Pollutant removal	17.7 ± 15.5 [0.6, 55.3]	16.1 ± 14.4 [0.6, 53.6]	2.29 ± 2.66 [0.5, 8.51]

3.4. BHM results

The model selection process favors models that incorporate spatially structured priors in their structure (M1-M29) over the simple linear regression model that does not account for spatial autocorrelation (M0) (Table A.1). Based on the results obtained from the LGOCV score, two distinct groups are observed. (1) Models with multivariate structure in spatial components (M1-M18), which show the best results in the LGOCV score. The collinearity among ES (Fig. A.1.) leads to perfect predictions when spatial patterns are shared (SCMs) between the response variables. In such cases, the probabilities assigned by the model can be very high for all correct responses, resulting in a positive log-score. Furthermore, as the collinearity between response variables arises from the deterministic iTree models, these correlation patterns persist even in unseen data, thus explaining the performance of the model, even in test sets. Based on these considerations and given the lack of interest in studying whether ES provision shares spatial patterns due to these being model-determined rather than observed data, models with univariate structures in the spatial component were preferred (M19-M29).

(2) M19-M29 show very similar LGOCV scores (Table A.1). This suggests that the choice of spatially structured prior (Besag, Besag proper, or BYM2) has little influence on the model's predictive performance. Given the highest LGOCV score for M29 and based on previous experiences, where the choice of prior for the spatially structured random effect had no significant impact on the estimates of the fixed effects, this model was selected. M29 shows an R^2 of 1 and a relatively low RMSE for the various ES (Table A.2), which can be attributed to the incorporation of an i.i.d. effect in the BYM2 structure, allowing the model to capture the specificities or anomalies in the provision of the different census tracts. As a result, the model can capture 100 % of the observed values for the provision of the various ES in the posterior distributions of the linear predictor (p-LP-HPDI) and predictive (p-PPD-HPDI), while also generating low average values for the ranges of the PPD across the entire dataset (Table A.3).

Model estimates of the standard deviations for the likelihoods of the response variables show very low values, indicating minimal unexplained variability by the model's linear predictor (Table 3). The i.i.d. component of the BYM2 model captures deviations and particularities not accounted for by the remaining components of the linear predictor, resulting in highly precise estimates. The random effects of the covariates originate from population distributions with standard deviations estimated with low uncertainty, except for the covariates ILR Asians, ILR 0–29, and ILR >65, where the estimation uncertainty spans several orders of magnitude. This is related to the fact that these variables are not related to the response variables. The model estimates that the spatially structured random effects are generated by distributions with variability exceeding one standard deviation, except in the case of oxygen

Table 3
Hyperparameters for model M29. The mean value, along with the 0.025 and 0.975 quantiles of its posterior distribution, are presented.

Hyperparameters	Global	Response variable specific			
		Avoided runoff	COV production	Oxygen production	Pollutant removal
Std. of the Gaussian Likelihood	–	0,008 [0,003, 0,019]	0,009 [0,003, 0,021]	0,009 [0,003, 0,024]	0,009 [0,003, 0,021]
Std. of Gini varying slope	0,341 [0,154, 0,687]	–	–	–	–
Std. of Log Income varying slope	0,294 [0,119, 0,637]	–	–	–	–
Std. of Log Population varying slope	0,464 [0,206, 0,912]	–	–	–	–
Std. of ILR Africans varying slope	0,432 [0,186, 0,844]	–	–	–	–
Std. of ILR Americans varying slope	0,337 [0,151, 0,669]	–	–	–	–
Std. of ILR Europeans varying slope	0,115 [0,026, 0,322]	–	–	–	–
Std. of ILR Asians varying slope	0,006 [0,002, 0,236]	–	–	–	–
Std. of ILR 0–29 varying slope	0,011 [0,004, 0,223]	–	–	–	–
Std. of ILR >65 varying slope	0,044 [0,017, 0,305]	–	–	–	–
Marginal Std. of BYM2 model (τ)	–	1212 [1,107, 1326]	1580 [1,434, 1739]	0,996 [0,908, 1090]	1212 [1,106, 1326]
Contribution of spatial structure in BYM2 (ψ)	–	0,510 [0,320, 0,697]	0,528 [0,328, 0,727]	0,489 [0,293, 0,687]	0,510 [0,320, 0,697]

production. The purely spatial effects of the BYM2 model contribute approximately 50 % of the variability in the effect, while the remaining portion is explained by the i.i.d. effect.

In general, the propagation mechanism of trait uncertainty leads to estimates of the effects, i.e. inequities in access to ES and exposure to BVOCs, with greater uncertainty estimated and mean values differing from those obtained using a single model fitted to average ES provision (Table A.6). Due to the high similarity of the sensitivity analysis results, a summarized version of the findings presented in Table A.6 is shown in the text (Table 4). Results show that the patterns of inequity among ES, specifically the sign and magnitude of the varying slopes estimated by the model, remain consistent across ES after controlling for urban tree density at the census tract level.

Positive relationships between population size in a census tract and the provision of each ES was found. In terms of elasticity, census tracts with a 5 % larger population tend to exhibit, on average, approximately 2 % higher provision of every ES. This suggests an equitable distribution of urban trees across census tracts relative to the number of residents.

Census tracts with higher average household income exhibit greater accessibility to ES. The uncertainty propagation method estimates stronger effects for this relationship, where, except for oxygen production, the mean difference for the remaining ES is close to 1 % in comparison to not propagating uncertainty. Similarly, credibility intervals are wider under the uncertainty propagation approach. While the inequity patterns associated with income levels remain qualitatively similar, the mean estimates indicate that exposure to BVOCs production and oxygen production are the most related to increasing income levels. In terms of elasticity, census tracts with a 5 % higher income level, holding all other variables constant, tend to have urban forests with, on average, 1.9 % greater annual runoff reduction, 2.18 % higher BVOCs production, 1.76 % higher oxygen production, and 1.97 % greater annual pollutant removal capacity.

The model estimates a strong positive effect of the Gini index on ES accessibility across census tracts, where census tracts with a less equitable income distribution show higher ES accessibility and BVOCs exposure. No differences are observed between uncertainty propagation framework and without it. In general, census tracts with a one-point increase in the Gini index tend to exhibit, on average, a 6.5 %–9.3 % higher provision of ES, with BVOCs production being the most positively associated and oxygen production showing the smallest estimated effect. Notably, while these estimates exclude zero from their 95 % credibility intervals, they still reflect considerable uncertainty, with intervals ranging from 7.4 % to 14.3 % in width.

Inequities in ES accessibility based on the region of origin indicate discrimination against minority ethnic groups, specifically Africans, Americans, and Europeans. Notably, within the latter two groups, individuals from the American continent are composed of 66.2 % ± 10.5 % Latin Americans compared to U.S. nationals across census tracts, while 82.5 % ± 9.91 % of Europeans in this category originate from Eastern Europe. Therefore, these two regional origins primarily reflect populations from countries with worse socioeconomic conditions than Spain, making them disadvantaged ethnic minorities. The model estimates a negative effect of an increasing ratio of Africans relative to the Spanish population in a given census tract, as well as of an increase in South Americans relative to the geometric mean of Spaniards and Africans, and of an increase in Eastern Europeans relative to the geometric mean of the previous three groups. Given that the Spanish-origin population dominates census tracts (87 % ± 6.4 %), these findings can be approximated as follows: census tracts with a higher proportion of Africans, South Americans, or Eastern Europeans relative to Spaniards tend to have lower accessibility to ES provision and BVOC exposure. The inequity in accessibility is estimated to be highest in census tracts with a greater proportion of Africans relative to Spaniards, compared to the other two groups. When comparing uncertainty propagation methods to those without propagation, estimates for inequities in ES accessibility for Africans remain largely similar. However, for Americans, estimates vary more significantly between methods, with uncertainty propagation reducing the estimated effect by more than two points while shifting and widening credibility intervals closer to zero. The no-uncertainty propagation approach does not estimate inequities in accessibility in census tracts with a higher proportion of Eastern Europeans. In contrast, uncertainty propagation does estimate a negative effect, with a 95 % credibility interval that does not include zero.

No significant differences are estimated in ES accessibility or BVOC exposure between census tracts with a higher proportion of the population of Asian or Oceanian origin. Similarly, no inequities in access appear to exist between age groups.

Fixed effects of increasing tree density in the census tract are of particular interest for land-use planning, as they help justify the benefits of increasing urban trees. As expected, an increase in urban tree density is positively correlated with both the provision of ES and the generation of BVOCs, regardless of the method employed. The estimated effect presents the same value since a fixed effect was estimated using complete pooling. The NU method produces slightly lower average fixed effect estimates compared to the uncertainty propagation methods, although the latter do not lead to an increase in the credibility interval ranges. It is estimated that census tracts with one additional tree per hectare tend to show about a 4.25 % [3.89 %, 4.58 %] higher provision of any of the studied ES.

4. Discussion

4.1. Ecosystem services provision

This study highlights the significant contribution of urban trees in Málaga to ES provision. The urban forest carbon storage (17,769 Tn C [13,437–22,191]) is comparable to results obtained by Nowak and Crane (2000). They observed carbon storage ranging from 1.2 million Tn

Table 4
M29 Fixed Effect Estimates. For each response variable, the fixed effect of each covariate is represented by its mean value along with the 0.025 and 0.975 quantiles. ΔX denotes the increase in the covariate that results in a ΔY change in the response variable. Covariates with a unit type (−) indicate they are dimensionless. No uncertainty propagation (NU), Uncertainty propagation with uniform marginals in the gaussian copula framework (UN).

