Capitalizing on opportunistic citizen science data to monitor urban biodiversity: a multi-taxa framework

Corey T. Callaghan[1,2,*], Ian Ozeroff[3], Colleen Hitchcock[4], Mark Chandler[3]

[1]Centre for Ecosystem Science; School of Biological, Earth and Environmental Sciences; UNSW Sydney; Sydney, NSW, AU

[2]Ecology & Evolution Research Centre; School of Biological, Earth and Environmental Sciences; UNSW Sydney; Sydney, NSW, Australia

[3]Earthwatch Institute; 1380 Soldiers Field Rd, Boston, MA, USA

[4]Biology Department; Brandeis University; 415 South Street, Waltham, MA, USA

*Corresponding author:
Corey T. Callaghan
e: c.callaghan@unsw.edu.au
p: +61 421 601 388
UNSW Sydney, Sydney, NSW, Australia

NOTE: This is a pre-print, and the final published version of this manuscript can be found here: https://doi.org/10.1016/j.biocon.2020.108753

ABSTRACT

Monitoring urban biodiversity is increasingly important, given the increasing anthropogenic pressures on biodiversity in urban areas. While the cost of broad-scale monitoring by professionals may be prohibitive, citizen science (also referred to as community science) will likely play an important role in understanding biodiversity responses to urbanization into the future. Here, we present a framework that relies on broad-scale citizen science data — collected through iNaturalist — to quantify (1) species-specific responses to urbanization on a continuous scale, capitalizing on globally-available VIIRS night-time lights data; and (2) community-level measures of the urbanness of a given biological community that can be aggregated to any spatial unit relevant for policy-decisions. We demonstrate the potential utility of this framework in the Boston metropolitan region, using > 1,000 species aggregated across 87 towns throughout the region. Of the most common species, our species-specific urbanness measures highlighted the expected difference between native and non-native species. Further, our biological community-level urbanness measures — aggregated by towns — negatively correlated with enhanced vegetation indices within a town and positively correlated with the area of impervious surface within a town. We conclude by demonstrating how towns can be 'ranked' promoting a framework where towns can be compared based on whether they over- or under-perform in the urbanness of their community relative to other towns. Ultimately, biodiversity conservation in urban environments will best succeed with robust, repeatable, and interpretable measures of biodiversity responses to urbanization, and involving the broader public in the derivation and tracking of these responses will likely result in increased bioliteracy and conservation awareness.

Keywords: citizen science; community science; participatory science; community-based monitoring; urban ecology; urban tolerance; sampling biases; iNaturalist, species occurrence data

INTRODUCTION

We are currently facing the 6th mass extinction event in the Anthropocene, and biodiversity is increasingly at risk from various anthropogenic pressures (Ceballos and Ehrlich 2002). Monitoring how biodiversity responds to both threats (e.g., pollution, habitat loss, invasive species, climate change, and other anthropogenically-derived pressures) as well as interventions for enhancement (e.g., habitat restoration, green infrastructure) is essential to understand how best to preserve and manage our collective biodiversity. Biodiversity plays a key role in regulating ecosystem processes, and as acts as an ecosystem service in itself, subject to valuation (Mace et al. 2012). This, combined with the increased recognition that human well-being is positively linked with increased biodiversity highlight the necessity of monitoring changes in biodiversity (Young and Potschin 2010; Davies et al. 2019). But current funding for conservation science is failing to keep pace with the increased necessity to fully understand and monitor biodiversity change in response to varied anthropogenic pressures (Bakker et al. 2010). So, how then can we monitor biodiversity cost-effectively, with the aim of understanding how biodiversity responds to anthropogenic changes?

Broad-scale citizen science or community science projects likely provide necessary data to monitor biodiversity into the future (Bonney et al. 2009; Chandler et al. 2017; McKinley et al. 2017). Citizen science — the collaboration between members of the public, regardless of citizen status in a particular jurisdiction, with professional scientists — projects are increasingly used in natural resource management, ecology, and conservation biology (McKinley et al. 2017), and the number of such projects is simultaneously increasing (Pocock et al. 2017). For example, citizen science data have been used to increase the accuracy and specificity of threat levels of endemic birds in the Western Ghats (Ramesh et al. 2017), identify the important role temperature plays in sexual coloration in a dragonfly (Moore et al.

2019), identify new records and range extensions (Rosenberg et al. 2017), and quantify biodiversity changes in space and time (Cooper et al. 2014). These are only a select few examples. Despite the increasing prevalence of citizen science data (Pocock et al. 2017), there is still reluctance to fully adapt such data in wide-spread monitoring of biodiversity (e.g., Burgess et al. 2017). This is, in part, likely due to the biases generally associated with citizen science data (Boakes et al. 2010). Such biases include increased sampling on weekends (Courter et al. 2012), taxonomic preferences for 'charismatic' fauna and flora (Ward 2014), and generally skewed data collections to areas with large human populations (Kelling et al. 2015). This latter bias is generally problematic for any citizen science project with semi-structured or unstructured data collection (Kelling et al. 2019).

While this sampling bias towards urban areas can limit our inferences surrounding biodiversity in natural, remote regions (Callaghan et al. 2020a), it offers opportunities to better understand urban ecological and conservation questions (Cooper et al. 2007) and can complement biases ecologists have in sampling predominantly protected areas (Martin et al. 2012). Indeed, citizen science data have recently been leveraged to understand many aspects of urban ecology (e.g., Boukili et al. 2017; Li et al. 2019; Leong and Trautwein 2019). And citizen science data may provide a relatively cost-effective method to monitor biodiversity in urban areas (Callaghan et al. 2019b), including private lands which are often only accessible to private landowners (e.g., Li et al. 2019). This is critical, given the fact that urbanization is an intense anthropogenic pressure, and habitat loss and fragmentation associated with urban land transformation has generally negative impacts on biodiversity (Cincotta et al. 2000; McKinney 2006). Further, the importance of fully using citizen science data in urban areas is made clear because: (1) urban areas are where many people experience nature, and thus involving urban residents in citizen science projects can have flow-on effects for conservation

(Lepczyk et al. 2017), because people are more likely to take conservation action when they have direct experiences with nature (i.e., the pigeon paradox; Dunn et al. 2006); (2) citizen science biodiversity research provides education benefits to participants (Jordan et al. 2011) with the potential to increase bioliteracy, benefitting biodiversity inside and outside of cities (Ballard et al. 2017); (3) urban areas can act as vessels for conservation (Dearborn and Kark 2010); and (4) urban areas can even protect threatened species (Ives et al. 2016).

Given the importance of understanding urban biodiversity, and the potential for citizen science data to enhance this understanding and increase bioliteracy, the use of citizen science data needs to be validated to better understand how these data can be used in future monitoring of urban biodiversity. By increasing the bioliteracy of participants in citizen science projects a positive feedback cycle can be initiated, leading to an increase in the quality of the data (i.e., people become better at identifying and finding specific species) as the project continues. Many people have quantified the relationship between citizen science data and 'professional' data (Kosmala et al. 2016; Aceves-Bueno et al. 2017). But most comparisons have been from semi-structured citizen science datasets (e.g., eBird). Opportunistic citizen science projects (e.g., iNaturalist) likely have their own sets of biases (Brown and Williams 2018), but are showing promise in helping to deduce patterns of biodiversity across urban environments (Leong and Trautwein 2019; Li et al. 2019). The development of repeatable and robust methods that harness the power of citizen science data may not only help monitor biodiversity responses to urbanization but potentially help bridge the translation gap from science to urban planning and conservation action (Norton et al 2017).

iNaturalist is one of the most popular global biodiversity recording platforms with over 33 million observations of 250,000 species made by more than 800,000 observers. Moreover, data from iNaturalist is the second most downloaded source of data from the Global Biodiversity Information Facility. Here, we use opportunistic (i.e., generally collected in an unstructured format) iNaturalist data from the metropolitan region of Boston, USA to detect and understand patterns in biodiversity across an urban to rural gradient. Urban environments differ from natural landscapes in many ways, and efforts to understand these differences (e.g., land use, fragmentation, disturbance) often rely on land use analyses (e.g., Pearse et al. 2018; Li et al. 2019; Leong and Trautwein 2019). A global dataset of night-time lights has allowed for an approach to analyze the response of organisms to urbanization on a continuous scale, and has thus far been used to understand patterns in urban bird biodiversity at local and regional scales using eBird citizen science data (Callaghan et al. 2019a, b, 2020b). Here we look to test whether opportunistically-collected iNaturalist data can similarly help to detect patterns in biodiversity across urbanization gradients, scaling from species-specific responses to town-specific measures of the urbanness of the biological community within that town. Our approach highlights how directed efforts of sampling such as the City Nature Challenge hold potential for building both a robust dataset to understand patterns of biodiversity responses to urbanization and increase public awareness of surrounding urban biodiversity.

First, we assess the sampling biases of participants contributing opportunistic citizen science iNaturalist observations, as it pertains to a continuous gradient of urbanization –defined using night-time lights – available to sample across. We hypothesized that the degree of urbanization in a town would be positively correlated with the degree of urbanization of the observations in that town (i.e., more urban towns would have more urban observations). We then use these citizen science data to assign species-specific measures of urban tolerance, defined as the median night-time lights value of all observations for a species. From this, we produce town-specific measures of the urbanness of the collective species found therein, defined as the median of all species-specific measures of urban tolerance. We hypothesized that the relationship between the underlying degree of urbanization in a town and the cumulative town-specific urban tolerance of the species found therein would be positively correlated. We then demonstrate how these town-specific measures of urbanness can be used in an ecological context by showing the relationship between the town-specific urbanness and its ecological attributes (i.e., tree cover and enhanced vegetation index). And lastly, we provide a forward-looking approach to compare individual planning units (e.g., towns) among one another in regards to the total urbanness of their biodiversity. Ultimately, we highlight a framework that is robust and uses globally-available datasets (i.e., VIIRS night-time lights and iNaturalist citizen science data) to better understand how to fully realize the potential of citizen science data to understand urban biodiversity. Because of the ubiquity of iNaturalist data in cities and availability of a global night-time lights data set, we expect this approach can be successfully applied to increase awareness of and manage urban environments worldwide.

METHODS

Study area

We used the Boston metropolitan region (Figure S1) as a case study to demonstrate the applicability of using citizen science data to monitor the urbanization of species and communities. This region was chosen because it has been documenting urban biodiversity since 2017 as part of the City Nature Challenge (hereafter CNC; https://citynaturechallenge.org/) — an annual challenge begun in 2016 by the California Academy of Sciences and the Natural History Museum of Los Angeles. The CNC focuses on encouraging city residents to document biodiversity during a 4-day bioblitz where cities are challenged to celebrate urban biodiversity on a global scale. The Boston CNC area includes a both urban and rural habitats and starts from the city of Boston extending to the outer limits, bounded by highway 495 — a large ring road that circumnavigates the City of Boston approximately 50 kms from Boston City centre. The Boston CNC area is made up of varied habitats, including varying degrees of residential, commercial, and industrial land use, upland forests, wetlands, lakes and ponds, and coastlines (Figure S1). It offers a wide range of taxa that have been observed and submitted to iNaturalist with over 8,000 species currently observed at least once. Our analyses are restricted to the Boston CNC area and data were split into the different municipal towns within this region to aggregate observations, using the town shapefile downloaded here: https://docs.digital.mass.gov/dataset/massgis-data-community-boundaries-towns-survey-points. The resulting area includes 147 towns (or parts of towns for towns which were intersected by the Boston CNC area) that met our minimum surface area (5 km^2) requirements for analyses.

iNaturalist citizen science data

iNaturalist (www.inaturalist.org) is a multi-taxa opportunistic citizen science project hosted by the California Academy of Sciences and National Geographic Society. Participants contribute observations (e.g., photos, recordings) of any living organism through a smart-phone or web-portal with location and date assigned. Records are then tagged and identified to the lowest possible taxonomic resolution by other iNaturalist community members. iNaturalist provides a coordinate uncertainty for each observation location – which can be adjusted to obscure sensitive data. To allow for fine-grain spatial analysis, we limited our dataset to a coordinate uncertainty less than 30 meters. (For more details on the iNaturalist methodology, see here: https://www.inaturalist.org/pages/getting+started.) Those observations with sufficient community agreement on taxon identity meet the "research grade" criterion, and are regularly uploaded to the Global Biodiversity Information Facility (GBIF). We downloaded iNaturalist observations from the Global Biodiversity Information Facility for the period between 07/22/1922 (the first observation in our dataset) and 08/28/2019 for the contiguous United States (GBIF Download 2019). Accordingly, the taxonomy in our analysis follows that of GBIF (see: https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c).

iNaturalist samples across all taxa, but we restricted our analysis to species observed within the Boston CNC area at least once. Fish were removed taxonomically (*Myxini*, *Petromyzontida*, *Hyperoartia*, *Chondrichthyes*, *Actinopterygii*, or *Sarcopterygii*), and marine species were excluded through cross-referencing with the World Registry of Marine Species (WoRMS Editorial Board 2020), as there was no a priori expectation that they would be impacted by terrestrial urbanization measures (see below). Additionally, we excluded birds (*Aves)* as others have previously investigated the relationship between birds and urbanization (e.g., Callaghan et al. 2019a), because birds are highly seasonal in nature compared with other taxa, and other sources of data (e.g., eBird) would better represent bird occurrence than iNaturalist data. A full list of taxa investigated in our analyses is available in Table S1. We classified species as either native or non-native as defined by the Go Botany New England website (https://gobotany.nativeplanttrust.org/) for plants and iNaturalist for other taxa.

Species-specific urban scores

Our goal was to assign a species-specific measure of urbannesss (i.e., urban score) for each species, creating a continuum of urban tolerance across species from the most urban tolerant to the least urban tolerant species (*sensu* Callaghan et al. 2020b). These species-specific urbanness scores were first derived from a regional dataset incorporating all observations of that species throughout a larger region than the Boston CNC area. This region was constructed using the Commission for Environmental Cooperation (CEC)'s North American ecoregion designations (https://www.epa.gov/eco-research/ecoregions), and outlining the ecoregions that make up the Boston CNC area with a bounding box (Figure S2).

