



Constructing ecological indices for urban environments using species distribution models

ARIEL LEVI SIMONS¹ · STEVIE CALDWELL¹ · MICHELLE FU¹ · JOSE GALLEGOS¹ · MICHAEL GATHERU¹ · LAURA RICCARDELLI¹ · NHI TRUONG¹ · VALERIA VIERA¹

Accepted: 27 June 2022 / Published online: 16 July 2022
© The Author(s) 2022

Abstract

In an increasingly urbanized world, there is a need to study urban areas as their own class of ecosystems as well as assess the impacts of anthropogenic impacts on biodiversity. However, collecting a sufficient number of species observations to estimate patterns of biodiversity in a city can be costly. Here we investigated the use of community science-based data on species occurrences, combined with species distribution models (SDMs), built using MaxEnt and remotely-sensed measures of the environment, to predict the distribution of a number of species across the urban environment of Los Angeles. By selecting species with the most accurate SDMs, and then summarizing these by class, we were able to produce two species richness models (SRMs) to predict biodiversity patterns for species in the class Aves and Magnoliopsida and how they respond to a variety of natural and anthropogenic environmental gradients.

We found that species considered native to Los Angeles tend to have significantly more accurate SDMs than their non-native counterparts. For all species considered in this study we found environmental variables describing anthropogenic activities, such as housing density and alterations to land cover, tend to be more influential than natural factors, such as terrain and proximity to freshwater, in shaping SDMs. Using a random forest model we found our SRMs could account for approximately 54% and 62% of the predicted variation in species richness for species in the classes Aves and Magnoliopsida respectively. Using community science-based species occurrences, SRMs can be used to model patterns of urban biodiversity and assess the roles of environmental factors in shaping them.

Keywords Species distribution modeling · Maxent · Random forest · Ecological assessments · Ecological indices

✉ ARIEL LEVI SIMONS
levisimons@gmail.com

STEVIE CALDWELL
stevie.km.caldwell@gmail.com

MICHELLE FU
mjfu@g.ucla.edu

JOSE GALLEGOS
josegallegos@g.ucla.edu

MICHAEL GATHERU
MGATHERU@gmail.com

LAURA RICCARDELLI
lariccardelli@ucla.edu

NHI TRUONG
NHTRUO2@gmail.com

VALERIA VIERA
VALERIAVIERA12@gmail.com

¹ University of California - Los Angeles Institute of the Environment and Sustainability Los Angeles, 90024 Los Angeles, CA, USA

Introduction

The expansion of cities, both in their geographic scope as well as their use of resources, can result in ecological degradation (Elmqvist et al. 2016; Johnson and Munshi-South 2017). However, urban environments are not necessarily just degraded forms of pre-existing ecosystems but may be more accurately considered as a class of anthropogenic biome (Pincetl 2015; Fleming and Bateman 2018; Teixeira and Fernandes 2020). Urban environments are shaped by a combination of variables both natural, such as bioclimate and topography (Qian et al. 2020; Kendal et al. 2018;), and anthropogenic, such as land use and artificial illumination (Johnson et al. 2018; Pauwels et al. 2019). Through a variety of human activities, ranging from transport to habitat modification, urban environments contain a mix of native and non-native species (Helden et al. 2012; Gaertner et al. 2017; Godefroid and Ricotta 2018; Gruver and CaraDonna

2021). While heavily modified, these environments can contain diverse and functioning ecosystems (Baldock et al. 2019; Beller et al. 2019; Wenzel et al. 2020; Casanelles-Abella et al. 2021). This in turn has motivated interest in the study and management of urban ecosystems in their own right, and how they are shaped by environmental gradients (Grêt-Regamey et al. 2017; Shaffer 2018; Montero 2020; Uchida et al. 2021).

Research has been carried out in the potential of various measures of biodiversity to be used as environmental indicators in urban environments (Godefroid 2001; Llop et al. 2012; Guiland et al. 2018; Alquezar et al. 2020). Though the presence or abundance of various species have been used as environmental indicators, there is still the ill-defined problem of selecting such species for a given set of environmental assessment criteria (Siddig et al. 2016). Furthermore, while some urban ecosystems have been systematically assessed (Baldock et al. 2019; Planillo et al. 2021; Casanelles-Abella et al. 2021), there have often been significant limitations in obtaining a sufficient number of observations with which to build and test assessment models for most cities (Cappa et al., 2021).

To address these limitations we propose using machine learning, in combination with community science-based species observations, to select and evaluate both the accuracy and behavior of species for use as indicators of environmental quality. Community science, the collection or analysis of data by non-professional scientists, has shown promise in enabling the collection of sets of species observations on a far larger scale than from individual research projects (McCaffrey 2005; Silvertown 2009; Kobori et al. 2016; Ballard et al. 2017; Spear et al. 2017), especially on private lands which are typically undersampled in urban environments (Ballard et al. 2017). Machine learning, in particular species distribution models (SDMs), can then further extend the geographic extent of our understanding of species distributions by generating predictions on how the presence of species will vary in response to environmental conditions, even from a relatively small set of observations (Elith and Leathwick 2009). The accuracy of SDMs has been used to investigate the potential of various species, both native and non-native, to act as environmental indicators (Sergio and Newton 2003; Grown et al., 2013; Vallecillo et al. 2016). In urban environments, SDMs have been investigated as a means for assessing patterns of biodiversity at spatial resolutions not typically possible with point-based sampling (Milanovich et al. 2012; Stas et al. 2020; Wellmann et al. 2020; Casanelles-Abella et al. 2021; Planillo et al. 2021). Urban SDMs have also enabled comparisons of the effects of socio-ecological factors, driven by anthropogenic activity, versus natural variations in the environment (Rhodes et al. 2006; Le Louarn et al. 2018; Liu

et al. 2019). Additionally, models predicting the richness of various groups of species, species richness models (SRMs), have also been studied in a similar fashion as SDMs for investigating the impacts of environmental conditions on urban biodiversity (Gavier-Pizarro et al. 2010; Perillo et al. 2017; Fröhlich and Ciach 2019).

In order to investigate the potential of combining the use of SDMs and SRMs with community science-based observations of species, we used the city of Los Angeles as our test case. We selected Los Angeles as there is current interest in assessing it ecologically (Jenerette et al. 2016; Spear et al. 2017; McGlynn et al. 2019; Avolio et al. 2020; Rausser 2021), covers significant variations in both elevation and microclimates (Tayyebi and Jenerette, 2016), lies within one of the 36 most biodiverse terrestrial ecosystems in the world (Myers et al. 2000), and contains a large number of observations from community scientists (Vendetti et al. 2018; Leong and Trautwein 2019; Callaghan et al., 2020). As Los Angeles is a heavily urbanized area within a designated biodiversity ‘hotspot’ (Gillespie et al. 2018), it can also serve as a model city for assessing the potential of various species to act as environmental indicators in an urban context.

Within Los Angeles we then propose to investigate the following in an urban environment:

- (1) Compare the accuracy of SDMs constructed from native and non-native species.
- (2) Assess the importance of anthropogenic and natural environmental variables in shaping urban biodiversity patterns as described by SDMs.
- (3) Identify individual species, given the accuracy of their SDMs, which may act as accurate environmental indicators.
- (4) Construct and assess the reliability of SRMs, assembled from species with the most accurate SDMs, to predict species richness across an urban landscape.

Methods

Software and workflow

Our analysis was conducted in R v4.1.2 using RStudio 2021.09.1+372 “Ghost Orchid” Release (RStudio Team, 2021). Processing of geospatial data involving the use of Geospatial Data Abstraction Library (GDAL) was done using version 3.4.1 of that software. A diagram of our analysis workflow is illustrated in Fig. 1.

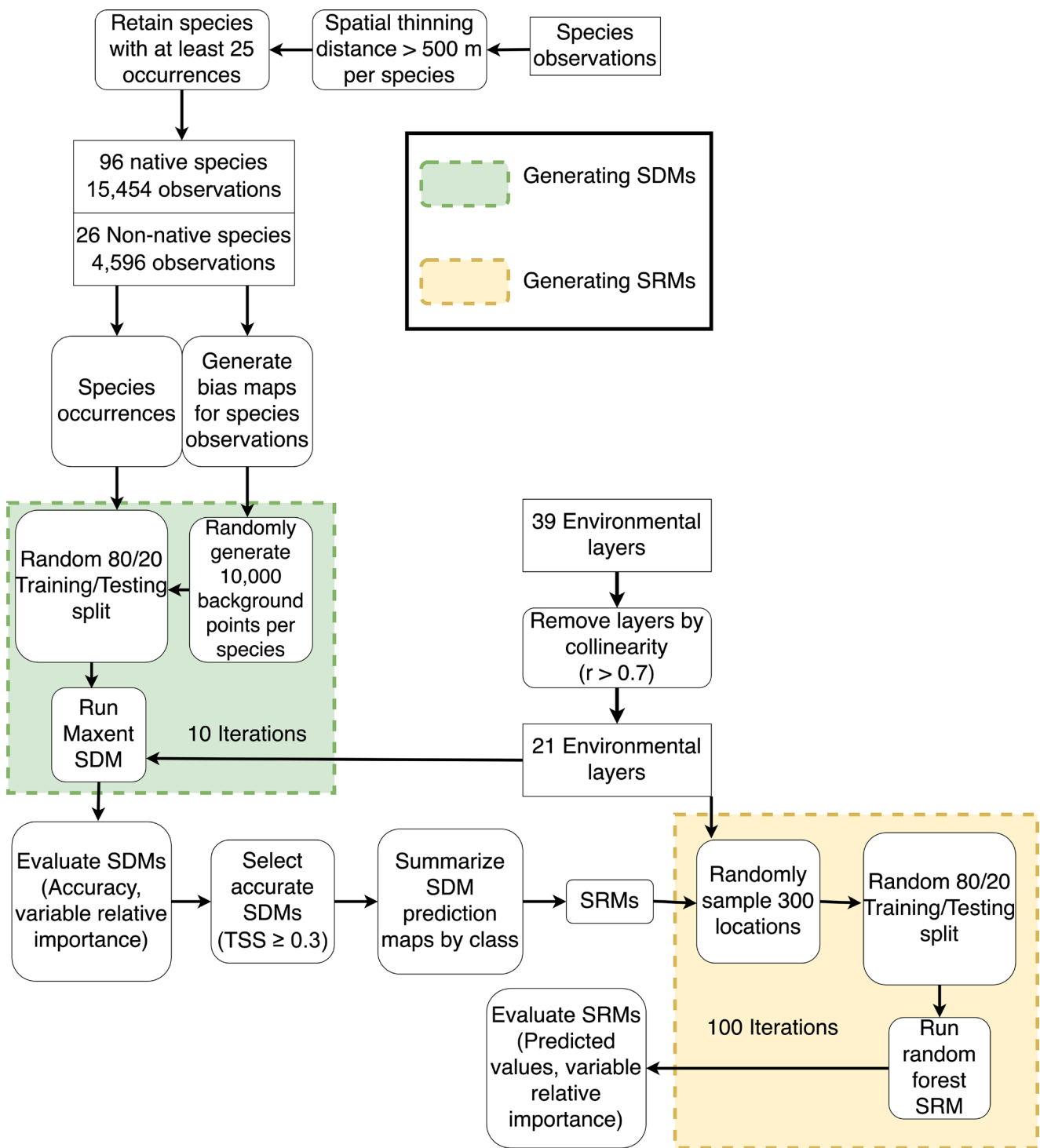


Fig. 1 Workflow diagram for generating and evaluating SDMs and SRMs

Study area

Our study area covered the city of Los Angeles, an urban environment centered at 34.05° N, 118.24° W with approximately 4 million inhabitants covering over 1200 km² (Kawabata and Shen 2006). The city is part of the greater

Los Angeles area, an urban agglomeration containing over 18 million people covering approximately 88,000 km² of southern California biomes ranging from chaparral to coastal oak forests (Tayyebi and Jenerette, 2016). This region is within a Mediterranean climatic zone, receiving the bulk of its annual rainfall (40 cm) during the winter season (Hill

Table 1 Number of species, number of species with fairly performing SDMs (Mean TSS ≥ 0.3), their native / non-native status, organized by class

Class	Status	Number of species	Number of species with fair performing SDMs
Agaricomycetes	Native	1	0
Aves	Native	30	9
Insecta	Native	13	0
Liliopsida	Native	1	1
Magnoliopsida	Native	41	13
Mammalia	Native	6	2
Reptilia	Native	4	1
Arachnida	Non-Native	3	0
Aves	Non-Native	3	1
Gastropoda	Non-Native	2	1
Insecta	Non-Native	10	0
Liliopsida	Non-Native	2	0
Magnoliopsida	Non-Native	3	0
Malacostraca	Non-Native	1	0
Mammalia	Non-Native	2	0

et al. 2016). While the environment of this region is heavily modified through the expansion of impervious surfaces, irrigation, river channelization and transportation networks, it has been found to contain a diverse, albeit heavily altered, set of species (Pataki et al. 2013; Li et al. 2019; Adams et al. 2020; Rogers et al. 2020).