	Avoided runoff (L/y) (ΔY = %)	COV production (g/y) (ΔY = %)	Oxygen production (Kg/y) (ΔY = %)	Pollutant removal (g/y) (ΔY = %)
Gini (−) (ΔX = 1)				
NU	**8.2 [3.78, 12.73]**	**9.24 [3.11, 15.25]**	**6.56 [2.7, 10.17]**	**8.2 [3.68, 12.82]**
UN	**7.98 [2.82, 13.33]**	**9.17 [2.33, 16.37]**	**6.49 [2.14, 10.96]**	**7.89 [2.6, 13.38]**
Log Income (Log €) (ΔX = 5 %)				
NU	**1.08 [0.27, 1.92]**	**1.1 [0.07, 2.15]**	**1.07 [0.36, 1.78]**	**1.08 [0.26, 1.88]**
UN	**1.9 [0.95, 2.86]**	**2.18 [0.88, 3.47]**	**1.76 [0.95, 2.57]**	**1.97 [0.99, 2.96]**
Log Population (Log n) (ΔX = 5 %)				
NU	**1.79 [1.24, 2.37]**	**2.04 [1.3, 2.75]**	**1.9 [1.46, 2.36]**	**1.79 [1.21, 2.37]**
UN	**2.03 [1.37, 2.7]**	**2.31 [1.42, 3.21]**	**2.17 [1.59, 2.74]**	**2.06 [1.37, 2.76]**
ILR Africans (−) (ΔX = 0.1)				
NU	**−6.86 [-9.54, -4]**	**−7.61 [-11.03, -3.71]**	**−5.21 [-7.8, -2.65]**	**−6.76 [-9.83, -3.75]**
UN	**−6.93 [-10.31, -3.46]**	**−7.66 [-12.1, -3.06]**	**−5.88 [-8.8, -2.89]**	**−6.94 [-10.42, -3.36]**
ILR Americans (−) (ΔX = 0.1)				
NU	**−6.89 [-9.77, -3.74]**	**−7.97 [-11.44, -4.16]**	**−5.68 [-8.28, -3.25]**	**−6.84 [-9.79, -3.83]**
UN	**−4.76 [-8.06, -1.37]**	**−5.58 [-9.84, -1.17]**	**−3.55 [-6.41, -0.62]**	**−4.66 [-8.05, -1.16]**
ILR Europeans (−) (ΔX = 0.1)				
NU	−0.48 [-3.57, 1.15]	−1.75 [-5.94, 0.81]	−1.08 [-4.06, 0.35]	−0.56 [-3.76, 1.2]
UN	**−4.17 [-7.7, -0.54]**	**−6.56 [-11.1, -1.88]**	**−4.74 [-7.73, -1.68]**	**−4.47 [-8.1, -0.74]**
ILR Asian (−) (ΔX = 0.1)				
NU	0.08 [-0.6, 0.72]	0.04 [-0.71, 0.76]	0.11 [-0.46, 0.78]	0.05 [-0.58, 0.7]
UN	0 [-0.44, 0.45]	0 [-0.54, 0.53]	0.03 [-0.5, 0.59]	0 [-0.45, 0.45]
ILR 0–29 (−) (ΔX = 0.1)				
NU	−0.74 [-7.29, 6.48]	−0.69 [-7.49, 6.74]	−0.97 [-7.34, 5.51]	−0.97 [-7.52, 5.86]
UN	2.01 [-5.53, 11.59]	1.85 [-6.62, 12.35]	1.74 [-5.36, 10.39]	2.36 [-5.4, 12.29]
ILR >65 (−) (ΔX = 0.1)				
NU	0.76 [-1.21, 3.54]	0.87 [-1.25, 4.21]	0.14 [-1.77, 2.12]	0.74 [-1.27, 3.44]
UN	0.44 [-1.85, 3.2]	0.55 [-1.93, 3.73]	−0.2 [-2.5, 2.04]	0.4 [-1.92, 3.17]
Tree density (trees/ha) (ΔX = 1)				
NU	**3.82 [3.47, 4.21]**	**3.82 [3.47, 4.21]**	**3.82 [3.47, 4.21]**	**3.82 [3.47, 4.21]**
UN	**4.22 [3.89, 4.55]**	**4.22 [3.89, 4.55]**	**4.22 [3.89, 4.55]**	**4.22 [3.89, 4.55]**

C to 19,300 Tn C for ten U.S. cities. When standardizing carbon storage values by tree cover area, Nowak and Crane (2002) reported median values of 9.25 kg C/m^2 [4.4–36.1]. In this study, carbon storage distributions based on 100 m^2 grid cells yielded an average of 2.5 kg C/m^2 ± 3.5 [0.08–12.35]. Regarding gross sequestration rates, the urban forest of Málaga is estimated to sequester 1285 Tn C/year [975–1604], placing it at the lower end of the range estimated by Nowak and Crane (2002), which varied from 42,100 Tn C/year in Atlanta to 800 Tn C/year in Jersey City.

The estimated annual air pollutant removal is 38.72 Tn/year [30.59–47.00], a value lower than those reported in other scientific studies. Selmi et al. (2016) found that urban trees in Strasbourg, France, removed 65 Tn of pollutants (PM_{10} and O_3) between July 2012 and June 2013. Bottalico et al. (2017), using a local-scale model instead of iTree, estimated that the urban forest in Florence, Italy, removed 249 Tn of pollutants in 2013. Similarly, Nowak et al. (2006) estimated pollutant removal across 48 U.S. cities, with a wide range from 8 to 14,900 Tn/year, largely attributed to differences in city size and canopy cover. More recently, Kofel et al. (2024) estimated that urban trees in Geneva removed 14 Tn/year of PM_{10} and 52 Tn/year of O_3. The gross results of this study are consistent with those of Baró et al. (2019) for Barcelona, a city with similar characteristics to Málaga, where pollutant removal was estimated at 28 Tn in 2018. On a per-area basis, we found that the distribution of annual pollutant removal rates has an average value of 5.45 g/m^2 [0.08–26.06]. This variability across different plots aligns with existing literature. Escobedo and Nowak (2009) reported annual removal rates ranging from 4.8 to 21.5 g/m^2 in the Santiago Metropolitan Region, Chile. Nowak et al. (2006) estimated pollution removal values per unit of canopy cover ranging from 6.2 to 23.1 g/m^2, with a median value of 10.8 g/m^2 across 48 U.S. cities. The urban forest of Geneva was estimated to have removal rates ranging from 0 to 8.4 g/m^2 for PM_{10} and 0–32.7 g/m^2 for O_3 (Kofel et al., 2024).

The contribution of the urban forest to runoff reduction is estimated at 0.011 hm^3/year [0.009–0.014], a value considerably lower than those reported in the literature. In Dalian, Liaoning Province, China, an i-Tree analysis estimated that urban trees intercept approximately 0.217 hm^3/year of precipitation annually (Wang et al., 2018). Similarly, Soares et al. (2011) estimated that urban trees in Lisbon, Portugal, mitigate 0.18 hm^3/year of stormwater runoff annually. Peper et al. (2007), using i-Tree, found that municipal tree cover in another study site helped avoid approximately 3.37 hm^3/year of runoff per year. In a more comparable setting, Baró et al. (2019) estimated 0.84 hm^3/year of avoided runoff in Barcelona, Spain.

Oxygen production was estimated at 3429 T O_2/year [2602–4278]. However, given the vast and relatively stable amount of oxygen in the atmosphere and the significant contribution of aquatic systems, this benefit is relatively minor compared to other ES (Suresh Ramanan et al., 2021). While Ginzburg et al. (2014) reported lower local oxygen concentrations in urban areas compared to those near green spaces, Suresh Ramanan et al. (2021) argue that current research remains inconclusive on this matter. Therefore, we do not consider oxygen production to be a key ES for guiding urban environmental equity efforts. Given its predominantly global nature and the relative stability of atmospheric oxygen concentrations, urban planning strategies should prioritize other ES that have a more direct and localized impact on environmental quality and public health.

BVOCs are recognized as a significant ecosystem disservice in urban environments due to their role in ozone and secondary organic aerosol (SOA) formation in areas with high nitrogen oxide (NOx) concentrations (Escobedo et al., 2011). Their emissions vary widely among tree species and are influenced by both physiological and environmental factors (Benjamin and Winer, 1998). The urban forest of Málaga generates an estimated 15.38 Tn BVOCs/year [11.98–18.83], significantly lower than the 50 Tn/year reported for Geneva (Kofel et al., 2024). Per unit area, BVOC production averages 2.17 g/m^2 [0–11.75], compared to Geneva's broader range of 0–16.5 g/m^2.

