Leveraging publicly available data to facilitate urban ecology and evolution

Shawn Arreguin[1], Joseph F. Walker[1*] Natalie L.R. Love[2*]

[1]Department of Biological Sciences, the University of Illinois at Chicago, Illinois, 60607, USA
[2]Negaunee Institute for Plant Conservation Science and Action, Chicago Botanic Garden, Glencoe, Illinois 60022 USA
Authors for correspondence: nlove@chicagobotanic.org and jfw52@uic.edu
*Denotes equal contribution
ORCID IDs:
Shawn Arreguin: https://orcid.org/0009-0003-9302-9979
Joseph F. Walker: https://orcid.org/0000-0003-2928-8899
Natalie L.R. Love: https://orcid.org/0000-0002-5013-5478

Abstract

The intensifying pace of urbanization has prompted researchers to investigate its ecological and evolutionary consequences more deeply, yet logistic and monetary challenges can impose a barrier to research. Publicly available databases offer scientists from broad backgrounds an opportunity to circumvent those barriers, enhancing participation in urban ecology and evolution. In conjunction with computational advances, these datasets allow research to be done entirely with a computer connected to the internet, which we term Digital Eco-Evo. Throughout this review, we highlight how Digital Eco-Evo approaches allow researchers to test their own hypotheses, thereby radically advancing our understanding of urban evolutionary and ecological processes that shape life in cities. By collating presently used databases, covering emerging tools and resources, and presenting biases and limitations of these data, we hope that this review will act as a guide for interested researchers and highlight the strengths and weaknesses of this approach. Opportunities to take part in this exciting field of research will continue to grow as new databases and associated tools are developed.

Keywords: Open Source; Urbanization; Databases; Inclusive Research; Digital Research

1) Introduction

The development of large scale community science databases, the emerging emphasis on open data initiatives, and advances in computational capacity have ushered in a new era of research. Increasingly, scientific inquiry can be conducted entirely with a computer connected to the internet, which we term Digital Eco-Evo. Rooted in the principles of synthesis-based research (Carpenter et al., 2009), this alternative approach to traditional field or experimental work leverages and integrates existing, publicly available data into new datasets capable of answering questions beyond that of their original intention. Digital Eco-Evo projects remove cost barriers imposed by generating new data, making them an excellent resource for both student learning and obtaining preliminary data. This facilitates broad accessibility in academic settings and allows standalone big-data projects to be conducted cost-effectively. Using novel and creative approaches, Digital Eco-Evo research has already begun to yield new and unique insights in the rapidly growing field of urban ecology and evolution (Schmidt et al., 2020; Cosentino & Gibbs 2022; Marín-Gómez et al., 2022; Estien et al., 2024).

Public repositories and databases of biological data are diverse, data-rich and importantly, freely accessible online. They span a wide range of taxonomic, ecological, and spatial scales, and include many types of data. For example, iNaturalist and eBird house species occurrence data; the Landsat program and Sentinel satellite missions provide remotely sensed data; MacroPopGen and GenBank create access to genetic information; and databases such as TRY, Amniote, and EltonTrait maintain trait data; among others. These datasets are an increasingly valuable resource for answering questions in urban ecology and evolution, especially when facing uncertain funding climates that can prevent other approaches to data collection. However, datasets are often scattered across disparate internet repositories, and researchers may need to synthesize data found across several sources to conduct a study. Thus, there is a need to summarize and collate resources as well as provide examples of existing studies to facilitate future Digital Eco-Evo research.

Publicly available community science generated data tends to be clustered in urban areas (Di Cecco et al., 2021). As a result, this data is particularly useful for conducting studies of ecological and evolutionary processes in urban landscapes. Urbanization causes widespread land cover and environmental change (Grimm et al., 2008), resulting in areas characterized by several interrelated factors, including increased temperatures (e.g., urban heat islands), high levels of pollution, fragmented habitats, and habitat loss. These changes create novel environments, inducing strong selective pressures that shape evolutionary trajectories and ecological interactions. For example, urbanization can drive rapid adaptive evolution (Cheptou et al., 2008; Nacci et al., 2010; Winchell, Campbell-Staton, et al., 2023) or shifts in phenology (D. Li et al., 2021; Merckx et al., 2021; Olsen et al., 2020), creating downstream effects on ecosystem-level processes. While ubiquitous globally, urbanization is not uniform, and its impacts on biological processes may differ spatially, temporally, and taxonomically. The pace of urbanization is rapidly expanding, with urban land area predicted to more than quadruple by 2050 from levels in 2000 (Angel et al., 2011). Given these patterns of urbanization, it is critical to assess these effects at broad spatial scales and across levels of biological organization, thereby clarifying general patterns and improving the predictability of responses to urbanization.

Traditional fieldwork in urban landscapes presents several challenges. Privately owned land, theft or vandalism of field equipment (Stadnicki et al., 2024), and high urban heterogeneity (Cadenasso et al., 2007) can impede gathering enough data to make city-wide generalizations. These factors are further exacerbated when investigating trends over multiple cities (Magle et al., 2019). Platforms, such as iNaturalist, contain biological data collected across diverse land ownership and land use classes in urban areas (Di Cecco et al., 2021). Many datasets are also available at broad spatial and temporal scales, facilitating multi-city comparisons and investigations into change through time (Lawrence et al., 2019; Shultz et al., 2020; Santini et al., 2024). When combined with freely accessible data that may act as predictor variables (e.g., climate, land cover, and sociodemographics), publicly available biological data provide a powerful platform to readily overcome many existing challenges and conduct hypothesis-driven research entirely on a computer. Digital Eco-Evo research does come with its own challenges, and it cannot act as a substitute to traditional urban field work. Hyper-local and mechanistic questions often require traditional approaches due to the coarse scale of publicly available data and its lack of experimental qualities. Additionally, these data may be non-randomly sampled leading to biases that should be considered and addressed (Boyd, Powney, et al., 2023; see Section 4)

This paper aims to provide a thorough guide for conducting Digital Eco-Evo research in an urban context. This is done by: (1) summarizing the existing data repositories currently available and highlighting the research questions that have been addressed by innovative approaches that leverage these data across multiple biological questions (Table 1, Table 2, Figure 1), (2) highlighting emerging resources and computational tools that will maximize the utility of publicly available data for Digital Eco-Evo research*,* and (3) discussing biases and limitations in these data and demonstrating how this data can be used to enhance participation in urban ecology and evolution research. By outlining both resources and example studies, this paper aims to provide a roadmap for researchers interested in addressing ecological and evolutionary dynamics in cities using entirely publicly available resources and demonstrate key findings from these studies.

2) Data sources and examples of novel studies leveraging publicly available data 99

In this section, we summarize publicly available data sources and highlight specific studies that have leveraged these resources to answer novel questions in urban landscapes, often taking advantage of urbananization gradients (Figure 1). Importantly, we focus on research that has been conducted mostly or entirely on a computer. For clarity, this section is arranged by biological research subject. This is not meant to be a comprehensive review; rather, the studies chosen highlight innovative approaches and emphasize the utility of these data for advancing the field of urban ecology and evolution. Our review focuses on studies that cover a range of biological questions and are entirely reliant on publicly available datasets. When few qualifying studies were found, we supplemented subsections with hybrid approaches.

To facilitate future Digital Eco-Evo research, we present two summary tables to assist in study design. **Table 1** summarizes research questions that have been addressed and the sources of biological data used. **Table 2** compiles commonly used non-biological data that often serve as covariates or predictor variables, such as measures of urban intensity (e.g., percent of a surface that is impervious). Together, these resources illustrate the depth and breadth of questions and data that can be synthesized for urban ecology and evolution studies.

2.1) Urban Ecosystems and Communities

2.1.1) Digital Eco-Evo insights into community assemblage and biodiversity distribution

Occurrence data (e.g., through iNaturalist, eBird, GBIF) have been used to examine how biodiversity varies along urbanization gradients. With these data, researchers can test the influence urbanization has on species assemblages within cities and identify biodiversity hotspots in urban landscapes (**Table 1**). These studies span a wide range of taxonomic clades, including birds (Callaghan, Bino, et al., 2019; Hensley et al., 2019; Kinnunen et al., 2024), insects (Fitch, Wilson, et al., 2019; Prudic et al., 2018; Ružanović & Mičetić Stanković, 2023), mammals

(Border et al., 2017; Santini et al., 2019), and plants (Ives et al., 2016; Velasquez-Camacho et al., 2024). Several studies have also compared patterns of biodiversity and community assembly across clades (Beninde et al., 2023; Estien et al., 2024; Leong & Trautwein, 2019; E. Li et al., 2019; Curti et al., 2024). This work has helped uncover the effects of urbanization on species diversity and richness, which may depend on various factors, such as local landscape structure or evolutionary history. Callaghan, Bino, et al., (2019) leveraged eBird to obtain over four million occurrence records throughout the contiguous USA, finding that bird richness and diversity were highest in urban greenspaces relative to natural green areas or medium/high developed areas. They attributed this response to increased habitat heterogeneity within urban greenspaces relative to either natural or developed areas, suggesting habitat structure influences urban bird diversity. Fitch, Wilson, et al., (2019) repurposed data from previous studies (Fitch, Glaum, et al., 2019; Jamieson et al., 2019) to find that urbanization positively influences introduced bee species, but not native species in Michigan, USA. These findings provide insights into how origin may predict species responses to urbanization.

Recent work has revealed that historical housing policies in the USA, such as redlining, the systematic exclusion of minority neighborhoods from receiving the same loans as wealthier, often white neighborhoods, continue to influence ecological and evolutionary processes in cities (Schell et al., 2020). For example, Estien et al., (2024) integrated over 120,000 iNaturalist observations, digitized HOLC (home-owners loan corporation) historical maps, and data from the National Land Cover Database (NLCD) to assess whether species richness and community assembly of birds, insects, arachnids, reptiles, and amphibians differ among historical HOLC grades that range from grade A (“best” or “greenlined” neighborhoods) to D (hazardous or redlined). Greenlined neighborhoods had higher species richness than redlined neighborhoods, indicating that the legacy of discriminatory housing practices shapes contemporary patterns of urban biodiversity. The authors hypothesized that this pattern can be attributed to spatial differences in environmental quality caused by systematic disinvestment in redlined neighborhoods. Importantly, Estien et al. (2024) addressed the biases present in observational data that occur across HOLC grades (Ellis-Soto et al. 2023), demonstrating the care that must be taken when inferring patterns using these datatypes.

2.1.2) Understanding urban ecosystem function with publicly available data

Freely available remotely-sensed satellite data (e.g., Landsat, MODIS, Sentinel, etc.) allows researchers to study landscape-scale responses of ecosystems to urbanization (Cârlan et al., 2020; Dong et al., 2019; Khikmah et al., 2024; Leisenheimer et al., 2024). Satellite programs collect imagery with high spatial resolutions (e.g., 10-30 meters) and temporal frequencies (e.g., twice monthly), covering a global extent that can span decades. For example, Landsat, a satellite program developed by NASA in the USA, has been collecting 30 meter resolution imagery at a global scale for more than 50 years (Crawford et al., 2023). These satellites are capable of capturing multispectral imagery, such as near-infrared light (NIR). NIR is reflected by healthy, photosynthetically active vegetation, and can be used to calculate the normalized difference vegetation index (NDVI) by using red reflectance (see Huang et al., 2021 for a review and recommendations for using NDVI in scientific studies) among other vegetation indices. NDVI can be used to monitor vegetative responses, including plant stress. When combined with climate data, researchers can begin to understand how altered environmental conditions, like heat or drought, influence responses of urban vegetation, especially along sociodemographic and economic gradients. Dong et al., (2019) used this metric, via MODIS data, to assess the sensitivity of urban vegetation to drought over a 20-year period in Los Angeles, California, USA. They found that vegetation was more sensitive to drought in economically disadvantaged communities, leading to less healthy vegetation and greater exposure to heatwaves under drought conditions.

Historically, analyzing remotely-sensed satellite data imposed a computational burden that limited its accessibility. With cloud-based computing such as Google Earth Engine (GEE), remote sensing workflows have undergone radical transformations. At present, GEE provides access to >80 petabytes of publicly available data for noncommercial use and affords researchers the computational power to analyze it (Cardille et al., 2024). This makes an internet connection the sole constraint for these projects. With tools such as the R package *rgee* (Aybar, 2023), researchers can integrate GEE remote sensing data and functionality into an R-based workflow via a user-friendly interface. These new tools enable multi-city comparisons, providing novel opportunities to assess whether cities differ in vegetation response to climate, drought, or urbanization and identify the potential drivers of such differences.

2.1.3) Phenological changes reflected in publicly available data

Mounting evidence suggests that phenology - the timing of life history events such as autumn senescence, seasonal migration, or flowering period - is sensitive to temperature and is an important bioindicator of climatic shifts (Fitter & Fitter, 2002; Walther et al., 2002; Visser & Both, 2005). Cities tend to be warmer than surrounding rural areas due to the absorption and maintenance of heat by impermeable surfaces, a phenomenon known as the urban heat island effect. These increased temperatures advance phenology and lengthen periods of activity in urban areas relative to rural landscapes (D. Li et al., 2021; Merckx et al., 2021; Wohlfahrt et al., 2019). Phenological data may be sourced from iNaturalist (D. Li et al., 2021), digitized museum collections (e.g., herbarium specimens; Willis et al., 2017), dedicated phenological monitoring networks (Merckx et al., 2021; Wohlfahrt et al., 2019), and remotely-sensed data sources (Li et al., 2017). For example, Wohlfahrt et al., (2019) demonstrated urbanization in Europe leads to earlier leaf out, flowering, and fruiting as well as delayed senescence by leveraging data from the Pan European Phenological Project. The use of exclusively publicly available data demonstrated a consistent effect of the urban heat island on plant phenology.