Using all observations for each species within that ecoregion, we calculated the underlying VIIRS night-time lights value (Elvidge et al. 2017) for every observation using Google Earth Engine (Gorelick et al. 2017). VIIRS night-lights values are available at the 500 m^2 scale. VIIRS night-time lights is a continuous proxy for urbanization, and uses a number of algorithms to exclude background noise including solar and lunar contamination, data degraded by cloud cover, and features unrelated to electric lighting such as wildfires (Elvidge et al. 2017). These night-time lights data have previously been used to track human population at many different scales (Zhang and Seto 2013). We acknowledge that although we use VIIRS night-time lights as a proxy for urbanization, species are differentially impacted by ambient light pollution (e.g., Longcore and Rich 2016), and it may be difficult to distinguish between whether species are responding to urbanization or night-time lights itself (i.e., ambient light pollution). Species respond differently to the intensity, direction, and duration of ambient light (Longcore and Rich 2016); most of which is not captured in the measurement of VIIRS night-time lights. And because intensity, direction, and duration of the night-time lights varies temporally and seasonally, by taking the mean VIIRS of many nights (and across years), we likely are producing a measure that better corresponds with urbanization at a 500 m^2 scale than it does the possible influences of ambient light pollution on specific species.

After each observation was assigned the VIIRS night-time lights value, a species was accordingly left with a continuous distribution (e.g., Figure 1). We defined a species-specific measure of urbanness as the median VIIRS value across a species' entire regional distribution of observations. Theoretically, a species with a negatively-skewed distribution would be a species which prefers and is well-adapted to urban areas, whereas a species with a positively-skewed distribution is a species which prefers non-urbanized areas, and there are many generalist distributions possible accounting for the continuum of species-specific responses to urbanization (see Callaghan et al. 2020b for details).

After the taxonomic filtering of the data, we included only species which had at least 100 regional observations to help ensure sufficient observations for a species to accurately represent its urbanness (Callaghan et al. 2019a, 2020b). We were then left with 1,004 species from the Boston CNC area with regional urban scores (Table S1). In order to test whether the regional urbanness scores were representative of species' scores within the Boston CNC area, we calculated a "local Boston urbanness" measure for the 97 species with >50 observations only using the VIIRS night-time lights values for each species within the Boston CNC area. There was a strong agreement between the regional and Boston specific approaches (Figure S3; R^2=0.64, p-value < 0.001), demonstrating that regional scores are a good representation of how biodiversity responds at a local scale (e.g., Callaghan et al. 2020b). By using the regional scores, we were accordingly able to incorporate more species into our downstream analyses.

Community-level urban scores

Using these regional species-specific urban scores, we then developed town-specific measures of how urban the community of species observed was for any given town —

subsequently referred to as the "Town Biodiversity Urbanness Index" This was done by taking the list of distinct species observed in a given town (that we had sufficient species-specific urbanness measures for) and taking the median of this distribution of urban tolerance scores (e.g., Callaghan et al. 2019b). But because many towns within the Boston CNC area have been relatively poorly sampled (Figure S4; Figure S5), we only investigated towns with a minimum of 30 observations (chosen based on exploratory analysis in the variance based on a priori local knowledge of species in the region), leaving us with a total of 87 towns used to make comparisons among. Across these 87 towns used in the analysis, the median species richness was 69 and the minimum species richness was 18.

Assessing sampling biases related to urbanization

In order to interpret our Town Biodiversity Urbanness Indices we investigated biases associated with these scores. To do so, we calculated two additional distributions specific to a given town: (1) the distribution of VIIRS night-time lights value for all observations (regardless of species observed) in a town (Figure S6) — which we call the "Opportunistic Observation Index"; and (2) the distribution of VIIRS across all underlying pixels in a town as an index of town urbanization — which we call the "Town Underlying Urbanness Index" (Figure S7). The first two were calculated by using the VIIRS values already assigned to all observations as described above, whereas the latter was done by extracting the pixels from within each town from the VIIRS night-time lights in Google Earth Engine. The gradient of Town Underlying Urbanness Index across the 87 towns used in downstream analyses ranged (examples in S7) from highly rural (26 towns had Town Underlying Urbanness Index< 2; e.g., Concord Town Underlying Urbanness Index=1.4), to urbanized (25 towns with a Town Underlying Urbanness Index>10), to highly urbanized (7 towns with a Town Underlying Urbanness Index > 20; e.g., Boston Town Underlying Urbanness Index=44). These three distributions (Town Biodiversity Urbanness Index, Town Underlying Urbanness Index, Opportunistic Observation Index ; e.g., Figure S8) for each town allowed us to draw comparisons between a town's measure of urbanness (i.e., Town Underlying Urbanness Index), where iNaturalist observations occurred (i.e., Opportunistic Observation Index), and the degree of urbanness of the species assemblage observed in that town (i.e., Town Biodiversity Urbanness Index).

First, we tested whether where people sample changes depending on the level of urbanization within a town by comparing the relationship between the observations in a town with a town's underlying urbanness index hypothesizing that as a town became more urban (i.e., higher Town Underlying Urbanness Index) the observations within that town would also become more urban (i.e., higher Opportunistic Observation Index . Second, we compared the median urbanness of all species found in a town (Town Biodiversity Urbanness Index) with the town's underlying urbanness (Town Underlying Urbanness Index), hypothesizing that as a town became more urban (i.e., higher Town Underlying Urbanness Index, the mix of species found there would comprise a greater fraction of urban tolerant species (i.e., higher Town Biodiversity Urbanness Index). These relationships were quantified using linear models for the 87 towns with > 30 observations where the respective distributions were collapsed as the median of that distribution (Figure S8), and both the predictor variables (i.e., Opportunistic Observation Index and Town Biodiversity Urbanness Index, respectively) and the response variables (Town Underlying Urbanness Index in both instances) were log-transformed.

Ecological attributes influencing the species assemblage of a town

After assessing the relationship between the species median and the underlying urbanization of a town, we demonstrated how Town Biodiversity Urbanness Index can be used to test ecological predictions using macro-ecological characteristics for each town. The characteristics we used were the percent of tree cover, mean Enhanced Vegetation Index (EVI), and mean impervious surface within a town (*sensu* Callaghan et al. 2018). We fitted a linear regression model to test the relationship between Town Biodiversity Urbanness Index values for the towns and the macro-ecological characteristics associated with each town. The response variable was log-transformed Town Biodiversity Urbanness Index and the predictor variables were tree cover, mean EVI of a town, and mean impervious surface of a town. We also included Town Underlying Urbanness Index (i.e., the median of the town's underlying pixels of VIIRS night-time lights) in the model as a covariate, and we used weights where weights were the number of observations originating from a town, providing more confidence to that town's relationship in the model-fitting procedure. Variables showed minimal multi-collinearity prior to modelling.

RESULTS

We used a total of 643,000 iNaturalist observations from the regional scale (Figure S2), and 20,292 observations from the Boston CNC area contributed by 2,085 observers (mean observations per observer: 9.7; range:1-788; sd: 40.7). A total of 2,023 species from the regional scale met our criteria, with 1,004 of these corresponding with at least 100 observations, and thus being used in our local-level analyses (Table S1). The 1,019 species not included in our analyses accounted for < 10 % of all research grade observations submitted within the Boston CNC area. Taxonomically, the 1,004 species used in analyses corresponded with a total of 9 phyla, 27 classes, 95 orders, and 280 families. Tracheophyta comprised 63% of observations, followed by Arthropoda (21%), Chordata (13%), Basidiomycota (2%), while Ascomycota, Mollusca, Mycetozoa, Annelida, and Bryophyta all comprised <1% of all observations.

The species-specific urban scores followed a log-normal distribution, with the mean urbanness being 5.07 ± 7.85 SD (Figure 2a). The three most urban species from the regional urban scores were Japanese creeper *Parthenocissus tricuspidata* (55.51), tree-of-heaven *Ailanthus altissima* (50.15), northern seaside goldenrod *Solidago sempervirens* (48.37). Conversely, the three least urban species in the regional urban scores were Canadian bunchberry *Cornus canadensis* (0.21), threeleaf goldthread *Coptis trifolia* (0.21), and frosted whiteface *Leucorrhinia frigida* (0.22). Native species dominated the species commonly observed within the Boston metropolitan region: of the 223 species with at least 20 observations, 142 were native and 81 were non-native species. While some of the non-native species found in this study in the more urbanized towns are commonly thought of as synathropes (American cockroach *Periplaneta americana*, common dandelion *Taraxacum officinale*), those species with the highest urbanness scores were lawn/yard plants (e.g., Broadleaf plantain *Plantago major*, common woodwort *Artemisia vulgaris*) or common to disturbed sites such as road sites or park entrances (e.g., garlic mustard *Allaria petiolata*, Japanese knotweed *Reynoutria japonica*, tree of heaven *Ailanthus altissima*). Several native species also had high urban scores including some common synanthropes (e.g., grey squirrel *Sciurus carolinensis*), lawn/yard taxa (e.g., fleabane *Erigeron Canadensis*) and species which exploit disturbances (e.g., American pokeweed *Phytolacca americana*). Native species also tended to be less urban tolerant than non-native species (i.e., native species' observations corresponded with lower VIIRS night-time lights values than non-native species). The mean urbanness of natives was 3.67 ± 5.36 compared with 10.9 ± 10.0 for non-native species ($t=8.923$, p-value < 0.001; Figure 2b). Whereas 78% of the 99 species with an urban score less than two were native, only 22% of the 45 species with an urban score greater than ten were native.

We found that the Opportunistic Observation Index (i.e., the median night-time lights value of all observations in a town) correlated very closely with Town Underlying Urbanness Index (i.e., the median of the VIIRS night-time lights value of underlying pixels for that town) (Figure 3). While there may be variation from town to town, as a whole there was not strong bias towards or against more natural (or disturbed) areas in towns in the Boston region (Figure 3; Figure S9; R^2=0.73, p-value <0.001). This suggests that the Boston iNaturalist community does not show a strong bias in where they sample with respect to the degree of urbanness found in a town: users are not preferentially choosing lighter or darker areas among towns to make their observations. This relationship appeared to be invariant to the number of observations in a town (Figure S10) — suggesting that the patterns observed would not change by increasing sample size. Furthermore, the Town Biodiversity Urbanness Index did not appear to move towards the town's underlying median urbanness score as the number of observations in a town increases, suggesting that simply increasing opportunistic sampling would not alter the Town Biodiversity Urbanness Index for a town (Figure S11). Towns that are more urbanized (i.e., higher Town Underlying Urbanness Index also were shown to have species with higher urbanness scores (i.e., with higher Town Biodiversity Urbanness Index) but there was significant variability in this relationship (Figure 3). For the more rural towns — with a Town Underlying Urbanness Index of 3 or less (e.g. Concord see Figure S8) — the median values for those species found had a similar degree of urbanness (i.e., Town Biodiversity Urbanness Index) to the town itself. However, as the towns became more urban — with a Town Underlying Urbanness Index above 3 (e.g., Waltham) — Town Biodiversity Urbanness Index did not track at the same pace as the increasing Town

Underlying Urbanness Index; as fewer species matched the increasing urbanness values of the towns (Figure S8).

Town Biodiversity Urbanness Indexwas negatively related to the mean EVI in a town and was positively associated with the mean impervious surface in a town (Figure 4; Table 1), and unsurprisingly was significantly related to the Town Underlying Urbanness Indexof a town (Figure 3; Table 1). Towns with more vegetation and/or trees also had an observed species assemblage that was less urban (i.e., lower Town Biodiversity Urbanness Index) and conversely towns with greater area of impervious surface had an observed species assemblage that was more urban (i.e., higher Town Biodiversity Urbanness Index; Figure 4).

We took the residuals from the relationship between Town Underlying Urbanness Indexand Town Biodiversity Urbanness Index (e.g., Figure 3), allowing each town to be ranked by the degree to which they have relatively more or fewer urban tolerant species found there (Figure 5). Towns that underperform (i.e., have relatively fewer urban species than predicted) include several coastal towns north of Boston (e.g., Newburyport, Duxbury), but also include towns that are considered more urbanized (e.g., Arlington, Salem and even Somerville — considered the most densely populated city in the United States). Conversely, towns that overperform (i.e., have more than the predicted assemblage of urbanized species recorded) included surbuban towns such as Winchester and more rural towns such as Littleton. No obvious geographic patterns emerged by mapping these towns (Figure 5), suggesting that local-level influences (i.e., habitat characteristics) likely lead to over- or under-performance of a given town.

DISCUSSION

We used data from iNaturalist — a successful citizen science project — to highlight the utility and practicality of opportunistic citizen science data to understand species and biological community-level responses to urbanization. First, the approach of assigning species-specific measures of urbanness based on underlying distributional response to VIIRS night-time lights can clearly highlight and differentiate species-specific responses to urbanization on a continuous scale (Callaghan et al. 2020b). This was clearly highlighted when considering the most abundant 223 species from the Boston CNC region, where we expectedly found the mean urbanness scores of non-native species to be more than twice that of natives. Such continuous information at the species-level is informative for understanding species' traits that predict presence in urban environments (Duncan et al. 2011; Lepczyk et al. 2017; Pearse et al. 2018; Borowy and Swan 2020), and understanding which species may deserve critical conservation attention in urban areas (Duncan et al. 2011; Lepczyk et al. 2017). Second, we were able to scale our species-specific approach to community-level metrics, quantifying the urbanness of a given community within a geopolitical region (i.e., towns). While traditional community-level measures of biodiversity (e.g., species richness, Shannon diversity) are certainly informative, a measure of the biological community's response to urbanization (i.e., the Town Biodiversity Urbanness Index) — derived from species-specific urbanness scores — can properly capture how a biological community is responding to urbanization. For example, two communities could have "equal" species richness values, but one of these communities could be dominated by synathropic species adapted to urbanization, whereas the other community could comprise less urbanized species which should be encouraged to persist in urban areas (Callaghan et al. 2019b). Of course, there will always be species that have a predisposition to persist in urban environments, whereas not all species can be expected to become 'urban species' (i.e., Moose are not expected to persist in downtown Boston). Importantly, our framework can help to understand the complex set of barriers and threats in the urban matrix by providing organismal-level responses to urbanization (e.g., native vs. non-native species), combined with local interpretation of which species could be targeted for persistence based on detailed natural history knowledge (see Figure 6). Moreover, as urban environments are managed or change – we can assess the species response to these interventions to better understand the impact of our activities on local biodiversity.