The differences in ES provision estimates between cities, and even within a single city, i.e., their spatial distribution as shown in Figures A.2–A.13, depend on multiple factors. (1) City size and urban forest coverage. As observed by Nowak and Crane (2002), total carbon storage and sequestration tend to increase with larger urban areas. (2) Differences in the standardization surface for ES provision. Most of the studies discussed standardized ES provision by tree canopy cover, reporting values in terms of provision per unit of canopy area. However, in the present study, standardization of ES provision is based on 100 m^2 grid cells, to understand the spatial distribution of ES provision per unit area of the urban matrix where urban trees are present. Except for Kofel et al. (2024), which also reports results using a pixel-based approach, differences in the reference unit of surface area may influence the

discrepancies observed between studies. Standardizing results by urban matrix area rather than canopy area leads to lower ES provision per unit area, as the latter exclusively represents surface with full provisioning capacity. (3) Differences in annual precipitation and background pollution levels across cities. Escobedo and Nowak (2009) observed that in areas with higher pollution levels, particularly PM10, the annual removal rate per square meter of tree canopy was higher. They suggested that lower vegetation cover in these areas might contribute to higher PM10 suspension rates, leading to increased ambient concentrations. Consequently, higher pollution availability results in greater potential for vegetation to remove pollutants, leading to higher estimated removal rates. A similar principle applies to annual precipitation. In cities with lower accumulated annual precipitation, less water is available to contribute to runoff formation, thereby reducing the total avoided runoff potential of urban vegetation. (4) Differences in urban tree density and dominant diameter classes. Higher tree density and a greater proportion of large-diameter trees generally lead to an increase in overall carbon storage density (Nowak and Crane, 2002). This difference in tree size distribution and overall tree density can significantly influence the carbon storage and sequestration capacity of urban forests, contributing to variations in ES provision across cities and regions. (5) Differences in species composition and structural traits. Species composition is a key factor influencing both the provision of ES and the production of BVOCs. The spatial distribution of high-emitting and low-emitting species is a critical factor driving exposure inequalities to BVOCs across different parts of the city (Vivaldo et al., 2017). In addition, species, along with management practices, also determine the Leaf Area Index (LAI). LAI varies across urban forests due to differences in foliation periods and canopy structure among species, which introduces spatial and temporal variability in the avoided runoff service (Dowtin et al., 2023). For evergreen species, the LAI, and consequently the interception of precipitation, is consistently higher throughout the year compared to deciduous species, which lose their foliage for a significant portion of the year (Clapp et al., 2014). This temporal variability in foliation for deciduous trees reduces their canopy interception storage compared to evergreen species over the course of a year (Yang et al., 2019; Yue et al., 2021). Similarly, the ability to remove pollutants is directly dependent on tree leaf area (Kofel et al., 2024).

4.2. Uncertainty estimates

Uncertainty arises both randomly and systematically, being associated with sampling, measurement, or systematic errors in data collection; epistemic uncertainty, which stems from a lack of knowledge; or structural uncertainty regarding the model itself (Hastie et al., 2009). In i-Tree Eco, only the sampling error from field plot data is evaluated, while other sources of uncertainty are not accounted for.

Uncertainty is gaining prominence in urban forest ES modeling, as researchers recognize the importance of quantifying variability and error to improve the reliability of estimates. Lin et al. (2021) assessed input, sampling, and model structure uncertainties, integrating these three sources to derive an estimator of total uncertainty. Their study focused on leaf area, biomass, carbon storage and sequestration, and BVOC emissions, reporting aggregated values across 15 U.S. cities. For carbon storage, they found a total coefficient of variation (CV) of 13.4 % $\pm$ 3.4, while for carbon sequestration, the CV was 11.1 % $\pm$ 2.6. Their findings highlighted that sampling uncertainty played the dominant role, model uncertainty had a minor influence, and input uncertainty contributed negligibly (less than 0.1 % on average). In contrast, for BVOC emissions, total uncertainty was significantly higher, with a CV of 40.7 % $\pm$ 7.9 for isoprene and 25.0 % $\pm$ 5.1 for monoterpenes. Unlike carbon-related ES, all three uncertainty sources played substantial roles in BVOC estimates, with input uncertainty emerging as a major contributor. Kofel et al. (2024) revealed uncertainty values of 40 % for BVOC production and a 83 % for PM10 removal based on the variances of each input variable.

Together with the work by Westfall et al. (2021), both prior studies represent the only research to date that examines the impact of input data uncertainty in i-Tree estimates. However, none of these studies propose an integrated methodology for addressing scenarios where data exhibit trait uncertainty (i.e., tree traits recorded as ranges) or where certain tree traits are missing for some observations. These situations are common in UTIs, as range-based measurement approaches improve data collection efficiency, and the absence of specific measurements often results from limitations in available materials or resources on the part of the organizations or managers responsible for UTI implementation.

In the literature, trait uncertainty has led to employing the range mean value as measurement for use in i-Tree Eco (e.g., Baró et al., 2019) and in the case of missing data, discarding incomplete observations or using genus-level averages is a common practice (e.g., Kofel et al., 2024). These approaches, while pragmatic, represent oversimplifications that may propagate systematic biases and underestimate the true uncertainty associated with ES quantifications, particularly when applied to heterogeneous urban forest contexts where measurement precision and data completeness vary considerably across different management sections and inventory protocols.

Our simulation-based approach enabled the quantification of uncertainty related to trait uncertainty at the individual tree scale, avoiding heuristic approaches in data preprocessing such as deletion or imputation based on averages. A key component of our framework involves ML models for imputing crown diameter and trunk height. ML models, specifically Random Forest (Breiman, 2001) in this study, a model extensively employed across numerous scientific fields for data imputation (Shah et al., 2014; Tang and Ishwaran, 2017; Kim and Violi, 2022; Zhou et al., 2024), enable the imputation of unmeasured values of crown diameter and trunk height by exploiting the capacity to model non-linear relationships and interactions within the feature space of selected covariates. The selected covariates, chosen considering factors that can model the patterns of these missing traits, contribute to more informed imputation of missing values due to existing correlations between functional traits and species, as well as the management type applied to each tree (Baines et al., 2020). Furthermore, a fundamental advantage of using ML models is that for each tree, both the mean imputed value and its standard error are predicted, allowing the integration and propagation of prediction uncertainty in subsequent steps, a characteristic not available by the previously mentioned heuristic imputation methods.

Although the validation metrics may initially appear low, it is important to note that these were computed under conditions that forced the models to extrapolate to unobserved datasets (Roberts et al., 2017). Specifically, the segregation in feature-space in a nested cross-validation resampling scheme allowed the generation of tree blocks as different as possible from each other and evaluated the prediction capacity on these blocks without the model being informed with these specific data patterns. This type of validation prevents overfitting and evaluates the generalization capacity of the imputation model to UTIs from other cities. The results indicate that while the models exhibit limited generalization capacity to entirely new UTIs, or more precisely, to UTIs with species and/or structural traits combinations not observed in the Malaga UTI, they do produce satisfactory results in cases of interpolation within the same dataset, i.e. the model is reliable for use on the Malaga inventory. Another virtue of ML models is the existing facility to improve their transferability to new datasets. As the model is applied and trained with data from other UTIs, patterns will be estimated more precisely, capturing greater variability, and a priori, greater generalization capacity to new inventories would be expected (Jin et al., 2018; Balestriero et al., 2021; Bjerre et al., 2022; Priyatikanto et al., 2023)

The integration of trait uncertainty and uncertainty in predictions from ML imputation models within a copula framework enables the explicit modeling of the dependence structure between ES estimates (Fig. A.1), capturing the correlations that exist between urban tree traits

and between ES (Ghosh et al., 2020; Lin et al., 2021). This approach offers a more flexible and comprehensive framework for uncertainty propagation compared with traditional univariate methods, which assume independence and are limited in specifying marginal distributions, thereby providing a robust tool for analyzing complex urban forest systems. Furthermore, the copula framework allows for the adaptation of marginal probability distributions, which serve to characterize the type of uncertainty existing for each variable. In this study, we assumed that no systematic bias exists, neither through overestimation nor underestimation of each ES based on trait measurements, establishing the means of marginal distributions at zero. However, if bias exists, based on prior information provided by urban forest managers, this could be perfectly incorporated into the model. The selection of probability distributions additionally allows for defining how uncertainty behaves, enabling the establishment of more informative distributions around the central value such as Gaussian or Laplace distributions, or non-informative distributions such as uniform distributions. This provides greater control and fine-tuning according to the needs and particularities of the UTI of the city where the method is intended to be applied.

The developed method is influenced by the Central Limit Theorem (CLT), particularly as the scale of analysis increases. This theorem states that if a sufficiently large number of i.i.d. random variables are summed, their distribution tends to approximate a normal distribution, regardless of the original distribution (Fischer, 2011). In this study, each random variable corresponds to the marginal distribution of ES provision incorporated into the copula for each individual tree. As the spatial scale of analysis increases, a greater number of random variables are summed, leading to a greater cancellation of individual fluctuations associated with independent simulations of ES provision at the tree level. At larger scales (i.e., census tracts), simulated values tend to converge toward the mean of the simulations, and uncertainty tends to decrease due to the cancellation of individual variations. This could also partially explain why any differences were observed in the sensitivity analysis regarding the selection of marginal probability distributions in the copula.

This same pattern has been reported by both Lin et al. (2021) and Westfall et al. (2021), who concluded that the effects of measurement uncertainty of input variables on ES estimation uncertainty are negligible at landscape scale or for computing aggregated values and their uncertainty at city scale. In any case, we consider that it is not that measurement uncertainty is unreliable, but rather that a process exists—in this case, we hypothesize it is the CLT—that leads to a cancellation of simulated uncertainty as the number of entities in an area increases. This aligns with the behavior of the uncertainty pattern across spatial scales, where at the highest level of detail (individual trees), uncertainty is maximum and decreases as area increases from pixel-support to census tract support, up to the scale of Malaga as a whole. This statistical phenomenon explains why landscape-scale estimates of ES can maintain reasonable precision despite substantial uncertainties in individual tree measurements. The hierarchical nature of uncertainty propagation in spatial ES assessments demonstrates that while individual tree predictions may carry significant uncertainty, the aggregation process inherently reduces the relative impact of these uncertainties on landscape-scale estimates.