Although plants have been the focal group for most urban phenological studies using publicly available data, occurrence data also exist for a broad range of other groups (Olsen et al., 2020). For example, Owen et al., (2024) used iNaturalist records to document earlier activity of the invasive spotted lanternfly (*Lycorma delicatula*) in response to increased temperatures in cities. Their results suggest that the urban heat island effect could facilitate species movements into cities in colder climates. Similarly, Merckx et al., (2021) used community science data from insect monitoring networks in Europe to investigate the phenology of two lepidopteran species. The authors found that urban populations had longer periods of flight activity relative to their rural counterparts. Follow-up experiments confirmed that these shifts were driven by changes in life-history plasticity in both species. Research expanding beyond plants will facilitate our understanding of how urbanization affects ecological interactions. As suggested by Bell et al. (2021), occurrence data such as those from iNaturalist and museum collections are well-suited to fill this knowledge gap.

2.1.4) Uncovering urban species interactions using* in-silico *approaches

The high habitat heterogeneity in urban environments can influence species behaviors and interactions (Diamond & Martin, 2021; Irwin et al., 2020). Occurrence data have been widely used to document spatial patterns of biodiversity (e.g., using the spatial coordinates of these records; Beninde et al., 2023), but their associated photographic data remain an underutilized research asset (Pernat et al., 2022). These photographs, sometimes termed “secondary data”, are a rich source of information that can be creatively leveraged to quantify ecological interactions in cities. Importantly, these data help overcome some of the challenges associated with monitoring secretive species in urban areas, such as accessing private land and the logistic difficulty of building adequately sized datasets (Putman et al., 2021).

Marín-Gómez et al., (2022) developed a framework for these studies by compiling a dataset of hummingbird-plant interactions across an urbanization gradient. The authors leveraged 742 photographs available through iNaturalist and eBird to construct plant-pollinator networks. The findings demonstrated that in Mexico City, Mexico, the hummingbird-plant networks were more generalized compared to the surrounding natural areas. They also found that in urban settings, non-native plant species and non-bird pollinated plants were important network components. Thereby, emphasizing the vital role that urban pollinator gardens play in increasing functional connectivity between plants and pollinators.

Image data has also been used to infer how species interactions change through time. Meineke & Davies, (2019) used 576 herbarium records to document changes in plant-herbivore interactions over a 100-year time period by assessing leaf herbivore damage. They found that although rising temperatures increased herbivory, urbanization disrupted this trend, potentially due to a lower abundance of herbivores in cities. Subsequent work has provided a framework for studying long-term plant-herbivore dynamics using Digital Eco-Evo approaches by replicating the findings using digitized herbarium specimens (Meineke et al., 2020).

Photographic data has also been used to study animal-animal interactions. Putman et al., (2021) used images from iNaturalist to study predation and parasitism of the elusive Southern alligator lizard in California, USA. From the images, they quantified the risk of predation, using lizard tail breaks as a proxy, and parasitism, based on the presence and abundance of ticks. Their results demonstrated that urbanization has opposing effects of predation and parasitism on lizards. Urbanization correlates with increased predation, attributed to increased house cat abundance, but also decreased parasitism, likely due to the lower abundance of viable tick hosts in cities.

In sum, these studies demonstrate the power of crowd-sourced image data and how secondary data may be used for understanding species interactions. This relatively untapped resource allows large datasets to be obtained that would otherwise be arduous for an individual researcher to collect. Digital Eco-Evo approaches that leverage secondary photographic data will be further facilitated by machine-learning techniques that are specialized for large-scale image analyses (see section 3.2).

2.2) Urban Species and Populations

2.2.1) Assessing urban tolerance across cities

Urban environments impose novel conditions that constrain which species can survive in cities. Species are often divided between those that succeed in urban environments – urban exploiters or urban adapters – and those that are excluded and threatened by urban expansion – urban avoiders (Blair, 1996, 2001; Faeth et al., 2011). Urban adapters are hypothesized to possess exaptations that facilitate urban success or tolerance (Winchell, Losos, et al., 2023).

Digital Eco-Evo approaches have combined occurrence data (e.g., GBIF, iNaturalist, and eBird) with trait databases to test hypotheses about urban tolerance, such as the Functional Biotic Homogenization hypothesis (Olden et al., 2004). This hypothesizes that within taxonomic groups, species traits are predicted to converge in response to urban conditions. Such analyses can uncover which species are most tolerant, which traits are most beneficial, and whether urban tolerance patterns are city-specific or generalizable across urban areas (see Williams et al. 2008 for a conceptual framework to understand filters specifically in urban flora).

Winchell et al. (2020) used a phylogenetic framework to understand the role evolutionary history plays in urban tolerance among Caribbean anoles (*Anolis* spp.). Using occurrence data from GBIF and iNaturalist, they found strong phylogenetic signals of urban tolerance in the *Anolis* genus. Additionally, locomotor traits such as rear lamellae number and hindlimb length were positively associated with urban tolerance. Warm temperature specialists were more likely to be urban tolerant, indicating that the urban heat island effect may impose an ecological filter for urban species composition. Previously, Callaghan, Major, et al., (2019) found a similar pattern in birds. The authors classified eBird observations using a continuous urbanization index (**Table 2**). The results showed generalism in feeding and breeding habitat, clutch size, and residual brain and body size were important for urban bird tolerance. Neate-Clegg et al., (2023) also found that select functional traits are associated with urban tolerance on a global scale. Latitude and human population density modulated this association for some traits. For example, the relationship between urban tolerance and body mass changed with latitude.

City-specific tolerance has been found as well, especially when accounting for city-level characteristics, such as the socioeconomic status and structure of the built environment. Hensley et al., (2019), using previously published bird counts, and Kinnunen et al., (2022), using eBird observations, combined bird counts and species-level trait data (Amniote Life-History database, BirdLife, Birds of the World) to investigate patterns of city-specific tolerance. Both studies found that urban tolerance of bird species is influenced by inter-city variability, resulting in differential trait-based filtering. Traits that varied across cities include body size, clutch size, longevity, diet guild, habitat preference, and migratory status.

Evidenced by the studies in this section, Digital Eco-Evo approaches can be used to compare cities around the globe, explaining how cultural, historical, and structural characteristics of cities influence species filters. As publicly available datasets grow, our understanding of urban filters will extend to other taxonomic groups and geographical ranges.

2.2.2) Publicly available data reveal urban evolutionary adaptations

Research assessing evolutionary processes in urban landscapes has seen a marked increase in the last 15 years (Diamond & Martin, 2021). This body of work has uncovered how trait evolution is shaped by urban conditions, such as elevated levels of impervious surfaces, increased temperature, changes in resource availability, and pollution (Cheptou et al. 2008; Campbell-Staton et al., 2021; Oziolor et al., 2014). Increasingly, multi-city comparisons have become a useful system for studying parallel evolution (Reid et al., 2016; Santangelo et al., 2022; Winchell, Campbell-Staton, et al., 2023). A typical multi-city project would involve significant logistical challenges and financial costs, but the abundance of photographic data from community science initiatives alleviates these issues. This makes Digital Eco-Evo approaches an advantageous avenue to understand how urbanization may elicit adaptive morphological changes. However, researchers should be careful when interpreting patterns observed via publicly available data. Without experimental means, comparisons across many datasets, or prior knowledge, determining whether changes in a trait are the result of phenotypic plasticity or evolutionary change can be impossible.

Cosentino & Gibbs (2022) demonstrated the power of Digital Eco-Evo by asking how urbanization influences melanization patterns in eastern gray squirrels. This is a known Mendelian trait in eastern grey squirrels, allowing for genotyping by photographic data. The authors compiled >60,000 iNaturalist photos across 43 cities, and community scientists scored squirrel coat color using Zooniverse, a platform to crowd source data collection for scientific research. Melanism increased with the percent impervious surfaces, especially in large, heavily forested, northern cities. The authors theorize that the inferred parallel shifts in melanism are in response to similar selective pressures, such as predation or selection on correlated traits.

Similarly, Fukano et al. (2023) combined iNaturalist photos of the flowering plant *Oxalis corniculata* with the percent impervious surface as a metric of urbanization. The authors found higher proportions of red individuals in urban areas, supporting their hypothesis that red leaf color is adaptive within an urban context. Further experimentation supported this, as red coloration potentially helps plants cope with the urban heat island effect.

Other digitized repositories have extended urban ecology and evolution studies beyond the traits that may be captured in photographs. In one such example, Hantak et al. (2021) used publicly available mammalian datasets to test Bergmann's rule, which states that in cooler climates populations tend to have larger body sizes (Bergmann, 1847). A dataset of over 100,000 size measurement records was assembled using VertNet, NEON, and NACSM, and correlated with human population density. The authors found that, despite the urban heat island effect making cities warmer (Phelan et al., 2015), Bergmann's rule did not apply. Their results provide further avenues for study, such as whether patterns in body size are better explained by access to novel food sources.

In summary, these studies demonstrate the effectiveness of Digital Eco-Evo approaches for compiling multi-city datasets of photographic data and identifying broad trends of parallel evolution. Integrating digitized traits and other specimen data will undoubtedly provide further avenues for study and help clarify how urban environments drive evolutionary change.

2.2.3) Inferring plastic responses to urbanization using publicly available data

Phenotypic plasticity, the ability for a single genotype to express differing phenotypes due to environmental conditions, can facilitate success in urban areas by buffering organisms against stressful urban conditions (Bressler et al., 2020). However, plasticity may also be maladaptive, leading to rapid adaptation or local extinction (Gomulkiewicz & Holt, 1995; Chevin et al., 2010; Ghalambor et al., 2015; Diamond & Martin, 2021). Studying plasticity (adaptive or maladaptive) in urban organisms can inform researchers of the genetic mechanisms underlying population persistence in novel environments.

Digital Eco-Evo approaches allow scientists to ask whether urban environmental cues elicit known plastic responses (**Table 1**). For example, Batool et al., (2024) tested how temperature differences, due to the urban heat island and latitude, affect cuticular melanin of monarch larvae, a known plastic trait. Using iNaturalist and BugGuide, they quantified melanism in larvae and correlated it with human population density and latitude. Larvae at warmer, lower latitudes had decreased levels of melanin. Urbanization, however, did not elicit a significant effect on melanism, indicating that increased temperatures due to the urban heat island effect were insufficient to trigger a plastic response.

Trait databases such as the TRY Plant Trait Database (**Table 1**) have also proven valuable for studying plasticity. Using a partially Digital Eco-Evo approach, Ibsen et al., (2023) compared water use and carbon gain strategies between 30 species of trees in urban Los Angeles and in their respective native habitats. Urban trees were irrigated, and trait data were collected *in situ,* while trait data for unirrigated trees in their native range were obtained from TRY. To collect such a trait dense dataset would require global fieldwork and be logistically challenging; however, the TRY database allowed Ibsen et al., (2023) to adapt several studies into a robust comparative analysis. Their analysis revealed a decoupling of the relationship between carbon gain and water use among urban trees relative to those in their native habitat. The effects were magnified in arid locations, presumably as a result of high resource availability.

Although these examples show that plasticity may be inferred from publicly available datasets, caution should be taken to avoid conflating plasticity and evolutionary adaptation. The outlined examples involved researchers using datasets of species with known plastic responses to make their inference (Ibsen et al., 2023; Batool et al., 2024), and this is key to designing a robust study using publicly available data.

2.2.4) Investigating how urbanization influences genetic diversity using Digital Eco-Evo

Urbanization is hypothesized to reduce genetic variation due to habitat loss and increased fragmentation, which decreases effective population size and disrupts gene flow (Johnson & Munshi-South, 2017). Understanding the influence urbanization has on genetic diversity is imperative for developing conservation strategies and predicting the persistence of urban populations.

Traditionally, samples are collected and sequenced, which can be laborious, time consuming, expensive, and logistically infeasible. These challenges can be amplified for multi-city or multi-species studies. However, there is a rich history of depositing genetic data into repositories such as the National Center of Biotechnology Information's (NCBI) GenBank which may alleviate some of these barriers. In addition, microsatellite data, commonly used for Digital Eco-Evo research, can be collected through the DataOne repository network or the MacroPopGen dataset (**Table 1**). Although mitochondrial DNA has seen reuse for macrogenetic work, its use is not without controversy (Galtier et al., 2009; Paz-Vinas et al., 2021; Schmidt et al., 2023) and results have been less consistent, even across the same molecular markers (Miraldo et al., 2016; Millette et al., 2020).