Here, we briefly highlighted the utility of our framework by correlating Town Biodiversity Urbanness Index with macro-ecological characteristics. We found that the mean EVI and percent tree cover (to a lesser extent) was, unsurprisingly, negatively correlated with Town Biodiversity Urbanness Index and the impervious surface area was positively correlated with Town Biodiversity Urbanness Index (Figure 4). Clearly, supporting green infrastructure in urban areas will have significant effects on the species that can persist there. We also showed how towns may "perform" with respect to the degree of urbanness of the species present (Figure 5) — with some towns underperforming (e.g., Marshfield; see interactive figure here) by having more urbanized species recorded than would have been predicted based on the town's underlying degree of urbanness. Town managers and community members might be able to use the relative "naturalness" of their biological community — as recorded by the public — to boost civic pride and take action to protect and build awareness of its biodiversity value. Conversely, towns which overperform by having fewer urban species found there than would be predicted by the underlying degree of urbanness for that town (e.g., Winchester) could be motivated to protect or enhance the remaining green areas and reduce threats to limit more natural species.

Several approaches have emerged to address the need to understand how biodiversity responds to urbanization. These approaches include comparisons across urban to rural gradients (e.g., McKinney 2006), comparisons across a series of networked patches such as lawns or parks (Zipperer 2002; Rega-Brdolsky et al. 2015; Locke et al. 2018) or hierarchical landscape units (Breuste et al. 2008; Li et al. 2019), and taxonomic comparisons across cities (e.g., Duncan et al. 2011; Pearse et al. 2018; Borowy and Swan 2020). Like much of New England, the urban region surrounding Boston is losing open greenspace — with active scenario-planning about how best to protect greenspace (Kittredge et al. 2015; Foster et al 2017; Ricci et al. 2020). Most of this effort in New England revolves around promoting forest or greenspace conservation using traditional metrics such as extent of protected land, habitat connectivity, and presence of rare and endangered species, among others (Kittredge et al. 2015; Foster et al. 2017; Ricci et al. 2020), without much integration of a fuller description of the response of biodiversity to those greenspaces. The development of a biodiversity urbanness index such as that proposed here can complement the existing habitat and rarity indices and help to inform conservation planning frameworks and bridge the acknowledged gap between the information derived by scientists and practitioners who plan and manage the environment (e.g., Norton et al 2016).

Developing pro-conservation attitudes by many small land-owners is critical in building the needed social capital to avoid loss of important natural habitats (Kittredge et al. 2015). The choice of monitoring tools and who is engaged in the process will not only influence the data collected, but also the uptake of the outcomes by policy makers as well as community members who will ultimately decide the fate of those lands. Making sure the research is accessible and relevant is essential to its uptake by planners and the broader community (Theobald et al. 2000; Norton et al. 2016). Theobald et al. (2000) articulates this well:

"Probably the most important of these is the idea that ecological data and analysis must be understood by those who will be affected by the decisions. In other words, citizens participating in planning processes 'will not support what they do not understand and cannot understand that in which they are not involved' (FEMAT 1993, II-80)." By increasing the bioliteracy of participants in iNaturalist – and other citizen science projects – it may be possible for changes in actions and attitudes towards urban biodiversity, and conservation more generally. As Heberlein (2012) highlights, norms will be necessary to influence behavioral change to overcome environmental problems, and by encouraging citizen science data collection and collaboration, iNaturalist may be able to enact positive behavioral change for conservation; although understanding the relation between actions and attitude in citizen science will require greater social science research (Sandbrook et al. 2013).

A collaborative approach between the participants of a citizen science project, project managers, and conservation and/or restoration projects will help to maximize the value of increasingly popular citizen science data (Figure 6). Local project managers (e.g., City Nature Challenge) can use our framework to encourage 'best practice' sampling of urban biodiversity. For instance, participants could be encouraged to reflect on where they are sampling (more or less urban), what they are sampling (i.e. the "urbanness" of the species they are observing), or by encouraging 'competition' among event organizers or towns to identify which places have the least urban community of plants and animals (e.g., Figure 5 and Figure 6). We demonstrated that there is currently a strong relationship between the underlying urbanization value of a town and the observations submitted from that town (Figure 3). But there is clearly variation in this, and this can likely be spurred on by individual efforts, where participants are encouraged to sample "where the wild things are". For example, participants may be encouraged to sample deeper into parks, fields, or local forests rather than simply at the parking lot or alongside roads. Our framework is also easily adaptable to other cities throughout the world, given the prevalence of iNaturalist data and the growing contributions to the project. We used local towns as grouping factors throughout the Boston CNC area, but these could be grouped by grids, different geo-political boundaries, or through spatial clustering approaches to better understand biological community-level urbanness responses within cities. We have used towns as a way that both the participants and municipal government agencies perceive their activities. Towns, municipalities, and other policy bodies could use these species-specific scores to help identify species which can be targeted for restoration and conservation projects (i.e., by targeting species with low urban scores; Callaghan et al. 2019b). Lastly, and importantly, these same policy bodies can then use the citizen science data to track how their restoration targets are performing in a positive feedback loop (Figure 6).

While we have demonstrated the power of broad-scale citizen science data, these data are not without their flaws and biases. We used strict filters to remove species and observations from potential inclusion (e.g., only included species with <30 m accuracy in their observation, removed marine species) in order to ensure we minimized the possibility of mismatch between a species' location and its measure of VIIRS night-time lights. This leaves many missing data from our current framework and future work should further investigate the complex trade-offs in quantity versus quality of data. These missing data include species that were excluded based on our criteria but more importantly many undetected species that have yet to be submitted to iNaturalist or that are hard to verify using photo identification (e.g., grasses, sedges, flies, ants). In our approach, we assumed that these data are missing at random with respect to the urbanness of a given species (Nakagawa and Freckleton 2008). That is, it is equally likely for a nonurban species and an urban species to be missing. We assume this because there are many taxonomic biases within citizen science data (Wei et al. 2016) that are likely driving the missingness of species, and even within a specific taxonomic group there are likely biases which influence the likelihood a species being detected, submitted, and identified in iNaturalist. For example, some charismatic species may be over-represented, or common species could be less frequently reported because many people use iNaturalist to learn identifications and once a species is known, a user may be less likely to submit records of that species. Another critique of this approach might be that most data from iNaturalist included in our analysis has been sampled "conveniently" (Anderson 2001); the observations are generally collected at a time and place convenient for the observer to record that observation (e.g., by their house, at a parking lot, along a trail). We might, for example, expect that it would be less convenient to sample in rural areas because there are fewer trails or less access there. However, at a town-level we did not find that the location of observations (i.e., Opportunistic Observation Index) deviated significantly from what was available (i.e., Town Underlying Urbanness Index) to observers; people sampled in urban locations in proportion to that which was available. We suspect the bias of convenience sampling might become more problematic when comparing regions that have different ease of access. And more convenient locations may also lead to more easily-detected species being submitted to iNaturalist. Future work should look to test how our approach interacts with missing data, and understand the biases in behaviour patterns that may influence the urbanness of species submitted to iNaturalist, likely by relying on simulations. For example, do participants show preferences for less urban species compared to more urban species? Importantly, our examples here illustrate only one way that these community-level scores could be used to understand biodiversity responses to urbanization. We envision these scores being updated regularly, given the near real-time nature of many citizen science projects, including iNaturalist (Callaghan et al. 2019c), and as these data continue to increase in quantity and quality, so too will the applicability of our proposed framework.

Conclusions

We demonstrated a framework that uses citizen science data to understand patterns of biodiversity at the town level — the relevant socio-economic unit that make policy-decisions about local investment, including zoning and building ordinances and restrictions. It remains to be tested whether planners or managers at the town or regional level will take-up a more integrated measure of the response of biodiversity to urbanness such as Town Biodiversity Urbanness Index, but it may provide a simple index to understand and communicate how a town compares to others in terms of the nature found there. Importantly, people's experience with nature will increasingly come from cities, with potential benefits for human well-being and biodiversity conservation both within and outside of cities (Soga and Gaston 2016; Prévot et al. 2018). Citizen science offers one mechanism in which we can better understand biodiversity responses to urbanization, encourage people to interact with the nature within their cities (Cooper et al. 2007; Li et al. 2019), and simultaneously increase scientific and environmental literacy (Ballard et al. 2017). Ultimately, citizen science data are dynamic: hundreds to thousands of observations are submitted every day. For our study area only, for example, there is clearly an exponential increase of observations through time (Figure S12). Collectively, we need to maximize the effectiveness of citizen science data in conservation, ecology, and natural resource management (McKinley et al. 2017), ensuring that the immense quantities of data being submitted to citizen science projects are appropriately used to inform biodiversity conservation.

ACKNOWLEDGEMENTS

We thank the countless contributors to citizen science projects who contribute both observations and community validation of species identifications to iNaturalist and the global scientific community to better understand urban biodiversity. Photos used in Figures 1, 2, and 6 were sourced from iNaturalist and were CC-BY. Contributions are as follows: Spring Peeper (Cullen Hanks); Eastern Skunk Cabbage (Clare Dellwo Cole); Common Milkweed (beautifulcataya); Common Eastern Bumble Bee (Patrick Randall); Japanese Knotweed (pottercountyconservationdistrict); Eastern Gray Squirrel (Sandy Wolkenberg); Canada Mayflower (Timothy Valentine); Common Garter Snake (Mark Kluge).

REFERENCES

Aceves-Bueno, E., Adeleye, A.S., Feraud, M., Huang, Y., Tao, M., Yang, Y., Anderson, S.E., 2017. The Accuracy of Citizen Science Data: A Quantitative Review. The Bulletin of the Ecological Society of America 98, 278–290. doi:10.1002/bes2.1336

Anderson, D.R., 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29, 1294–1297.

Bakker, V.J., Baum, J.K., Brodie, J.F., Salomon, A.K., Dickson, B.G., Gibbs, H.K., Jensen, O.P., Mcintyre, P.B., 2010. The changing landscape of conservation science funding in the United States. Conservation Letters 3, 435–444. doi:10.1111/j.1755-263x.2010.00125.x

Ballard, H.L., Dixon, C.G., Harris, E.M., 2017. Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. Biological Conservation 208, 65–75. doi:10.1016/j.biocon.2016.05.024

Boakes, E.H., Mcgowan, P.J.K., Fuller, R.A., Chang-Qing, D., Clark, N.E., O'connor, K., Mace, G.M., 2010. Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. PLoS Biology 8. doi:10.1371/journal.pbio.1000385

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J., 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience 59, 977–984. doi:10.1525/bio.2009.59.11.9

Borowy, D., Swan, C.M., 2020. A Multi-Trait Comparison of an Urban Plant Species Pool Reveals the Importance of Intraspecific Trait Variation and Its Influence on Distinct Functional Responses to Soil Quality. Frontiers in Ecology and Evolution 8. doi:10.3389/fevo.2020.00068

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

Breuste, J., Niemelä, J., Snep, R.P.H., 2008. Applying landscape ecological principles in urban environments. Landscape Ecology 23, 1139–1142. doi:10.1007/s10980-008-9273-0

Brown, E.D., Williams, B.K., 2018. The potential for citizen science to produce reliable and useful information in ecology. Conservation Biology 33, 561–569. doi:10.1111/cobi.13223

Burgess, H., Debey, L., Froehlich, H., Schmidt, N., Theobald, E., Ettinger, A., Hillerislambers, J., Tewksbury, J., Parrish, J., 2017. The science of citizen science: Exploring barriers to use as a primary research tool. Biological Conservation 208, 113–120. doi:10.1016/j.biocon.2016.05.014

Callaghan, C.T., Roberts, J.D., Poore, A.G.B., Alford, R.A., Cogger, H., Rowley, J.J.L., 2020. Citizen science data accurately predicts expert-derived species richness at a continental scale when sampling thresholds are met. Biodiversity and Conservation 29, 1323–1337. doi:10.1007/s10531-020-01937-3

Callaghan, C.T., Major, R.E., Wilshire, J.H., Martin, J.M., Kingsford, R.T., Cornwell, W.K., 2019. Generalists are the most urban-tolerant of birds: a phylogenetically controlled analysis of ecological and life history traits using a novel continuous measure of bird responses to urbanization. Oikos 128, 845–858. doi:10.1111/oik.06158

Callaghan, C.T., Poore, A.G.B., Major, R.E., Rowley, J.J.L., Cornwell, W.K., 2019. Optimizing future biodiversity sampling by citizen scientists. Proceedings of the Royal Society B: Biological Sciences 286, 20191487. doi:10.1098/rspb.2019.1487

Callaghan, C.T., Major, R.E., Lyons, M.B., Martin, J.M., Wilshire, J.H., Kingsford, R.T., Cornwell, W.K., 2019. Using citizen science data to define and track restoration targets in urban areas. Journal of Applied Ecology. doi:10.1111/1365-2664.13421

Callaghan, C.T., Major, R.E., Cornwell, W.K., Poore, A.G.B., Wilshire, J.H., Lyons, M.B., 2020. A continental measure of urbanness predicts avian response to local urbanization. Ecography 43, 528–538. doi:10.1111/ecog.04863

Callaghan, C.T., Major, R.E., Lyons, M.B., Martin, J.M., Kingsford, R.T., 2018. The effects of local and landscape habitat attributes on bird diversity in urban greenspaces. Ecosphere 9. doi:10.1002/ecs2.2347

Ceballos, G., 2002. Mammal Population Losses and the Extinction Crisis. Science 296, 904–907. doi:10.1126/science.1069349

Chandler, M., See, L., Copas, K., Bonde, A.M., López, B.C., Danielsen, F., Legind, J.K., Masinde, S., Miller-Rushing, A.J., Newman, G., Rosemartin, A., Turak, E., 2017. Contribution of citizen science towards international biodiversity monitoring. Biological Conservation 213, 280–294. doi:10.1016/j.biocon.2016.09.004

Cincotta, R.P., Wisnewski, J., Engelman, R. 2000. Human population in the biodiversity hotspots. Nature 404, 990-992.