These results are also consistent with findings reported in the forestry literature. McRoberts and Westfall (2015) demonstrated that although individual tree measurements may exhibit considerable variability, their impact on plot- and population-level forest estimates is minimal. This is primarily due to the statistical aggregation of random errors across multiple trees within plots, which effectively cancels out individual-level deviations. Paul et al. (2017) similarly confirmed that measurement errors in stem diameter, while significant at the individual tree level, have substantially reduced impacts on stand-scale estimates. As a result, measurement uncertainty contributes little to the overall variance in large-area estimates of forest attributes and ES. These studies point out the need to differentiate between random measurement error, which averages out during aggregation, and systematic bias, which persists and must be addressed through improved model calibration.

4.3. Inequities in urban tree ES provision accessibility

The BHM results highlight inequities in access to ES among social groups after accounting for census tract tree density, unaccounted covariables and spatial autocorrelation via random effects, and tree traits uncertainty. Riley and Gardiner (2020) recommend accounting for spatial structure when analyzing urban tree distribution and associated ES. Models incorporating spatial dependence, they use Spatial Autoregressive (SAR) models, outperform traditional linear regressions by providing more accurate and reliable estimates. By addressing spatial autocorrelation, these models reduce estimation errors and prevent misleading inferences about the relationships between socioeconomic covariables and response variables. Incorporating spatial structure thus not only enhances model performance but also improves our understanding of the complex dynamics within urban socio-ecological systems (Locke et al., 2016) Two fundamental advantages of using BHMs over frequentist models such as SAR or spatial lag models are: (1) the flexibility of BHMs, which can incorporate other latent structures into the model if the data require it, such as temporal random effects if spatio-temporal inequality patterns of accessibility data were available; (2) availability of posterior probability distributions of model effects and the possibility of performing combinations or operations between them following the rules of probability. The concatenating process applied for integrating uncertainty from different simulations in ES inequality access patterns, framed by the effects of covariates in the model, is ideally achievable under the Bayesian framework. This Bayesian approach enables comprehensive uncertainty propagation throughout the modeling framework, allowing for probabilistic statements about ES inequities that properly account for multiple sources of uncertainty. The hierarchical structure facilitates borrowing strength across spatial units while maintaining the ability to capture local variations in urban forest benefits distribution.

It was found that lower Gini index, median household income, and population size, as well as a higher proportion of African, South American, or Eastern European populations relative to the Spanish population at the census tract level, were associated with reduced accessibility to ES and increased exposure to BVOCs emitted by urban trees. The results in the literature are mixed, and this variability is largely due to the specific cities studied, each of which has its own unique socio-demographic patterns that ultimately shape the inequity patterns in access to ES. Cohen et al. (2012) and Escobedo and Nowak (2009) observed higher ES delivery in lower-income areas compared to wealthier zones in Paris and Santiago de Chile, respectively. Graça et al. (2017), in Porto, Portugal, observed that wealthier areas experience higher ES provision. Baró et al. (2019) in Barcelona found that ES provision was neither significantly associated with income nor ethnicity. However, they found a positive association with elderly residents and those with lower educational attainment, indicating that accessibility tends to favor vulnerable groups. Additionally, they reported that the most affluent neighborhoods in the city generally displayed intermediate to low ES provision, while the most disadvantaged neighborhoods tended to show intermediate to high values. Riley and Gardiner (2020), for nine cities in the United States, found that income levels and the percentage of the population with undergraduate education were generally positively related to ES provision. Conversely, the variables “percent without a high school degree” and “percent renters,” when significant, were negatively related. They observed that inequities related to population origin varied widely depending on the city; in the Midwest, significant relationships were positive, while the pattern was reversed when the variable was a significant predictor in East and West Coast cities. Nyelele and Kroll (2020) found that carbon storage, stormwater runoff reduction, air pollutant removal, and heat index reduction services are unequally and inequitably distributed across different socio-demographic and

socio-economic groups. They found that disadvantaged communities, particularly those with low median income and high poverty rates, receive disproportionately lower benefits from these services.

Income levels are the primary factor explaining the inequitable distribution of ES, which predominantly affects disadvantaged and marginalized socio-demographic neighborhoods. High-income earners have been shown to afford properties in neighborhoods with greener areas with greater tree cover, due to their higher purchasing power (Gerrish and Watkins, 2018). Urban trees represent an economic good that does not typically become a priority until residents have sufficient income to at least cover their basic needs. Consequently, demand for tree cover tends to increase significantly with rising income levels (Zhu and Zhang, 2008).

In Málaga, new urban development projects are increasingly incorporating higher urban canopy coverage in pedestrian-accessible areas. However, due to the high real estate prices in these areas, access is largely restricted to high-income families, thereby exacerbating the observed inequities in income levels and access to ES provided by urban trees. This pattern also explains the positive relationship observed with the Gini coefficient. In peripheral census tracts, where most new developments are taking place, pre-existing built-up areas often coexist with vacant lands. These areas are where most urban expansion is occurring. As a result, the increase in tree canopy coverage in these census tracts, coupled with the influx of higher-income families into these newly developed areas, leads to growing local income disparities, raising the Gini index. This process may explain the observed positive relationship between income inequality and ES accessibility; however, further research is necessary to confirm this association.

In most developed countries, ethnic minorities tend to have the lowest income levels (Makhlouf, 2023), which often translates into residing in neighborhoods with lower urban tree canopy coverage and, consequently, reduced access to the ES provided by urban trees. This pattern could explain the lower accessibility to ES estimated for ethnic minorities in the present study. In addition to having limited access to urban green spaces and their associated benefits, ethnic minorities are also disproportionately exposed to urban environmental hazards, such as higher levels of air pollution and reduced thermal comfort (Hsu et al., 2021; Aznarez et al., 2024; van den Brekel et al., 2024).

4.4. Limitations

While this study has quantified ES provision and BVOC production by urban trees and highlighted the inequities in accessibility and exposure among sociodemographic groups, several limitations warrant further investigation:

(1) The Monte Carlo simulation using Gaussian copulas provides accurate uncertainty estimates at the individual tree level. However, there are two limitations that warrant further investigation. First, the method is subject to the CLT due to the random sampling used to sample from the copula of each tree. This implies that estimates across simulations tend to converge towards the mean, reducing uncertainty, particularly at larger spatial scales. Further work is needed to systematically explore how this effect evolves with increasing scale in a systematic way, as well as to test different methods for sampling values. Second, the copula is used on the theoretical provision ranges derived from the combination of maximum and minimum traits. This represents a compromise solution that allows for provision ranges over which simulations can be conducted. Ideally, simulations would begin by adjusting copulas at the trait level. This approach would enable the generation of different UTI for each iteration, with trees of varying sizes corresponding to uncertainty in trait measurement. However, this is not feasible as iTree requires manual data entry, making it resource-intensive for thousands of simulations. Therefore, implementing iTree models within programming environments such as R or Python would facilitate the integration of uncertainty propagation frameworks, enabling more automated simulation and estimation processes.

(2) In this study, BVOC generation was the only disservice considered. However, another important disservice is pollen exposure. Living near urban parks and trees can pose health risks due to increased exposure to allergenic pollen. Similar to BVOC spatial patterns, pollen allergenicity riskscapes in different cities are shaped by the dominant tree species in the urban forest (Sousa-Silva et al., 2021). For example, the widespread use of plane trees (*Platanus* spp.) as ornamental species in many Mediterranean cities has been linked to the emergence of new allergies among residents (Picornell et al., 2024). Future research should integrate pollen allergenicity riskscapes to assess potential socio-demographic inequalities in pollen exposure, providing a more comprehensive evaluation of the benefits and disservices of urban greening.

(3) According to scientific literature, inequalities in access to ES provided by urban trees, exposure to BVOCs generated by urban trees, and exposure to pollutants across social groups vary significantly depending on the city. An important avenue for future research would involve extending the current study, along with the previously outlined limitations, to the major provincial capitals in Andalusia and, more generally, to include all municipalities with a complete UTI available. Incorporating multilevel structures into the model presented in this article will allow for the estimation of the specificities of different cities and help identify whether common patterns exist in the relationships between inequities and the provision of ES.

5. Conclusion

This study introduces an integrated uncertainty propagation framework that combines machine learning imputation of missing tree trait values, copula-based integration of imputation and trait measurement uncertainty, and Monte Carlo simulation. The framework effectively quantifies the influence of measurement uncertainty on i-Tree Eco estimates of urban forest ecosystem services. Across spatial scales, uncertainty decreases progressively, reflecting aggregation effects as areas expand from individual trees to entire cities.

The BHM results reveal systematic inequities in ecosystem service access among socioeconomic groups, even after controlling for census tract-level tree density, unobserved covariates, spatial autocorrelation, and tree trait uncertainty. While trait uncertainty propagation in the BHM generally preserves the posterior distribution means of socioeconomic variable effects, it systematically increases credibility interval widths. This widening reflects successful uncertainty propagation through the modeling framework. Importantly, inequality patterns remain largely consistent between propagation and non-propagation scenarios, though uncertainty propagation enables detection of credible relationships at the 95 % level that would otherwise be missed.

From a methodological perspective, the results suggest that the value of propagating trait-level uncertainty into second-stage models depends critically on the spatial scale of analysis. At coarse aggregation levels such as census tracts, while error cancellation effects reduce the impact on point estimates, uncertainty propagation remains valuable for assessing parameter reliability, detecting previously unidentified relationships, and providing comprehensive credible intervals for decision-making.