For example, Schmidt et al., (2020) compiled a dataset of microsatellites to assess whether urbanization and human land use negatively affect population genetic parameters. DataOne and Dryad were used to obtain data from 66 species of North American mammals and non-migratory birds. Using these microsatellites, they calculated effective population size, gene diversity, allelic richness, and fixation index (F_{st}). These metrics were then correlated with measures of urbanization, including urban vs. rural land use classifications, human population density (U.S. Census Bureau and Statistics Canada), and the Human Footprint index. With increasing urbanization, mammals exhibited consistent and significant decreases in effective population size, allelic richness, and gene diversity, as well as an increase in population differentiation. Digital Eco-Evo approaches using microsatellites have also been applied to amphibians and reptiles (Schmidt & Garroway, 2021) and marine fishes (Karachaliou et al., 2025), showing neutral and negative effects, respectively.

Habrich et al., (2021) reused microsatellite data from MacroPopGen (**Table 1**) to ask how road networks and human population density (**Table 2**) affect the genetic diversity of North American terrestrial mammals. They found genera-specific effects that are weakly modulated by life history traits and behaviors. For example, genera with large home range sizes, such as Caribou *(Rangifer*), were most negatively affected by increasing road density. Furthermore, the authors found that human density has a greater negative effect on genetic diversity than road impacts do, corroborating the results of Schmidt et al., (2020).

Public data also allows researchers to explore how past and present social policies and practices impact the genetic landscape of cities. Schmidt and Garroway (2022) collected microsatellite data through the DataOne network for 40 terrestrial animal taxa. They tested whether socioeconomic legacies influence population genetic metrics, and found that wildlife in predominantly white neighborhoods maintained slightly larger effective population sizes, higher genetic diversity, and lower genetic differentiation. These results highlight how past and contemporary societal processes contribute to neighborhood-specific results (Schell et al., 2020).

Collectively, the studies outlined demonstrate that urbanization and its influences on population genetics has the potential to be inferred using Digital Eco-Evo approaches, but more data (taxonomically and geographically) is needed to fully realize its impacts (Paz-Vinas et al., 2025). As consistent patterns emerge across studies, these critical insights can inform city planning and urban conservation efforts more effectively. However, researchers must be cautious with publicly available genetic data. Errors or contamination in sequence data and metadata are not uncommon (X. Li et al., 2018; Schmidt et al., 2020; van den Burg & Vieites, 2023) and, if clustering sequences into artificial populations, care must be taken to ensure those populations remain biologically relevant (Paz-Vinas et al., 2021).

3) Resources for use in Urban Ecology and Evolution

3.1) Emerging and Underutilized Datasets

In this section, we present a range of emerging and underutilized resources that show promise for studying urban ecology and evolution. Like all resources, the extent to which they may be leveraged varies. Importantly, scientists and the general public can contribute to these repositories by depositing their own original data, allowing data to be repurposed and fully utilized in new and creative ways. Larger repositories and Digital Eco-Evo approaches offer the ability to combine and synthesize data, producing new insights into the broader patterns of urbanization. In addition to the datasets and repositories covered here, The Opportunistic Database of Biodiversity Databases (https://earthskysea.org/biodiversity-databases/) and The Urban Eco-Evo Trait Mapping Toolkit (Savage et al., 2024 - https://www.urbanevoecotools.org/) document publicly available databases useful for Digital Eco-Evo research which may have relevance to urban systems.

3.1.1) Digitized Museum Specimens

Natural history collections house a plethora of information relevant to urban ecology and evolution (Shultz et al., 2020), but access to physical specimens can be costly and logistically challenging. The rapid digitization of collections allows access to specimens entirely online, furthering research that has already leveraged this data source for urban ecology and evolution (e.g., Meineke & Davies, 2019; Meineke et al., 2020; Pearson et al., 2020; Hantak et al., 2021).

Digitization efforts go beyond images. Digital 3D models of specimens are becoming a prevalent data type, adding another dimension to current research. Many 3D structures can be accessed on Morphosource, a data repository that houses 3D and 2D media of physical objects (e.g., bones, whole specimens, cultural artifacts).

Through Morphosource, researchers can access media from research bodies such as oVert and interact with files in the browser.

iDigBio, a repository aggregating digitization efforts across predominantly U.S. based institutions, facilitates access to digitized specimens. The Global Registry of Scientific Collections (GRSciColl), through GBIF, helps connect iDigBio with global digitized collections. Together, these platforms increase the accessibility of digitized data for urban ecology and evolution research.

3.1.2) Spatial Resources

In large urban systems where the vast majority of land is privately owned, spatial resources are especially valuable. Google Street View (GSV) has images for all roadways in an urban area, often generated every few years. This data may be used to analyze the distribution of urban vegetation and changes in the urban landscape over time. GSV has already shown potential in investigating diversity in urban areas (Velasquez-Camacho et al., 2024; Ye et al., 2025), surveying plants growing in private yards (Ringland et al., 2021), and monitoring invasive species (Closa-Tarres et al., 2025). With the rapid development of AI tools for image recognition, this resource may become more prominent in urban ecology and evolution.

Another valuable resource is the ArcGIS Living Atlas of the World, which contains a plethora of maps (e.g., iNaturalist observations and World Traffic Service) and deep-learning packages (e.g., Tree Point Classification and Building Footprint Extraction) that may be leveraged for urban research. While a membership is required, many institutions and universities have licenses for student use, making this widely accessible in an academic setting.

3.1.3) Ecological Survey Data

Several repositories of ecological survey data have been created to standardize data re-use, while also bypassing the need for extensive and time consuming literature reviews. Movebank collects, archives, and provides visualization for animal tracking data (Kays et al., 2021), including data from urban systems that may be used to study how urban organisms navigate through cities. Although not strictly ‘urban’, Tucker et al. (2018) used this dataset to analyze how mammalian movement patterns changed with human influence using the Human Footprint Index. They find that the extent of movement becomes more restricted as human influence increases.

Similarly, sPlotOpen provides a centralized resource for archiving vegetation plot surveys (Sabatini et al., 2021). Although sPlotOpen is mainly composed of surveys in natural or semi-natural sites, the addition of vegetation plot surveys will assist in urban ecology and evolution pursuits. For example, sPlotOpen has been used to investigate how species co-occurrence can be used for ecological predictions (Siefert et al., 2024), a concept relevant to urban systems.

Arctos (Cicero et al., 2024) is another integrative and unified platform that provides online access to physical data housed at museums and cultural institutions. Data is contributed by diverse organizations (e.g., NEON, LTER, and Museums) spanning many disciplines (e.g., geology, anthropology, botany, and mammalogy). For example, Arctos provides access to specimen measurements, mark/recapture data, species sighting surveys, documents (e.g., field notes), and media, among others. The collated data provide a broad foundation for urban ecology and evolution studies.

Investigating population density of organisms can be facilitated with the TetraDENSITY 2.0 database (Santini et al., 2024). The database contains more than 18,000 population estimates for vertebrates spanning between 1926 and 2017. Researchers can expand the temporal range of this database to include more recent estimates by extracting data from literature, as performed by Tucker et al. (2020). Tucker et al. (2020) were able to leverage this database and their expanded estimates to ask how mammalian population size changes with human modification on a global scale.

3.1.4) Genetic Data

Publicly available genetic data, aside from microsatellite and mitochondrial DNA, have thus far been underutilized in urban ecology and evolution, likely due to a lack of geolocated sequence information. Presently available sequence data can be accessed through NCBI, which houses data ranging from genes to genomes. Unfortunately, filtering for applicable sequence data is difficult due to inconsistent documentation of metadata (e.g., location information). The Global Genome Biodiversity Network (GGBN) and the Genomic Observatories Metadatabase (GeOMe) aim to remediate this issue for sequence information by linking genetic data to valuable metadata (location, collector, date, etc.). GGBN specifically links museum collections with NCBI accession numbers while GeOMe functions more broadly across current and past sequencing projects.

3.2) Utilizing Artificial Intelligence for Urban Ecology and Evolution Research

Given the large number of publicly available photographic datasets (e.g., iNaturalist records), image processing to automate the extraction of relevant information from photographs has been the most applied method of AI in urban ecology and evolution. Convolutional neural networks (CNNs) have been a cornerstone for accomplishing this. Tools utilizing CNNs have been developed to extract trait data from photographic museum collections of plants (LeafMachine2 by Weaver & Smith, 2023; Love et al., 2021; Davis et al., 2020; Pearson et al., 2020; Meineke et al., 2020) and animal skeletons (Skelevision; Weeks et al., 2023) as well as phenological data from iNaturalist photographs (PhenoVision by Dinnage et al., 2025; Reeb et al., 2022; Stewart et al., 2025). These tools provide detailed workflows, open source code, and trained models, allowing researchers to leverage them for their own studies. Similarly, tools have been developed to analyze Google Street View (GSV) images using CNNs and existing plant identification software (Closa-Tarres et al., 2025; Velasquez-Camacho et al., 2024; Ringland et al., 2021). Automated image selection may allow this technology to be more widely used for urban ecology and evolution research, as most studies choose images manually or based on prior information (but see Velasquez-Camacho et al., 2024). When building new tools, developers should ensure their code is well documented and open source so that researchers can modify programs for specific use cases. Although CNNs are not new to biological analyses, these and similar algorithms such as the rapidly emerging transformer architectures will likely see increased development as AI models advance. Outside of pattern analyses in photographic images, AI tools may be a promising solution to analyzing and handling biases in publicly available data. Tools such as Fink et al.'s (2023) double machine learning model for navigating biases in community science data may become more common as models develop. For further discussion exploring the uses of AI in ecology, see Cipriano et al.'s (2025) review of AI models and their applicability in terrestrial ecology. Their review presents a detailed introduction to the different types of AI models, their uses, and challenges.

4) Dealing with biases and limitations of publicly available datasets

Although publicly available datasets provide an invaluable resource for ecology and evolution, they are not without limitations. Several studies highlighted in this review have outlined how this data can be biased, incomplete, or lacking key pieces of metadata. For example, the distribution of iNaturalist occurrence data can be geographically biased within cities (Carlen et al., 2024), or incomplete metadata can result in unusable genetic data (Schmidt et al., 2020). Acknowledging these limitations and ensuring they are addressed is an essential aspect of performing reliable Digital Eco-Evo research. While we provide an overview of biases and limitations, there are several papers that discuss this topic in detail (Boyd, Powney, et al., 2023) and provide tools to help researchers address biases when conducting Digital Eco-Evo research (Boyd et al., 2022).

4.1) Addressing sampling bias

Spatial sampling biases are thoroughly documented across publicly available datasets (see Beck et al., 2014; Di Cecco et al., 2021; Johnston et al., 2022; Backstrom et al., 2025). These biases are often amplified in urban areas, as sociological variables influence sampling accessibility and effort. For example, wealthier, predominantly white neighborhoods tend to be better sampled, likely due to a history of unjust land-use policies (Ellis-Soto et al., 2023; Carlen et al., 2024). This form of sampling, known as non-probability sampling, introduces systematic bias because not all populations are sampled equally (e.g., HOLC grades) and is common amongst community science initiatives. When the reason for missingness is known, these data are missing at random (MAR) and can be accounted for more robustly (though caution is still required) than when data is missing not at random (MNAR; Boyd, Stewart, et al., 2023; Bowler et al., 2024). MNAR data is missing in a way that is directly related to the data missing and the reason for this missingness is not known. Ignoring this quality can perpetuate severe bias, but Bowler et al (2024) present several methods that may function well for these data (see also Bird et al., 2014; Beck et al., 2014; and Fink et al., 2023). Backstrom et al (2025) provide a framework for estimating spatial and temporal biases which may help ameliorate problematic results, but complex biases can still result in poor inference. This is demonstrated by Boyd, Steward, et al (2023) where several adjustment methods were applied to citizen data and compared to the true distribution trend of the plant *Calluna vulgaris.* Their models proved closer to the truth, but were still biased in the inferred biodiversity trend.

Missing data issues, regardless of the type, will likely never disappear from publicly available datasets, but that does not mean they are to be discarded. Rather, we need to clearly communicate biases that are at play and completing formal risk-of-bias assessments may aid in making those biases known (Boyd et al., 2022; Boyd, Powney, et al., 2023). Importantly, these issues are not new and have been documented in other disciplines such as survey data in social sciences or medical research. Learning from and collaborating with statisticians in other fields may be a path towards more accurate inferences based on these data (Boyd, Powney, et al., 2023). In sum, these tools provide researchers with an array of approaches to circumvent biases, making Digital Eco-Evo studies more reliable.

4.2) Barriers to data reuse

Open data initiatives and centralized repositories have facilitated easier access to data, yet barriers to data reuse still exist. For example, Schmidt et al., (2020) excluded 36 of the 313 initially found microsatellite datasets from their analysis due to missing or improperly deposited metadata (e.g., spatial information, population delineation). Some of these issues were resolved by contacting the research groups that generated the data, but this quickly becomes infeasible as datasets are added. It is imperative to standardize archival procedures to ensure that data can be effectively reused. A recent set of guidelines by Leigh et al., (2024) provides best practices for archiving genomic data, and the FAIR principles (Findable, Accessible, Interoperable, and Reusable) offer broader guidance and are applicable to all scientific data types (Wilkinson et al., 2016). In some cases, the creation of new centralized data repositories that facilitate the linkage between data and its metadata may be necessary for effective reuse (Schmidt et al., 2023). As noted, GeOMe has taken steps toward improving the metadata landscape by creating a centralized database that links metadata to sequence data. In turn, such initiatives may spur the reuse of new datatypes, further facilitating our understanding of urban processes (e.g., large genomic datasets for macrogenetics). The adoption of standard archival procedures will create more operable data repositories, greatly facilitating the novel insights generated from Digital Eco-Evo studies.