Cooper, C.B., Dickinson, J., Phillips, T., Bonney, R., 2007. Citizen Science as a Tool for Conservation in Residential Ecosystems. Ecology and Society 12. doi:10.5751/es-02197-120211

Cooper, C.B., Shirk, J., Zuckerberg, B., 2014. The invisible prevalence of citizen science in global research: Migratory birds and climate change. PLoS One 9, e106508. doi.10.1371/journal.pone.0106508

Courter, J.R., Johnson, R.J., Stuyck, C.M., Lang, B.A., Kaiser, E.W., 2012. Weekend bias in Citizen Science data reporting: implications for phenology studies. International Journal of Biometeorology 57, 715–720. doi:10.1007/s00484-012-0598-7

Davies, Z.G., Dallimer, M., Fisher, J.C., Fuller, R.A., 2019. Biodiversity and Health: Implications for Conservation. Biodiversity and Health in the Face of Climate Change 283–294. doi:10.1007/978-3-030-02318-8_12

Dearborn, D.C., Kark, S., 2010. Motivations for Conserving Urban Biodiversity. Conservation Biology 24, 432–440. doi:10.1111/j.1523-1739.2009.01328.x

Duncan, R.P., Clemants, S.E., Corlett, R.T., Hahs, A.K., Mccarthy, M.A., Mcdonnell, M.J., Schwartz, M.W., Thompson, K., Vesk, P.A., Williams, N.S.G., 2011. Plant traits and extinction in urban areas: a meta-analysis of 11 cities. Global Ecology and Biogeography 20, 509–519. doi:10.1111/j.1466-8238.2010.00633.x

Dunn, R.R., Gavin, M.C., Sanchez, M.C., Solomon, J.N., 2006. The pigeon paradox: dependence of global conservation on urban nature. Conservation Biology 20, 1814-1816. doi: 10.1111/j.1523-1739.2006.00533.x

Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. International Journal of Remote Sensing 38, 5860–5879. doi:10.1080/01431161.2017.1342050

Foster, D., Lamber, K.F., Kittredge, D., Donahue, B., Hart, C., Labich, W., Meyer, S., Thompson, J., Buchanan, M., Levitt, J., Perschel, R., Ross, K., Elkins, G., Daigle, C., Hall, B., Faison, E., D'Amato, A., Roman, R., Del Tredici, P., Irland, L., et al. 2017. Wildlands and Woodlands Farmlands and Communities Broadening the Vision for New England. Available here: https://wildlandsandwoodlands.org/sites/default/files/Wildlands%20and%20Woodlands%202017%20Report.pdf. Accessed April 20th, 2020.

GBIF Download. 2019. 28 August 2019, GBIF Occurrence Download. https://doi.org/10.15468/dl.ggjmnk

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment 202, 18–27. doi:10.1016/j.rse.2017.06.031

Haines-Young, R., Potschin, M., n.d. The links between biodiversity, ecosystem services and human well-being. Ecosystem Ecology 110–139. doi:10.1017/cbo9780511750458.007

Heberlein, T.A., 2012. Navigating Environmental Attitudes. New York, NY: Oxford University Press. 228 pp. ISBN: 978-0-19-977333-6

Ives, C.D., Lentini, P.E., Threlfall, C.G., Ikin, K., Shanahan, D.F., Garrard, G.E., Bekessy, S.A., Fuller, R.A., Mumaw, L., Rayner, L., Rowe, R., Valentine, L.E., Kendal, D., 2015. Cities are hotspots for threatened species. Global Ecology and Biogeography 25, 117–126. doi:10.1111/geb.12404

Jordan, R.C., Gray, S.A., Howe, D.V., Brooks, W.R., Ehrenfeld, J.G., 2011. Knowledge Gain and Behavioral Change in Citizen-Science Programs. Conservation Biology 25, 1148–1154. doi:10.1111/j.1523-1739.2011.01745.x

Kelling, S., Fink, D., Sorte, F.A.L., Johnston, A., Bruns, N.E., Hochachka, W.M., 2015. Taking a ‘Big Data’ approach to data quality in a citizen science project. Ambio 44, 601–611. doi:10.1007/s13280-015-0710-4

Kelling, S., Johnston, A., Bonn, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., Fernandez, M., Hochachka, W.M., Julliard, R., Kraemer, R., Guralnick, R., 2019. Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity. BioScience 69, 170–179. doi:10.1093/biosci/biz010

Kittredge, D.B., Gianotti, A.G.S., Hutyra, L.R., Foster, D.R., Getson, J.M., 2015. Landowner conservation awareness across rural-to-urban gradients in Massachusetts. Biological Conservation 184, 79–89. doi:10.1016/j.biocon.2015.01.001

Kosmala, M., Wiggins, A., Swanson, A., Simmons, B., 2016. Assessing data quality in citizen science. Frontiers in Ecology and the Environment 14, 551–560. doi:10.1002/fee.1436

Leong, M., Trautwein, M.D., 2019. A citizen science approach to evaluating US cities for biotic homogenization. doi:10.7287/peerj.preprints.27472v1

Lepczyk, C.A., Aronson, M.F.J., Evans, K.L., Goddard, M.A., Lerman, S.B., Macivor, J.S., 2017. Biodiversity in the City: Fundamental Questions for Understanding the Ecology of Urban Green Spaces for Biodiversity Conservation. BioScience 67, 799–807. doi:10.1093/biosci/bix079

Li, E., Parker, S.S., Pauly, G.B., Randall, J.M., Brown, B.V., Cohen, B.S., 2019. An Urban Biodiversity Assessment Framework That Combines an Urban Habitat Classification Scheme and Citizen Science Data. Frontiers in Ecology and Evolution 7. doi:10.3389/fevo.2019.00277

Locke, D.H., Avolio, M., Trammell, T.L.E., Chowdhury, R.R., Grove, J.M., Rogan, J., Martin, D.G., Bettez, N., Cavender-Bares, J., Groffman, P.M., Hall, S.J., Heffernan, J.B., Hobbie, S.E., Larson, K.L., Morse, J.L., Neil, C., Ogden, L.A., O'Neil-Dunne, J.P.M., Pataki, D., Pearse, W.D., Polsky, C., Wheeler, M.M., 2018. A multi-city comparison of front and backyard differences in plant species diversity and nitrogen cycling in residential landscapes. Landscape and Urban Planning 178, 102–111. doi:10.1016/j.landurbplan.2018.05.030

Longcore, T., Rich, C. 2016. Ecological and organismic effects of light pollution. eLS, 1-8. doi:10.1002/9780470015902.a0026328

Mace, G.M., Norris, K., Fitter, A.H. 2012. Biodiversity and ecosystem services: a multilayered relationship. Trends in Ecology & Evolution 27, 19-26. doi:10.1016/j.tree.2011.08.006

Martin, L.J., Blossey, B., Ellis, E., 2012. Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. Frontiers in Ecology and the Environment 10, 195–201. doi:10.1890/110154

Mckinley, D.C., Miller-Rushing, A.J., Ballard, H.L., Bonney, R., Brown, H., Cook-Patton, S.C., Evans, D.M., French, R.A., Parrish, J.K., Phillips, T.B., Ryan, S.F., Shanley, L.A., Shirk, J.L., Stepenuck, K.F., Weltzin, J.F., Wiggins, A., Boyle, O.D., Briggs, R.D., Chapin, S.F., Hewitt, D.A., Preuss, P.W., Soukup, M.A., 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. Biological Conservation 208, 15–28. doi:10.1016/j.biocon.2016.05.015

Mckinney, M.L., 2006. Urbanization as a major cause of biotic homogenization. Biological Conservation 127, 247–260. doi:10.1016/j.biocon.2005.09.005

Nakagawa, S., Freckleton, R.P., 2008. Missing inaction: the dangers of ignoring missing data. Trends in Ecology & Evolution 23, 592–596. doi:10.1016/j.tree.2008.06.014

Norton, B.A., Evans, K.L., Warren, P.H., 2016. Urban Biodiversity and Landscape Ecology: Patterns, Processes and Planning. Current Landscape Ecology Reports 1, 178–192. doi:10.1007/s40823-016-0018-5

Pearse, W.D., Cavender-Bares, J., Hobbie, S.E., Avolio, M.L., Bettez, N., Chowdhury, R.R., Darling, L.E., Groffman, P.M., Grove, J.M., Hall, S.J., Heffernan, J.B., Learned, J., Neill, C., Nelson, K.C., Pataki, D.E., Ruddell, B.L., Steele, M.K., Trammell, T.L.E., 2018. Homogenization of plant diversity, composition, and structure in North American urban yards. Ecosphere 9. doi:10.1002/ecs2.2105

Pocock, M.J.O., Tweddle, J.C., Savage, J., Robinson, L.D., Roy, H.E., 2017. The diversity and evolution of ecological and environmental citizen science. Plos One 12. doi:10.1371/journal.pone.0172579

Prévot, A.-C., Cheval, H., Raymond, R., Cosquer, A., 2018. Routine experiences of nature in cities can increase personal commitment toward biodiversity conservation. Biological Conservation 226, 1–8. doi:10.1016/j.biocon.2018.07.008

Ramesh, V., Gopalakrishna, T., Barve, S., Melnick, D.J., 2017. IUCN greatly underestimates threat levels of endemic birds in the Western Ghats. Biological Conservation 210, 205–221. doi:10.1016/j.biocon.2017.03.019

Rega, C.C., Nilon, C.H., Warren, P.S., 2015. Avian Abundance Patterns in Relation to the Distribution of Small Urban Greenspaces. Journal of Urban Planning and Development 141. doi:10.1061/(asce)up.1943-5444.0000279

Ricci, E.H., Collins, J., Clarke, J., Dolci, P., de la Parra, L., 2020. Losing Ground. Massachusetts Mass Audubon. Available here: https://www.massaudubon.org/content/download/41477/1007612/file/Losing-Ground-VI_2020_final.pdf. Accessed April 20th, 2020.

Rosenberg, K.V., Blancher, P.J., Stanton, J.C., Arvind, P., 2017. Use of North American Breeding Bird Survey data in avian conservation assessments. The Condor 119, 594–606.

Sandbrook, C., Adams, W.M., Büscher, B, Vira, B., 2013. Social research and biodiversity conservation. Conservation Biology 27, 1487-1490.

Soga, M., Gaston, K.J., 2016. Extinction of experience: the loss of human-nature interactions. Frontiers in Ecology and the Environment 14, 94–101. doi:10.1002/fee.1225

Theobald, D.M., Hobbs, N.T., Bearly, T., Zack, J.A., Shenk, T., Riebsame W.E., 2000. Incorporating biological information in local land-use decision making: designing a system for conservation planning. Landscape Ecology 15, 35–45.

Ward, D.F., 2014. Understanding sampling and taxonomic biases recorded by citizen scientists. Journal of Insect Conservation 18, 753–756. doi:10.1007/s10841-014-9676-y

Wei, J.W., Lee, B.P.Y-H., Wen, L.B., 2016. Citizen science and the urban ecology of birds and butterflies – A systematic review. PLoS One 11, e0156425. doi:10.1371/journal.pone.0156425

WoRMS Editorial Board. 2020. World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2020-03-10. doi:10.14284/170

Zhang, Q., Seto, K., 2013. Can Night-Time Light Data Identify Typologies of Urbanization? A Global Assessment of Successes and Failures. Remote Sensing 5, 3476–3494. doi:10.3390/rs5073476

Zipperer, W.C., 2002. Species composition and structure of regenerated and remnant forest patches within an urban landscape. Urban Ecosystems 6, 271–290. doi:10.1023/b:ueco.0000004827.12561.d4

FIGURES

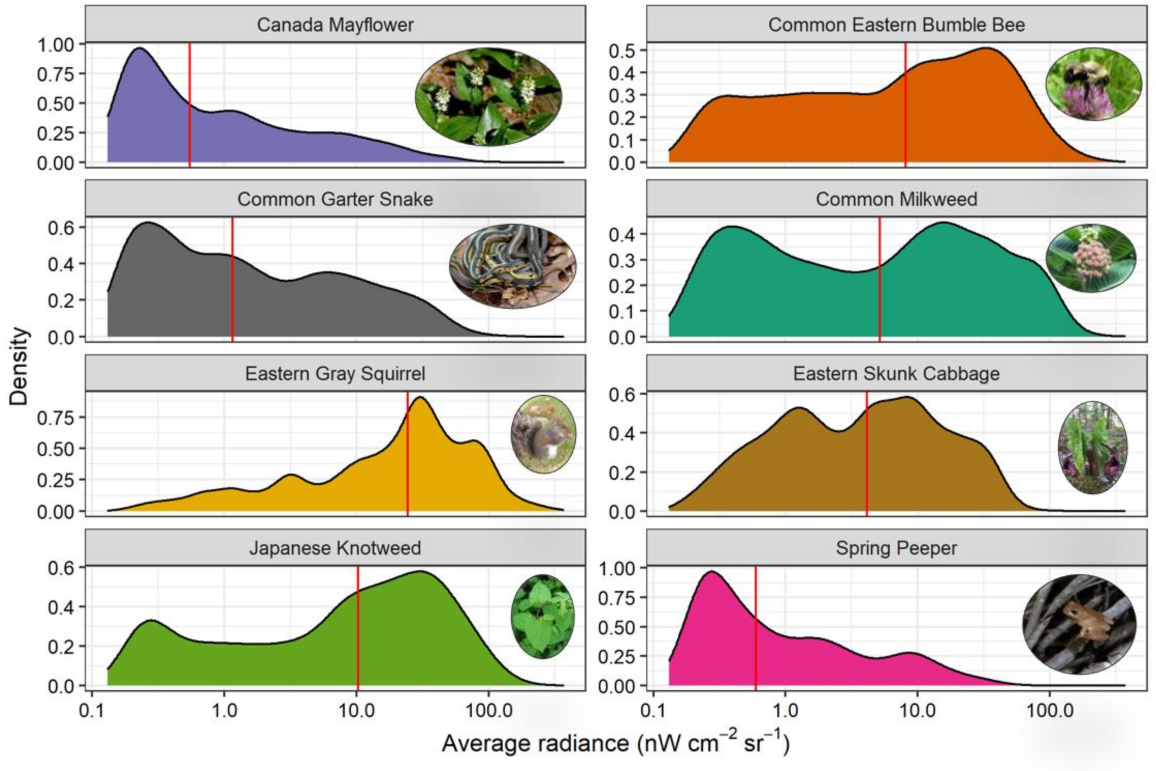


Figure 1. Eight example species — chosen based on their prevalence in the Boston CNC area — and their distributional response to VIIRS night-time lights (on a log-scale), showing an example of the differences among species. The red line represents the median. This was repeated for every species with >100 observations in the continental region (Fig. S2).