Occurrence data

We sought to obtain as diverse an array of species as possible from public databases in order to assess their behavior as potential environmental indicators. Our initial set of species observational data was obtained from the Global Biodiversity Information Facility (GBIF) (GBIF.org, 2021) with the following query: (1) observations made within the period 2010–2020, (2) within the spatial extent of Los Angeles County, and (3) with a spatial uncertainty of less than 30 m. We selected GBIF data within the decade 2010–2020 as it both provided for species observations covered by the range of time when our environmental data sets were collected (SI: File 4). This initial set of 120,713 observations was then filtered using the function `st_intersection` from the R package `sf` v1.0-5 (Pebesma 2018) to contain only observations within the political boundaries of the city of Los Angeles and with a spatial uncertainty of less than 10 m.

To reduce the effects of spatial clustering of our occurrence data we performed spatial thinning, using the function `thin` within the R package `spThin` v0.2.0 (Aiello-Lammens et al. 2015), with a minimum separation distance of 500 m. Species were retained if at least 25 presence points remained after spatial thinning, as this is a conservative minimum

number for generating accurate SDMs using MaxEnt (van Proosdij et al. 2016), leaving 20,050 observations covering 122 species for analysis. These species were assigned a native or non-native status using a CalFlora list of plants and fungi native to Los Angeles County (SI: File 1), and a corresponding list of animal species curated by the Los Angeles Sanitation & Environment (LASAN, 2021) (SI: File 2), producing a split of 96 native and 26 non-native species across 10 classes (Table 1).

The number of remaining presences used to estimate sampling bias for each species ranged from 25 to 242. This method allows for the number of sampling of background points for each species to grow or shrink in proportion to its sampling effort. This method of sampling is often used to account for spatial bias in non-systematic sampling efforts (Phillips et al. 2009; Syfert et al. 2013; Molloy et al. 2017).

Environmental data

We obtained our initial set of 39 environmental layers, 25 natural and 14 anthropogenic, from a variety of sources (<https://doi.org/10.5068/D1W988>). We selected these initial layers (SI: File 4), as they cover both variations in bioclimate and topography as well as anthropogenic modifications to the environment such as housing density and exposure to light pollution, which have been found to influence the spatial distributions of a wide variety of species (Davies et al. 2008; Santorufo et al. 2012; Chong et al., 2014; Norton et al. 2016; Lin et al. 2021; Simons et al. 2021). We also selected a number of composite measures of anthropogenic disturbance, specifically habitat quality (Brown, 2019), the global human modification index (gHM) (Kennedy et al. 2019), the Calenviroscreen pollution exposure score (PollutionS) and its composite with a human population vulnerability index (CIscore) (Faust et al. 2017). These measures integrate data on land cover, habitat connectivity, and anthropogenic disturbance, impacts to human health, and interactions between human population characteristics and environmental contamination, all of which have been found to influence urban biodiversity (Table 2).

This set of environmental layers was then clipped and aligned to the city boundaries of Los Angeles using GDAL (Warmerdam 2008) with the project coordinate reference system (EPSG:2229) and resolution (30 ft / 9.1 m). From this initial set of environmental layers we retained 21 after removing those with a Pearson's correlation greater than 0.7 to other layers (Barber et al., 2021). This was done using the function `removeCollinearity` within the R package `virtualespecies` v1.5.1 (Leroy et al. 2016) with 100,000 randomly selected points. Of the layers we retained, 12 described anthropogenic variables and 9 natural environmental variables (Table 2). We calculated the variability in

Table 2 Environmental variables used in this project, with attributions

Anthropogenic environmental variables					
Variable name	Variable description	Proxy	Source resolution	Source	Evidence of ecological relevance
CIscore	CalEnviroScreen Score, Pollution Score multiplied by Population Characteristics Score	Ecological disturbance: terrestrial, atmospheric, water	30 ft / 9.1 m	(Faust et al. 2017)	(Brown, 2019) (Cushing et al. 2015) (Faust et al. 2017)
gHM	Global Human Modification index	Ecological disturbance: terrestrial	1000 m	(Kennedy et al. 2019)	(Liu et al. 2021)
gwthreats	Groundwater threats, sum of weighted GeoTracker leaking underground storage tank sites within buffered distances to populated blocks of census tracts	Ecological disturbance: water	30 ft / 9.1 m	(Faust et al. 2017)	(Ferreira et al. 2018) (Hassall 2014)
HabitatQuality	A 10 point habitat quality scale for Los Angeles native species	Ecological disturbance: terrestrial	30 ft / 9.1 m	(Brown, 2019)	(Roeland et al. 2019) (Beninde et al. 2015)
haz	Sum of weighted hazardous waste facilities and large quantity generators within buffered distances to populated blocks of census tracts	Ecological disturbance: terrestrial	30 ft / 9.1 m	(Faust et al. 2017)	(Santorufo et al. 2012)
HousingDensity	Number of housing units per square kilometer	Ecological disturbance: terrestrial, Habitat availability	30 m / 9.1 m	(Radeloff et al. 2018)	(Davies et al. 2008) (Kolbe et al. 2016) (Gagné et al., 2011) (Aronson et al. 2014) (Sol et al. 2014)
iwb	Impaired water bodies, sum of number of pollutants across all impaired water bodies within buffered distances to populated blocks of census tracts	Ecological disturbance: water	30 ft / 9.1 m	(Faust et al. 2017)	(Ferreira et al. 2018) (Datry et al. 2006)
LightPollution	Estimated exposure to artificial night sky illumination in microcandela per square meter.	Ecological disturbance: terrestrial	30 arcseconds / 776.7 m	(Falchi et al. 2016)	(Hölker et al. 2010) (Ciach et al., 2017) (Lian et al. 2021)
pesticides	Total pounds of selected active pesticide ingredients (filtered for hazard and volatility) used in production-agriculture per square mile in the census tract	Ecological disturbance: terrestrial	30 ft / 9.1 m	(Faust et al. 2017)	(Bertoncini et al. 2012) (Aronson et al. 2017)
PollutionS	Pollution Burden variable scaled with a range of 0–10. (Used to calculate CalEnviroScreen Score)	Ecological disturbance: terrestrial, atmospheric, water	30 ft / 9.1 m	(Faust et al. 2017)	(Ferreira et al. 2018)
swis	Sum of weighted solid waste sites and facilities (SWIS) within buffered distances to populated blocks of census tracts	Ecological disturbance: terrestrial	30 ft / 9.1 m	(Faust et al. 2017)	(Santorufo et al. 2012) (Ferreira et al. 2018)
traffic	Traffic density, in vehicle-kilometers per hour per road length, within 150 m of the census tract boundary	Ecological disturbance: terrestrial	30 ft / 9.1 m	(Faust et al. 2017)	(Davies et al. 2008) (Kolbe et al. 2016) (Proppe et al. 2013)
Natural environmental variables					
Aspect	Direction with which the landscape faces. North is at 0 degrees, and increases clockwise.	Habitat topography	1 arcsecond / 25.9 m	(Maune 2006)	(Davies et al. 2008) (Rastandeh et al. 2019)
bio06	Minimum Temperature of Coldest Month	Climate	30 arcseconds / 776.7 m	(Fick and Hijmans 2017)	(Kendal et al. 2018) (Polidori et al. 2021) (Heringer et al. 2021) (Beninde et al. 2015) (Harrison and Winfree 2015)

Table 2 (continued)

Anthropogenic environmental variables					
bio11	Mean Temperature of Coldest Quarter	Climate	30 arcseconds / 776.7 m	(Fick and Hijmans 2017)	(Kendal et al. 2018) (Polidori et al. 2021) (Heringer et al. 2021) (Beninde et al. 2015) (Harrison and Winfree 2015)
bio12	Annual Precipitation	Climate	30 arcseconds / 776.7 m	(Fick and Hijmans 2017)	(Kendal et al. 2018) (Polidori et al. 2021) (Heringer et al. 2021) (Beninde et al. 2015) (Harrison and Winfree 2015)
bio17	Precipitation of Driest Quarter	Climate	30 arcseconds / 776.7 m	(Fick and Hijmans 2017)	(Kendal et al. 2018) (Polidori et al. 2021) (Heringer et al. 2021) (Beninde et al. 2015) (Harrison and Winfree 2015)
LACoastlineProximity	Distance to coastline	Climate	30 ft / 9.1 m	(Hapke et al. 2006)	(Sperandii et al. 2019) (Orlando et al. 2020)
LALakeProximity	Distance to nearest open body of freshwater	Water availability	30 ft / 9.1 m	(California Department of Fish and Game, 2012)	(Hassall 2014) (Callaghan et al. 2018) (de Camargo Barbosa et al. 2020)
LAStramsProximity	Distance to nearest stream	Water availability	30 ft / 9.1 m	(Simley and Carswell 2009)	(Chin and Kupfer 2020) (Houlahan and Findlay 2003) (Melles et al. 2003)
Slope	Percent slope	Habitat topography	1 arcsecond / 25.9 m	(Maune 2006)	(Davies et al. 2008) (Rastandeh et al. 2019)

these remaining environmental layers by calculating their coefficients of variation using the function `cellStats` in raster (SI: Table 5).

Accounting for spatial bias in species observations

To visualize geographic biases in observational data we converted our spatially thinned occurrences into a heatmap using the function `stat_density_2d` in the R package `ggplot2` v3.3.5 (Wickham et al. 2016). Prior to running species distribution models, spatial biases in species observations were accounted for by modifying the sampling of background points using a probability function corresponding to sampling effort for each species (Filazzola et al. 2018; von Takach et al. 2020). This probability density function was generated for each species using a two-dimensional kernel density estimate, with the function `kde2d` in the R package `MASS` v7.3-54 (Ripley et al. 2013), using the density of spatially thinned observations for each species. The spatial bias raster of each species was then generated, along with the spatially thinned data set as a whole (SI: Fig. 4), in the project coordinate reference system and resolution, from this kernel density estimate using the function `raster` in the R package `raster` v2.5-2 (Hijmans and van Etten., 2015). Each of these spatial bias rasters were then clipped and aligned to the boundaries of Los Angeles using the function `gdalwarp`

within the R package `gdalUtils` v2.0.3.2 (Greenberg and Mattiuzzi 2015).

Building SDMs

In order to construct SDMs we used the machine learning technique of Maximum Entropy (MaxEnt). We selected MaxEnt for constructing our SDMs as it can work with presence-only data for species observations (Elith et al., 2011), which are commonly found in databases such as the GBIF (Edwards 2004). The accuracy of MaxEnt-based SDMs have also been used in assessing the potential of species to act as environmental indicators (Jose V, 2020), that is, species whose likelihood of presence can be accurately predicted by a set of environmental conditions.

For each species a set of 10 MaxEnt models were run using the function `maxent`, within the R package `dismo` v1.3-5 (Hijmans et al. 2017), with its default settings. Species input data consisted of both presence points, as well as 10,000 background points sampled using the function `xyFromCell` within the package `raster` and a sampling probability determined using a species specific spatial bias raster. To enable an assessment of accuracy, each model utilized a randomly sampled set of 80% of presence and background points for model training, while the remaining 20% were used for testing.

Evaluating SDMs

Comparing native to non-native SDM accuracy: To assess the accuracy of these models rates of true and false positives and negatives were found using the function `evaluate` within the `dismo` package. These rates were then used to calculate the True Skill Statistic (TSS) for each of the 10 models run per species. The TSS was used as it has been found to be a metric with little dependence on species prevalence (Allouche et al. 2006). The value of the TSS ranges from -1 to 1 , with a value of 1 corresponding to a perfect agreement between predicted and actual distributions, and negative values indicating a model's predictions are no better than random (Allouche et al. 2006; Liu et al. 2009). For each species we calculated the mean and standard on the TSS scores for its maxent models (SI: File 3).