Furthermore, researchers should be aware that thorough quality inspections are needed when using publicly available data. This is especially relevant for repositories such as NCBI where incorrect species assignment or the use of taxonomic synonyms, sequence contamination, or sequencing errors are known to occur (X. Li et al., 2018; van den Burg & Vieites, 2023). Ignoring thorough quality inspections can prevent effective reuse and undermine the interpretation of any results.

4.3) Caveats to research questions and data interpretation

Although Digital Eco-Evo approaches help circumvent issues pertaining to urban field or experimental work, this strategy cannot act as a replacement for it and several caveats do exist. For example, research questions are often restricted to large-scale patterns of urban ecology and evolution given the coarse scale of many publicly available environmental and biological data types. This makes Digital Eco-Evo research excellent for investigating patterns across multiple cities, but the fine-scale heterogeneity of urban areas can make local analyses less robust. Generally, specific research questions pertaining to a single species on a local scale are likely best approached with field-based tools. In addition, questions must be tailored to the data that exists and investigating specific processes may be difficult. Biases in taxonomic groups and the geographic range of data may also limit research questions and statistical power. Particularly, data is biased to geographic regions, habitat types, neighborhoods, and taxonomic groups (Leigh et al., 2021; Tyska et al., 2025; Di Cecco; Johnston et al, 2022; Beck et al.,2014; Ellis-Soto et al., 2023; Shultz et al 2020, Paz-Vinas et al., 2025). Accounting for these biases is crucial for proper interpretation and filling these gaps should be a priority for original data collection efforts.

Importantly, though, a foundational research question should be devised before analyses take place to prevent fishing for significant patterns and hypothesizing after results are produced (Bissonette, 2021). Careful consideration of your research question and the biological relevancy of publicly available data is non-negotiable. Assessing the biological relevancy of available data is especially important when it is derived from varying sources. This is highlighted by Paz-Vinas et al., (2021) who found that biologically unrealistic populations were created when artificially grouping sequence data into populations for a macrogenetic study (Millette et al., 2020), potentially creating biased results.

Furthermore, ensuring that rigorous statistical analyses are employed and results are carefully interpreted is necessary to prevent misrepresentation. Statistical analyses over large datasets are often able to detect very small and statistically significant effects, but these effects may be so small that they lack any biological relevancy (Steel et al., 2013). Developing a deep familiarity with your study system or collaborating with domain experts can prevent misinterpretation of these negligible effects. Thorough familiarity with your research system is also necessary to prevent inaccurate interpretations. For example, Batool et al. (2024) were familiar with the biological basis of melanism in monarch caterpillars. If they were not, they could have attributed differences in melanistic patterns to adaptive evolution when plasticity is the basis of the trait.

In the burgeoning field of Digital Eco-Evo research, practitioners need to stay keenly aware that solid research questions are foundational for research and that results need to be interrogated for biological relevancy and not over extrapolated (Steel et al., 2013). Fortunately, this creates an exceptional opportunity to foster collaboration across disciplines. Often, research questions and data interpretation can benefit from the input of taxonomic, ecosystem, and field experts. These collaborations can thereby inspire questions from the Digital Eco-Evo and field-based communities and inform how, where, and what data needs to be further collected.

5) Enhancing Participation in Urban Ecology and Evolution with Publicly Available Datasets

Publicly available datasets have the power to broaden participation in urban ecology and evolution by removing financial, structural, and logistical barriers. Because these studies only require an internet connection, they are accessible to all students. This accessibility is especially relevant to non-traditional students, who are often balancing employment, family, and other responsibilities that hinder their engagement in traditional lab work. Expanding Digital Eco-Evo approaches can improve participation, by promoting a more equitable and inclusive research environment, enabling the participation of individuals historically underrepresented in science. We believe that urban ecological and evolutionary research may be particularly valuable for universities located within urban areas. The ability for students to experience urban ecosystems in their daily lives makes these datasets and concepts relatable, offering an avenue to capture students' interest and to engage meaningfully in ecology and evolution.

5.1) Improving the undergraduate research experience

Publicly available data are particularly well suited for course-based undergraduate research experiences (CUREs), which expose students to authentic scientific pursuits that generate novel findings. These in class research experiences, regardless of the data used, have been shown to improve retention, persistence, and confidence among students from underrepresented groups (Shapiro et al., 2015; Eagan et al., 2013; Hunter et al., 2007; Lopatto, 2007; Russell et al., 2007). Because acquiring original data for these programs can be logistically and financially challenging, publicly available databases provide a free and accessible resource. This makes them a low-cost option for instructors seeking to obtain data for authentic research in the classroom (see Arreguin et al., 2025 for a CURE utilizing publicly available genetic data). As outlined throughout this review, this data has relevance across a range of biological questions and topics, making these datasets suitable for a broad range of classes or research programs. Moreover, the large size of these datasets exposes students to valuable data science and programming experience. This is highly translational across fields, and early training can support and influence students' future careers.

Practical examples are the CUREs developed by Gastreich (2021), centered around assessing urban bird diversity using eBird data, and Heard et al., (2025), which investigates the impacts of urbanization on lizards using iNaturalist. Such courses can offer a useful example for instructors and be readily adapted and modified to explore urban ecology and evolution across spatial scales and taxa.

Although this research structure provides hands-on authentic research experiences, it may not be suited for all classes and cannot provide laboratory or field-specific training. In addition, it lacks experimental manipulation and the datasets are primarily observational in nature, creating the need for careful analyses and data filtration. As such, this research may be best suited for courses where the focus is observing patterns over large scales (landscape/macroecology), programmatic in nature, or statistically based.

5.2) Facilitating graduate and faculty research

Beyond undergraduate training, graduate students and research faculty can successfully employ publicly available data for their own research goals. The low-cost nature of these data make for an excellent resource when gathering preliminary data for research and grant proposals or performing standalone research. The use of publicly available datasets can alleviate barriers associated with original data collection. Individuals with disabilities may find aspects of field or laboratory based studies to be challenging, impeding their ability to participate in urban ecology and evolution projects. Furthermore, the high percentage of private land ownership in cities often hinders urban field collections, and the monetary cost associated with data generation may be prohibitively expensive.

Publicly available datasets allow researchers to perform rigorous, hypothesis driven research that contributes meaningfully to the field of urban ecology and evolution.

This type of research also presents an opportunity for collaboration among research groups. Often, analysis of large-scale datasets may benefit from the input of experts familiar with the taxonomic groups or ecosystems under investigation. Interdisciplinary working groups can be created to refine research questions and interpret the biological relevancy of the results, given the datasets at hand. Such collaborations may be especially useful for early career scientists, who can benefit from expert mentorship.

6) Concluding Remarks

In this review, we have collated presently used and developing datasets, along with potentially useful tools, with the intention that this work will act as a guide for utilizing Digital Eco-Evo approaches. This research strategy has provided powerful insights in urban ecology and evolution, but it's worth noting that it is not a cure-all. Field-based approaches offer unique insights that can be rarely replicated via Digital Eco-Evo approaches. In fact, research groups should consider using these two approaches together, as they often inform one another (Fukano et al., 2023; McCleery et al., 2023). It is important to note that this work is only possible thanks to databases that make data freely available and researchers who work hard to ensure that their data is easily accessible. As tools and datasets are developed, it is imperative that data remains publicly accessible and code is open-source. In doing so, we allow scientists to adjust tools as needed and use data to its full extent, furthering the range of questions that can be answered.

Data Accessibility

Datasets highlighted in this manuscript are accessible via the in-text tables and the appendix.

Conflicts of Interests Statements

The authors declare no conflicts of interests

Author Contributions Statement

Shawn Arreguin: Conceptualization, writing - original draft, writing - review and editing, visualizations (supporting). Joseph F. Walker: Conceptualization, writing - original draft, writing - review and editing. Natalie L.R. Love: Conceptualization, visualizations (lead), writing - original draft, writing - review and editing.

Acknowledgements

We would like to thank Erin Grady for helping us find relevant literature. Roberta Mason-Gamer, Beatriz Dybas de Natividade, Hector Palma, Alexa Tyszka, Jiale Zhang, Jordan Kressman, Karolis Ramanaskus, Eric Bretz, Timi Okeseni, and Linda Mansour all provided valuable comments and suggestions to improve our manuscript and pointed us to relevant databases. We would also like to thank two anonymous reviewers for their insightful comments which greatly improved our manuscript. This manuscript is supported by start-up funding through the Walker Lab at the University of Illinois Chicago.

su

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data, 5(1), 170191. https://doi.org/10.1038/sdata.2017.191

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning*, *75*(2), 53–107. https://doi.org/10.1016/j.progress.2011.04.001

Arreguin, S., Walker-Hale, N., Casagrande, M. R., Bishop, S., Pugacewicz, E., Ramizuddin, M., Savitzky, C., Morales, N., Dhan, A., Natividade, B. D. da, Madrid, V., Simmons, C., Diaz, A., Ahmed, D., Ho, Z., Pellecer, J. N., Diaz, S., Mathew, T., Mirshed, T., … Walker, J. F. (2025). *Going green: Recycling transcriptomes to infer evolutionary relationships, gene duplication, gene tree conflict, and patterns of molecular evolution in the Apocynaceae* (p. 2025.06.20.660724). bioRxiv. https://doi.org/10.1101/2025.06.20.660724

Aybar et al., (2020). rgee: An R package for interacting with Google Earth Engine. Journal of Open Source Software, 5(51), 2272, https://doi.org/10.21105/joss.02272

Backstrom, L. J., Callaghan, C. T., Worthington, H., Fuller, R. A., & Johnston, A. (2025). Estimating sampling biases in citizen science datasets. Ibis, 167(1), 73–87. https://doi.org/10.1111/ibi.13343

Batool, R., Jenan, J., Schuster, M., & Murray, R. L. (2024). How do latitude and urban heat islands affect larval melanization in monarch butterflies (Danaus plexippus) ? Canadian Journal of Zoology, 102(3), 315–321. https://doi.org/10.1139/cjz-2023-0099

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, *19*, 10–15. https://doi.org/10.1016/j.ecoinf.2013.11.002

Bell, K. C., Adams, B. J., Brown, B. V., Ludt, W. B., Pauly, G. B., Shultz, A. J., & Vendetti, J. E. (2021).

Museums Are Critical Infrastructure for Studying Urban Biodiversity. *BioScience*, *71*(5), 433. https://doi.org/10.1093/biosci/biab038

Beninde, J., Delaney, T. W., Gonzalez, G., & Shaffer, H. B. (2023). Harnessing iNaturalist to quantify hotspots of urban biodiversity: The Los Angeles case study. Frontiers in Ecology and Evolution, 11, 983371. https://doi.org/10.3389/fevo.2023.983371

Bergmann, C. (1847). About the relationships between heat conservation and body size of animals. Goett. Stud. (original Ger.) 1, 595–708.

Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J. F., Pecl, G. T., Barrett, N., & Frusher, S. (2014). Statistical solutions for error and bias in global citizen science datasets. Biological Conservation, 173, 144–154. https://doi.org/10.1016/j.biocon.2013.07.037

Bissonette, J. A. (2021). Big Data, Exploratory Data Analyses and Questionable Research Practices: Suggestion for a Foundational Principle. *Wildlife Society Bulletin (2011-)*, *45*(3), 366–370.