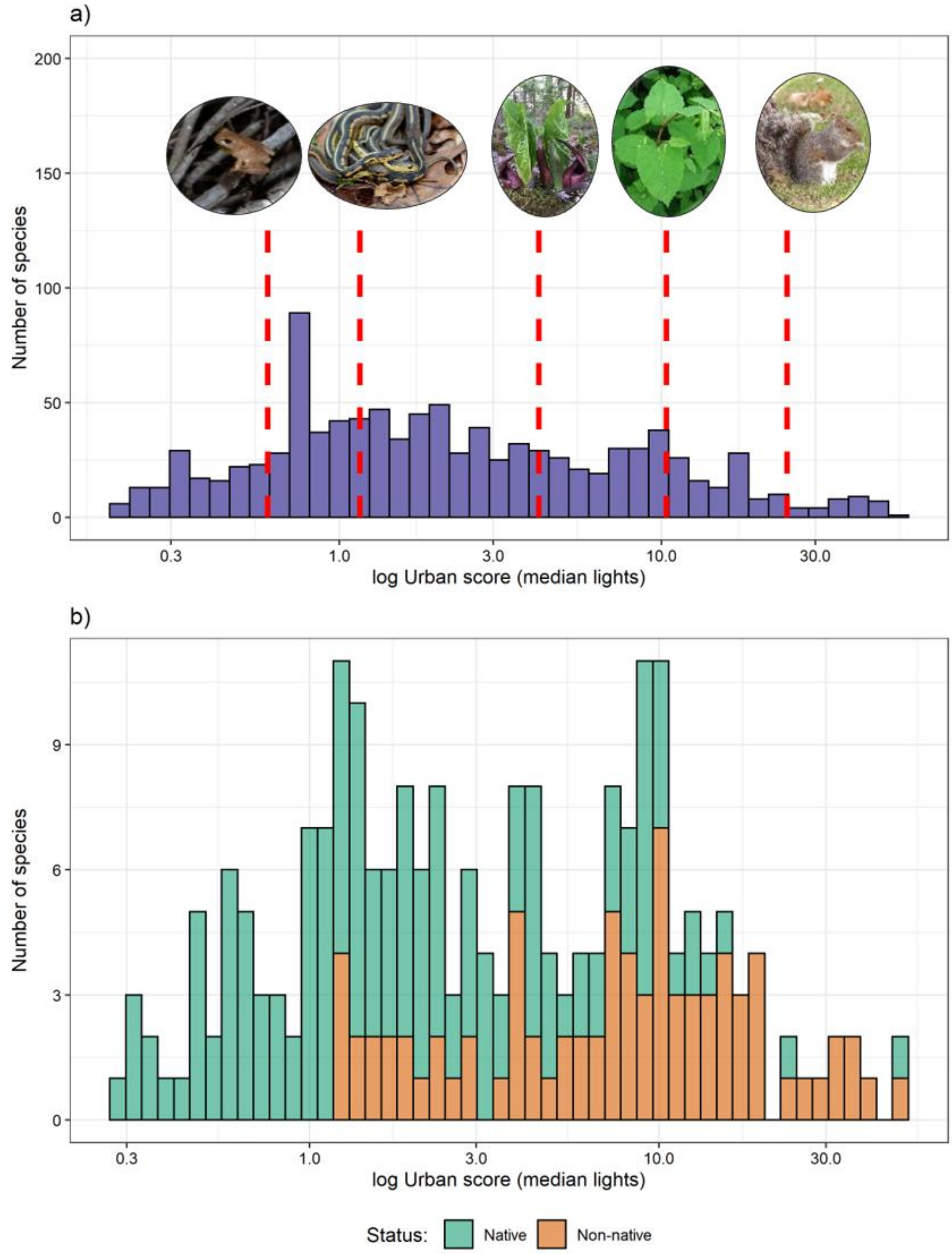


Figure 2. a) The species-specific regional urban scores for 1,004 species found in the Boston region; the distribution follows a log-normal distribution with some species being very urban (e.g., Eastern Gray Squirrel) and others being less urban (e.g., Common Eastern Bumble Bee), compared with the majority of species which are distributed between. The y-axis represents the number of species which fall into the specific bin corresponding with the x-axis. Five example species, chosen based on their prevalence in the Boston CNC area are displayed. b) The 223 species with > 20 observations in the Boston CNC area and their species-specific regional urban scores (as in a) stratified to their status as native or non-native.

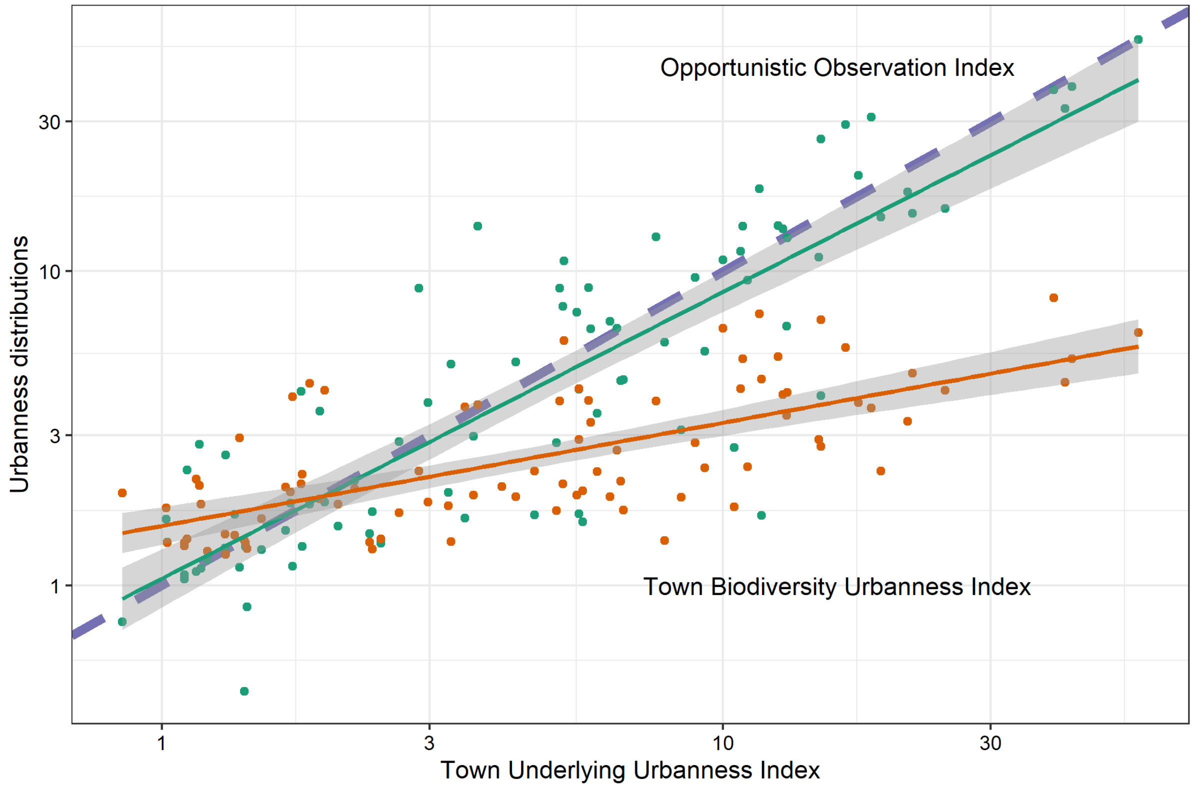
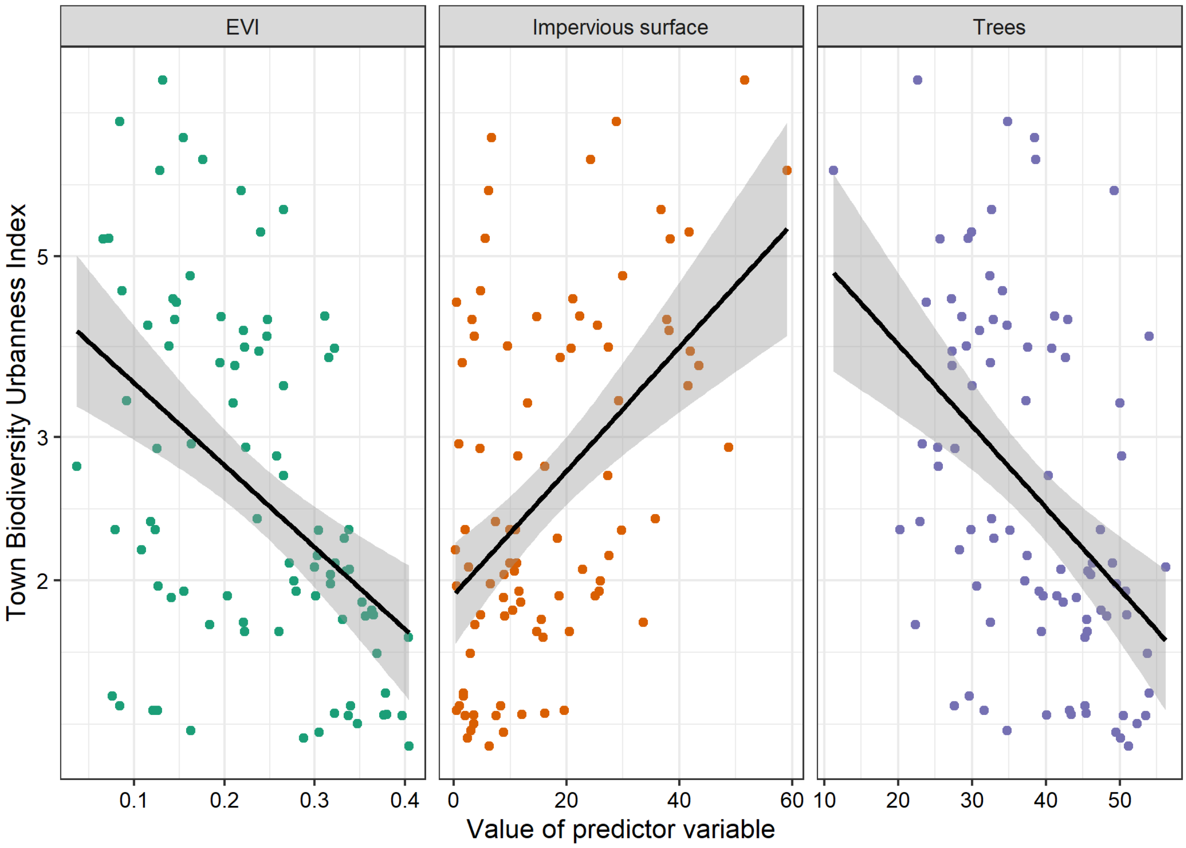


Figure 3. Relationship between the log-transformed Town Underlying Urbanness Index (x-axis) and both the log-transformed Opportunistic Observation Index and the Town Biodiversity Urbanness Index (y-axis). Blue is a one-to-one line. And linear regressions are shown for each variable. The residuals between the Town Underlying Urbanness Index and Town Biodiversity Urbanness Index were extracted for further analyses.



822
823 **Figure 4**. The relationship between three macroecological variables (EVI=Enhanced
824 Vegetation Index) extracted from each town (N=87) where there were at least 30 iNaturalist
825 observations and the log-transformed Town Biodiversity Urbanness Index for each town (i.e.,
826 community-level urbanness).
827

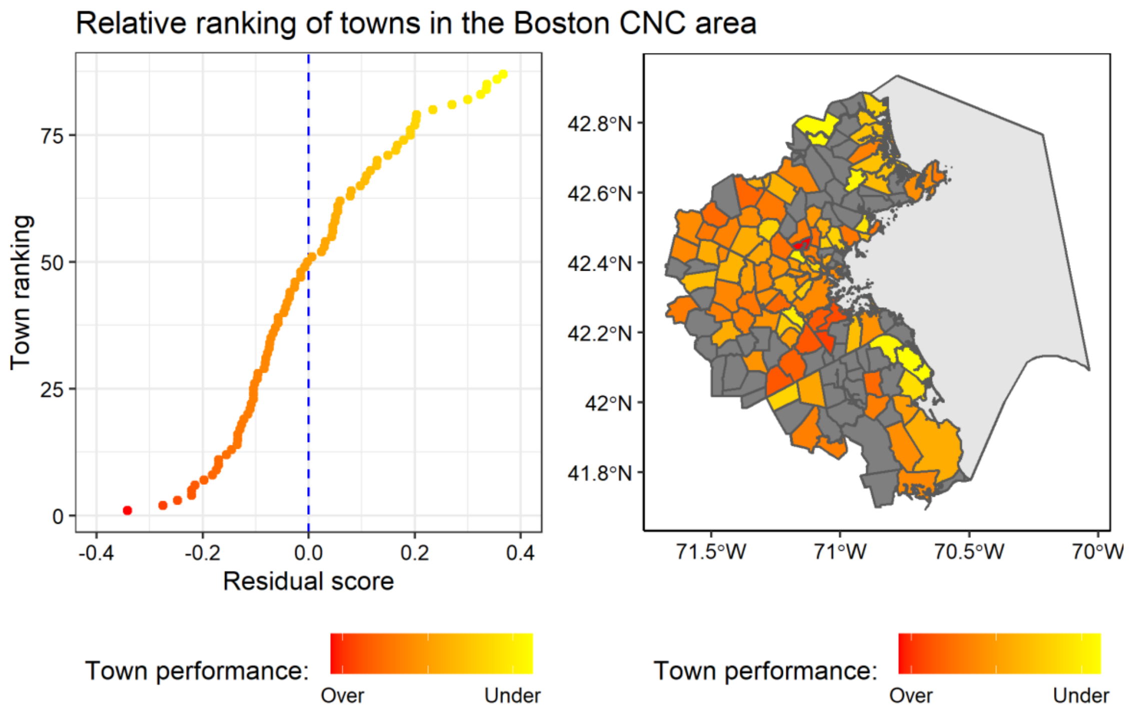


Figure 5. The residuals between Town Biodiversity Urbanness Index and Town Underlying Urbanness Index extracted and plotted based on the relative ranking of over- and under-performance for each of the 87 towns considered for analyses. Over-performing towns are towns that have less urban tolerant species than would be expected based on their Town Underlying Urbanness Index, and vice versa for underperforming towns. These residuals are plotted based on the ranking (left) and spatially (right). An interactive version of the left-hand panel is [here](#).