To visualize the distribution of mean TSS scores for each species, and how they differ between native and non-native species, the functions `geom_violin` and `facet_grid` were used from `ggplot2`. These functions were also used to visualize the distributions of the percent relative importances of both anthropogenic and natural environmental variables in our SDMs. To test if the distribution of TSS scores differed significantly between SDMs made for native and non-native species, we used a Kruskal-Wallis (K-W) test implemented with the function `kruskal.wallis` in the R package `stats` where α was set to 0.05 .

Comparing the importance of natural and anthropogenic environmental variables: The percent relative importance of the 21 environmental variables used to generate each SDM were calculated, as a percent contribution to each model, using the function `var.importance` within the R package `ENMeval` v2.0.2 (Muscarella et al. 2014). The mean and standard deviation on these relative importance values were calculated for each species (SI: File 3).

To visualize the rank mean relative importance of environmental variables for each SDM the function `geom_tile`, within the R package `ggplot2`, was used to visualize a heatmap. To test if the distribution of the percent relative importance for natural and anthropogenic environmental variables differed significantly, for both native and non-native SDMs, we again used a K-W where α was set to 0.05 .

Selecting environmental indicator species

Species were selected for constructing our SRMs following a cutoff on the mean TSS scores for their SDMs (≥ 0.3). While a TSS cutoff of 0.4 has been used to classify SDMs as accurate (Thuiller et al. 2019), only one of our species exceeded this threshold (*Ardea herodias*, mean TSS of 0.43), and so we used a more relaxed threshold of 0.3 to select SDMs with a fair level of accuracy (Landis and Koch

1977). This selection criterion produced a list of 9 native species in the class Aves, and 13 in the class Magnoliopsida (Table 1). For species in both lists, SDM maps were generated using the function `predict` within the R package `raster`. Each SDM map was generated using the `maxent` function with all available spatially thinned presence points, 10,000 background points, and a presence threshold set as the maximum sum of the specificity and sensitivity.

Constructing and evaluating SRMs

The SDMs from these two classes were then summed, using the `calc` function in `raster`, to construct a SRM for species in each class (Fig. 2). These two SRMs represent the richness of nine species within the class Aves and another representing the richness of thirteen species within the class Magnoliopsida (SI: Table 6). Visualization of these SRM maps, which illustrate the predicted species richness per $900 \text{ ft}^2 / 83.6 \text{ m}^2$ map call, was done using the function `leaflet` in the R package `leaflet` v2.0.4.1 (Cheng et al. 2018).

To evaluate the accuracy of our SRMs to respond to our environmental variables (Table 2) we used a random forest model, which has been shown to be an accurate method for predicting the spatial distribution of other ecological data sets (Rodriguez-Galiano et al. 2012). For each SRM we constructed a set of 100 random forest models using training data sets and the function `tuneRF`, within the R package `randomForest` v4.6-14 (Liaw and Wiener 2002), with `stepFactor` set to a value of 1 and `doBest` set as 'true'. Data for each random forest model of a SRM was extracted from 300 randomly selected locations within our study area using the function `extract` within `raster`. We chose 300 locations as this was found to be the minimum number needed to consistently model significant (Pearson correlation, $p < 10^{-4}$) predictions of both SRMs using random forest. A training set was generated by randomly sampling 80% of these extracted data using the function `kfold` within `dismo`. Predicted SRM values were calculated using the function `predict` within `randomForest`, which were then compared against their actual counterparts with a Pearson correlation coefficient. The root mean square error (RMSE) for each random forest model was calculated using the function `rmse` in the R package `Metrics` v0.1.4 (Hamner et al. 2018).

The mean and standard deviation on the 100 Pearson correlation coefficients, generated in evaluating each random forest model of a SRM, were then calculated for each class' SRM. This was done by first performing a Fisher transformation, using the `FisherZ` function within the R package `DescTools` v0.99.44 (Signorelli, 2020), on the set of Pearson correlation coefficients. We used a Fisher transformation as it has been found to produce less biased summary statistics for a set of Pearson correlation coefficients (Corey

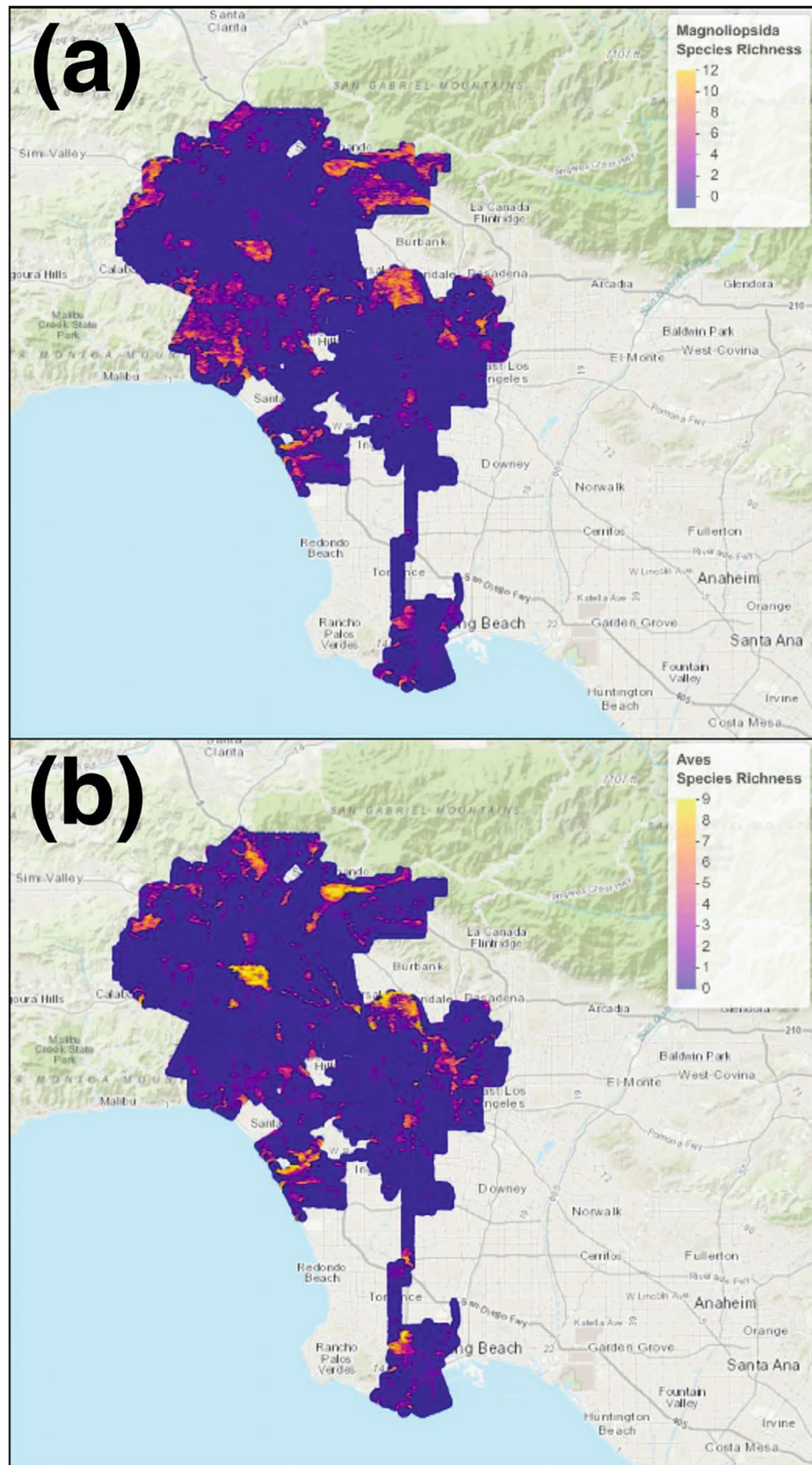


Fig. 2 SRMs covering the city boundaries of Los Angeles, for native species in the class Aves (a) and Magnoliopsida (b), constructed using SDMs with mean TSS scores exceeding 0.3. Color scales reflect the predicted number of species per map cell

et al., 1998). The average and standard deviation were then inverse Fisher transformed, using the function `FisherZInv` within the `DescTools` package, to produce a single statistic per SRM. For each SRM the median and variance on the RMSE values was also calculated. To account for the percent variance explained by our SRMs we used the following formula for each SRM:

$$\% \text{Variance explained} = \frac{\sum (Richness_{actual} - Richness_{predicted})^2}{\sum (Richness_{actual} - \text{mean}(Richness_{actual}))^2} \quad (1)$$

To compare the relative importance of environmental variables within our models we used the function `importance` within `randomForest`. The relative importance of our environmental variables was calculated as their mean decrease in node impurity as quantified by their Gini indices, with a larger value denoting greater relative importance. To generate individual partial dependence plots we used the function `partialPlot` within `randomForest`.

To visualize a heat map of the 100 iterations of the partial dependence plots for our environmental variables for each SRM, we then used the function `geom_bin2d` within `ggplot2`. For all our partial dependence heat maps, we divided our axes into 20 bins and generated a best-fit curve using the function `stat_smooth` within `ggplot2`.

Results

Evaluating SDMs

Comparing native to non-native SDM accuracy We found native species had SDMs which tended to be more accurate, as assessed by their mean TSS scores (Fig. 3a) (K-W test, $\chi^2 = 7.57$, $p < 10^{-2}$), than their non-native counterparts. For all of our SDMs, the relative importance of anthropogenic environmental variables tended to exceed their natural counterparts (K-W test, $\chi^2 = 16.54$, $p < 10^{-4}$). Even within native and non-native species groups (Fig. 3B), anthropogenic environmental variables tended to have a greater relative importance for both native (K-W test, $\chi^2 = 11.58$, $p < 10^{-3}$) and non-native species (K-W test, $\chi^2 = 5.23$, $p < 0.05$).

Comparing the importance of natural and anthropogenic environmental variables: For a majority of species we found their spatial distributions to be strongly influenced by the density of housing units and slope (SI: Fig. 5). Both habitat quality and proximity to streams were found to be of high relative importance in influencing the SDMs of some plants, birds, and insects (SI: Fig. 5a-c). For members of

Magnoliopsida and Liliopsida the aspect of the local terrain also tended to be influential on the spatial distribution of most species (SI: Fig. 5a and h).

Selecting environmental indicator species

Of the 122 species in this study, 28 were found to have SDMs considered accurate (Table 1), that is those with a TSS score of at least 0.3 (Landis and Koch 1977). Of these species 26 were native to Los Angeles, and were predominantly in the class Aves (9 species) or Magnoliopsida (13 species) (Table 1 and SI: Table 6).

Constructing and evaluating SRMs

SRMs were assembled by summing SDMs for Los Angeles native species in the class Aves and Magnoliopsida (Fig. 2), which were the only two sets of species in this study to provide for much variation in richness (Table 1 and SI: Table 6). The richness of species in both classes could be reliably predicted across Los Angeles, from the environmental variables used in this study, albeit with a moderate level of uncertainty as described by the RMSEs of SRMs from each class (Table 3).

For both SRMs the density of housing units was the most influential variable considered (Table 4), which was found to generally have a negative relationship with species richness in both classes (SI: Figures 6a and 7a). Both slope and the CIscore, which is a composite of various measures of environmental contamination along with both human health and socioeconomic factors, tended to be highly important in predicting the local richness of species in the class Aves or Magnoliopsida (Table 4). For both SRMs a decline in species richness tended to be associated with an increase in slope or CIscore (SI: Figures 6d-e and 7c and e). The richness of bird species, more than species in the class Magnoliopsida, tended to be more strongly influenced by proximity to either streams or lakes (Table 4), with species richness tending to decline with distance to either type of freshwater (SI: Fig. 6b-c). The richness of species in Magnoliopsida were found to be strongly influenced by both habitat quality and the gHM (Table 4), a measure of anthropogenic modification to the landscape, with a positive trend associated with habitat quality and a negative one with the gHM (SI: Fig. 7b and d).