Blair, R. B. (1996). Land Use and Avian Species Diversity Along an Urban Gradient. Ecological Applications, 6(2), 506–519. https://doi.org/10.2307/2269387

Blair, R. B. (2001). Birds and Butterflies Along Urban Gradients in Two Ecoregions of the United States: Is Urbanization Creating a Homogeneous Fauna? In J. L. Lockwood & M. L. McKinney (Eds.), Biotic Homogenization (pp. 33–56). Springer US. https://doi.org/10.1007/978-1-4615-1261-5_3

Border, J. A., Newson, S. E., White, D. C. J., & Gillings, S. (2017). Predicting the likely impact of urbanisation on bat populations using citizen science data, a case study for Norfolk, UK. Landscape and Urban Planning, 162, 44–55. https://doi.org/10.1016/j.landurbplan.2017.02.005

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J. B., & Pocock, M. J. O. (2024). Treating gaps and biases in biodiversity data as a missing data problem. Biological Reviews of the Cambridge Philosophical Society, 100(1), 50–67. https://doi.org/10.1111/brv.13127

Boyd, R. J., Powney, G. D., Burns, F., Danet, A., Duchenne, F., Grainger, M. J., Jarvis, S. G., Martin, G., Nilsen, E. B., Porcher, E., Stewart, G. B., Wilson, O. J., & Pescott, O. L. (2022). ROBITT: A tool for assessing the risk-of-bias in studies of temporal trends in ecology. *Methods in Ecology and Evolution*, *13*(7), 1497–1507. https://doi.org/10.1111/2041-210X.13857

Boyd, R. J., Powney, G. D., & Pescott, O. L. (2023). We need to talk about nonprobability samples. *Trends in Ecology & Evolution*, *38*(6), 521–531. https://doi.org/10.1016/j.tree.2023.01.001

Boyd, R. J., Stewart, G. B., & Pescott, O. L. (2023). Descriptive inference using large, unrepresentative nonprobability samples: An introduction for ecologists. *Ecology*, *105*(2), e4214. https://doi.org/10.1002/ecy.4214

Bressler, S. A., Diamant, E. S., Tingley, M. W., & Yeh, P. J. (2020). Nests in the cities: Adaptive and non-adaptive phenotypic plasticity and convergence in an urban bird. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1941), 20202122. https://doi.org/10.1098/rspb.2020.2122

Cadenasso, M. L., Pickett, S. T. A., & Schwarz, K. (2007). Spatial heterogeneity in urban ecosystems: Reconceptualizing land cover and a framework for classification. *Frontiers in Ecology and the Environment*, *5*(2), 80–88. https://doi.org/10.1890/1540-9295(2007)5%255B80:SHIUER%255D2.0.CO;2

Callaghan, C. T., Bino, G., Major, R. E., Martin, J. M., Lyons, M. B., & Kingsford, R. T. (2019). Heterogeneous urban green areas are bird diversity hotspots: Insights using continental-scale citizen science data. Landscape Ecology, 34(6), 1231–1246. https://doi.org/10.1007/s10980-019-00851-6

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(6), 845–858. https://doi.org/10.1111/oik.06158

Campbell-Staton, S. C., Velotta, J. P., & Winchell, K. M. (2021). Selection on adaptive and maladaptive gene expression plasticity during thermal adaptation to urban heat islands. *Nature Communications*, *12*(1), 6195. https://doi.org/10.1038/s41467-021-26334-4

Cardille, J. A., Crowley, M. A., Saah, D., & Clinton, N. E. (Eds.). (2024). Cloud-Based Remote Sensing with Google Earth Engine: Fundamentals and Applications. Springer International Publishing. https://doi.org/10.1007/978-3-031-26588-4

Cârlan, I., Mihai, B.-A., Nistor, C., & Große-Stoltenberg, A. (2020). Identifying urban vegetation stress factors based on open access remote sensing imagery and field observations. Ecological Informatics, 55, 101032. https://doi.org/10.1016/j.ecoinf.2019.101032

Carlen, E. J., Estien, C. O., Caspi, T., Perkins, D., Goldstein, B. R., Kreling, S. E. S., Hentati, Y., Williams, T. D., Stanton, L. A., Des Roches, S., Johnson, R. F., Young, A. N., Cooper, C. B., & Schell, C. J. (2024). A framework for contextualizing social-ecological biases in contributory science data. People and Nature, 6(2), 377–390. https://doi.org/10.1002/pan3.10592

Carpenter, S. R., Armbrust, E. V., Arzberger, P. W., Chapin, F. S., Elser, J. J., Hackett, E. J., Ives, A. R., Kareiva, P. M., Leibold, M. A., Lundberg, P., Mangel, M., Merchant, N., Murdoch, W. W., Palmer, M. A., Peters, D. P. C., Pickett, S. T. A., Smith, K. K., Wall, D. H., & Zimmerman, A. S. (2009). Accelerate Synthesis in Ecology and Environmental Sciences. *BioScience*, *59*(8), 699–701. https://doi.org/10.1525/bio.2009.59.8.11

Cheptou, P.-O., Carrue, O., Rouifed, S., & Cantarel, A. (2008). Rapid evolution of seed dispersal in an urban environment in the weed Crepis sancta. Proceedings of the National Academy of Sciences, 105(10), 3796–3799. https://doi.org/10.1073/pnas.0708446105

Chevin, L.-M., Lande, R., & Mace, G. M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. PLOS Biology, 8(4), e1000357. https://doi.org/10.1371/journal.pbio.1000357

Cicero, C., Koo, M. S., Braker, E., Abbott, J., Bloom, D., Campbell, M., Cook, J. A., Demboski, J. R., Doll, A. C., Frederick, L. M., Linn, A. J., Mayfield-Meyer, T. J., McDonald, D. L., Nachman, M. W., Olson, L. E., Roberts, D., Sikes, D. S., Witt, C. C., & Wommack, E. A. (2024). Arctos: Community-driven innovations for managing natural and cultural history collections. PLOS ONE, 19(5), e0296478. https://doi.org/10.1371/journal.pone.0296478

Cipriano, C., Noce, S., Mereu, S., & Santini, M. (2025). Algorithms going wild – A review of machine learning techniques for terrestrial ecology. *Ecological Modelling*, *506*, 111164. https://doi.org/10.1016/j.ecolmodel.2025.111164

Closa-Tarres, A., Rojano, F., & Strager, M. P. (2025). Geospatial Data and Google Street View Images for Monitoring Kudzu Vines in Small and Dispersed Areas. Earth, 6(2), 40. https://doi.org/10.3390/earth6020040

Cosentino, B. J., & Gibbs, J. P. (2022). Parallel evolution of urban–rural clines in melanism in a widespread mammal. *Scientific Reports*, *12*, 1752. https://doi.org/10.1038/s41598-022-05746-2

Crawford, C. J., Roy, D. P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Gerace, A., Choate, M., Engebretson, C., Micijevic, E., Schmidt, G., Anderson, C., Anderson, M., Bouchard, M., Cook, B., Dittmeier, R., Howard, D., Jenkerson, C., Kim, M., … Zahn, S. (2023). The 50-year Landsat collection 2 archive. Science of Remote Sensing, 8, 100103. https://doi.org/10.1016/j.srs.2023.100103

Curti, J. N., Barton, M., Flores, R. G., Lechner, M., Lipman, A., Montgomery, G. A., Park, A. Y., Rochel, K., & Tingley, M. W. (2024). Using unstructured crowd-sourced data to evaluate urban tolerance of terrestrial native animal species within a California Mega-City. *PLOS ONE*, *19*(5), e0295476. https://doi.org/10.1371/journal.pone.0295476

Davis, C. C., Champ, J., Park, D. S., Breckheimer, I., Lyra, G. M., Xie, J., Joly, A., Tarapore, D., Ellison, A. M., & Bonnet, P. (2020). A New Method for Counting Reproductive Structures in Digitized Herbarium Specimens Using Mask R-CNN. *Frontiers in Plant Science*, *11*, 1129. https://doi.org/10.3389/fpls.2020.01129

Deck, J., Gaither, M. R., Ewing, R., Bird, C. E., Davies, N., Meyer, C., Riginos, C., Toonen, R. J., & Crandall, E. D. (2017). The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. *PLOS Biology*, *15*(8), e2002925. https://doi.org/10.1371/journal.pbio.2002925

de Magalhães, J. P., Abidi, Z., Dos Santos, G. A., Avelar, R. A., Barardo, D., Chatsirisupachai, K., Clark, P., De-Souza, E. A., Johnson, E. J., Lopes, I., Novoa, G., Senez, L., Talay, A., Thornton, D., &

To, P. K. P. (2024). Human Ageing Genomic Resources: Updates on key databases in ageing research. *Nucleic Acids Research*, *52*(D1), D900–D908. https://doi.org/10.1093/nar/gkad927

Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for Biodiversity Science. BioScience, 71(11), 1179–1188. https://doi.org/10.1093/biosci/biab093

Diamond, S. E., & Martin, R. A. (2021). Evolution in Cities. Annual Review of Ecology, Evolution, and Systematics, 52(Volume 52, 2021), 519–540. https://doi.org/10.1146/annurev-ecolsys-012021-021402

Dinnage, R., Grady, E., Neal, N., Deck, J., Denny, E., Walls, R., Seltzer, C., Guralnick, R., & Li, D. (2025). PhenoVision: A framework for automating and delivering research-ready plant phenology data from field images. *Methods in Ecology and Evolution*, *16*(8), 1763–1780. https://doi.org/10.1111/2041-210X.70081

Dong, C., MacDonald, G., Okin, G. S., & Gillespie, T. W. (2019). Quantifying Drought Sensitivity of Mediterranean Climate Vegetation to Recent Warming: A Case Study in Southern California. Remote Sensing, 11(24), 2902. https://doi.org/10.3390/rs11242902

Droege, G., Barker, K., Seberg, O., Coddington, J., Benson, E., Berendsohn, W. G., Bunk, B., Butler, C., Cawsey, E. M., Deck, J., Döring, M., Flemons, P., Gemeinholzer, B., Güntsch, A., Hollowell, T., Kelbert, P., Kostadinov, I., Kottmann, R., Lawlor, R. T., … Zhou, X. (2016). The Global Genome Biodiversity Network (GGBN) Data Standard specification. *Database: The Journal of Biological Databases and Curation*, *2016*, baw125. https://doi.org/10.1093/database/baw125

Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a Difference in Science Education: The Impact of Undergraduate Research Programs. 50(4), 656–850. https://doi.org/10.3102/0002831213482038

Ellis-Soto, D., Chapman, M., & Locke, D. H. (2023). Historical redlining is associated with increasing geographical disparities in bird biodiversity sampling in the United States. *Nature Human Behaviour*, *7*(11), 1869–1877. https://doi.org/10.1038/s41562-023-01688-5

Estien, C. O., Fidino, M., Wilkinson, C. E., Morello-Frosch, R., & Schell, C. J. (2024). Historical redlining is associated with disparities in wildlife biodiversity in four California cities. Proceedings of the National Academy of Sciences, 121(25), e2321441121. https://doi.org/10.1073/pnas.2321441121

Faeth, S. H., Bang, C., & Saari, S. (2011). Urban biodiversity: Patterns and mechanisms. Annals of the New York Academy of Sciences, 1223(1), 69–81. https://doi.org/10.1111/j.1749-6632.2010.05925.x

Fang, Y., & Jawitz, J. W. (2018). High-resolution reconstruction of the United States human population distribution, 1790 to 2010. *Scientific Data*, *5*(1), 180067. https://doi.org/10.1038/sdata.2018.67

Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., Oldham Jaromczyk, L., Robinson, O., Wood, C., Kelling, S., & Rodewald, A. D. (2023). A Double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, *14*(9), 2435–2448. https://doi.org/10.1111/2041-210X.14186

Fitch, G., Glaum, P., Simao, M.-C., Vaidya, C., Matthijs, J., Iuliano, B., & Perfecto, I. (2019). Changes in adult sex ratio in wild bee communities are linked to urbanization. Scientific Reports, 9, 3767. https://doi.org/10.1038/s41598-019-39601-8

Fitch, G., Wilson, C. J., Glaum, P., Vaidya, C., Simao, M.-C., & Jamieson, M. A. (2019). Does urbanization favour exotic bee species? Implications for the conservation of native bees in cities. Biology Letters, 15(12), 20190574. https://doi.org/10.1098/rsbl.2019.0574

Fitter, A. H., & Fitter, R. S. R. (2002). Rapid Changes in Flowering Time in British Plants | Science. Science, 296(5573), 1689–1691. https://doi.org/10.1126/science.1071617

Fukano, Y., Yamori, W., Misu, H., Sato, M. P., Shirasawa, K., Tachiki, Y., & Uchida, K. (2023). From green to red: Urban heat stress drives leaf color evolution. *Science Advances*, *9*(42), eabq3542. https://doi.org/10.1126/sciadv.abq3542

Galtier, N., Nabholz, B., Glémin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, *18*(22), 4541–4550.

https://doi.org/10.1111/j.1365-294X.2009.04380.x

Gastreich, K. R. (2021). *Assessing Urban Biodiversity With the eBird Citizen Science Project: A Course-Based Undergraduate Research Experience (CURE) Module*. https://doi.org/10.24918/cs.2020.18

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., & Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature, 525(7569), 372–375. https://doi.org/10.1038/nature15256

Gomulkiewicz, R., & Holt, R. D. (1995). When does Evolution by Natural Selection Prevent Extinction? Evolution, 49(1), 201–207. https://doi.org/10.2307/2410305

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global Change and the Ecology of Cities. Science, 319(5864), 756–760. https://doi.org/10.1126/science.1150195

Habrich, A. K., Lawrence, E. R., & Fraser, D. J. (2021). Varying genetic imprints of road networks and human density in North American mammal populations. Evolutionary Applications, 14(6), 1659. https://doi.org/10.1111/eva.13232

Hantak, M. M., McLean, B. S., Li, D., & Guralnick, R. P. (2021). Mammalian body size is determined by interactions between climate, urbanization, and ecological traits. Communications Biology, 4(1), 1–10. https://doi.org/10.1038/s42003-021-02505-3

Heard, M. J., Putman, B. J., Riskin, S. H., & Thawley, C. J. (2025). Collecting Novel Data from Inaturalist Photos to Teach Students About the Impacts of Urbanization on Animal Behavior & Ecological Interactions. The American Biology Teacher, 87(4–5), 286–291. https://doi.org/10.1525/abt.2025.87.4-5.286

Hensley, C. B., Trisos, C. H., Warren, P. S., MacFarland, J., Blumenshine, S., Reece, J., & Katti, M. (2019). Effects of Urbanization on Native Bird Species in Three Southwestern US Cities. Frontiers in Ecology and Evolution, 7. https://doi.org/10.3389/fevo.2019.00071

Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of

Forestry Research, 32(1), 1–6. https://doi.org/10.1007/s11676-020-01155-1

Hunter, A.-B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. Science Education, 91(1), 36–74. https://doi.org/10.1002/sce.20173

Ibsen, P. C., Santiago, L. S., Shiflett, S. A., Chandler, M., & Jenerette, G. D. (2023). Irrigated urban trees exhibit greater functional trait plasticity compared to natural stands. Biology Letters, 19(1), 20220448. https://doi.org/10.1098/rsbl.2022.0448

Irwin, R. E., Youngsteadt, E., Warren, P. S., & Bronstein, J. L. (2020). The Evolutionary Ecology of Mutualisms in Urban Landscapes. In M. Szulkin, J. Munshi-South, & A. Charmantier (Eds.), Urban Evolutionary Biology (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198836841.003.0007

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. (2016). Cities are hotspots for threatened species. Global Ecology and Biogeography, 25(1), 117–126. https://doi.org/10.1111/geb.12404

Jamieson, M. A., Carper, A. L., Wilson, C. J., Scott, V. L., & Gibbs, J. (2019). Geographic Biases in Bee Research Limits Understanding of Species Distribution and Response to Anthropogenic Disturbance. Frontiers in Ecology and Evolution, 7. https://doi.org/10.3389/fevo.2019.00194

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*(7424), 444–448. https://doi.org/10.1038/nature11631

Johnson, M. T. J., & Munshi-South, J. (2017). Evolution of life in urban environments. Science, 358(6363), eaam8327. https://doi.org/10.1126/science.aam8327

Johnston, A., Matechou, E., & Dennis, E. B. (2022). Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, *14*(1), 103–116. https://doi.org/10.1111/2041-210X.13834

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., … Purvis, A. (n.d.). *PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals*. https://doi.org/10.1890/08-1494.1

Karachaliou, E., Schmidt, C., de Greef, E., Docker, M. F., & Garroway, C. J. (2025). Urbanisation Is Associated With Reduced Genetic Diversity in Marine Fish Populations. Molecular Ecology, 34(7), e17711. https://doi.org/10.1111/mec.17711

Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B., Wright, I. J., Cornelissen, J. H. C., Violle, C., Harrison, S. P., Van Bodegom, P. M., Reichstein, M., Enquist, B. J., Soudzilovskaia, N. A., Ackerly, D. D., Anand, M., … Wirth, C. (2011). TRY – a global database of plant traits. *Global Change Biology*, *17*(9), 2905–2935. https://doi.org/10.1111/j.1365-2486.2011.02451.x

Kays, R., Davidson, S. C., Berger, M., Bohrer, G., Fiedler, W., Flack, A., Hirt, J., Hahn, C., Gauggel, D., Russell, B., Kölzsch, A., Lohr, A., Partecke, J., Quetting, M., Safi, K., Scharf, A., Schneider, G., Lang, I., Schaeuffelhut, F., … Wikelski, M. (2022). The Movebank system for studying global animal movement and demography. Methods in Ecology and Evolution, 13(2), 419–431. https://doi.org/10.1111/2041-210X.13767

Kelso, N.V. and Patterson, T. (2012). World Urban Areas, LandScan, 1:10 million (2012). Made with Natural Earth, online at http://www.naturalearthdata.com

Khikmah, F., Sebald, C., Metzner, M., & Schwieger, V. (2024). Modelling Vegetation Health and Its Relation to Climate Conditions Using Copernicus Data in the City of Constance. Remote Sensing, 16(4), 691. https://doi.org/10.3390/rs16040691

Kinnunen, R. P., Fraser, K. C., Schmidt, C., & Garroway, C. J. (2022). The socioeconomic status of cities covaries with avian life-history strategies. Ecosphere, 13(2), e3918. https://doi.org/10.1002/ecs2.3918

Kinnunen, R. P., Fraser, K. C., Schmidt, C., & Garroway, C. J. (2024). Structural and socioeconomic features of cities predict migratory bird species richness. *Journal of Avian Biology*, *2025*(1), e03189. https://doi.org/10.1111/jav.03189

Lawrence, E. R., Benavente, J. N., Matte, J.-M., Marin, K., Wells, Z. R. R., Bernos, T. A., Krasteva, N., Habrich, A., Nessel, G. A., Koumrouyan, R. A., & Fraser, D. J. (2019). Geo-referenced population-specific microsatellite data across American continents, the MacroPopGen Database. *Scientific Data*, *6*(1), 14. https://doi.org/10.1038/s41597-019-0024-7

Leigh, D. M., van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., Segelbacher, G., Hoban, S., & Paz-Vinas, I. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics*, *22*(12), 791–807. https://doi.org/10.1038/s41576-021-00394-0

Leigh, D. M., Vandergast, A. G., Hunter, M. E., Crandall, E. D., Funk, W. C., Garroway, C. J., Hoban, S., Oyler-McCance, S. J., Rellstab, C., Segelbacher, G., Schmidt, C., Vázquez-Domínguez, E., & Paz-Vinas, I. (2024). Best practices for genetic and genomic data archiving. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-024-02423-7

Leisenheimer, L., Wellmann, T., Jänicke, C., & Haase, D. (2024). Monitoring drought impacts on street trees using remote sensing—Disentangling temporal and species-specific response patterns with Sentinel-2 imagery. Ecological Informatics, 102659. https://doi.org/10.1016/j.ecoinf.2024.102659

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

Li, D., Barve, N., Brenskelle, L., Earl, K., Barve, V., Belitz, M. W., Doby, J., Hantak, M. M., Oswald, J. A., Stucky, B. J., Walters, M., & Guralnick, R. P. (2021). Climate, urbanization, and species traits interactively drive flowering duration. Global Change Biology, 27(4), 892–903. https://doi.org/10.1111/gcb.15461

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*. https://doi.org/10.3389/fevo.2019.00277

Li, X., Shen, X., Chen, X., Xiang, D., Murphy, R. W., & Shen, Y. (2018). Detection of Potential Problematic Cytb Gene Sequences of Fishes in GenBank. Frontiers in Genetics, 9. https://doi.org/10.3389/fgene.2018.00030

Li, X., Zhou, Y., Asrar, G. R., Mao, J., Li, X., & Li, W. (2017). Response of vegetation phenology to urbanization in the conterminous United States. Global Change Biology, 23(7), 2818–2830. https://doi.org/10.1111/gcb.13562

Lopatto, D. (2007). Undergraduate Research Experiences Support Science Career Decisions and Active Learning. CBE—Life Sciences Education, 6(4), 297–306. https://doi.org/10.1187/cbe.07-06-0039

Love, N. L. R., Bonnet, P., Goëau, H., Joly, A., & Mazer, S. J. (2021). Machine Learning Undercounts Reproductive Organs on Herbarium Specimens but Accurately Derives Their Quantitative Phenological Status: A Case Study of Streptanthus tortuosus. *Plants*, *10*(11), Article 11. https://doi.org/10.3390/plants10112471

Magle, S. B., Fidino, M., Lehrer, E. W., Gallo, T., Mulligan, M. P., Ríos, M. J., Ahlers, A. A., Angstmann, J., Belaire, A., Dugelby, B., Gramza, A., Hartley, L., MacDougall, B., Ryan, T., Salsbury, C., Sander, H., Schell, C., Simon, K., St Onge, S., & Drake, D. (2019). Advancing urban wildlife research through a multi-city collaboration. *Frontiers in Ecology and the Environment*, *17*(4), 232–239. https://doi.org/10.1002/fee.2030

Marín-Gómez, O. H., Rodríguez Flores, C., & Arizmendi, M. del C. (2022). Assessing ecological interactions in urban areas using citizen science data: Insights from hummingbird–plant meta-networks in a tropical megacity. *Urban Forestry & Urban Greening*, *74*, 127658. https://doi.org/10.1016/j.ufug.2022.127658

McCleery, R., Guralnick, R., Beatty, M., Belitz, M., Campbell, C. J., Idec, J., Jones, M., Kang, Y., Potash, A., & Fletcher, R. J. (2023). Uniting Experiments and Big Data to advance ecology and conservation. *Trends in Ecology & Evolution*, *38*(10), 970–979. https://doi.org/10.1016/j.tree.2023.05.010

Meineke, E. K., & Davies, T. J. (2019). Museum specimens provide novel insights into changing plant–herbivore interactions. Philosophical Transactions of the Royal Society B: Biological Sciences, 374(1763), 20170393. https://doi.org/10.1098/rstb.2017.0393

Meineke, E. K., Tomasi, C., Yuan, S., & Pryer, K. M. (2020). Applying machine learning to investigate long-term insect–plant interactions preserved on digitized herbarium specimens. Applications in Plant Sciences, 8(6). https://doi.org/10.1002/aps3.11369

Merckx, T., Nielsen, M. E., Heliölä, J., Kuussaari, M., Pettersson, L. B., Pöyry, J., Tiainen, J., Gotthard, K., & Kivelä, S. M. (2021). Urbanization extends flight phenology and leads to local adaptation of seasonal plasticity in Lepidoptera. Proceedings of the National Academy of Sciences, 118(40), e2106006118. https://doi.org/10.1073/pnas.2106006118

Millette, K. L., Fugère, V., Debyser, C., Greiner, A., Chain, F. J. J., & Gonzalez, A. (2020). No consistent effects of humans on animal genetic diversity worldwide. Ecology Letters, 23(1), 55–67. https://doi.org/10.1111/ele.13394

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K. A., & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. Science, 353(6307), 1532–1535. https://doi.org/10.1126/science.aaf4381

Murdock, A.P., Harfoot, A.J.P., Martin, D., Cockings, S. and Hill, C. (2015) OpenPopGrid: an open gridded population dataset for England and Wales. GeoData, University of Southampton.

Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., & Ernest, S. K. M. (2015). An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles. *Ecology* https://doi.org/10.1890/15-0846R.1

Nacci, D. E., Champlin, D., & Jayaraman, S. (2010). Adaptation of the Estuarine Fish Fundulus heteroclitus (Atlantic Killifish) to Polychlorinated Biphenyls (PCBs). Estuaries and Coasts, 33(4), 853–864. https://doi.org/10.1007/s12237-009-9257-6

Neate-Clegg, M. H. C., Tonelli, B. A., Youngflesh, C., Wu, J. X., Montgomery, G. A., Şekercioğlu, Ç. H., & Tingley, M. W. (2023). Traits shaping urban tolerance in birds differ around the world. Current Biology, 33(9), 1677-1688.e6. https://doi.org/10.1016/j.cub.2023.03.024

Nelson, Robert K., LaDale Winling, et al. "Mapping Inequality: Redlining in New Deal America." Edited by Robert K. Nelson. American Panorama: An Atlas of United States History, 2023. https://dsl.richmond.edu/panorama/redlining.

Olden, J. D., Poff, N. L., Douglas, M. R., Douglas, M. E., & Fausch, K. D. (2004). Ecological and evolutionary consequences of biotic homogenization. Trends in Ecology & Evolution, 19(1), 18–24. https://doi.org/10.1016/j.tree.2003.09.010

Olsen, K., Holm, T. E., Pape, T., & Simonsen, T. J. (2020). Natural history museum collection and citizen science data show advancing phenology of Danish hoverflies (Insecta: Diptera, Syrphidae) with increasing annual temperature. PLOS ONE, 15(5), e0232980. https://doi.org/10.1371/journal.pone.0232980

Owen, H. L., Meng, F., & Winchell, K. M. (2024). Urbanization and environmental variation drive phenological changes in the spotted lanternfly, Lycorma delicatula (Hemiptera: Fulgoridae). Biological Journal of the Linnean Society, 143(4), blae099. https://doi.org/10.1093/biolinnean/blae099

Oziolor, E. M., Bigorgne, E., Aguilar, L., Usenko, S., & Matson, C. W. (2014). Evolved resistance to PCB- and PAH-induced cardiac teratogenesis, and reduced CYP1A activity in Gulf killifish (*Fundulus grandis*) populations from the Houston Ship Channel, Texas. *Aquatic Toxicology*, *150*, 210–219. https://doi.org/10.1016/j.aquatox.2014.03.012

Paz-Vinas, I., Jensen, E. L., Bertola, L. D., Breed, M. F., Hand, B. K., Hunter, M. E., Kershaw, F., Leigh, D. M., Luikart, G., Mergeay, J., Miller, J. M., Van Rees, C. B., Segelbacher, G., & Hoban, S. (2021). Macrogenetic studies must not ignore limitations of genetic markers and scale. *Ecology Letters*, *24*(6), 1282–1284. https://doi.org/10.1111/ele.13732

Paz-Vinas, I., Vandergast, A. G., Schmidt, C., Leigh, D. M., Blanchet, S., Clark, R. D., Crandall, E. D., De Kort, H., Falgout, J., Garroway, C. J., Karachaliou, E., Kershaw, F., O'Brien, D., Pinsky, M. L., Segelbacher, G., Toczydlowski, R. H., & Hunter, M. E. (2025). Sparse genetic data limit biodiversity assessments in protected areas globally. *Frontiers in Ecology and the Environment*, *23*(8), e2867. https://doi.org/10.1002/fee.2867