Future work should systematically evaluate uncertainty attenuation patterns across incremental spatial scales, establish guidelines for optimal aggregation levels in ecosystem service studies, and explore trait uncertainty propagation in analyses conducted at finer resolutions (e.g., when ecosystem service estimates and their uncertainty are employed in pixel-grid support models or when individual trees constitute observational units), where the high quantified uncertainties

would likely result in substantial propagation effects in second-stage models. Validation across diverse urban contexts will be essential to confirm the framework's broader applicability and refine recommendations for uncertainty propagation in urban forest research.

CRediT authorship contribution statement

Jaime Francisco Pereña-Ortiz: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation. **Ángel Enrique Salvo-Tierra:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Funding acquisition, Conceptualization. **Pablo Cozano-Pérez:** Visualization, Resources, Investigation, Formal analysis. **Ángel Ruiz-Valero:** Writing – review & editing, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

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

Acknowledgements

Angel Ruiz Valero was supported by a pre-doctoral grant financed by the Ministry of Education, Professional Formation and Sport of Spain, in the Program of University Teaching Program (Formación de Profesorado Universitario, FPU) (FPU22/00067). Ángel Ruiz-Valero, Jaime Francisco Pereña-Ortiz, Pablo Cozano-Pérez and Ángel Enrique Salvo Tierra are part of the research team RNM-262: Biogeography, Diversity and Conservation of Junta de Andalucía, Spain. The authors specially want to thank the Málaga City Council for the provision of urban tree inventory.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.indic.2025.100864.

Data availability

The data related to the urban tree inventory requires prior consultation and subsequent acceptance by the City Council of Malaga for its distribution.

References

Adin, A., Krainski, E.T., Lenzi, A., Liu, Z., Martínez-Minaya, J., Rue, H., 2024. Automatic cross-validation in structured models: is it time to leave out leave-one-out? Spatial Statistics, 100843. https://doi.org/10.1016/j.spasta.2024.100843.

Aguaron, E., McPherson, E.G., 2012. Comparison of methods for estimating carbon dioxide storage by sacramento's urban forest. In: Lal, R., Augustin, B. (Eds.), Carbon Sequestration in Urban Ecosystems. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2366-5_3.

Aznarez, C., Kumar, S., Marquez-Torres, A., Pascual, U., Baró, F., 2024. Ecosystem service mismatches evidence inequalities in urban heat vulnerability. Sci. Total Environ. 922, 171215. https://doi.org/10.1016/j.scitotenv.2024.171215.

Baines, O., Wilkes, P., Disney, M., 2020. Quantifying urban forest structure with open-access remote sensing data sets. Urban Forestry. Urban Greening 50, 126653. https://doi.org/10.1016/j.ufug.2020.126653.

Balestriero, R., Pesenti, J., LeCun, Y., 2021. Learning in high dimension always amounts to extrapolation. arXiv preprint arXiv:2110.09485.

Ballester, J., Quijal-Zamorano, M., Méndez Turrubiates, R.F., Pegenaute, F., Herrmann, F.R., Robine, J.M., Basagaña, X., Tonne, C., Antó, J.M., Achebak, H., 2023. Heat-related mortality in Europe during the summer of 2022. Nature medicine 29 (7), 1857–1866. https://doi.org/10.1038/s41591-023-02419-z.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2014. Hierarchical Modeling and Analysis for Spatial Data, second ed. Chapman and Hall/CRC. https://doi.org/10.1201/b17115.

Bancks, N., North, E., Johnson, G.R., 2018. An analysis of agreement between volunteer- and researcher-collected urban tree inventory data. Arboriculture. Amp; Urban Forestry 44 (2), 73–86. https://doi.org/10.48044/jauf.2018.007.

Barriopedro, D., Fischer, E.M., Luterbacher, J., Trigo, R.M., García-Herrera, R., 2011. The hot summer of 2010: redrawing the temperature record map of Europe. Science 332 (6026), 220–224. https://doi.org/10.1126/science.12012.

Baró, F., Calderón-Argelich, A., Langemeyer, J., Connolly, J.J., 2019. Under one canopy? Assessing the distributional environmental justice implications of street tree benefits in Barcelona. Environ. Sci. Pol. 102, 54–64. https://doi.org/10.1016/j.envsci.2019.08.016.

Benjamin, M.T., Winer, A.M., 1998. Estimating the ozone-forming potential of urban trees and shrubs. Atmos. Environ. 32 (1), 53–68. https://doi.org/10.1016/S1352-2310(97)00176-3.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Stat. Math. 43, 1–20. https://doi.org/10.1007/BF00116466.

Binder, M., Pfisterer, F., Lang, M., Schneider, L., Kotthoff, L., Bischl, B., 2021. mlr3pipelines - flexible machine learning pipelines in R. J. Mach. Learn. Res. 22 (184), 1–7. https://jmlr.org/papers/v22/21-0281.html.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., Deng, D., Lindauer, M., 2023. Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. Wiley Interdisciplinary Reviews: Data Min. Knowl. Discov. 13 (2), e1484. https://doi.org/10.1002/widm.1484.

Bischl, B., Sonabend, R., Kotthoff, L., Lang, M. (Eds.), 2024. Applied Machine Learning Using mlr3 in R. CRC Press. https://mlr3book.mlr-org.com.

Bjerre, E., Fienen, M.N., Schneider, R., Koch, J., Højberg, A.L., 2022. Assessing spatial transferability of a random forest metamodel for predicting drainage fraction. J. Hydrol. 612, 128177. https://doi.org/10.1016/j.jhydrol.2022.128177.

Bottalico, F., Travaglini, D., Chirici, G., Garfì, V., Giannetti, F., De Marco, A., Fares, S., Marchetti, M., Nocentini, S., Paoletti, S., Salbitano, F., Sanesi, G., 2017. A spatially-explicit method to assess the dry deposition of air pollution by urban forests in the city of florence, Italy. Urban For. Urban Green. 27, 221–234. https://doi.org/10.1016/j.ufug.2017.08.013.

Boukili, V., Bebber, D.P., Mortimer, T., Venicx, G., Lefcourt, D., Chandler, M., et al., 2017. Assessing the performance of urban forest carbon sequestration models using direct measurements of tree growth. Urban Forestry. Urban Greening 24, 212–221. https://doi.org/10.1016/j.ufug.2017.03.015.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Breugel, M.v., Ransijn, J., Craven, D., Bongers, F., Hall, J.S., 2011. Estimating carbon stock in secondary forests: decisions and uncertainties associated with allometric biomass models. For. Ecol. Manag. 262 (8), 1648–1657. https://doi.org/10.1016/j.foreco.2011.07.018.

Chapman, S., Watson, J.E., Salazar, A., Thatcher, M., McAlpine, C.A., 2017. The impact of urbanization and climate change on urban temperatures: a systematic review. Landsc. Ecol. 32 (10), 1921–1935. https://doi.org/10.1007/s10980-017-0561-4.

Clapp, J.C., Ryan III, H.D.P., Harper, R.W., Bloniarz, D.V., 2014. Rationale for the increased use of conifers as functional green infrastructure: a literature review and synthesis. Arboric. J.: The International Journal of Urban Forestry 36 (3), 161–178. https://doi.org/10.1080/03071375.2014.950861.

Cohen, M., Baudoin, R., Palibrk, M., Persyn, N., Rhein, C., 2012. Urban biodiversity and social inequalities in built-up cities: new evidences, next questions. The example of paris, France. Landsc. Urban Plann. 106 (3), 277–287. https://doi.org/10.1016/j.landurbplan.2012.03.007.

Coville, R.C., Kruegler, J., Selbig, W.R., Hirabayashi, S., Loheide, S.P., Avery, W.A., et al., 2022. Loss of street trees predicted to cause 6000 l/tree increase in leaf-on stormwater runoff for Great Lakes urban sewershed. Urban Forestry. Urban Greening 74, 127649. https://doi.org/10.1016/j.ufug.2022.127649.

Domeisen, D.I., Eltahir, E.A., Fischer, E.M., Knutti, R., Perkins-Kirkpatrick, S.E., Schär, C., Seneviratne, S.I., Wernli, H., 2023. Prediction and projection of heatwaves. Nat. Rev. Earth Environ. 4 (1), 36–50. https://doi.org/10.1038/s43017-022-00371-z.

Dowtin, A.L., Cregg, B.C., Nowak, D.J., Levia, D.F., 2023. Towards optimized runoff reduction by urban tree cover: a review of key physical tree traits, site conditions, and management strategies. Landsc. Urban Plann. 239, 104849. https://doi.org/10.1016/j.landurbplan.2023.104849.

Escobedo, F.J., Nowak, D.J., 2009. Spatial heterogeneity and air pollution removal by an urban forest. Landsc. Urban Plann. 90 (3–4), 102–110. https://doi.org/10.1016/j.landurbplan.2008.10.021.

Escobedo, F.J., Kroeger, T., Wagner, J.E., 2011. Urban forests and pollution mitigation: analyzing ecosystem services and disservices. Environ. Pollut. 159 (8–9), 2078–2087. https://doi.org/10.1016/j.envpol.2011.01.010.

Fischer, H., 2011. A History of the Central Limit Theorem: from Classical to Modern Probability Theory, vol. 4. Springer, New York.