While the relative importance of environmental variables declines quickly with rank order for both SRMs (Table 4), we do note the predicted behavior of both models to be somewhat counterintuitive with a variety of the environmental variables considered. For our avian SRM we predict a general increase in species richness with various measures of anthropogenic disturbance ranging from nearby vehicular

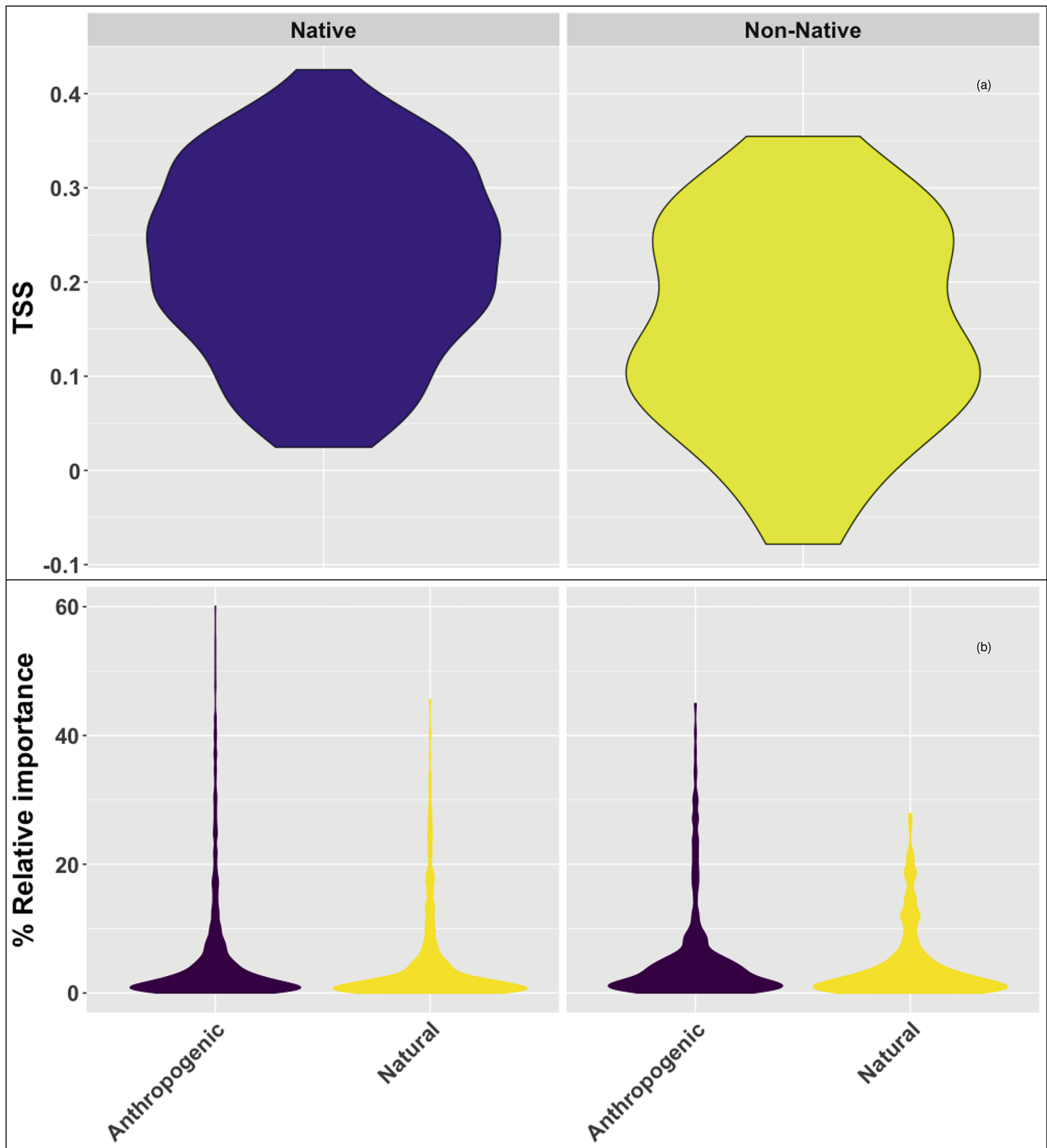


Fig. 3 Violin plots of the mean TSS scores per SDM split by native / non-native status (a), and the mean percent relative importance of environmental variables for SDMs split both by native / non-native status and anthropogenic / natural category (b)

traffic density, to soil and water contamination, pesticide use, and light pollution (SI: Fig. 6f-i, o-p, s-t). For our Magnoliopsida SRM, species richness was also predicted to increase with various measures of soil and water contamination and pesticide use (SI: Fig. 7n-o, q-s). However,

both SRMs tend to increase with variables which are composite metrics of anthropogenic disturbance. In particular, we modeled species richness declining with the gHM and CIScore and increasing with habitat quality for both SRMs (SI: Figures 6e, j and m and 7 b, d-e).

Table 3 Summary statistics on the accuracy of 100 SRMs, for species in class Aves and Magnoliopsida, generated via random forest. Pearson correlation coefficients calculated between training and testing values of species richness. Percent variance explained is calculated as how well predicted richness explains the target variance of the training richness. RMSE is calculated as the deviance of actual to predicted species richness

Class	Mean (Standard deviation) Pearson correlation coefficient	Mean (Standard deviation) % Variance explained	Median (Variance) RMSE	Variance on RMSE
Aves	0.78 (0.19)	54 (13)	1.16 (0.06)	0.06
Magnoliopsida	0.83 (0.21)	62 (12)	1.54 (0.06)	0.06

Discussion

Evaluating SDMs

Comparing native to non-native SDM accuracy For the species we studied, SDMs derived from native species tended to be significantly more accurate than SDMs derived from non-native species (Fig. 3a). This may reflect the tendency of non-native species, especially in urban environments, to be able to occupy a broader set of environmental conditions than native species (Le Viol et al. 2012; Cervelli et al., 2013; Concepción et al., 2016; Callaghan et al., 2019; Colléony and Shwartz, 2020). The result of this ecological tendency being less expected variation in the likelihood of a species being present in response to variations in the environment.

We found that membership in our list of potential environmental indicator species (Table 1), particularly those native

to Los Angeles, were biased towards members of the classes Aves and Magnoliopsida. This reflects a bias in the community science data sets found in GBIF towards observations of both classes (Troutet et al. 2017; Petersen et al. 2021). Despite this bias, species from both classes have been used as environmental indicators in a variety of studies of urban areas. For example, avian diversity has been found to respond to human modification of landscapes (Callaghan et al., 2021), the extent and density of the built environment (Pinho et al. 2016), as well as habitat quality and groundwater contamination (Mekonen 2017), while the diversity of flowering plants has been found to respond to the impact of road traffic on air and soil quality (Philips et al., 2021) as well as landscape modifications (Ioja and Breuste 2020).

Comparing the importance of natural and anthropogenic environmental variables: For both native and non-native species, we tend to find the relative importance of anthropogenic environmental variables to be greater in our SDMs than natural variables (Fig. 3b and SI: Fig. 5). This

Table 4 Relative rank importance of environmental variables to our SRMs with variable importance listed in descending order, along with the mean and standard deviation of their relative importance, as measured by the mean decrease in their Gini indices. Values recorded as the mean value (standard deviation on the mean value), with higher values indicating greater importance of the variable to the model

Rank importance Aves SRM	Relative importance Aves SRM	Rank importance Magnoliopsida SRM	Relative importance Magnoliopsida SRM
HousingDensity	131.31 (38.95)	HousingDensity	339.52 (86.15)
LASStreamsProximity	114.44 (45.42)	HabitatQuality	208.40 (82.43)
LALakeProximity	63.89 (35.99)	Slope	184.45 (57.87)
Slope	41.06 (12.42)	gHM	144.35 (45.86)
CIscore	36.70 (16.68)	CIscore	108.93 (56.62)
pesticides	33.90 (23.98)	LASStreamsProximity	91.00 (34.74)
PollutionS	32.92 (13.80)	LACoastlineProximity	56.98 (16.80)
traffic	31.82 (13.04)	bio12	53.83 (16.94)
iwb	29.53 (15.79)	LightPollution	47.71 (14.54)
HabitatQuality	28.77 (12.32)	LALakeProximity	44.38 (22.18)
LACoastlineProximity	27.76 (9.10)	Aspect	39.45 (12.78)
bio12	27.46 (9.71)	bio11	38.29 (10.95)
gHM	24.69 (6.70)	bio06	35.58 (12.41)
Aspect	24.46 (10.00)	PollutionS	34.20 (11.75)
LightPollution	20.84 (7.17)	iwb	32.11 (14.90)
haz	19.20 (6.78)	traffic	31.15 (10.00)
bio06	18.38 (4.71)	gwthreats	26.45 (7.65)
bio11	18.20 (5.11)	haz	24.82 (7.71)
gwthreats	17.38 (4.92)	pesticides	24.40 (14.39)
swis	13.83 (5.40)	swis	20.08 (7.43)
bio17	5.93 (1.90)	bio17	18.16 (9.04)

pattern has been observed in other cities (Aronson et al. 2014; Liu et al. 2017), and the greater importance of anthropogenic factors in shaping urban biodiversity (Faeth et al. 2011; Li et al. 2019) may simply reflect the fact of cities being some of the most anthropogenically modified habitats on Earth (Chase and Chase, 2016).

Many of the species we selected as potential environmental indicators also have large geographic ranges, often extending well beyond California (Table 1). A large geographic extent for a species can be associated with a wide climatic tolerance (Slatyer et al. 2013), although this may only weakly hold for various plant species (Bocsi et al. 2016). With many of our potential environmental indicator species having ranges which cover multiple climatic zones, this may partly explain the low relative importance of bioclimatic variables in their SDMs (SI: Fig. 5), which would reduce the overall relative importance of natural environmental variables in our overall set of SDMs. The lower relative importance for natural environmental variables, may in part also result from their lower levels of variation across the landscape of Los Angeles (SI: Table 5).

Selecting environmental indicator species

We find our species with the most environmentally responsive SDMs to be strongly biased toward native species in the class Aves or Magnoliopsida (Table 1). While this in part reflects taxonomic biases in our source data, the presence of species responsive enough to environmental variations to be reliable as indicators in an urban environment may stem from traits conducive to tolerance of anthropogenic environments such as small size and high dispersal (McKinney and Lockwood 1999; Lizée et al. 2011).

Constructing and evaluating SRMs

Our SRMs were able to capture a number of ecological relationships involving simple environmental variables, such as the distance to freshwater or housing density, indicating their potential utility in assessing environmental conditions in an urban environment such as Los Angeles. For example, housing density has been associated with a decline in avian species richness (Gagné and Fahrig 2011), which has also been more broadly observed in relation to the overall fraction of an urban area dedicated to buildings (Aronson et al. 2014). Both this decline in urban species richness, as well as the importance of the role of building density, has also been observed in a variety of plant species (Godefroid and Koedam 2007). Our Avian SRM also appears to capture, at least on the scale of approximately 10 km (SI: Fig. 6b–c), the expected decline in species richness with distance to freshwater (Chin and Kupfer 2020; de Camargo Barbosa

et al. 2020). At a similar geographic scale our Magnoliopsida SRM predicts a decline with the distance to freshwater (SI: Fig. 7f and j), in line with the observed decline of plant species richness with distance to freshwater observed in Mediterranean climates such as Los Angeles (Hawkins et al. 2003).

Less intuitive relationships between species richness and individual anthropogenic variables we also observed. For example, variables describing the number of unique pollutants within nearby impaired water bodies (iwb) and the areal density of pesticide use (pesticides), were both positively associated with a rise in Avian and Magnoliopsidan species richness (SI: Figures 6f and i and 7 o and s). While pesticide use or impaired water bodies would be expected to reduce biodiversity, in urban environments they are often associated with irrigated and fertilized landscapes, which are often more species rich as a result (Clarke et al., 2013; Avolio et al. 2020). Another measure of anthropogenic disturbance, the density of road traffic (traffic), was also associated with a positive response in species richness with our Aves SRM (SI: Fig. 6 h). Nonetheless, this response may be a result of bias in community scientists tending to take bird observations in proximity to roads (Keller and Scallan 1999; Mair and Reute, 2016).

We find additional potential in our method for constructing SRMs for urban environmental assessments given the responses of our two SRMs to more integrated measures of anthropogenic disturbance: an index of the quality of habitat for species native to Los Angeles (HabitatQuality), a measure of anthropogenic modification (gHM), an integrated measure of exposure to pollution (PollutionS), and an integrated measure of exposure to pollution weighted by local measures of socioeconomic vulnerability and health outcomes (CIScore). The value of HabitatQuality integrates data on land cover and vegetation found to support native biodiversity in Los Angeles, along with a measure of geographic connectivity between habitats supportive of native biodiversity (Brown, 2019). That our two SRMs both predict a positive relationship between HabitatQuality and species richness (SI: Figures 6j and 7b) is not surprising, with urban species richness generally found to increase with related metrics such as the area or connectivity of vegetated habitats (Aronson et al. 2014; Beninde et al. 2015; Callaghan et al. 2018). In a similar fashion the value of the gHM, in integrating the areal fraction dedicated to human activities such as impervious surfaces and electricity transmission (Kennedy et al. 2019), would also correspond to habitat fragmentation and decline in species richness (SI: Fig. 6 m and 7d).