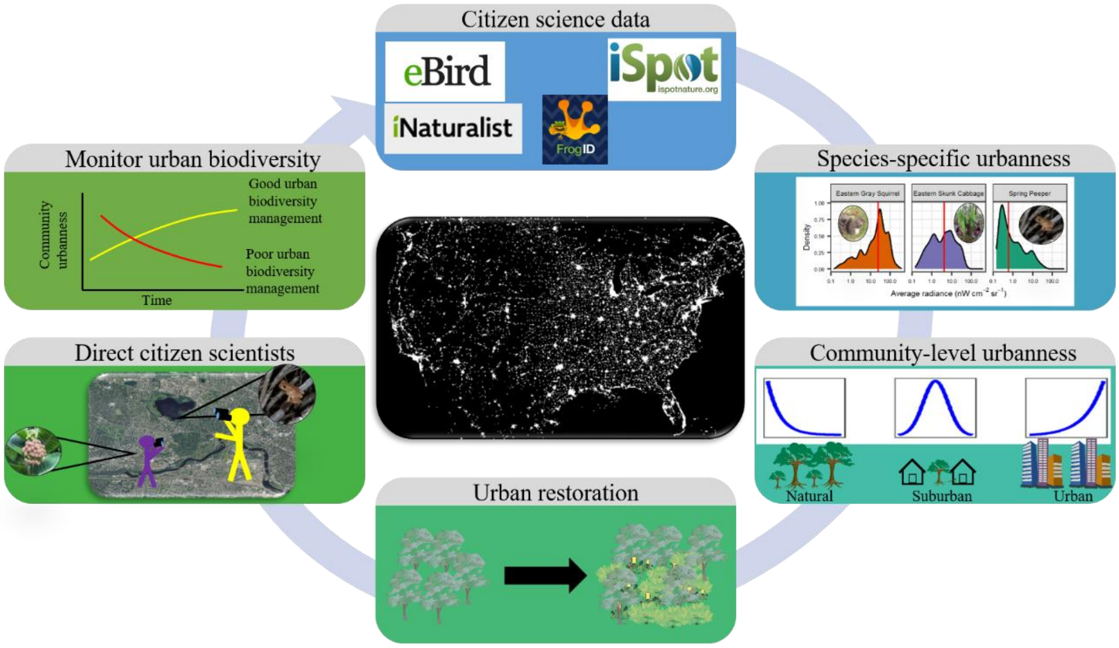


Figure 6. The theoretical positive-feedback loop that can be implemented through our proposed framework. Species-specific urbanness can be derived from citizen science data, and then community-level urbanness values can be derived across multiple taxa. These provide baseline data for future urban restoration projects, and local citizen science project managers can direct participants to sample meaningfully to help monitor urban biodiversity through citizen science projects.

TABLES

Table 1. Results of a multiple linear regression model where the town-specific community urbanness measure (i.e., Town Biodiversity Urbanness Index) was the response variable, log-transformed. The urbanness of a town (Town Underlying Urbanness Index was included as a covariate as this was correlated with the Town Biodiversity Urbanness Index (Figure 3). Significant variables are in bold.

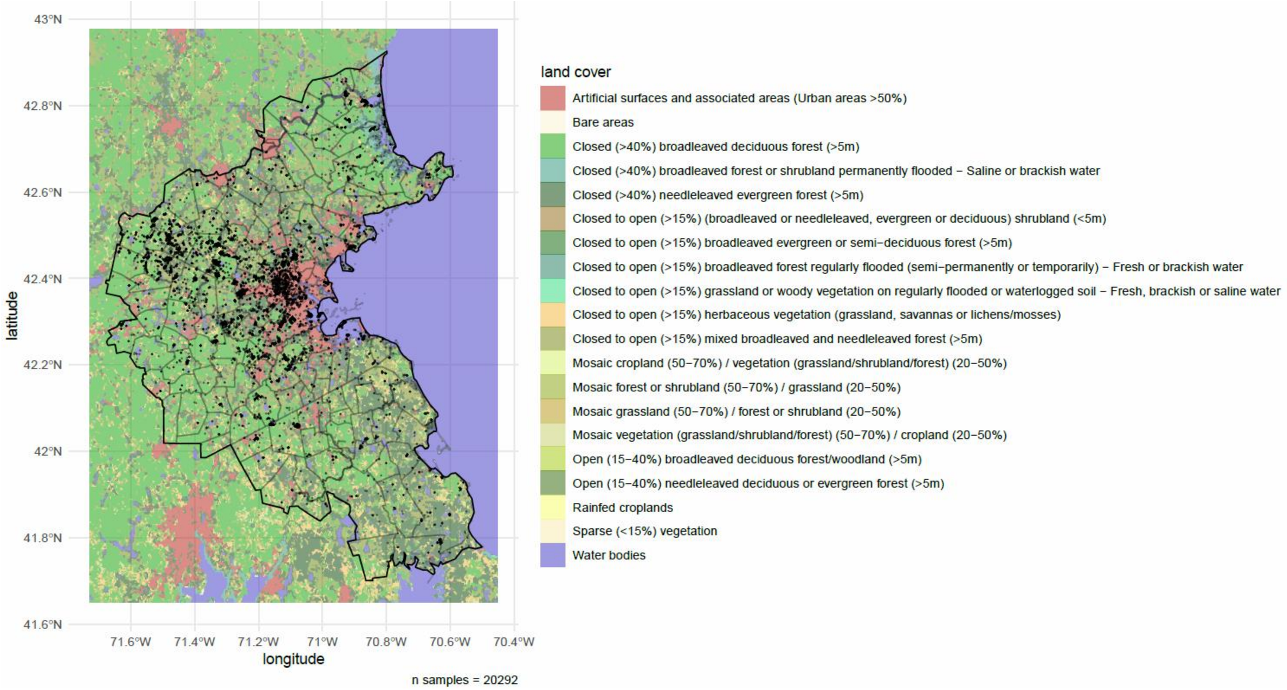
Term	Estimate	Standard error	t-value	p-value
Intercept	**0.443**	**0.087**	**5.109**	**<0.001**
Town urbanness	**0.005**	**0.001**	**3.553**	**<0.001**
Trees	0.000	0.002	0.005	0.996
EVI	**-0.612**	**0.221**	**-2.763**	**0.007**
Impervious surface	**0.004**	**0.001**	**3.159**	**0.002**

Supporting information for:

Callaghan et al. Capitalizing on opportunistic citizen science data to monitor urban biodiversity: A multi-taxa framework. Biological Conservation.
https://doi.org/10.1016/j.biocon.2020.108753

SUPPLEMENTARY FIGURES

A)



B)

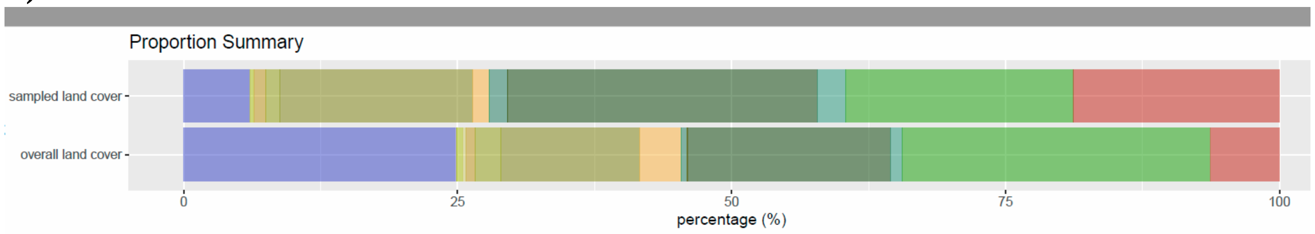


Figure S1. A) The Boston area, and underlying landcover categories derived from GlobCover, as well as B) the proportion of GlobCover categories and the proportion of sampled categories.

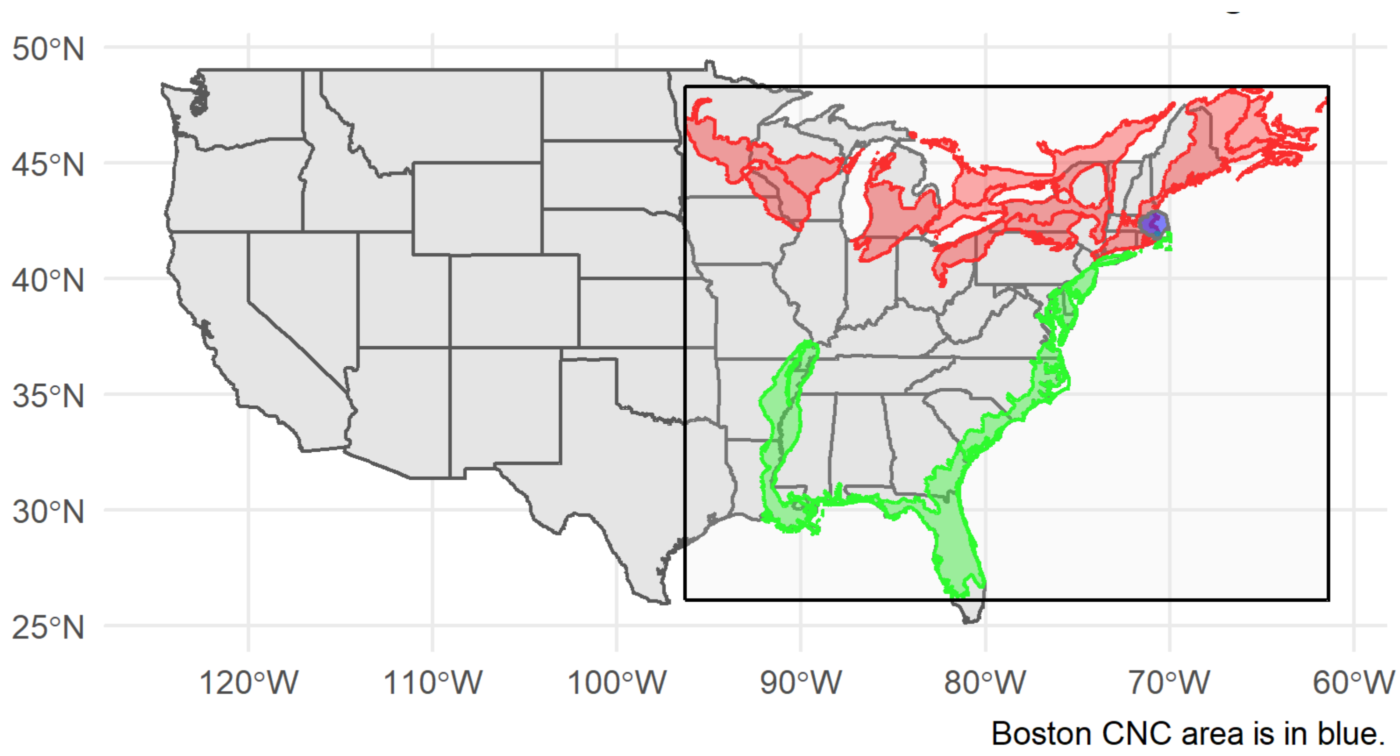


Figure S2. The Boston CNC area (blue region) and the CEC Level II Ecoregions used to derive species-specific urbanness scores. In order to query GBIF we bound these regions with a bounding box surrounding these regions. This approach worked to include as many species as possible in the analyses, and only species with at least 100 regional observations were included in further analyses. There was strong correlation between the regional urbanness and the local-scale urbanness (i.e., Boston CNC area) demonstrating the robustness of this approach (see Figure S3 for more details).