Pearson, K. D., Nelson, G., Aronson, M. F. J., Bonnet, P., Brenskelle, L., Davis, C. C., Denny, E. G., Ellwood, E. R., Goëau, H., Heberling, J. M., Joly, A., Lorieul, T., Mazer, S. J., Meineke, E. K., Stucky, B. J., Sweeney, P., White, A. E., & Soltis, P. S. (2020). Machine Learning Using Digitized Herbarium Specimens to Advance Phenological Research. BioScience, 70(7), 610–620. https://doi.org/10.1093/biosci/biaa044

Pernat, N., August, T., Groom, Q. J., Memedemin, D., & Reyserhove, L. (2022). An iNaturalist-Pl@ntNet-workflow to identify plant-pollinator interactions – a case study of Isodontia mexicana. https://doi.org/10.37044/osf.io/em3rk

Phelan, P. E., Kaloush, K., Miner, M., Golden, J., Phelan, B., Iii, H. S., & Taylor, R. A. (2015). Urban Heat Island: Mechanisms, Implications, and Possible Remedies. *Annual Review of Environment and Resources*, *40*(Volume 40, 2015), 285–307. https://doi.org/10.1146/annurev-environ-102014-021155

Prudic, K., Oliver, J., Brown, B., & Long, E. (2018). Comparisons of Citizen Science Data-Gathering Approaches to Evaluate Urban Butterfly Diversity. Insects, 9(4), 186. https://doi.org/10.3390/insects9040186

Putman, B. J., Williams, R., Li, E., & Pauly, G. B. (2021). The power of community science to quantify ecological interactions in cities. Scientific Reports, 11(1), 3069. https://doi.org/10.1038/s41598-021-82491-y

Reeb, R. A., Aziz, N., Lapp, S. M., Kitzes, J., Heberling, J. M., & Kuebbing, S. E. (2022). Using Convolutional Neural Networks to Efficiently Extract Immense Phenological Data From Community Science Images. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.787407

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., Karchner, S. I., Hahn, M. E., Nacci, D., Oleksiak, M. F., Crawford, D. L., & Whitehead, A. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, *354*(6317), 1305–1308. https://doi.org/10.1126/science.aah4993

Ringland, J., Bohm, M., Baek, S.-R., & Eichhorn, M. (2021). Automated survey of selected common plant species in Thai homegardens using Google Street View imagery and a deep neural network. Earth Science Informatics, 14(1), 179–191. https://doi.org/10.1007/s12145-020-00557-3

Russell, S. H., Hancock, M. P., & McCullough, J. (2007). Benefits of Undergraduate Research Experiences. Science, 316(5824), 548–549. https://doi.org/10.1126/science.1140384

Ružanović, L., & Mičetić Stanković, V. (2023). Urban Beetle Diversity in Natural History Collections—A Hundred-Year Perspective. Diversity, 15(12), 1224. https://doi.org/10.3390/d15121224

Sabatini, F. M., Lenoir, J., Hattab, T., Arnst, E. A., Chytrý, M., Dengler, J., De Ruffray, P., Hennekens, S. M., Jandt, U., Jansen, F., Jiménez-Alfaro, B., Kattge, J., Levesley, A., Pillar, V. D., Purschke, O., Sandel, B., Sultana, F., Aavik, T., Aćić, S., … Bruelheide, H. (2021). sPlotOpen – An environmentally balanced, open-access, global dataset of vegetation plots. Global Ecology and Biogeography, 30(9), 1740–1764. https://doi.org/10.1111/geb.13346

Santangelo, J. S., Ness, R. W., Cohan, B., Fitzpatrick, C. R., Innes, S. G., Koch, S., Miles, L. S., Munim, S., Peres-Neto, P. R., Prashad, C., Tong, A. T., Aguirre, W. E., Akinwole, P. O., Alberti, M., Álvarez, J., Anderson, J. T., Anderson, J. J., Ando, Y., Andrew, N. R., … Johnson, M. T. J. (2022). Global urban environmental change drives adaptation in white clover. *Science*, *375*(6586), 1275–1281. https://doi.org/10.1126/science.abk0989

Santini, L., González-Suárez, M., Russo, D., Gonzalez-Voyer, A., von Hardenberg, A., & Ancillotto, L. (2019). One strategy does not fit all: Determinants of urban adaptation in mammals. Ecology Letters, 22(2), 365–376. https://doi.org/10.1111/ele.13199

Santini, L., Mendez Angarita, V. Y., Karoulis, C., Fundarò, D., Pranzini, N., Vivaldi, C., Zhang, T., Zampetti, A., Gargano, S. J., Mirante, D., & Paltrinieri, L. (2024). TetraDENSITY 2.0—A

Database of Population Density Estimates in Tetrapods. *Global Ecology and Biogeography*, *33*(12), e13929. https://doi.org/10.1111/geb.13929

Savage, A. M., Willmott, M. J., Moreno-García, P., Jagiello, Z., Li, D., Malesis, A., Miles, L. S., Román-Palacios, C., Salazar-Valenzuela, D., Verrelli, B. C., Winchell, K. M., Alberti, M., Bonilla-Bedoya, S., Carlen, E., Falvey, C., Johnson, L., Martin, E., Kuzyo, H., Marzluff, J., … Gotanda, K. M. (2024). Online toolkits for collaborative and inclusive global research in urban evolutionary ecology. *Ecology and Evolution*, *14*(6), e11633. https://doi.org/10.1002/ece3.11633

Schell, C. J., Dyson, K., Fuentes, T. L., Des Roches, S., Harris, N. C., Miller, D. S., Woelfle-Erskine, C. A., & Lambert, M. R. (2020). The ecological and evolutionary consequences of systemic racism in urban environments. Science, 369(6510), eaay4497. https://doi.org/10.1126/science.aay4497

Schmidt, C., Domaratzki, M., Kinnunen, R. P., Bowman, J., & Garroway, C. J. (2020). Continent-wide effects of urbanization on bird and mammal genetic diversity. Proceedings of the Royal Society B: Biological Sciences, 287(1920), 20192497. https://doi.org/10.1098/rspb.2019.2497

Schmidt, C., & Garroway, C. J. (2021). The population genetics of urban and rural amphibians in North America. Molecular Ecology, 30(16), 3918–3929. https://doi.org/10.1111/mec.16005

Schmidt, C., & Garroway, C. J. (2022). Systemic racism alters wildlife genetic diversity. Proceedings of the National Academy of Sciences, 119(43), e2102860119. https://doi.org/10.1073/pnas.2102860119

Schmidt, C., Hoban, S., & Jetz, W. (2023). Conservation macrogenetics: Harnessing genetic data to meet conservation commitments. *Trends in Genetics*, *39*(11), 816–829. https://doi.org/10.1016/j.tig.2023.08.002

Şekercioğlu, Ç. H., Kittelberger, K. D., Mota, F. M. M., Buxton, A. N., Orton, N., DeNiro, A., Buechley, E. R., Horns, J. J., Blount, J. D., Socci, J., & Neate-Clegg, M. H. C. (2025). BIRDBASE: A Global Dataset of Avian Biogeography, Conservation, Ecology and Life History Traits. *Scientific Data*, *12*(1), 1558. https://doi.org/10.1038/s41597-025-05615-3

Shapiro, C., Moberg-Parker, J., Toma, S., Ayon, C., Zimmerman, H., Roth-Johnson, E. A., Hancock, S.

P., Levis-Fitzgerald, M., & Sanders, E. R. (2015). Comparing the Impact of Course-Based and Apprentice-Based Research Experiences in a Life Science Laboratory Curriculum. Journal of Microbiology & Biology Education, 16(2), 186–197. https://doi.org/10.1128/jmbe.v16i2.1045

Shultz, A. J., Adams, B. J., Bell, K. C., Ludt, W. B., Pauly, G. B., & Vendetti, J. E. (2020). Natural history collections are critical resources for contemporary and future studies of urban evolution. Evolutionary Applications, 14(1), 233–247. https://doi.org/10.1111/eva.13045

Siefert, A., Laughlin, D. C., & Sabatini, F. M. (2024). You shall know a species by the company it keeps: Leveraging co-occurrence data to improve ecological prediction. Journal of Vegetation Science, 35(6), e13314. https://doi.org/10.1111/jvs.13314

Stadnicki, I., Corsini, M., & Szulkin, M. (2024). Application of criminology in urban ecology and evolution: Routine Activity Theory and field equipment disappearance dynamics. *Ecological Indicators*, *165*, 112095. https://doi.org/10.1016/j.ecolind.2024.112095

Steel, E. A., Kennedy, M. C., Cunningham, P. G., & Stanovick, J. S. (2013). Applied statistics in ecology: Common pitfalls and simple solutions. *Ecosphere*, *4*(9), art115. https://doi.org/10.1890/ES13-00160.1

Stewart, R. D., Bard, N. W., van der Bank, M., & Davies, T. J. (2025). Leveraging machine learning and citizen science data to describe flowering phenology across South Africa. *PLANTS, PEOPLE, PLANET*, *n/a*(n/a). https://doi.org/10.1002/ppp3.70059

Templ, B., Koch, E., Bolmgren, K., Ungersböck, M., Paul, A., Scheifinger, H., Rutishauser, T., Busto, M., Chmielewski, F.-M., Hájková, L., Hodzić, S., Kaspar, F., Pietragalla, B., Romero-Fresneda, R., Tolvanen, A., Vučetič, V., Zimmermann, K., & Zust, A. (2018). Pan European Phenological database (PEP725): A single point of access for European data. *International Journal of Biometeorology*, *62*(6), 1109–1113. https://doi.org/10.1007/s00484-018-1512-8

Tobias, J. A., Sheard, C., Pigot, A. L., Devenish, A. J. M., Yang, J., Sayol, F., Neate-Clegg, M. H. C., Alioravainen, N., Weeks, T. L., Barber, R. A., Walkden, P. A., MacGregor, H. E. A., Jones, S. E. I., Vincent, C., Phillips, A. G., Marples, N. M., Montaño-Centellas, F. A., Leandro-Silva, V., Claramunt, S., … Schleuning, M. (2022). AVONET: Morphological, ecological and geographical data for all birds. Ecology Letters, 25(3), 581–597. https://doi.org/10.1111/ele.13898

Tucker, M. A., Böhning-Gaese, K., Fagan, W. F., Fryxell, J. M., Van Moorter, B., Alberts, S. C., Ali, A. H., Allen, A. M., Attias, N., Avgar, T., Bartlam-Brooks, H., Bayarbaatar, B., Belant, J. L., Bertassoni, A., Beyer, D., Bidner, L., van Beest, F. M., Blake, S., Blaum, N., … Mueller, T. (2018). Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science*, *359*(6374), 466–469. https://doi.org/10.1126/science.aam9712

Tucker, M. A., Santini, L., Carbone, C., & Mueller, T. (2020). Mammal population densities at a global scale are higher in human-modified areas. *Ecography*, *44*(1), 1–13. https://doi.org/10.1111/ecog.05126

Tyszka, A. S., Larson, D. A., & Walker, J. F. (2025). Sequencing historical RNA: Unrealized potential to increase understanding of the plant tree of life. *Trends in Plant Science*, *30*(7), 705–711. https://doi.org/10.1016/j.tplants.2024.11.004

van den Burg, M. P., & Vieites, D. R. (2023). Bird genetic databases need improved curation and error reporting to NCBI. Ibis, 165(2), 472–481. https://doi.org/10.1111/ibi.13143

Velasquez-Camacho, L., Merontausta, E., Etxegarai, M., & de-Miguel, S. (2024). Assessing urban forest biodiversity through automatic taxonomic identification of street trees from citizen science applications and remote-sensing imagery. International Journal of Applied Earth Observation and Geoinformation, 128, 103735. https://doi.org/10.1016/j.jag.2024.103735

Visser, M. E., & Both, C. (2005). Shifts in phenology due to global climate change: The need for a yardstick. Proceedings of the Royal Society B: Biological Sciences, 272(1581), 2561–2569. https://doi.org/10.1098/rspb.2005.3356

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. Nature, 416(6879), 389–395. https://doi.org/10.1038/416389a

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially

Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6), e0156720. https://doi.org/10.1371/journal.pone.0156720

Weaver, W. N., & Smith, S. A. (2023). From leaves to labels: Building modular machine learning networks for rapid herbarium specimen analysis with LeafMachine2. *Applications in Plant Sciences*, *11*(5), e11548. https://doi.org/10.1002/aps3.11548

Weeks, B. C., Zhou, Z., O'Brien, B. K., Darling, R., Dean, M., Dias, T., Hassena, G., Zhang, M., & Fouhey, D. F. (2023). A deep neural network for high-throughput measurement of functional traits on museum skeletal specimens. *Methods in Ecology and Evolution*, *14*(2), 347–359. https://doi.org/10.1111/2041-210X.13864

Willis, C. G., Ellwood, E. R., Primack, R. B., Davis, C. C., Pearson, K. D., Gallinat, A. S., Yost, J. M., Nelson, G., Mazer, S. J., Rossington, N. L., Sparks, T. H., & Soltis, P. S. (2017). Old Plants, New Tricks: Phenological Research Using Herbarium Specimens. Trends in Ecology & Evolution, 32(7), 531–546. https://doi.org/10.1016/j.tree.2017.03.015

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Wilman, H., Belmaker, J., Simpson, J., Rosa, C. de la, Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology*. https://doi.org/10.1890/13-1917.1

Winchell, K. M., Campbell-Staton, S. C., Losos, J. B., Revell, L. J., Verrelli, B. C., & Geneva, A. J. (2023). Genome-wide parallelism underlies contemporary adaptation in urban lizards. Proceedings of the National Academy of Sciences, 120(3), e2216789120. https://doi.org/10.1073/pnas.2216789120

Winchell, K. M., Losos, J. B., & Verrelli, B. C. (2023). Urban evolutionary ecology brings exaptation back into focus. Trends in Ecology & Evolution, 38(8), 719–726. https://doi.org/10.1016/j.tree.2023.03.006

Winchell, K. M., Schliep, K. P., Mahler, D. L., & Revell, L. J. (2020). Phylogenetic signal and evolutionary correlates of urban tolerance in a widespread neotropical lizard clade*. Evolution, 74(7), 1274–1288. https://doi.org/10.1111/evo.13947

Wohlfahrt, G., Tomelleri, E., & Hammerle, A. (2019). The urban imprint on plant phenology. Nature Ecology & Evolution. https://doi.org/10.1038/s41559-019-1017-9

Ye, Y., Qiu, H., Zhou, X., & Li, X. (2025). Urban landscape pattern affects plant diversity in green spaces: Evidence from street view imagery. Global Ecology and Conservation, 57, e03395. https://doi.org/10.1016/j.gecco.2024.e03395

Figure 1. From genes to ecosystems, examples of publicly available datasets that have enabled urban ecology and evolution studies to be conducted across multiple biological questions. Here, we illustrate examples of studies that have assessed the effects of urbanization using publicly available datasets. The schematic depicts the dominant pattern that emerges with increasing urbanization. The corresponding research questions addressed by these studies can be found in **Table 1**.