Another unexpected result were discrepancies between the predicted species richness of our SRMs and the values of PollutionS and the CIScore. The value of PollutionS

incorporates exposure to measures of air pollution, such as ground-level ozone concentration, human health effects, such as asthma hospitalization rates, and additional measures of environmental degradation, such as the density of hazardous waste sites (Meehan August et al. 2012). However, the value of PollutionS is not associated with a predicted decline in species richness in either of our SRMs (SI: Fig. 6 g and 7n). When the pollution metrics summarized in PollutionS are multiplied against demographic and socioeconomic characteristics, such as the fraction of the population under 5 or over 65 or people over 25 without a secondary education, to produce the CIscore it is associated with a predicted decline in species richness (SI: Figures 6e and 7e). This difference in predicted responses of species richness between exposure to pollution and a composite measure of it with characteristics of the human population may reflect a relationship between the interactions of the physical environment, socioeconomic factors, and biodiversity in an urban context (Schell et al. 2020).

Limitations

Beyond the overrepresentation of species from a few classes in GBIF (Troudet et al. 2017; Petersen et al. 2021), there remain issues with geographic biases in the observational data used to construct SDMs. We found geographic biases in the locations of all the species used in this study (SI: Fig. 4a), as well those used to construct our SRMs (SI: Fig. 4b-c), and this type of observational bias has been observed in prior studies using community science data (Mair and Ruete 2016; Petersen et al. 2021). While we were able to account for these spatial biases in generating our SDMs, we found our data to be clustered in large urban parks (SI: Fig. 4). This geographic bias toward large urban parks may in part be a result of their accessibility, as well as the expectation of community scientists on finding more diversity to capture in park spaces versus more developed land (Bonney et al. 2009; Callaghan et al., 2020). In urban environments biodiversity tends to be relatively high in large park spaces (Matthies et al. 2013), with the richness of species in parks following a species-area relationship (Nielsen et al. 2014). Additionally, cities are often built in biodiverse areas and many of these park spaces may simply reflect this relic biodiversity (Kühn et al. 2004; Luck 2007).

Geographic biases in species observations may also stem from the skewed demographics of people collecting community science data, which is biased towards those who are under 65 years of age and have a post-secondary degree (Ganzevoort et al. 2017; Lopez et al. 2020). Such a bias may affect the number of species observations in areas with a higher proportion of the population below 5 or over 65

years of age, which would bias predicted species richness downwards. This may in part explain our negative relationship between pollution exposure, which factors in exposure of a variety of environmental pollutants to more vulnerable young and elderly populations, and species richness as predicted by our SRMs. Such biases in observer demographics may then tend to skew observed biodiversity downwards, enhancing any modeled negative response between species richness and various socio-ecological variables.

Beyond the potential confounding of observer biases with socioeconomic and anthropogenic environmental variables, is the issue of temporal gaps between the collection of species occurrences and environmental variables used to generate SDMs. First, our species observations are recent and only cover a decade, mirroring the rapid but recent growth in community science-based platforms for recording species observations (Di Cecco et al. 2021). We aggregated our observations across this window of time to provide for sufficient data density to generate SDMs and SRMs, although this coarse temporal resolution may obscure a number of ecological patterns. A similar issue exists with many of our environmental layers, such as those derived from the Calenviroscreen or bioclimate, which are updated only once every few years and may therefore reduce the predictive power of any SDM or SRM which incorporates them.

We also note that the spatial resolutions of a number of the remotely sensed environmental layers used in this study (SI: File 4) may not be fine enough to sufficiently capture their influence in a highly heterogeneous urban environment. For example, while urban areas tend to produce highly localized microclimates which influence local patterns of biodiversity (Fournier et al. 2020; Casanelles-Abella et al. 2021), the resolution the bioclimatic variables we used are near a kilometer in scale and may effectively obscure the signal any potential relationship. Exposure to artificial light at night is also mapped at the same spatial resolution (Falchi et al. 2016), although its illuminance has been found to vary by orders of magnitude at the scale of only hundreds of meters (Simons et al. 2020). Our species observation data then, while large in scale thanks to the efforts of numerous community scientists, may be mapped against a level of environmental variation artificially lowered by spatially coarse remote sensing data.

Future prospects

The volume of data captured through community science has grown rapidly in recent years. Although there are biases in the type of species observed, and where they're observed, these may be compensated for through the use of environmental DNA (eDNA). Using eDNA from soil, sediment,

water, or air sample inventories, we can complement existing species monitoring efforts through the identification of thousands of species at once, including plants, animals, and microbes (Stat et al. 2019; Lin et al. 2021; Nørgaard et al. 2021). Use of eDNA can greatly complement traditional species monitoring by enabling greater taxonomic resolution (Deiner et al. 2017; Ruppert et al. 2019), the detection of species which tend to avoid the presence of humans (Yonezawa et al. 2020; Mas-Carrió et al. 2021), or organisms such as bacteria and fungi which can be difficult to monitor using traditional observations (Frøslev et al. 2019; Liddicoat et al. 2022). Comparisons of eDNA with observational methods have also indicated their potential to help capture additional elements of ecologically relevant information, such as the functional diversity of various groups of species (Aglieri et al., 2021; Donald et al. 2021; Sigsgaard et al. 2021), particularly with regards to identifying ecological indicators (Yan et al. 2018; Blattner et al. 2021; Seymour et al. 2021).

With the recent developments in routine low-cost hyperspectral imaging there is the potential to overcome a number of these limitations in being able to acquire frequent, and high resolution, environmental data to improve models of urban biodiversity (Mozgeris et al. 2018; Zhang et al. 2020; Hartling et al. 2021). Such remotely sensed environmental data may be captured using cubesats (Kimm et al. 2020; Grøtte et al. 2021), small and low-cost satellites, as well as airborne drones (Räsänen et al. 2020; Dierssen et al. 2021). Of particular use to monitoring highly dynamic and heterogeneous urban environments, such data can be collected at resolutions under 3 m in scale (Salgado-Hernanz et al. 2021) and daily in time (Rhodes et al. 2022).

Underlying many SDMs there is the assumption that the likely geographic distribution of species is purely a function of the environmental gradients present. While environmental variations may be an important driver in the geographic distribution of species, a variety of ecological factors ranging from interspecies competition to dispersal will also have some degree of influence (Soberón 2007). One potential method for inferring the components of variation in SDMs which may be attributed to interspecies interactions, or unknown environmental factors, is the use of joint-SDMs (jSDMs) (Pollock et al. 2014). Although, as with our study, there is still a large degree of uncertainty in disentangling the contributions to variations in SDMs with a large number of potentially interacting species (Zurell et al. 2018).

Conclusions

As a predominantly urban species there is a need for humans to better understand the ecology, and condition, of their most common habitat. This presents particular difficulties

in developing ecological assessments for urban areas as they are both highly heterogeneous, and in a global environment rapidly responding to anthropogenic activity, ever more dynamic. Despite potential biases, both in the spatial distribution of sampling efforts and towards particular taxonomic groups, we demonstrate the potential use of species distribution modeling and community science-based observations to both identify potential environmental indicators and assess the response of biodiversity to environmental conditions in an urban environment. We found evidence that native SDMs tend to be more accurate than their non-native counterparts, and that biodiversity patterns in urban environments are driven more by anthropogenic activities than variations in the natural environment. In constructing SRMs from the most accurate SDMs, we were able to detect a number of plausible responses of urban biodiversity to environmental conditions. Of particular interest is the potential for our SRMs to detect declines in biodiversity associated with measures which integrate both exposure to various pollutants, as well socioeconomic characteristics, on a local basis. However, because of various biases associated with the backgrounds of community scientists, some of the environmental responses of our SRMs may be confounded with socioeconomic variables. We therefore recommend future development of this methodology to incorporate a broader initial set of environmental and socioeconomic variables in order to better correct for potential observer biases, and subsequently improve the accuracy of result SDMs and SRMs to capture meaningful environmental responses in urban environments.

Acknowledgements We would like to acknowledge Michelle Barton, Dr. Travis Longcore, Dr. Rachel Meyer, and Dr. Sabrina Drill for their insights in landscape ecology used in this project.

Authors' contributions All authors contribution to the original draft, as well as subsequent edits. Conceptualization of this project was performed by Ariel Levi Simons. Methodology was performed by Ariel Levi Simons, Jose Gallegos, and Michael Gatheru. Software development was performed by Ariel Levi Simons, Laura Riccardelli, and Jose Gallegos. Validation was performed by Valeria Viera. Data curation was performed by Michelle Fu. Supervision was performed by Stevie Caldwell. Visualization was performed by Nhi Truong.

Funding Not applicable.

Data Availability All map layers used in this analysis are available at <https://doi.org/10.5068/D1W988>. All species observations were obtained from <https://www.gbif.org>.

Code Availability All scripts used in this analysis are available here: <https://github.com/LASanitation/LASAN>.