Fletcher Jr., R.J., Hefley, T.J., Robertson, E.P., Zuckerberg, B., McCleery, R.A., Dorazio, R.M., 2019. A practical guide for combining data to model species distributions. Ecology 100 (6), e02710. https://doi.org/10.1002/ecy.2710.

Founda, D., Santamouris, M., 2017. Synergies between urban heat island and heat waves in Athens (greece), during an extremely hot summer (2012). Sci. Rep. 7 (1), 10973. https://doi.org/10.1038/s41598-017-11407-6.

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, third ed. Sage, Thousand Oaks CA https://www.john-fox.ca/Companion/.

Fuglstad, G.A., Simpson, D., Lindgren, F., Rue, H., 2019. Constructing priors that penalize the complexity of Gaussian random fields. J. Am. Stat. Assoc. 114 (525), 445–452. https://doi.org/10.1080/01621459.2017.1415907.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2013. Bayesian Data Analysis, third ed. Chapman and Hall/CRC. https://doi.org/10.1201/b16018.
Gerrish, E., Watkins, S.L., 2018. The relationship between urban forests and income: a meta-analysis. Landsc. Urban Plann. 170, 293–308. https://doi.org/10.1016/j.landurbplan.2017.09.005.
Ginzburg, A.S., Vinogradova, A.A., Fedorova, E.I., Nikitich, E.V., Karpov, A.V., 2014. Content of oxygen in the atmosphere over large cities and respiratory problems. Izvestiya Atmos. Ocean. Phys. 50, 782–792. https://doi.org/10.1134/S0001433814080040.
Ghosh, S., Sheppard, L.W., Holder, M.T., Loecke, T.D., Reid, P.C., Bever, J.D., et al., 2020. Copulas and their potential for ecology. Adv. Ecol. Res. 409–468. https://doi.org/10.1016/bs.aecr.2020.01.003.
Graça, M.S., Gonçalves, J.F., Alves, P.J., Nowak, D.J., Hoehn, R., Ellis, A., Farinha-Marques, P., Cunha, M., 2017. Assessing mismatches in ecosystem services proficiency across the urban fabric of Porto (portugal): the influence of structural and socioeconomic variables. Ecosyst. Serv. 23, 82–93. https://doi.org/10.1016/j.ecoser.2016.11.015.
Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2. springer, New York, pp. 1–758.
Held, L., Natário, I., Fenton, S.E., Rue, H., Becker, N., 2005. Towards joint disease mapping. Stat. Methods Med. Res. 14 (1), 61–82. https://doi.org/10.1191/0962280205sm389.
Hirabayashi, S., Kroll, C.N., Nowak, D.J., Endreny, T.A., 2012. i-Tree Eco Dry Deposition Model Descriptions. Citeseer, Princeton, NJ, USA.
Hirabayashi, S., 2013. i-Tree Eco Precipitation Interception Model Descriptions, vol. 1. US Department of Agriculture Forest Service, Washington, DC, USA, pp. 1–21.
Hirabayashi, S., 2015. i-Tree Eco United States County-based Hydrologic Estimates. US Department of Agriculture Forest Service, Pacific Southwest Research Station. Center for Urban Forest Research: Washington, DC, USA.
Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., 2025. Copula: multivariate dependence with copulas. R package version 1, 1–5. https://CRAN.R-project.org/package=copula.
Hsu, A., Sheriff, G., Chakraborty, T., Manya, D., 2021. Disproportionate exposure to urban heat island intensity across major US cities. Nat. Commun. 12 (1), 2721. https://doi.org/10.1038/s41467-021-22799-5.
Hu, P., Zhang, Q., Shi, P., Chen, B., Fang, J., 2018. Flood-induced mortality across the globe: spatiotemporal pattern and influencing factors. Sci. Total Environ. 643, 171–182. https://doi.org/10.1016/j.scitotenv.2018.06.197.
[dataset] Institute of Statistics and Cartography of Andalusia (ISCA), 2023. Indicadores trimestrales de actividad económica de la población de Andalucía. https://www.juntadeandalucia.es/institutodeestadisticaycartografia/badea/informe/anual?CodOper=b3_3155&idNode=104425 [dataset].
Jin, S., Su, Y., Gao, S., Hu, T., Liu, J., Guo, Q., 2018. The transferability of random forest in canopy height estimation from multi-source remote sensing data. Remote Sens. 10 (8), 1183. https://doi.org/10.3390/rs10081183.
Juginović, A., Vuković, M., Aranza, I., Biloš, V., 2021. Health impacts of air pollution exposure from 1990 to 2019 in 43 European countries. Sci. Rep. 11 (1), 22516. https://doi.org/10.1038/s41598-021-01802-5.
Kim, D., Violi, A., 2022. Uncertainty-based weight determination for surrogate optimization. Combust. Flame 237, 111850. https://doi.org/10.1016/j.combustflame.2021.111850.
Kofel, D., Bourgeois, I., Paganini, R., Pulfer, A., Grossiord, C., Schmale, J., 2024. Quantifying the impact of urban trees on air quality in Geneva, Switzerland. Urban For. Urban Green. 101, 128513. https://doi.org/10.1016/j.ufug.2024.128513.
Lang, M., Binder, M., Richter, J., Schratz, P., Pfisterer, F., Coors, S., Au, Q., Casalicchio, G., Kotthoff, L., Bischl, B., 2019. mlr3: a modern object-oriented machine learning framework in R. J. Open Source Softw. https://doi.org/10.21105/joss.01903. https://joss.theoj.org/papers/10.21105/joss.01903.
Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525 (7569), 367–371. https://doi.org/10.1038/nature15371.
Liang, D., Huang, G., 2023. Influence of urban tree traits on their ecosystem services: a literature review. Land 12 (9), 1699. https://doi.org/10.3390/land12091699.
Lin, J., Kroll, C.N., Nowak, D.J., Greenfield, E.J., 2019. A review of urban forest modeling: implications for management and future research. Urban Forestry. Urban Greening 43, 126366. https://doi.org/10.1016/j.ufug.2019.126366.
Lin, J., Kroll, C.N., Nowak, D.J., 2020. Ecosystem service-based sensitivity analyses of i-tree eco. Arboriculture. Amp; Urban Forestry 46 (4), 287–306. https://doi.org/10.48044/jauf.2020.021.
Lin, J., Kroll, C.N., Nowak, D.J., 2021. An uncertainty framework for i-Tree eco: a comparative study of 15 cities across the United States. Urban For. Urban Green. 60, 127062. https://doi.org/10.1016/j.ufug.2021.127062.
Liu, Z., Rue, H., 2022. Leave-group-out cross-validation for latent Gaussian models. arXiv preprint arXiv:2210.04482. https://doi.org/10.48550/arXiv.2210.04482.
Locke, D.H., Landry, S.M., Grove, J.M., Chowdhury, R.R., 2016. What's scale got to do with it? Models for urban tree canopy. J. Urban Econ. 2 (1), juw006. https://doi.org/10.1093/jue/juw006.
Lorenzo, N., Díaz-Poso, A., Royé, D., 2021. Heatwave intensity on the iberian peninsula: future climate projections. Atmos. Res. 258, 105655. https://doi.org/10.1016/j.atmosres.2021.105655.
Makhlouf, Y., 2023. Trends in income inequality: evidence from developing and developed countries. Soc. Indic. Res. 165 (1), 213–243. https://doi.org/10.1007/s11205-022-03010-8.
Manzini, J., Hoshika, Y., Carrari, E., Sicard, P., Watanabe, M., Tanaka, R., et al., 2023. Flortree: a unifying modelling framework for estimating the species-specific pollution removal by individual trees and shrubs. Urban Forestry. Urban Greening 85, 127967. https://doi.org/10.1016/j.ufug.2023.127967.
Martin, N.A., Chappelka, A.H., Somers, G.L., Loewenstein, E.F., Keever, G.J., 2013. Evaluation of sampling protocol for i-tree eco: a case study in predicting ecosystem services at auburn university. Arboriculture. Amp; Urban Forestry 39 (2). https://doi.org/10.48044/jauf.2013.008.
McPherson, E.G., Simpson, J.R., Xiao, Q., Wu, C., 2011. Million trees Los Angeles canopy cover and benefit assessment. Landsc. Urban Plann. 99 (1), 40–50. https://doi.org/10.1016/j.landurbplan.2010.08.011.
McPherson, E.G., van Doorn, N.S., Peper, P.J., 2016. Urban tree database and allometric equations. In: Gen. Tech. Rep. PSW-GTR-253. US Department of Agriculture, Forest Service, Albany, CA, p. 253. Pacific Southwest Research Station. 86.
McRoberts, R.E., Westfall, J.A., 2015. Propagating uncertainty through individual tree volume model predictions to large-area volume estimates. Ann. For. Sci. 73 (3), 625–633. https://doi.org/10.1007/s13595-015-0473-x.
Messner, J.W., Mayr, G.J., Zeileis, A., 2016. Heteroscedastic censored and truncated regression with crch. The R Journal 8 (1), 173–181. https://doi.org/10.32614/RJ-2016-012.
Miedema-Brown, L., Anand, M., 2022. Plant functional traits as measures of ecosystem service provision. Ecosphere 13 (2), e3930. https://doi.org/10.1002/ecs2.3930.
Morani, A., Nowak, D.J., Hirabayashi, S., Guidolotti, G., Medori, M., Muzzini, V.G., et al., 2014. Comparing i-tree modeled ozone deposition with field measurements in a periurban mediterranean forest. Environ. Pollut. 195, 202–209. https://doi.org/10.1016/j.envpol.2014.08.031.
Murray, C.J., Aravkin, A.Y., Zheng, P., Abbafati, C., Abbas, K.M., Abbasi-Kangevari, M., et al., 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. The lancet 396 (10258), 1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2.
[dataset] National Institute of Geography, (NIG), 2022. Plan Nacional de Ortofotografía Aérea (PNOA)/Plan Nacional de Observación del Territorio (PNOT). https://pnoa.ign.es/ [dataset].
[dataset] National Institute of Statistics, (NIS), 2022a. National institute of statistics population by cities. Malaga (spain), 2022. https://www.ine.es/dyngs/INEbase/en/categoria.htm?c=Estadistica_P&cid=1254734710990 [dateset].
[dataset] National Institute of Statistics, (NIS), 2022b. Atlas de Distribución de Renta de los Hogares. Año 2022. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177088&menu=ultiDatos&idp=1254735976608 [dataset].
[dataset] National Institute of Statistics, (NIS), 2023. Cifras oficiales de población de los municipios españoles en aplicación de la Ley de Bases del Régimen Local (Art. 17). Detalle municipal: Málaga. Población por municipios y sexo. Unidades: Personas. https://www.ine.es/dynt3/inebase/index.htm?padre=525 [dataset].
[dataset] National Institute of Statistics, (NIS), 2024. Censo anual de población. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176992&menu=ultiDatos&idp=1254735572981 [dataset].
Nelsen, R.B., 2005. Copulas and quasi-copulas: an introduction to their properties and applications. In: Logical, Algebraic, Analytic and Probabilistic Aspects of Triangular Norms. Elsevier Science BV, pp. 391–413. https://doi.org/10.1016/B978-044451814-9/50014-8.
Nowak, D.J., 1996. Estimating leaf area and leaf biomass of open-grown deciduous urban trees. For. Sci. 42 (4), 504–507. https://doi.org/10.1093/forestscience/42.4.504.
Nowak, D.J., Crane, D.E., 2000. The urban forest effects (UFORE) model: quantifying urban forest structure and functions. In: Hansen, Mark, Burk, Tom (Eds.), Integrated Tools for Natural Resources Inventories in the 21st Century. Gen. Tech. Rep. NC-212. St. Paul, MN: US Dept. of Agriculture, Forest Service, North Central Forest Experiment Station, pp. 714–720, 212.
Nowak, D.J., Crane, D.E., 2002. Carbon storage and sequestration by urban trees in the USA. Environ. Pollut. 116 (3), 381–389. https://doi.org/10.1016/S0269-7491(01)00214-7.
Nowak, D.J., Crane, D.E., Stevens, J.C., 2006. Air pollution removal by urban trees and shrubs in the United States. Urban For. Urban Green. 4 (3–4), 115–123. https://doi.org/10.1016/j.ufug.2006.01.007.
Nowak, D.J., Walton, J.T., Stevens, J.C., Crane, D.E., Hoehn, R.E., 2008. Effect of plot and sample size on timing and precision of urban forest assessments. Aboriculture & Urban Forestry 34 (6), 386–390, 34(6).
Nowak, D.J., Greenfield, E.J., Hoehn, R., LaPoint, E.B., 2013a. Carbon storage and sequestration by trees in urban and community areas of the United States. Environ. Pollut. 178, 229–236. https://doi.org/10.1016/j.envpol.2013.03.019.
Nowak, D.J., Hirabayashi, S., Bodine, A.R., Hoehn, R., 2013b. Modeled pm2.5 removal by trees in ten u.s. cities and associated health effects. Environ. Pollut. 178, 395–402. https://doi.org/10.1016/j.envpol.2013.03.050.
Nowak, D.J., Hirabayashi, S., Bodine, A., Greenfield, E., 2014. Tree and forest effects on air quality and human health in the United States. Environ. Pollut. 193, 119–129. https://doi.org/10.1016/j.envpol.2014.05.028.
Nowak, D.J., 2024. Understanding i-Tree: 2023 Summary of Programs and Methods. General Technical Report NRS-200-2023. US Department of Agriculture, Forest Service, Northern Research Station, Madison, WI, p. 103 [plus 14 appendixes]., 200.
Nyelele, C., Kroll, C.N., 2020. The equity of urban forest ecosystem services and benefits in the bronx, NY. Urban For. Urban Green. 53, 126723. https://doi.org/10.1016/j.ufug.2020.126723.
Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge university press. https://doi.org/10.1017/9781139016476.
Pace, R., Biber, P., Pretzsch, H., Grote, R., 2018. Modeling ecosystem services for park trees: sensitivity of i-tree eco simulations to light exposure and tree species classification. Forests 9 (2), 89. https://doi.org/10.3390/f9020089.