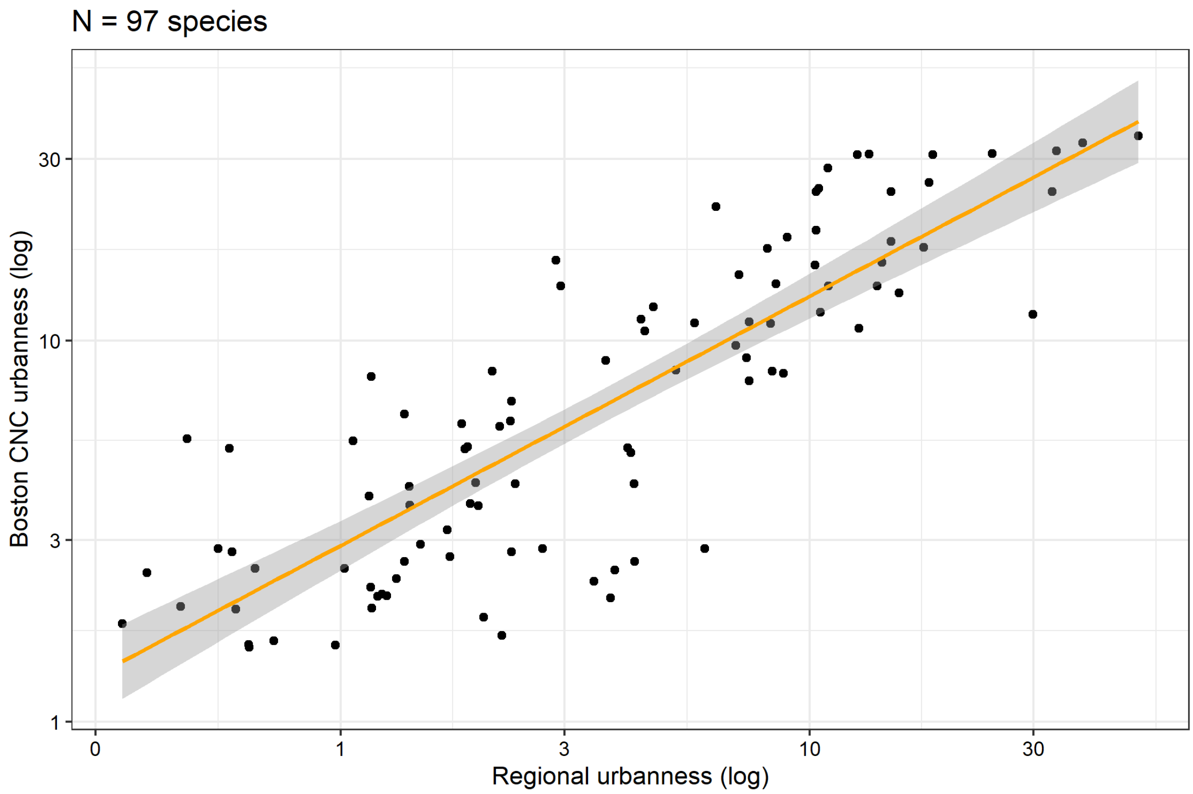


Figure S3. The relationship between the regional urban scores for 97 species, and the urban scores calculated from observations only within the Boston CNC area, both log-transformed. Only species with at least 50 observations from the CNC area were considered to assess this relationship. The positive relationship highlights that regional scores predict local-level scores and thus serve as a good representation of how species respond to urbanization at the local scale.

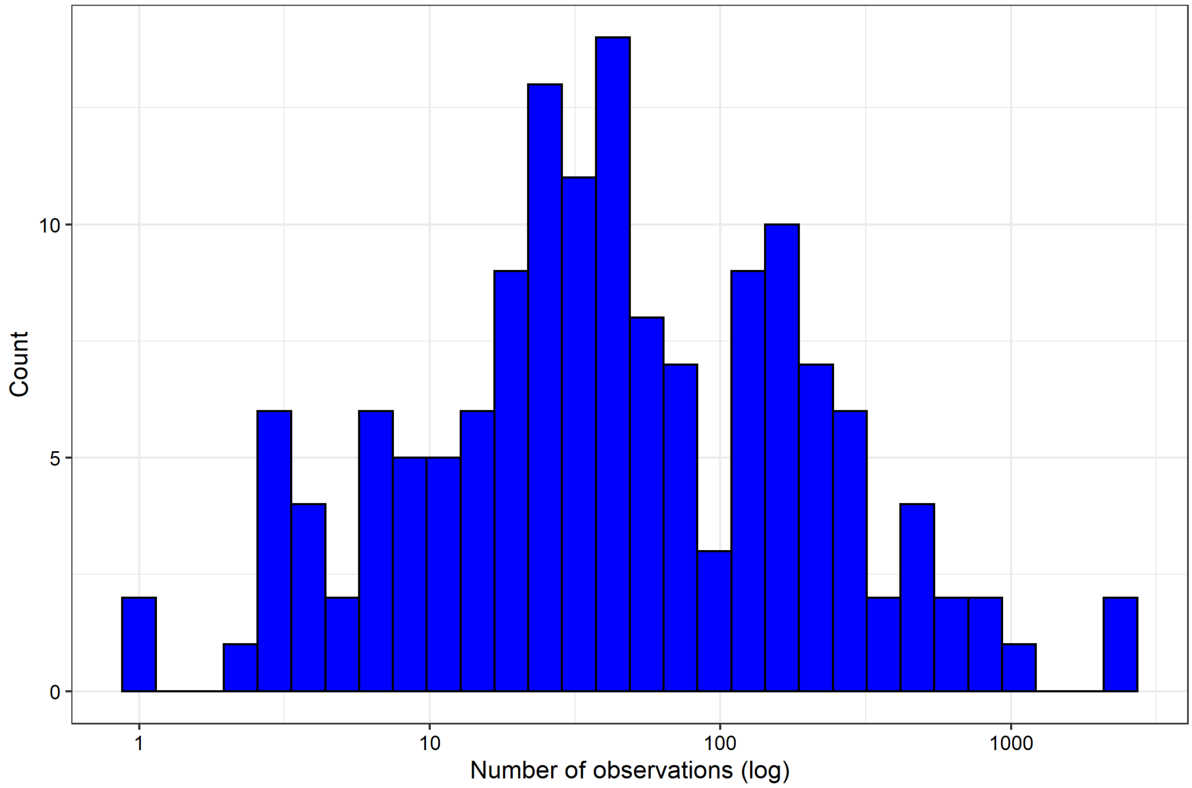


Figure S4. The number of observations from each of the potential 147 towns in the Boston CNC area. We only included towns which had at least 30 iNaturalist observations in analyses (N=87) based on a priori local knowledge of the variation of biodiversity that exists among towns.

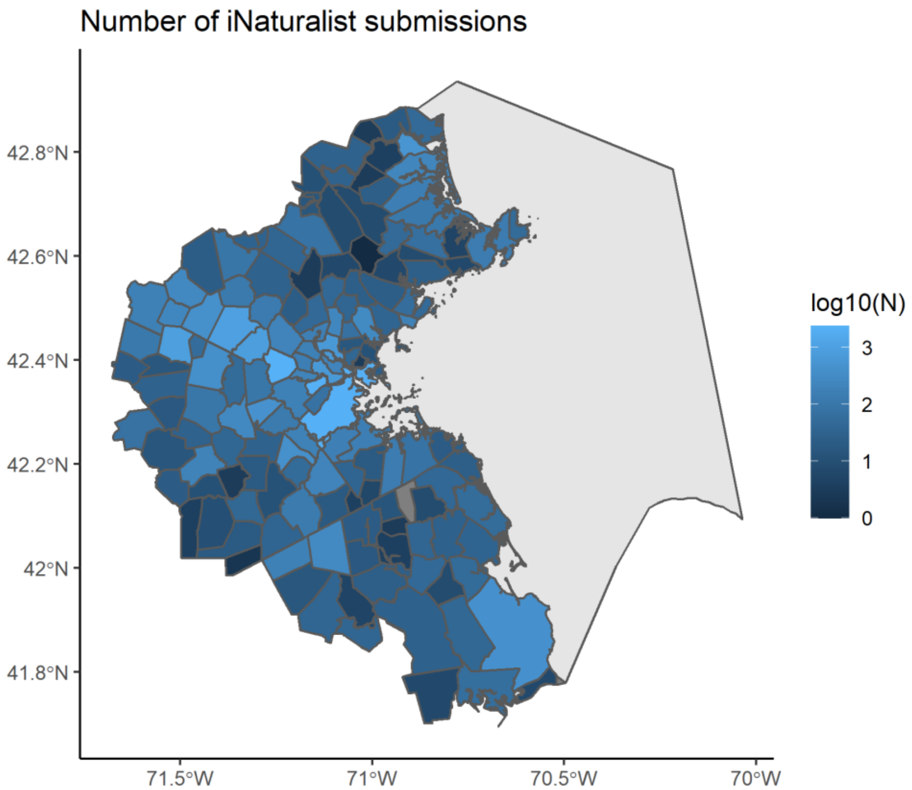


Figure S5. The spatial representation of the towns in the Boston CNC and the number of iNaturalist submissions on a logarithmic scale (e.g., Figure S4). There were clear biases in the number of observations from a town, and so we only included towns with at least 30 observations in further analyses.

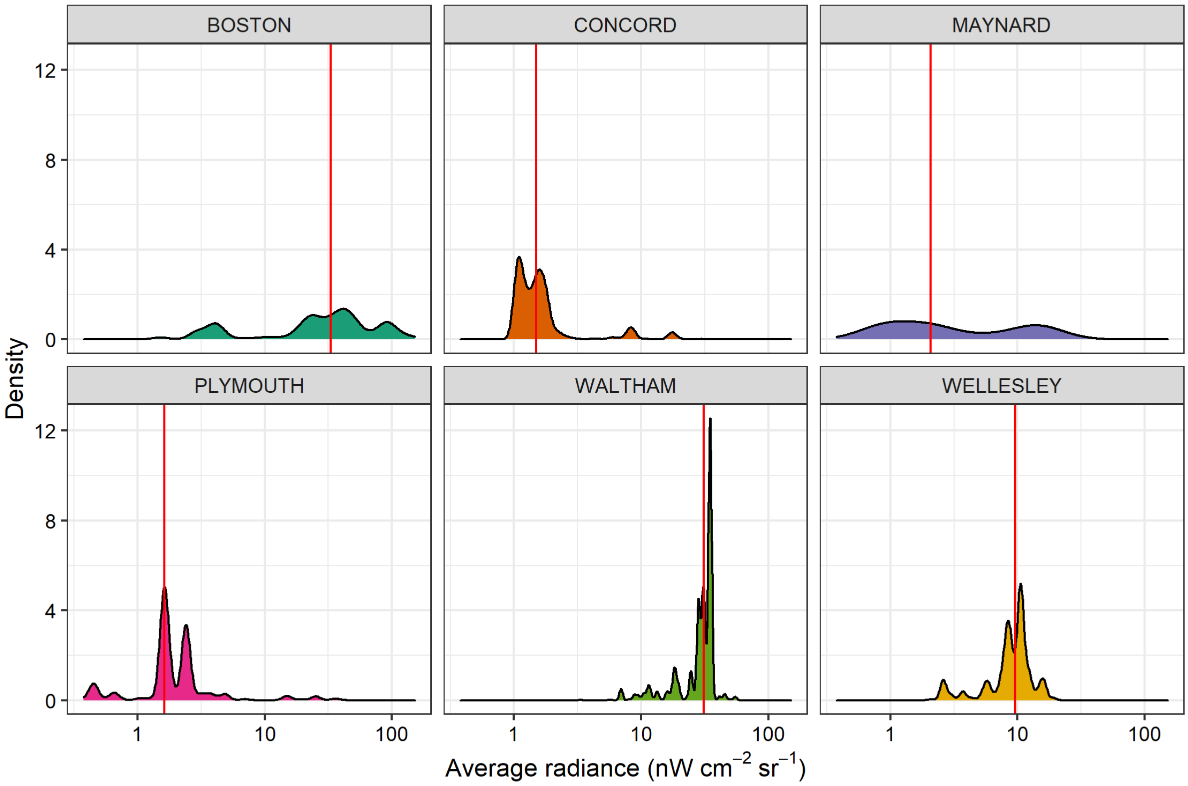


Figure S6. The distribution of observations (i.e., the Opportunistic Observation Index) in six example towns showing the differences in how towns sample, where the red line represents the median. Note the axes are the same.

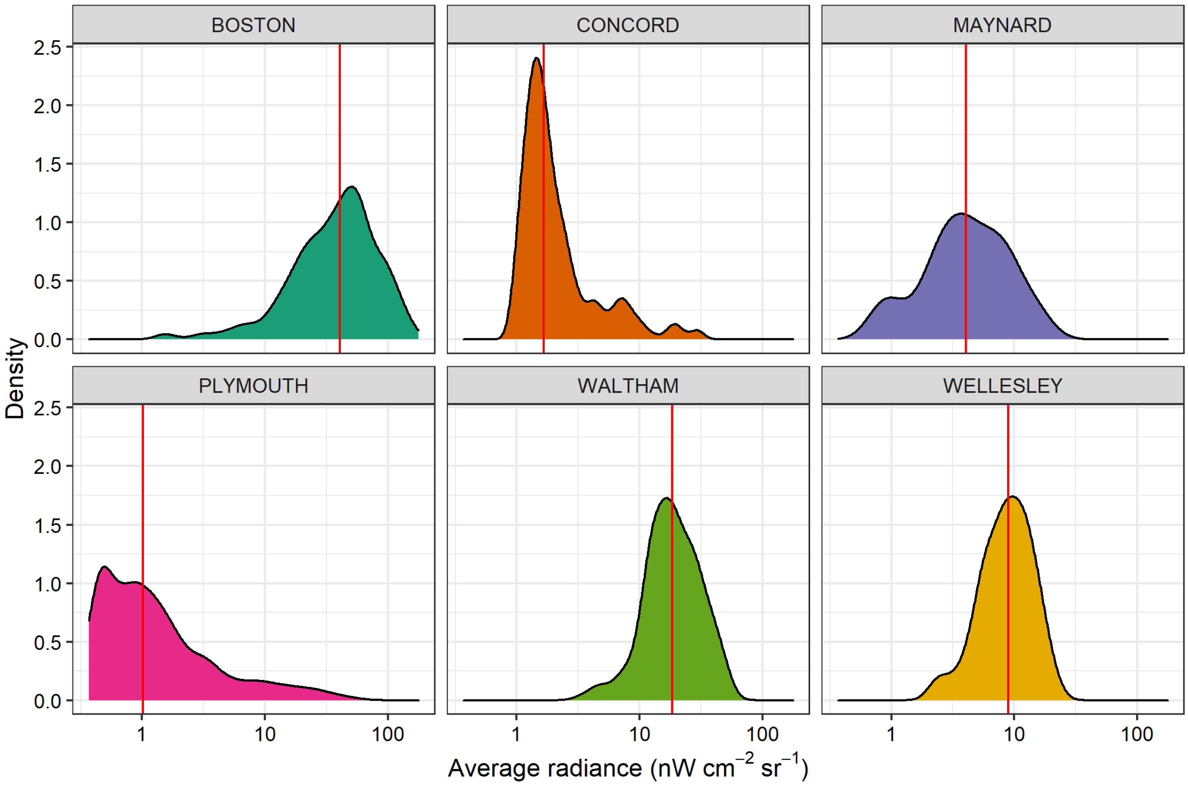


Figure S7. Six example towns (as above) and their underlying pixels of VIIRS night-time lights values (i.e., the Town Underlying Urbanness Index), showing the difference in how ‘urban’ a town is, based on a suite of variables (e.g., built up cover, natural area), where the red line represents the median.