Declarations

Conflicts of interest/Competing interests Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams BJ, Li E, Bahlai CA et al (2020) Local- and landscape-scale variables shape insect diversity in an urban biodiversity hot spot. *Ecol Appl* 30. <https://doi.org/10.1002/eap.2089>
- Aiello-Lammens ME, Boria RA, Radosavljevic A et al (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography (Cop)* 38:541–545. <https://doi.org/10.1111/ecog.01132>
- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Alquezar RD, Tolesano-Pascoli G, Gil D, Macedo RH (2020) Avian biotic homogenization driven by airport-affected environments. *Urban Ecosyst* 23:507–517. <https://doi.org/10.1007/s11252-020-00936-0>
- Aronson MFJ, Lepczyk CA, Evans KL et al (2017) Biodiversity in the city: key challenges for urban green space management. *Front Ecol Environ* 15:189–196
- Aronson MFJ, La Sorte FA, Nilon CH et al (2014) A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proc R Soc B Biol Sci* 281. <https://doi.org/10.1098/rspb.2013.3330>
- Avolio M, Pataki DE, Jenerette GD et al (2020) Urban plant diversity in Los Angeles, California: Species and functional type turnover in cultivated landscapes. *Plants People Planet* 2:144–156. <https://doi.org/10.1002/ppp3.10067>
- Baldock KCR, Goddard MA, Hicks DM et al (2019) A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nat Ecol Evol* 3:363–373. <https://doi.org/10.1038/s41559-018-0769-y>
- Ballard HL, Robinson LD, Young AN et al (2017) Contributions to conservation outcomes by natural history museum-led citizen science: Examining evidence and next steps. *Biol Conserv* 208:87–97. <https://doi.org/10.1016/j.biocon.2016.08.040>
- Barber RA, Ball SG, Morris RKA, Gilbert F (2022) Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Divers Distrib* 28:128–141. <https://doi.org/10.1111/ddi.13442>
- Beller EE, Spotswood EN, Robinson AH et al (2019) Building Ecological Resilience in Highly Modified Landscapes. *Bioscience* 69:80–92. <https://doi.org/10.1093/biosci/biy117>
- Beninde J, Veith M, Hochkirch A (2015) Biodiversity in cities needs space: A meta-analysis of factors determining intra-urban biodiversity variation. *Ecol Lett* 18:581–592
- Bertoncini AP, Machon N, Pavoine S, Muratet A (2012) Local gardening practices shape urban lawn floristic communities. *Landsc Urban Plan* 105:53–61
- Blattner L, Ebner JN, Zopfi J, von Fumetti S (2021) Targeted non-invasive bioindicator species detection in eDNA water samples to assess and monitor the integrity of vulnerable alpine freshwater environments. *Ecol Indic* 129:107916
- Bocsi T, Allen JM, Bellemare J et al (2016) Plants' native distributions do not reflect climatic tolerance. *Divers Distrib* 22:615–624. <https://doi.org/10.1111/ddi.12432>
- Bonney R, Cooper CB, Dickinson J et al (2009) Citizen science: A developing tool for expanding science knowledge and scientific literacy. *Bioscience* 59:977–984
- Bornstein C, Fross D et al (1953) (2005) California native plants for the garden
- Brown IT, Gillespie TW, Gold MA (2019) Managing Cities as Urban Ecosystems: Analysis Tools for Biodiversity Stewardship in Los Angeles. University of California, Los Angeles
- Bullock SH, Ceseña JMS, Rebman JP, Riemann H (2008) Flora and vegetation of an isolated mountain range in the desert of Baja California. *Southwest Nat* 53:61–73. [https://doi.org/10.1894/0038-4909\(2008\)53\[61:FAVOAI\]2.0.CO;2](https://doi.org/10.1894/0038-4909(2008)53[61:FAVOAI]2.0.CO;2)
- Callaghan CT, Major RE, Lyons MB et al (2018) The effects of local and landscape habitat attributes on bird diversity in urban greenspaces. *Ecosphere* 9. <https://doi.org/10.1002/ecs2.2347>
- Casanelles-Abella J, Chauvier Y, Zellweger F et al (2021) Applying predictive models to study the ecological properties of urban ecosystems: A case study in Zürich, Switzerland. *Landsc Urban Plan* 214:104137
- Cheng J, Karambelkar B, Xie Y (2018) Leaflet: Create interactive web maps with the javascript 'leaflet' library. R Packag version 2
- Chin EY, Kupfer JA (2020) Identification of environmental drivers in urban greenway communities. *Urban For Urban Green* 47:126549
- Ciach M, Fröhlich A (2017) Habitat type, food resources, noise and light pollution explain the species composition, abundance and stability of a winter bird assemblage in an urban environment. *Urban Ecosyst* 20:547–559
- Cronquist A, Holmgren AH, Holmgren NH et al (1984) Intermountain Flora: Vascular Plants of the Intermountain West, USA The New York Botanical Garden
- Cushing L, Faust J, August LM et al (2015) Racial/ethnic disparities in cumulative environmental health impacts in California: evidence from a statewide environmental justice screening tool (CalEnviroScreen 1.1). *Am J Public Health* 105:2341–2348
- Datry T, Malard F, Gibert J (2006) Effects of artificial stormwater infiltration on urban groundwater ecosystems. *Urban groundwater management and sustainability*. Springer, pp 331–345
- Davies RG, Barbosa O, Fuller RA et al (2008) City-wide relationships between green spaces, urban land use and topography. *Urban Ecosyst* 11:269–287
- de Camargo Barbosa KV, Rodewald AD, Ribeiro MC, Jahn AE (2020) Noise level and water distance drive resident and migratory bird species richness within a Neotropical megacity. *Landsc Urban Plan* 197:103769
- Deiner K, Bik HM, Mächler E et al (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol Ecol* 26:5872–5895
- Del Hoyo J, Elliot A, Sargatal J (1992) Handbook of the Birds of the World. Vol. 1., (Lynx Edicions: Barcelona, Spain)
- Di Cecco GJ, Barve V, Belitz MW et al (2021) Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for Biodiversity Science. *Bioscience* 71:1179–1188. <https://doi.org/10.1093/biosci/biab093>
- Dierrsen HM, Ackleson SG, Joyce KE et al (2021) Living up to the hype of hyperspectral aquatic remote sensing: science, resources and outlook. *Front Environ Sci* 9
- Donald J, Muriene J, Chave J et al (2021) Multi-taxa environmental DNA inventories reveal distinct taxonomic and functional

- diversity in urban tropical forest fragments. *Glob Ecol Conserv* 29:e01724
- Drilling N (2002) Mallard (*Anas platyrhynchos*). *Birds North Am*
- Edwards JL (2004) Research and societal benefits of the Global Biodiversity Information Facility. *Bioscience* 54:485–486
- Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu Rev Ecol Syst* 40:677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elmqvist T, Zipperer W, Güneralp B (2016) Urbanization, habitat loss, biodiversity decline: solution pathways to break the cycle. In: Seta, Karen; Solecki, William D; Griffith, Corrie A(eds) *Routledge Handb Urban Glob Environ Chang* London New York Routledge 2016:139–151
- Faeth SH, Bang C, Saari S (2011) Urban biodiversity: patterns and mechanisms. *Ann N Y Acad Sci* 1223:69–81. <https://doi.org/10.1111/j.1749-6632.2010.05925.x>
- Falchi F, Cinzano P, Duriscoe D et al (2016) The new world atlas of artificial night sky brightness. *Sci Adv* 2. <https://doi.org/10.1126/sciadv.1600377>
- Faust J, August L, Bangia K et al (2017) CalEnviroScreen 3.0 Report. Calif Environ Prot Agency, Off Environ Heal Hazard Assess Sacramento, CA
- Ferreira CSS, Walsh RPD, Ferreira AJD (2018) Degradation in urban areas. *Curr Opin Environ Sci Heal* 5:19–25
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315. <https://doi.org/10.1002/joc.5086>
- Filazzola A, Sotomayor DA, Lortie CJ (2018) Modelling the niche space of desert annuals needs to include positive interactions. *Oikos* 127:264–273. <https://doi.org/10.1111/oik.04688>
- Fleming PA, Bateman PW (2018) Novel predation opportunities in anthropogenic landscapes. *Anim Behav* 138:145–155. <https://doi.org/10.1016/j.anbehav.2018.02.011>
- Fournier B, Vázquez-Rivera H, Clappe S et al (2020) The spatial frequency of climatic conditions affects niche composition and functional diversity of species assemblages: the case of Angiosperms. *Ecol Lett* 23:254–264
- Fröhlich A, Ciach M (2019) Nocturnal noise and habitat homogeneity limit species richness of owls in an urban environment. *Environ Sci Pollut Res* 26:17284–17291. <https://doi.org/10.1007/s11356-019-05063-8>
- Froslev TG, Kjøller R, Bruun HH et al (2019) Man against machine: Do fungal fruitbodies and eDNA give similar biodiversity assessments across broad environmental gradients? *Biol Conserv* 233:201–212
- Gaertner M, Wilson JRU, Cadotte MW et al (2017) Non-native species in urban environments: patterns, processes, impacts and challenges. *Biol Invasions* 19:3461–3469. <https://doi.org/10.1007/s10530-017-1598-7>
- Gagné SA, Fahrig L (2011) Do birds and beetles show similar responses to urbanization? *Ecol Appl* 21:2297–2312. <https://doi.org/10.1890/09-1905.1>
- Ganzevoort W, van den Born RJG, Halfman W, Turnhout S (2017) Sharing biodiversity data: citizen scientists' concerns and motivations. *Biodivers Conserv* 26:2821–2837. <https://doi.org/10.1007/s10531-017-1391-z>
- Gavier-Pizarro GI, Radeloff VC, Stewart SI et al (2010) Housing is positively associated with invasive exotic plant species richness in New England, USA. *Ecol Appl* 20:1913–1925. <https://doi.org/10.1890/09-2168.1>
- GBIF.org (2021) GBIF.org. <https://doi.org/10.15468/dl.vs5s6n>. Accessed 29 Jan 2021
- Gillespie TW, Ostermann-Kelm S, Dong C et al (2018) Monitoring changes of NDVI in protected areas of southern California. *Ecol Indic* 88:485–494. <https://doi.org/10.1016/j.ecolind.2018.01.031>
- Godefroid S (2001) Temporal analysis of the Brussels flora as indicator for changing environmental quality. *Landsc Urban Plan* 52:203–224. [https://doi.org/10.1016/S0169-2046\(00\)00117-1](https://doi.org/10.1016/S0169-2046(00)00117-1)
- Godefroid S, Koedam N (2007) Urban plant species patterns are highly driven by density and function of built-up areas. *Landsc Ecol* 22:1227–1239. <https://doi.org/10.1007/s10980-007-9102-x>
- Godefroid S, Ricotta C (2018) Alien plant species do have a clear preference for different land uses within urban environments. *Urban Ecosyst* 21:1189–1198. <https://doi.org/10.1007/s11252-018-0792-4>
- Greenberg JA, Mattiuzzi M (2015) gdalUtils: wrappers for the geospatial data abstraction library (GDAL) utilities. R Packag version 2
- Greenwell BM (2017) pdp: An R Package for Constructing Partial Dependence Plots. *R J* 9:421. <https://doi.org/10.32614/RJ-2017-016>
- Grêt-Regamey A, Altwegg J, Sirén EA et al (2017) Integrating ecosystem services into spatial planning—A spatial decision support tool. *Landsc Urban Plan* 165:206–219. <https://doi.org/10.1016/j.landurbplan.2016.05.003>
- Grotte ME, Birkeland R, Honoré-Livermore E et al (2021) Ocean Color Hyperspectral Remote Sensing With High Resolution and Low Latency—The HYPSO-1 CubeSat Mission. *IEEE Trans Geosci Remote Sens* 60:1–19
- Gruver A, CaraDonna P (2021) Chicago Bees: Urban Areas Support Diverse Bee Communities but With More Non-Native Bee Species Compared to Suburban Areas. *Environ Entomol* 50:982–994
- Guilland C, Maron PA, Damas O, Ranjard L (2018) Biodiversity of urban soils for sustainable cities. *Environ Chem Lett* 16:1267–1282. <https://doi.org/10.1007/s10311-018-0751-6>
- Guin JD, Gillis WT, Beaman JH (1981) Recognizing the Toxicodendrons (poison ivy, poison oak, and poison sumac). *J Am Acad Dermatol* 4:99–114. [https://doi.org/10.1016/S0190-9622\(81\)70014-8](https://doi.org/10.1016/S0190-9622(81)70014-8)
- Guinan JA, Gowaty PA, Eltzroth EK (2000) Western Bluebird (*Sialia mexicana*). *birds North Am* 32
- Hamner B, Frasco M, LeDell E (2018) Metrics: Evaluation metrics for machine learning. R Packag version 01 4:2018
- Hancock J, Kushlan JA (2010) *The herons handbook*. A&C Black
- Hapke CJ, Reid D, Richmond BM et al (2006) National assessment of shoreline change Part 3: Historical shoreline change and associated coastal land loss along sandy shorelines of the California Coast. US Geol Surv Open File Rep 1219:79
- Harrison T, Winfree R (2015) Urban drivers of plant-pollinator interactions. *Funct Ecol* 29:879–888
- Hartling S, Sagan V, Maimaitijiang M (2021) Urban tree species classification using UAV-based multi-sensor data fusion and machine learning. *GIScience Remote Sens* 58:1250–1275. <https://doi.org/10.1080/15481603.2021.1974275>
- Hassall C (2014) The ecology and biodiversity of urban ponds. *Wiley Interdiscip Rev Water* 1:187–206
- Hawkins BA, Field R, Cornell HV et al (2003) ENERGY, WATER, AND BROAD-SCALE GEOGRAPHIC PATTERNS OF SPECIES RICHNESS. *Ecology* 84:3105–3117. <https://doi.org/10.1890/03-8006>
- Helden AJ, Stamp GC, Leather SR (2012) Urban biodiversity: comparison of insect assemblages on native and non-native trees. *Urban Ecosyst* 15:611–624. <https://doi.org/10.1007/s11252-012-0231-x>
- Heringer G, Del Bianco Faria L, Villa PM et al (2021) Urbanization affects the richness of invasive alien trees but has limited influence on species composition. *Urban Ecosyst* 1–11
- Hijmans RJ (2015) Raster: Geographic Data Analysis and Modeling
- Hijmans RJ, Phillips S, Leathwick J et al (2017) Package ‘dismo’. *Circles* 9:1–68
- Hill RA, Weber MH, Leibowitz SG et al (2016) The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. *JAWRA J Am Water Resour Assoc* 52:120–128. <https://doi.org/10.1111/1752-1688.12372>