Pace, R., Grote, R., 2020. Deposition and resuspension mechanisms into and from tree canopies: a study modeling particle removal of conifers and broadleaves in different cities. Frontiers in Forests and Global Change 3. https://doi.org/10.3389/ffgc.2020.00026.

Pace, R., Guidolotti, G., Baldacchini, C., Pallozzi, E., Grote, R., Nowak, D.J., et al., 2021. Comparing i-tree eco estimates of particulate matter deposition with leaf and canopy measurements in an urban mediterranean holm oak forest. Environmental Science Amp; Technology 55 (10), 6613–6622. https://doi.org/10.1021/acs.est.0c07679.

Palmí-Perales, F., Gómez-Rubio, V., Bivand, R.S., Cameletti, M., Rue, H., 2022. Bayesian inference for multivariate spatial models with R-INLA. arXiv preprint arXiv:2212.10976.

Pataki, D.E., Alberti, M., Cadenasso, M.L., Felson, A.J., McDonnell, M.J., Pincetl, S., Pouyat, R.V., Setälä, H., Whitlow, T.H., 2021. The benefits and limits of urban tree planting for environmental and human health. Frontiers in Ecology and Evolution 9, 603757. https://doi.org/10.3389/fevo.2021.603757.

Paul, K.I., Larmour, J.S., Roxburgh, S.H., England, J.R., Davies, M., Luck, H.D., 2017. Measurements of stem diameter: implications for individual- and stand-level errors. Environ. Monit. Assess. 189 (8). https://doi.org/10.1007/s10661-017-6109-x.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30 (7), 683–691. https://doi.org/10.1016/j.cageo.2004.03.012.

Peper, P.J., McPherson, E.G., 2003. Evaluation of four methods for estimating leaf area of isolated trees. Urban Forestry. Urban Greening 2 (1), 19–29. https://doi.org/10.1078/1618-8667-00020.

Peper, P.J., McPherson, E.G., Simpson, J.R., Gardner, S.L., Vargas, K.E., Xiao, Q., 2007. New York City, New York Municipal Forest Resource Analysis. US Department of Agriculture Forest Service, Pacific Southwest Research Station, Center for Urban Forest Research, p. 65 [Technical Report].

Picornell, A., Maya-Manzano, J.M., Fernández-Ramos, M., Hidalgo-Barquero, J.J., Pecero-Casimiro, R., Ruiz-Mata, R., de Gálvez-Montañez, E., Trigo, M.M., Fernández-Rodríguez, S., 2024. Effects of climate change on platanus flowering in Western mediterranean cities: current trends and future projections. Sci. Total Environ. 906, 167800.

Priyatikanto, R., Lu, Y., Dash, J., Sheffield, J., 2023. Improving generalisability and transferability of machine-learning-based maize yield prediction model through domain adaptation. Agric. For. Meteorol. 341, 109652. https://doi.org/10.1016/j.agrformet.2023.109652.

Probst, P., Wright, M.N., Boulesteix, A.L., 2019. Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Min. Knowl. Discov. 9 (3), e1301. https://doi.org/10.1002/widm.1301.

Raschka, S., 2020. Model evaluation, model selection, and algorithm selection in machine learning. (arXiv:1811.12808). arXiv. https://doi.org/10.48550/arXiv.1811.12808.

Rentschler, J., Salhab, M., Jafino, B.A., 2022. Flood exposure and poverty in 188 countries. Nat. Commun. 13 (1), 3527. https://doi.org/10.1038/s41467-022-30727-4.

Riebler, A., Sørbye, S.H., Simpson, D., Rue, H., 2016. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Stat. Methods Med. Res. 25 (4), 1145–1165. https://doi.org/10.1177/0962280216660421.

Riley, C.B., Gardiner, M.M., 2020. Examining the distributional equity of urban tree canopy cover and ecosystem services across United States cities. PLoS One 15 (2), e0228499. https://doi.org/10.1371/journal.pone.0228499.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.L., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40 (8), 913–929. https://doi.org/10.1111/ecog.02881.

Robine, J.M., Cheung, S.L.K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.P., Herrmann, F.R., 2008. Death toll exceeded 70,000 in Europe during the summer of 2003. Comptes rendus. Biologies 331 (2), 171–178. https://doi.org/10.1016/j.crvi.2007.12.001.

Roman, L.A., Scharenbroch, B.C., Östberg, J., Mueller, L.S., Henning, J.G., Koeser, A.K., et al., 2017. Data quality in citizen science urban tree inventories. Urban Forestry. Urban Greening 22, 124–135. https://doi.org/10.1016/j.ufug.2017.02.001.

Rosenzweig, C., Solecki, W., Romero-Lankao, P., Mehrotra, S., Dhakal, S., Bowman, T., et al., 2018. Climate Change and Cities: Second Assessment Report of the Urban Climate Change Research Network. Cambridge University Press, Cambridge, pp. 17–42. https://doi.org/10.1017/9781316563878.007.