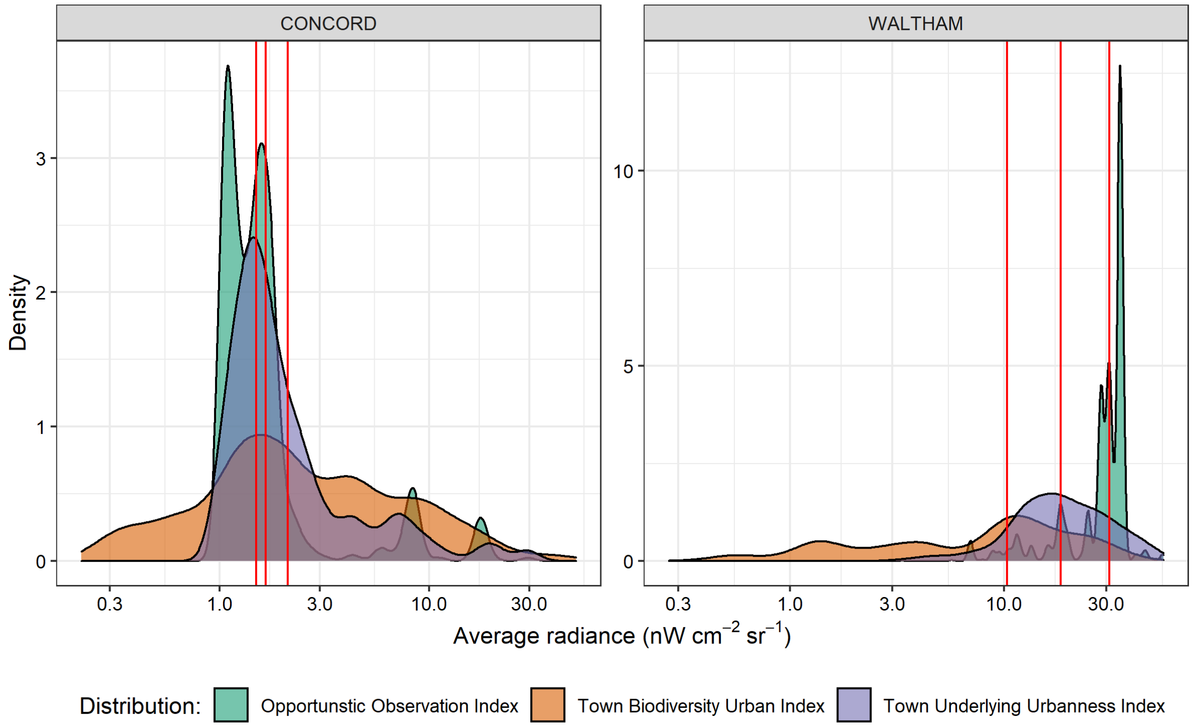


Figure S8. Two example towns (Concord and Waltham) showing the three distributions used and compared among one another in our analyses.

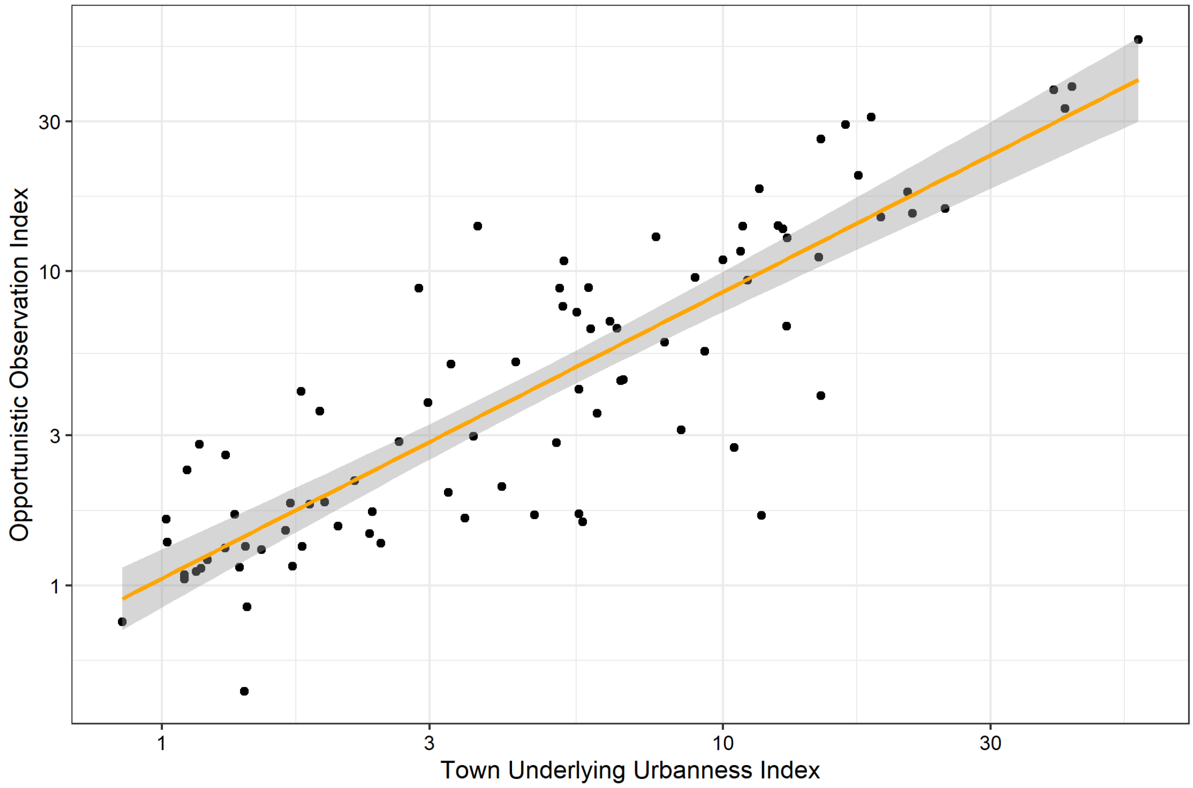


Figure S9. The relationship between the underlying urbanness of a town (i.e., TUUI) and the observations submitted from a town (i.e., OOI). As expected, the more ‘urban’ a town is, the more ‘urban’ their observations are. Citizen scientists are generally sampling around the median of the urbanness in a town. Only towns with at least 30 observations are included in this figure (N=87 towns).

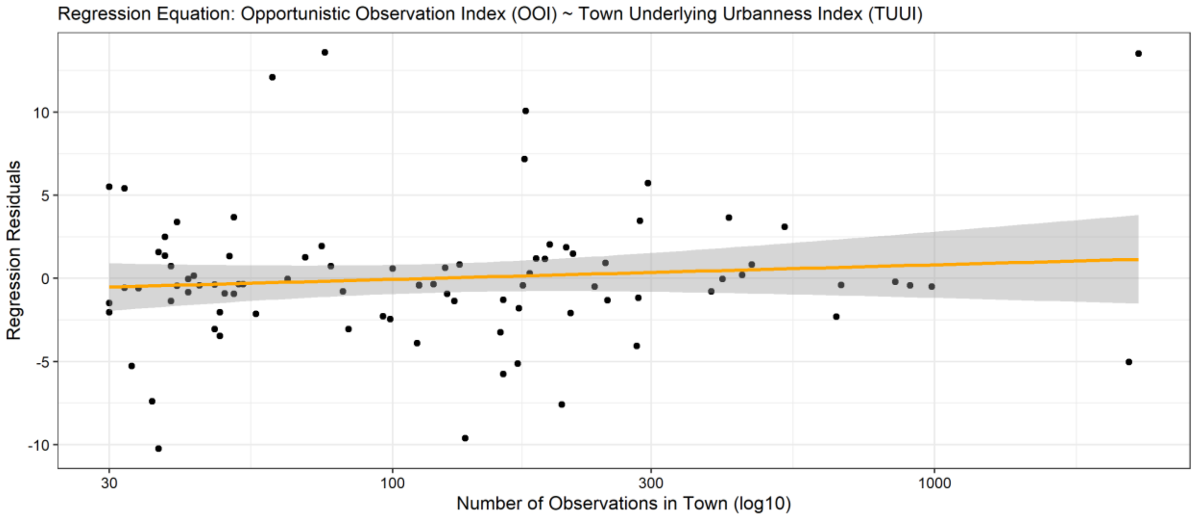


Figure S10. The x-axis shows the number of iNaturalist observations in the town (logarithmic scale), and on the y-axis the residual value of a simple linear regression between Median Town Observation Pixel Value and the Town Underlying Median Pixel Value (i.e., Figure S9). There is no significant relationship, suggesting that the distribution of observations across a town's urban gradient is not impacted by the increase in the volume of observations. Only towns with >=30 observations are included in this figure (N=87 towns).

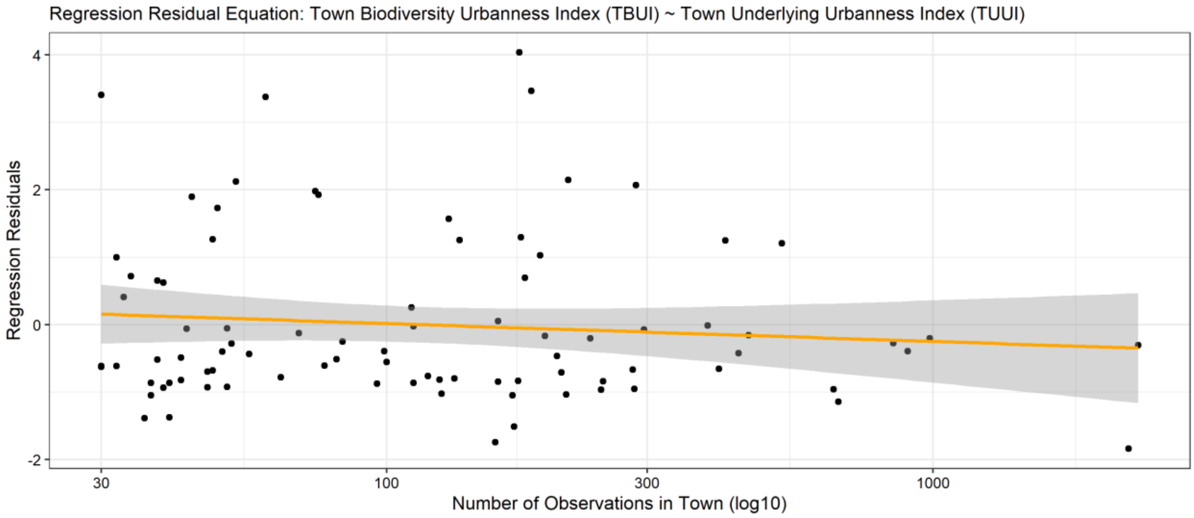


Figure S11. Parallel to above, on the x-axis we have the number of observations in the town (log10), and on the y-axis the residual value of a simple linear regression between Median Species Urbanness of a town (species counted distinctly), and the Town Underlying Median Pixel Value. There is no relationship here, indicating the urbanness of the species observed in a town does not approach the town urbanness median as observations increase.

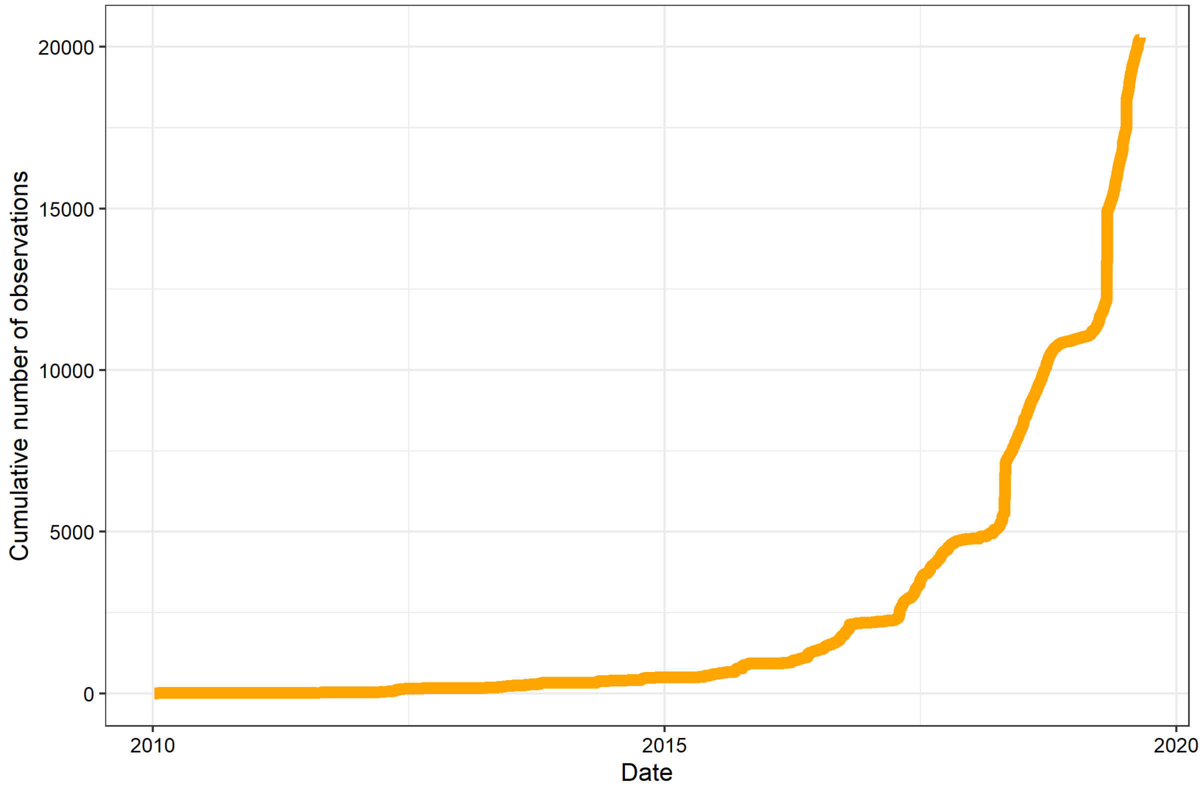


Figure S12. The cumulative observations from the Boston CNC area since 2010, submitted to iNaturalist.