- Hölker F, Wolter C, Perkin EK, Tockner K (2010) Light pollution as a biodiversity threat. *Trends Ecol Evol* 25:681–682
- Houlahan JE, Findlay CS (2003) The effects of adjacent land use on wetland amphibian species richness and community composition. *Can J Fish Aquat Sci* 60:1078–1094
- Joja C, Breuste J (2020) Urban Protected Areas and Urban Biodiversity. *Making Green Cities*. Springer, pp 289–398
- Jenerette GD, Clarke LW, Avolio ML et al (2016) Climate tolerances and trait choices shape continental patterns of urban tree biodiversity. *Glob Ecol Biogeogr* 25:1367–1376. <https://doi.org/10.1111/geb.12499>
- Johnson AL, Borowy D, Swan CM (2018) Land use history and seed dispersal drive divergent plant community assembly patterns in urban vacant lots. *J Appl Ecol* 55:451–460. <https://doi.org/10.1111/1365-2664.12958>
- Johnson MTJ, Munshi-South J (2017) Evolution of life in urban environments. *Sci* (80-) 358:eaam8327. <https://doi.org/10.1126/science.aam8327>
- Jose VS, Nameer PO (2020) The expanding distribution of the Indian Peafowl (*Pavo cristatus*) as an indicator of changing climate in Kerala, southern India: A modelling study using MaxEnt. *Ecol Indic* 110:105930. <https://doi.org/10.1016/j.ecolind.2019.105930>
- Kawabata M, Shen Q (2006) Job Accessibility as an Indicator of Auto-Oriented Urban Structure: A Comparison of Boston and Los Angeles with Tokyo. *Environ Plan B Plan Des* 33:115–130. <https://doi.org/10.1068/b31144>
- Keller CME, Scallan JT (1999) Potential roadside biases due to habitat changes along breeding bird survey routes. *Condor* 101:50–57
- Kendal D, Dobbs C, Gallagher RV et al (2018) A global comparison of the climatic niches of urban and native tree populations. *Glob Ecol Biogeogr* 27:629–637. <https://doi.org/10.1111/geb.12728>
- Kennedy CM, Oakleaf JR, Theobald DM et al (2019) Managing the middle: A shift in conservation priorities based on the global human modification gradient. *Glob Chang Biol* 25:811–826. <https://doi.org/10.1111/geb.14549>
- Kimm H, Guan K, Jiang C et al (2020) Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the US Corn Belt using Planet Labs CubeSat and STAIR fusion data. *Remote Sens Environ* 239:111615
- Kobori H, Dickinson JL, Washitani I et al (2016) Citizen science: a new approach to advance ecology, education, and conservation. *Ecol Res* 31:1–19. <https://doi.org/10.1007/s11284-015-1314-y>
- Kolbe SE, Miller AI, Cameron GN, Culley TM (2016) Effects of natural and anthropogenic environmental influences on tree community composition and structure in forests along an urban-wildland gradient in southwestern Ohio. *Urban Ecosyst* 19:915–938
- Kühn I, Brandl R, Klotz S (2004) The flora of German cities is naturally species rich. *Evol Ecol Res* 6:749–764
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 159–174
- LASAN (2021) Los Angeles Native Fauna. Los Angeles
- Lathrop EW, Thorne RF (1978) A flora of the Santa Ana Mountains, California. *Aliso A J Syst Florist Bot* 9:197–278
- Le Louarn M, Clergeau P, Strubbe D, Deschamps-Cottin M (2018) Dynamic species distribution models reveal spatiotemporal habitat shifts in native range-expanding versus non-native invasive birds in an urban area. *J Avian Biol* 49:jav-01527
- Le Viol I, Jiguet F, Brotons L et al (2012) More and more generalists: two decades of changes in the European avifauna. *Biol Lett* 8:780–782. <https://doi.org/10.1098/rsbl.2012.0496>
- Leong M, Trautwein M (2019) A citizen science approach to evaluating US cities for biotic homogenization. *PeerJ* 7:e6879. <https://doi.org/10.7717/peerj.6879>
- Leroy B, Meynard CN, Bellard C, Courchamp F (2016) virtualspecies, an R package to generate virtual species distributions. *Ecography (Cop)* 39:599–607. <https://doi.org/10.1111/ecog.01388>
- Li E, Parker SS, Pauly GB et al (2019) An Urban Biodiversity Assessment Framework That Combines an Urban Habitat Classification Scheme and Citizen Science Data. <https://doi.org/10.3389/fevo.2019.00277>. *Front Ecol Evol* 7:
- Lian X, Jiao L, Zhong J et al (2021) Artificial light pollution inhibits plant phenology advance induced by climate warming. *Environ Pollut* 291:118110
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R news* 2:18–22
- Liddicoat C, Krauss SL, Bissett A et al (2022) Next generation restoration metrics: Using soil eDNA bacterial community data to measure trajectories towards rehabilitation targets. *J Environ Manage* 310:114748
- Lin M, Simons AL, Harrigan RJ et al (2021) Landscape analyses using eDNA metabarcoding and Earth observation predict community biodiversity in California. *Ecol Appl* 31. <https://doi.org/10.1002/eap.2379>
- Liu C, White M, Newell G (2009) Measuring the accuracy of species distribution models: a review. In: *Proceedings 18th World IMACs/MODSIM Congress*. Cairns, Australia. p 4247
- Liu G, Rowley JLL, Kingsford RT, Callaghan CT (2021) Species' traits drive amphibian tolerance to anthropogenic habitat modification. *Glob Chang Biol* 27:3120–3132
- Liu J, Yu M, Tomlinson K, Slik JWF (2017) Patterns and drivers of plant biodiversity in Chinese university campuses. *Landsc Urban Plan* 164:64–70. <https://doi.org/10.1016/j.landurbplan.2017.04.008>
- Liu J, Bai H, Ma H, Feng G (2019) Bird diversity in Chinese urban parks was more associated with natural factors than anthropogenic factors. *Urban For Urban Green* 43:126358. <https://doi.org/10.1016/j.ufug.2019.06.001>
- Lizée M-H, Mauffrey J-F, Taton T, Deschamps-Cottin M (2011) Monitoring urban environments on the basis of biological traits. *Ecol Indic* 11:353–361. <https://doi.org/10.1016/j.ecolind.2010.06.003>
- Llop E, Pinho P, Matos P et al (2012) The use of lichen functional groups as indicators of air quality in a Mediterranean urban environment. *Ecol Indic* 13:215–221. <https://doi.org/10.1016/j.ecolind.2011.06.005>
- Lopez B, Minor E, Crooks A (2020) Insights into human-wildlife interactions in cities from bird sightings recorded online. *Landsc Urban Plan* 196:103742. <https://doi.org/10.1016/j.landurbplan.2019.103742>
- Luck GW (2007) A review of the relationships between human population density and biodiversity. *Biol Rev* 82:607–645. <https://doi.org/10.1111/j.1469-185X.2007.00028.x>
- Mair L, Ruete A (2016) Explaining Spatial Variation in the Recording Effort of Citizen Science Data across Multiple Taxa. *PLoS ONE* 11:e0147796. <https://doi.org/10.1371/journal.pone.0147796>
- Mas-Carrió E, Schneider J, Nasanbat B et al (2021) Assessing environmental DNA metabarcoding and camera trap surveys as complementary tools for biomonitoring of remote desert water bodies. *Environ DNA*
- Matthies S, Kopel D, Rüter S et al (2013) Vascular Plant Species Richness Patterns in Urban Environments: Case Studies from Hannover, Germany and Haifa, Israel. *Modeling of Land-Use and Ecological Dynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 107–118
- Maune DF (2006) Vertical accuracy assessment report, 2006 LiDAR Bare-Earth Dataset for Los Angeles Region Imagery Acquisition Consortium (LAR-IAC). Dewberry, Fairfax, VA
- McCaffrey RE (2005) Using citizen science in urban bird studies. *Urban habitats* 3:70–86
- McCrimmon DA, Ogden JC, Bancroft GT et al (2020) Great Egret (*Ardea alba*), version 1.0. *Birds World* (SM Billerman, Ed Cornell Lab Ornithol Ithaca, NY, USA <https://doi.org/10.2173/bowgregr1>

- McGlynn TP, Meineke EK, Bahlai CA et al (2019) Temperature accounts for the biodiversity of a hyperdiverse group of insects in urban Los Angeles. *Proc R Soc B Biol Sci* 286:20191818. <https://doi.org/10.1098/rspb.2019.1818>
- McKinney ML, Lockwood JL (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends Ecol Evol* 14:450–453. [https://doi.org/10.1016/S0169-5347\(99\)01679-1](https://doi.org/10.1016/S0169-5347(99)01679-1)
- Meehan August L, Faust JB, Cushing L et al (2012) Methodological considerations in screening for cumulative environmental health impacts: Lessons learned from a pilot study in California. *Int J Environ Res Public Health* 9:3069–3084
- Mekonen S (2017) Birds as biodiversity and environmental indicator. *J Nat Sci Res* 7:28–34
- Melles S, Glenn S, Martin K (2003) Urban bird diversity and landscape complexity: species–environment associations along a multiscale habitat gradient. *Conserv Ecol* 7
- Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography (Cop)* 36:1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Milanovich JR, Peterman WE, Barrett K, Hopton ME (2012) Do species distribution models predict species richness in urban and natural green spaces? A case study using amphibians. *Landsc Urban Plan* 107:409–418. <https://doi.org/10.1016/j.landurbplan.2012.07.010>
- Molloy SW, Davis RA, Dunlop JA, van Etten EJB (2017) Applying surrogate species presences to correct sample bias in species distribution models: a case study using the Pilbara population of the Northern Quoll. *Nat Conserv* 18:25–46. <https://doi.org/10.3897/natureconservation.18.12235>
- Montero S (2020) Leveraging Bogotá: Sustainable development, global philanthropy and the rise of urban solutionism. *Urban Stud* 57:2263–2281. <https://doi.org/10.1177/0042098018798555>
- Mowbray TB, Sedinger JS, Trost RE (2002) Canada goose (*Branta canadensis*). *Birds North Am*
- Mozgeris G, Juodkienė V, Jonikavičius D et al (2018) Ultra-Light Aircraft-Based Hyperspectral and Colour-Infrared Imaging to Identify Deciduous Tree Species in an Urban Environment. *Remote Sens* 10:1668. <https://doi.org/10.3390/rs10101668>
- Muscarella R, Galante PJ, Soley-Guardia M et al (2014) ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol Evol* 5:1198–1205. <https://doi.org/10.1111/2041-210X.12261>
- Myers N, Mittermeier RA, Mittermeier CG et al (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858. <https://doi.org/10.1038/35002501>
- Nell CS, Meza-Lopez MM, Croy JR et al (2018) Relative effects of genetic variation sensu lato and sexual dimorphism on plant traits and associated arthropod communities. *Oecologia* 187:389–400. <https://doi.org/10.1007/s00442-018-4065-y>
- Nesom GL (2013) Taxonomic notes on *Diplacus* (Phrymaceae). *Phytoneuron* 66:1–8
- Nielsen AB, Van Den Bosch M, Maruthaveeran S, van den Bosch CK (2014) Species richness in urban parks and its drivers: a review of empirical evidence. *Urban Ecosyst* 17:305–327
- Nørgaard L, Olesen CR, Trøjsgaard K et al (2021) eDNA metabarcoding for biodiversity assessment, generalist predators as sampling assistants. *Sci Rep* 11:1–12
- Norton BA, Evans KL, Warren PH (2016) Urban Biodiversity and Landscape Ecology: Patterns, Processes and Planning. *Curr Landsc Ecol Reports* 1:178–192. <https://doi.org/10.1007/s40823-016-0018-5>
- Orlando L, Ortega L, Defeo O (2020) Urbanization effects on sandy beach macrofauna along an estuarine gradient. *Ecol Indic* 111:106036
- Panetsos CA, Baker HG (1967) The origin of variation in “wild” *Raphanus sativus* (Cruciferae) in California. *Genetica* 38:243–274
- Parsons KC, Master TL (2020) Snowy Egret (*Egretta thula*). In: Poole AF, Gill FB (eds) *Birds of the World*. Cornell Lab of Ornithology
- Pataki DE, McCarthy HR, Gillespie T et al (2013) A trait-based ecology of the Los Angeles urban forest. *Ecosphere* 4:1–20. <https://doi.org/10.1890/ES13-00017.1>
- Pauwels J, Le Viol I, Azam C et al (2019) Accounting for artificial light impact on bat activity for a biodiversity-friendly urban planning. *Landsc Urban Plan* 183:12–25. <https://doi.org/10.1016/j.landurbplan.2018.08.030>
- Pebesma E (2018) Simple features for R: standardized support for spatial vector data. *R J* 10:439–446 doi: 10.32614. RJ-2018-009
- Perillo A, Mazzoni LG, Passos LF et al (2017) Anthropogenic noise reduces bird species richness and diversity in urban parks. *Ibis (Lond 1859)* 159:638–646. <https://doi.org/10.1111/ibi.12481>
- Petersen TK, Speed JDM, Grøtan V, Austrheim G (2021) Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection. *Ecol Solut Evid*. <https://doi.org/10.1002/2688-8319.12048.2>
- Phillips BB, Bullock JM, Gaston KJ et al (2021) Impacts of multiple pollutants on pollinator activity in road verges. *J Appl Ecol* 58:1017–1029. <https://doi.org/10.1111/1365-2664.13844>
- Phillips SJ, Dudik M, Elith J et al (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl* 19:181–197. <https://doi.org/10.1890/07-2153.1>
- Pincell S (2015) Cities as Novel Biomes: Recognizing Urban Ecosystem Services as Anthropogenic. *Front Ecol Evol* 3. <https://doi.org/10.3389/fevo.2015.00140>
- Pinho P, Correia O, Lecoq M et al (2016) Evaluating green infrastructure in urban environments using a multi-taxa and functional diversity approach. *Environ Res* 147:601–610. <https://doi.org/10.1016/j.envres.2015.12.025>
- Planillo A, Kramer-Schadt S, Buchholz S et al (2021) Arthropod abundance modulates bird community responses to urbanization. *Divers Distrib* 27:34–49. <https://doi.org/10.1111/ddi.13169>
- Polidori C, García-Gila J, Blasco-Aróstegui J, Gil-Tapetado D (2021) Urban areas are favouring the spread of an alien mud-dauber wasp into climatically non-optimal latitudes. *Acta Oecol* 110:103678
- Pollock LJ, Tingley R, Morris WK et al (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods Ecol Evol* 5:397–406. <https://doi.org/10.1111/2041-210X.12180>
- Prosdij ASJ, Sosef MSM, Wieringa JJ, Raes N (2016) Minimum required number of specimen records to develop accurate species distribution models. *Ecography (Cop)* 39:542–552. <https://doi.org/10.1111/ecog.01509>
- Proppe DS, Sturdy CB, St. Clair CC (2013) Anthropogenic noise decreases urban songbird diversity and may contribute to homogenization. *Glob Chang Biol* 19:1075–1084. <https://doi.org/10.1111/gcb.12098>
- Qian S, Qin D, Wu X et al (2020) Urban growth and topographical factors shape patterns of spontaneous plant community diversity in a mountainous city in southwest China. *Urban For Urban Green* 55:126814. <https://doi.org/10.1016/j.ufug.2020.126814>
- Radeloff VC, Helmers DP, Kramer HA et al (2018) Rapid growth of the US wildland-urban interface raises wildfire risk. *Proc Natl Acad Sci* 115:3314–3319. <https://doi.org/10.1073/pnas.1718850115>
- Rail S (1993) Black-Crowned Night Heron—*Nycticorax nycticorax*
- Räsänen A, Juutinen S, Kalacska M et al (2020) Peatland leaf-area index and biomass estimation with ultra-high resolution remote sensing. *GIScience Remote Sens* 57:943–964
- Rastandeh A, Pedersen Zari M, Brown D, Vale R (2019) Analysis of landform and land cover: Potentials for urban biodiversity conservation against rising temperatures. *Urban Policy Res* 37:338–349