Rue, H., Martino, S., Chopin, N., 2009. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. J. Roy. Stat. Soc. B Stat. Methodol. 71 (2), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x.

Schmidt, T., 2007. Coping with copulas. Copulas-From theory to application in finance 3, 1–34.

Schneider, L., Richter, J., Becker, M., Lang, M., Bischl, B., Pfisterer, F., Binder, M., Fischer, S., 2025. mlr3mbo: flexible bayesian optimization. R package version 0.2.9. https://github.com/mlr-org/mlr3mbo. https://mlr3mbo.mlr-org.com.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecol. Model. 406, 109–120. https://doi.org/10.1016/j.ecolmodel.2019.06.002.

Schwarz, K., Fragkias, M., Boone, C.G., Zhou, W., McHale, M., Grove, J.M., O'Neil-Dunne, T., McFadden, J.P., Buckley, G.L., Childers, D., Ogden, L., Pincetl, S., Pataki, D., Whitmer, A., Cadenasso, M.L., 2015. Trees grow on money: urban tree canopy cover and environmental justice. PLoS One 10 (4), e0122051. https://doi.org/10.1371/journal.pone.0122051.

Selmi, W., Weber, C., Rivière, E., Blond, N., Mehdi, L., Nowak, D., 2016. Air pollution removal by trees in public green spaces in strasbourg city, France. Urban For. Urban Green. 17, 192–201. https://doi.org/10.1016/j.ufug.2016.04.010.

Shah, A.D., Bartlett, J., Carpenter, J.R., Nicholas, O., Hemingway, H., 2014. Comparison of random forest and parametric imputation models for imputing missing data using mice: a caliber study. Am. J. Epidemiol. 179 (6), 764–774. https://doi.org/10.1093/aje/kwt312.

Selbig, W.R., Loheide, S.P., Shuster, W.D., Scharenbroch, B.C., Coville, R.C., Kruegler, J., et al., 2022. Quantifying the stormwater runoff volume reduction benefits of urban street tree canopy. Sci. Total Environ. 806, 151296. https://doi.org/10.1016/j.scitotenv.2021.151296.

Sklar, M., 1959. Fonctions de répartition à n dimensions et leurs marges. Annales de l'ISUP 8 (3), 229–231.

Szkop, Z., 2020. Evaluating the sensitivity of the i-tree eco pollution model to different pollution data inputs: a case study from warsaw, Poland. Urban Forestry. Urban Greening 55, 126859. https://doi.org/10.1016/j.ufug.2020.126859.

Soares, A.L., Rego, F.C., McPherson, E.G., Simpson, J.R., Peper, P.J., Xiao, Q., 2011. Benefits and costs of street trees in Lisbon, Portugal. Urban For. Urban Green. 10 (2), 69–78. https://doi.org/10.1016/j.ufug.2010.12.001.

Sousa-Silva, R., Smargiassi, A., Kneeshaw, D., Dupras, J., Zinszer, K., Paquette, A., 2021. Strong variations in urban allergenicity riskscapes due to poor knowledge of tree pollen allergenic potential. Sci. Rep. 11 (1), 10196. https://doi.org/10.1038/s41598-021-89353-7.

Suresh Ramanan, S., Osman, M., Shanker, A.K., Sridhar, K.B., 2021. Oxygen production potential of trees in urban areas: a reality check? Curr. Sci. 121 (5), 622–625. https://www.jstor.org/stable/27310664.

Szepannek, G., 2018. clustMixType: user-friendly clustering of mixed-type data in R. The R Journal 200–208. https://doi.org/10.32614/RJ-2018-048.

Tang, F., Ishwaran, H., 2017. Random forest missing data algorithms. Stat. Anal. Data Min.: The ASA Data Science Journal 10 (6), 363–377. https://doi.org/10.1002/sam.11348.

Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A., Slayback, D.A., 2021. Satellite imaging reveals increased proportion of population exposed to floods. Nature 596 (7870), 80–86. https://doi.org/10.1038/s41586-021-03695-w.

Timilsina, N., Beck, J., Eames, M.S., Hauer, R.J., Werner, L.P., 2017. A comparison of local and general models of leaf area and biomass of urban trees in USA. Urban Forestry. Urban Greening 24, 157–163. https://doi.org/10.1016/j.ufug.2017.04.003.

United Nations, UN, 2023. Independent group of scientists appointed by the secretary-general, global sustainable development report 2023: times of crisis, times of change: science for accelerating transformations to sustainable development. https://sustainabledevelopment.un.org/content/documents/24797GSDR_report_2019.pdf.

United Nations, UN, 2025. United nations trade & development (UNCTAD) data hub. Total and urban population, annual. https://unctadstat.unctad.org/datacentre/dataviewer/US.PopTotal. (Accessed 2 March 2025) [dataset].

USDA Forest Service, 2020. i-Tree hydro user's manual (V 6.3 beta). https://www.itreetools.org/documents/251/Hydro_Manual_v6.pdf.

van den Boogaart, K.G., Tolosana-Delgado, R., Bren, M., 2024. Compositions: compositional data analysis. R package version 2.0–8. https://CRAN.R-project.org/package=compositions.

van den Brekel, L., Lenters, V., Mackenbach, J.D., Hoek, G., Wagtendonk, A., Lakerveld, J., Grobbee, D.E., Vaartjes, I., 2024. Ethnic and socioeconomic inequalities in air pollution exposure: a cross-sectional analysis of nationwide individual-level data from the Netherlands. Lancet Planet. Health 8 (1), e18–e29. https://doi.org/10.1016/S2542-5196(23)00258-9.

Vivaldo, G., Masi, E., Taiti, C., Caldarelli, G., Mancuso, S., 2017. The network of plants volatile organic compounds. Sci. Rep. 7 (1), 11050. https://doi.org/10.1038/s41598-017-10975-x.

Wang, J., Endreny, T.A., Nowak, D.J., 2008. Mechanistic simulation of tree effects in an urban water balance model1. JAWRA J. American Water Res. Assoc 44 (1), 75–85. https://doi.org/10.1111/j.1752-1688.2007.00139.x.

Wang, X., Yao, J., Yu, S., Miao, C., Chen, W., He, X., 2018. Street trees in a Chinese forest city: structure, benefits and costs. Sustainability 10 (3), 674. https://doi.org/10.3390/su10030674.

Westfall, J.A., Henning, J.G., Edgar, C.B., 2021. Urban tree measurement variability and the contribution to uncertainty in estimates of ecosystem services. Urban Forestry. Urban Greening 64, 127302. https://doi.org/10.1016/j.ufug.2021.127302.

Xu, X., Zhang, Z., Bao, L., Li, M., Yu, X., Fan, D., et al., 2017. Influence of rainfall duration and intensity on particulate matter removal from plant leaves. Sci. Total Environ. 609, 11–16. https://doi.org/10.1016/j.scitotenv.2017.07.141.

Xu, X., Xia, J., Gao, Y., Zheng, W., 2020. Additional focus on particulate matter wash-off events from leaves is required: a review of studies of urban plants used to reduce airborne particulate matter pollution. Urban Forestry. Urban Greening 48, 126559. https://doi.org/10.1016/j.ufug.2019.126559.

Yang, Y., Endreny, T.A., Nowak, D.J., 2011. Itree-hydro: snow hydrology update for the urban forest hydrology model1. JAWRA J. American Water Res. Assoc 47 (6), 1211–1218. https://doi.org/10.1111/j.1752-1688.2011.00564.x.

Yang, B., Lee, D.K., Heo, H.K., Biging, G., 2019. The effects of tree characteristics on rainfall interception in urban areas. Landsc. Ecol. Eng. 15, 289–296. https://doi.org/10.1007/s11355-019-00383-w.

Yu, Z., Yang, G., Zuo, S., Jørgensen, G., Koga, M., Vejre, H., 2020. Critical review on the cooling effect of urban blue-green space: a threshold-size perspective. Urban For. Urban Green. 49, 126630. https://doi.org/10.1016/j.ufug.2020.126630.

Yue, K., De Frenne, P., Fornara, D.A., Van Meerbeek, K., Li, W., Peng, X., Ni, X., Peng, Y., Wu, F., Yang, Y., Peñuelas, J., 2021. Global patterns and drivers of rainfall partitioning by trees and shrubs. Glob. Change Biol. 27 (14), 3350–3357. https://doi.org/10.1111/gcb.15644.

Zardo, L., Geneletti, D., Pérez-Soba, M., Van Eupen, M., 2017. Estimating the cooling capacity of green infrastructures to support urban planning. Ecosyst. Serv. 26, 225–235. https://doi.org/10.1016/j.ecoser.2017.06.016.

Zhou, W., Yan, Z., Zhang, L., 2024. A comparative study of 11 non-linear regression models highlighting autoencoder, dbn, and svr, enhanced by shap importance analysis in soybean branching prediction. Sci. Rep. 14 (1). https://doi.org/10.1038/s41598-024-55243-x.

Zhu, P., Zhang, Y., 2008. Demand for urban forests in United States cities. Landsc. Urban Plann. 84 (3–4), 293–300. https://doi.org/10.1016/j.landurbplan.2007.09.005.

Ziegler, A., König, I.R., 2014. Mining data with random forests: current options for real-world applications. Wiley Interdisciplinary Reviews: Data Min. Knowl. Discov. 4 (1), 55–63. https://doi.org/10.1002/widm.1114.