- Rausser C (2021) UCLA Sustainable LA Grand Challenge: Innovatively Applying Research to Create the First Sustainable Megacity. In: World Scientific Encyclopedia of Climate Change: Case Studies of Climate Risk, Action, and Opportunity Volume 1. World Scientific, pp 105–114
- Rebman JP, Gibson J, Rich K (2016) Annotated checklist of the vascular plants of Baja California, Mexico. San Diego Society of Natural History
- Rhodes JR, Wiegand T, McAlpine CA et al (2006) Modeling species' distributions to improve conservation in semiurban landscapes: koala case study. *Conserv Biol* 20:449–459
- Rhodes MW, Bennie JJ, Spalding A et al (2022) Recent advances in the remote sensing of insects. *Biol Rev* 97:343–360
- Ripley B, Venables B, Bates DM et al (2013) Package 'mass'. *Cran r* 538:113–120
- Rodriguez-Galiano VF, Ghimire B, Rogan J et al (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J Photogramm Remote Sens* 67:93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Roeland S, Moretti M, Amorim JH et al (2019) Towards an integrative approach to evaluate the environmental ecosystem services provided by urban forest. *J For Res* 30:1981–1996
- Rogers JB, Stein ED, Beck MW, Ambrose RF (2020) The impact of climate change induced alterations of streamflow and stream temperature on the distribution of riparian species. *PLoS ONE* 15:e0242682. <https://doi.org/10.1371/journal.pone.0242682>
- RStudio Team (2021) RStudio: Integrated Development Environment for R
- Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob Ecol Conserv* 17:e00547
- Salgado-Hernanz PM, Bauzá J, Alomar C et al (2021) Assessment of marine litter through remote sensing: recent approaches and future goals. *Mar Pollut Bull* 168:112347
- Santorufu L, Van Gestel CAM, Rocco A, Maisto G (2012) Soil invertebrates as bioindicators of urban soil quality. *Environ Pollut* 161:57–63. <https://doi.org/10.1016/j.envpol.2011.09.042>
- Schell CJ, Dyson K, Fuentes TL et al (2020) The ecological and evolutionary consequences of systemic racism in urban environments. *Science* (80-) 369-. <https://doi.org/10.1126/science.aay4497>
- Sergio F, Newton IAN (2003) Occupancy as a measure of territory quality. *J Anim Ecol* 72:857–865
- Seymour M, Edwards FK, Cosby BJ et al (2021) Environmental DNA provides higher resolution assessment of riverine biodiversity and ecosystem function via spatio-temporal nestedness and turnover partitioning. *Commun Biol* 4:1–12
- Shaffer HB (2018) Urban biodiversity arks. *Nat Sustain* 1:725–727. <https://doi.org/10.1038/s41893-018-0193-y>
- Siddig AAH, Ellison AM, Ochs A et al (2016) How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. *Ecol Indic* 60:223–230. <https://doi.org/10.1016/j.ecolind.2015.06.036>
- Signorell A, Aho K, Alfons A et al (2019) DescTools: Tools for descriptive statistics. *R Packag version* 099 28:17
- Sigsgaard EE, Olsen K, Hansen MDD et al (2021) Environmental DNA metabarcoding of cow dung reveals taxonomic and functional diversity of invertebrate assemblages. *Mol Ecol* 30:3374–3389
- Silvertown J (2009) A new dawn for citizen science. *Trends Ecol Evol* 24:467–471. <https://doi.org/10.1016/j.tree.2009.03.017>
- Simley JD, Carswell WJ Jr (2009) The National Map—Hydrography: US Geological Survey Fact Sheet 2009–3054. US Geol Surv Natl Center, Reston, VA
- Simons AL, Martin KLM, Longcore T (2021) Determining the Effects of Artificial Light at Night on the Distributions of Western Snowy Plovers (*Charadrius nivosus nivosus*) and California Grunion (*Leuresthes tenuis*) in Southern California. *J Coast Res*
- Simons AL, Yin X, Longcore T (2020) High correlation but high scale-dependent variance between satellite measured night lights and terrestrial exposure. *Environ Res Commun* 2:021006
- Slatyer RA, Hirst M, Sexton JP (2013) Niche breadth predicts geographical range size: a general ecological pattern. *Ecol Lett* 16:1104–1114
- Soberón J (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecol Lett* 10:1115–1123. <https://doi.org/10.1111/j.1461-0248.2007.01107.x>
- Sol D, González-Lagos C, Moreira D et al (2014) Urbanisation tolerance and the loss of avian diversity. *Ecol Lett* 17:942–950. <https://doi.org/10.1111/ele.12297>
- Spear DM, Pauly GB, Kaiser K (2017) Citizen Science as a Tool for Augmenting Museum Collection Data from Urban Areas. *Front Ecol Evol* 5. <https://doi.org/10.3389/fevo.2017.00086>
- Sperandii MG, Bazzichetto M, Acosta ATR et al (2019) Multiple drivers of plant diversity in coastal dunes: A Mediterranean experience. *Sci Total Environ* 652:1435–1444
- Stas M, Aerts R, Hendrickx M et al (2020) An evaluation of species distribution models to estimate tree diversity at genus level in a heterogeneous urban-rural landscape. *Landsc Urban Plan* 198:103770. <https://doi.org/10.1016/j.landurbplan.2020.103770>
- Stat M, John J, DiBattista JD et al (2019) Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conserv Biol* 33:196–205
- Stocking KM (1955) Some taxonomic and ecological considerations of the genus *Marah* (Cucurbitaceae). *Madroño* 13:113–137
- Syfert MM, Smith MJ, Coomes DA (2013) The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS ONE* 8:e55158. <https://doi.org/10.1371/journal.pone.0055158>
- Tayyebi A, Darrel Jenerette G (2016) Increases in the climate change adaptation effectiveness and availability of vegetation across a coastal to desert climate gradient in metropolitan Los Angeles, CA, USA. *Sci Total Environ* 548–549:60–71. <https://doi.org/10.1016/j.scitotenv.2016.01.049>
- Teixeira CP, Fernandes CO (2020) Novel ecosystems: a review of the concept in non-urban and urban contexts. *Landsc Ecol* 35:23–39. <https://doi.org/10.1007/s10980-019-00934-4>
- Terres JK (1980) Audubon Society Encyclopedia of North American Birds. Distributed by Random House
- Thuiller W, Guéguen M, Renaud J et al (2019) Uncertainty in ensembles of global biodiversity scenarios. *Nat Commun* 10:1–9
- Troudet J, Grandcolas P, Blin A et al (2017) Taxonomic bias in biodiversity data and societal preferences. *Sci Rep* 7:9132. <https://doi.org/10.1038/s41598-017-09084-6>
- Uchida K, Blakey RV, Burger JR et al (2021) Urban Biodiversity and the Importance of Scale. *Trends Ecol Evol* 36:123–131. <https://doi.org/10.1016/j.tree.2020.10.011>
- Vallecillo S, Maes J, Polce C, Lavallo C (2016) A habitat quality indicator for common birds in Europe based on species distribution models. *Ecol Indic* 69:488–499. <https://doi.org/10.1016/j.ecolind.2016.05.008>
- Vendetti JE, Lee C, LaFollette P (2018) Five new records of introduced terrestrial gastropods in southern California discovered by citizen science. *Am Malacol Bull* 36:232–247
- von Takach B, Scheele BC, Moore H et al (2020) Patterns of niche contraction identify vital refuge areas for declining mammals. *Divers Distrib* 26:1467–1482. <https://doi.org/10.1111/ddi.13145>
- Warmerdam F (2008) The geospatial data abstraction library. Open source approaches in spatial data handling. Springer, pp 87–104
- Wellmann T, Lausch A, Scheuer S, Haase D (2020) Earth observation based indication for avian species distribution models using the spectral trait concept and machine learning in an

- urban setting. *Ecol Indic* 111:106029. <https://doi.org/10.1016/j.ecolind.2019.106029>
- Wenzel A, Grass I, Belavadi VV, Tschardt T (2020) How urbanization is driving pollinator diversity and pollination – A systematic review. *Biol Conserv* 241:108321. <https://doi.org/10.1016/j.biocon.2019.108321>
- Wickham H, Chang W, Wickham MH (2016) Package ‘ggplot2’ Create Elegant Data Visualisations Using the Grammar of Graphics; Version 2. 2016
- Williams K, Hobbs RJ, Hamburg SP (1987) Invasion of an annual grassland in Northern California by *Baccharis pilularis* ssp. *con-sanguinea*. *Oecologia* 72:461–465
- Yan D, Mills JG, Gellie NJC et al (2018) High-throughput eDNA monitoring of fungi to track functional recovery in ecological restoration. *Biol Conserv* 217:113–120
- Yonezawa S, Ushio M, Yamanaka H et al (2020) Environmental DNA metabarcoding reveals the presence of a small, quick-moving, nocturnal water shrew in a forest stream. *Conserv Genet* 21:1079–1084
- Young DA (1972) The reproductive biology of *Rhus integrifolia* and *Rhus ovata* (Anacardiaceae). *Evolution* (N Y) 406–414
- Zhang Y, Wu L, Ren H et al (2020) Mapping Water Quality Parameters in Urban Rivers from Hyperspectral Images Using a New Self-Adapting Selection of Multiple Artificial Neural Networks. *Remote Sens* 12:336. <https://doi.org/10.3390/rs12020336>
- Zurell D, Pollock LJ, Thuiller W (2018) Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? *Ecography* (Cop) 41:1812–1819

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11252-022-01265-0>.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com