Table 1. This table summarizes the types of research questions that have been addressed across various biological questions and provides representative examples. Questions in bold deserve greater exploration using publicly available datasets. It also summarizes the resources used to obtain the biological data analyzed in the studies. Often, these data serve as response variables in models evaluating the effect of urbanization on ecological and evolutionary processes.

Topic	Subject	Questions	Biological Data Sources	Examples
Ecosystems and Biodiversity	Community assemblage and biodiversity distribution	• How does species diversity vary across urbanization gradients?[1] • Are patterns of urban biodiversity related to socioeconomic or sociodemographic variables or historic redlining?[2] • Where are the biodiversity hotspots on urban landscapes?[3]	iNaturalist, Google Street View, eBird, Birds of the World, Norfolk Bat Survey, Global Biodiversity Information Facility (GBIF), Australian Department of the Environment	Beninde et al., 2023[3]; Border et al., 2017[1,3]; Callaghan, Bino, et al., 2019[1]; Estien et al., 2024[2]; Fitch, Wilson, et al., 2019[1]; Ives et al., 2016[3]; Kinnunen et al., 2024[1,2]; Leong and Trautwein, 2019[2]; E. Li et al., 2019[1,5]; Prudic et al., 2018[4]; RužanoviĆ and StankoviĆ, 2023[1]; Velasquez-Camacho et al., 2024[3]
	Ecosystem function	• How does urbanization interact with other stressors (e.g., drought) to affect urban vegetation health and resilience?[1] • Are patterns of vegetation health and resilience related to socioeconomic or sociodemographic variables?[2]	NASA (National Aeronautics and Space Administration), ESA (European Space Agency), Copernicus	Dong et al., 2023[1,2]; Leisenheimer et al., 2024[1];
	Phenological Change	• How does urbanization affect the timing and duration of community-level phenological events?[1] • Does urbanization result in phenological shifts?[2] • **Does tree species diversity explain growing season length in urban forests?** • **Do phenological shifts affect species-interactions in urban areas?**	iNaturalist, pan-European Phenology Project (PEP), national Phenology Network (NPN), GBIF, Previously published phylogeny, Open Tree of Life, Digitized Museum Specimens	D. Li et al., 2021[1]; Owen et al., 2024[2]; Wohlfahrt et al., 2019[1,2]; Merckx et al., 2021[2]; Li et al., 2017[1]
Species Interactions	Plant-animal and animal-animal interactions	• How does urbanization affect plant-pollinator networks?[1] • How does urbanization affect plant-herbivore interactions?[2] • How does urbanization affect predation?[3] • Do interactions change through time?[4] • **Do urban gardens provide pollinator resources to native insects?**	iNaturalist, eBird, GBIF, Digitized Museum Specimens	Marín-Gómez et al., 2022[1]; Meineke et al., 2019[2,4]; Putman et al., 2021[3]
Species and Populations	Urban Tolerance	• Does urbanization result in biotic homogenization of species composition?[1] • Does urbanization act as a filter to select for tolerant species?[2] • Is there phylogenetic signal in urban tolerance?[3] • Which species traits confer urban tolerance?[4] • Is urban tolerance generalizable across cities?[5] • **How does urban tolerance vary by taxonomic group?** • **Does urbanization filter for reproductive strategies in plants?**	Amniote Life-History database, AnAge, EtlonTrait, PanTHERIA, Previously published phylogeny, eBird, GBIF, iNaturalist, AVONET, BirdBase, Birds of the World, Bird Life	Hensley et al., 2019[1,2,4,5]; Winchell et al., 2020[3,4]; Callaghan, Major, et al., 2019[2,4]; Neate-Clegg et al., 2023[2,3,4]; Kinnunen et al., 2022[1,3,5]; Santini et al., 2019[3,4]; Curti et al., 2024[5]
	Adaptations	• Is there evidence of heat stress adaptation among urban populations?[1] • How does urbanization affect body size?[2] • How does urbanization affect color patterning?[3]	iNaturalist, VertNet, National Ecological Observatory Network (NEON), North American Census of Small Mammals (NACSM)	Cosentino and Gibbs, 2022[3]; Fukano et al., 2023[1,3]; Hantak et al., 2021[2,4.]
	Plastic responses	• Do traits respond plastically to urbanization?[1] • **Do traits exhibit a greater degree of plasticity in urban environments?**	iNaturalist, TRY, Nocturna, Swedish and Finnish Butterfly Monitoring Scheme, Finnish Biodiversity Info Facility, BugGuide	Batool et al., 2024[1]; Ibsen et al., 2023[1]; Merckx et al., 2021[1]
	Genetic Diversity	• How does urbanization affect population genetic parameters like genetic diversity, N_e, F_{ST}, allelic richness, etc.)[1] • Have population genetic parameters changed through time in urban areas?[2] • Is genetic diversity related to socioeconomic or sociodemographic variables?[3] • **How does urbanization affect inbreeding in wild populations?**	Dryad, GenBank, DataOne, BOLD, Sequence Read Archive, MacroPopGen	Millette et al., 2020[1,2]; Miraldo et al., 2016[1]; Schmidt et al., 2020, 2021[1]; Schmidt and Garroway, 2022[3]; Habrich et al., 2021[1]; Karachaliou et al., 2025[1]

Table 2. This table summarizes the data sources for common predictor variables used in models assessing urban ecological and evolutionary processes. It also lists the specific studies that the data sources were used in. Studies may use more than one data source to capture different aspects of urbanization.

Predictor Variable Category	**Sources**	**Examples**
Urbanization	Ocean Health Index Cumulative Human Impacts[1], National Land Cover Database Impervious Surface (NLCD)[2], Global Man-made Impervious Surface (GMIS)[3], 2010 Global Tree Cover Dataset[4], VIIRS Nighttime Light[5], Human Population Density[6], Human Footprint Index[7], Anthropogenic biomes or Anthromes[8], History database of the Global Environment (HYDE)[9], National Park Service Noise Pollution[10], US National Transportation Noise Map[11], Copernicus Land Monitoring Service (CLMS) impervious surface[12], European Settlement Map[13], Gridded Population of the World[14], Defense Meteorological Satellite Program Operational Linescan System[15]; OpenPopGrid[16], Australian Bureau of Statistics[17], OpenDataSoft[18]; World Urban Areas[19]	Batool et al., 2024[6]; Callaghan, Major, et al., 2019[5]; Cosentino and Gibbs, 2022[3,4]; Curti et al., 2024[2,5,10]; Estien et al., 2024[2]; Fitch, Wilson, et al., 2019[2]; Hantak et al., 2021[6]; Karachaliou et al., 2025; Leong and Trautwein, 2019[2]; D. Li et al., 2021[6]; E. Li et al., 2019[2, 6, 11]; Meineke et al., 2019[6]; Millette et al., 2020[9]; Miraldo et al., 2016[8]; Owen et al., 2024[2]; Putman et al., 2021[2]; Schmidt et al., 2020, 2021[7]; Schmidt and Garroway, 2022[7]; Wohlfahrt et al., 2019[12,13]; Habrich et al., 2021[14]; Li et al., 2017[15], Border et al., 2017[16], Neate-Clegg et al., 2023[5,14,18], Santini et al., 2019[19]
Climate (current and projected)	WorldClim[1], CHELSA[2], ClimateNA[3], PRISM[4], NASA NEX-DCP30[5], NASA NEX-GDDP[6], Basin Climate Characterization Model (BCM)[7], Environment Canda[8], TerraClimate, Palmer Drought Severity Index[9]	Batool et al., 2024[8]; Beninde et al., 2023[1,2,3]; Cosentino and Gibbs, 2022[1]; Hantak et al., 2021[4]; D. Li et al., 2021[4,5,7]; Meineke et al., 2020[4]; Owen et al., 2024[1]; Winchell et al., 2020[1]; Dong et al., 2019[4,9]
Land Use and Land Cover	National Land Cover Database (NLCD)[1], MODIS[2], Landsat[3], CORINE Land Cover[4], USGS Hydrography Dataset[5]; Copernicus Land Monitoring Service (CLMS)[6], Global Roads Open Access Dataset (gROADS)[7], Highway Performance Monitoring System[8], Mapping Inequality database (historical HOLC maps)[9]	Callaghan, Major, et al., 2019[2,3]; Callaghan, Bino, et al., 2019[1]; Estien et al., 2025[9]; E. Li et al., 2019[1, 3, 5]; RužanoviĆ and StankoviĆ, 2023[4]; Wohlfahrt et al., 2019[1,6], Habrich et al., 2021[7,8], RužanoviĆ and StankoviĆ, 2023[4]
Socio-demographic and Socio-economic	United States Census Bureau[1], Statistics Canada[2], CalEnviroScreen[3], CEJST, IPUMS[4]	Batool et al., 2024[2]; Kinnunen et al., 2025, 2022[1]; E. Li et al., 2019[1,3]; Schmidt et al., 2021, 2020[1,2]; Schmidt and Garroway, 2022[1,2,4]

Table 3. This table presents emerging and underutilized datasets that may prove useful to the field of urban ecology and evolution. It also summarizes the information housed in each repository and the corresponding link address or citation.

Repository or Dataset	**Data Type or Information**	**Citation/Link**
MorphoSource	A repository of 3D and 2D media of natural history specimens	https://www.morphosource.org/

Digitized Natural History Collections: iDigBio	Digitizied photographs, scans, CT-scans, etc., of preserved natural history specimens	https://www.idigbio.org/
Global Registry of Scientific Collections	A repository of information that connects researchers to data	https://scientific-collections.gbif.org/
Google Street View Images	Photographs of Roadways	https://developers.google.com/maps/documentation/streetview/get-api-key
ArcGIS Living Atlas of the World	Paid application with access to maps and deep-learning packages	https://livingatlas.arcgis.com/en/home/
Movebank	A database to store, collect, analyze, visualize, and archive tracking data of animals.	Kays et al., 2021 https://www.movebank.org/cms/movebank-main
sPlotOpen	Vegetation plot surveys across landscapes	Sabatini et al., 2021 https://www.idiv.de/research/projects/splot/splotopen-splot/
Arctos	Biological data includes specimen measurements, mark/recapture data, species sighting surveys, etc.	Cicero et al., 2024 https://arctosdb.org/
TetraDENSITY 2.0	Global population estimates for tetrapods	Santini et al., 2024 https://andrewzamp.github.io/TetraDENSITY/
National Center for Biotechnology Information	Repositories for genetic data, encompassing genes, transcriptomes, protein sequences, and genomes.	https://www.ncbi.nlm.nih.gov/
Global Genome Biodiversity Network	Facilitates sharing of physically archived DNA and tissue samples. Connects sequence data derived from collections to relevant metadata.	Droege et al., 2016 https://www.ggbn.org/ggbn_portal/
Genomic Observatories Metadatabase	Database that links metadata with sequence information from current and past projects.	Deck et al., 2017 https://geome-db.org/

Appendix 1

Biological Repository/Datasets covered		Type
BugGuide	https://bugguide.net	Citizen Science Platform
eBird	https://ebird.org/home	Citizen Science Platform
FrogID	https://www.frogid.net.au/	Citizen Science Platform

iNaturalist	https://www.inaturalist.org	Citizen Science Platform
iSpot	https://www.ispotnature.org/	Citizen Science Platform
Amniote Life-History database	Myhrvold et al., 2015 https://doi.org/10.6084/m9.figshare.c.3308127.v1	Database
AnAge	de Magalhães et al., 2024 https://genomics.senescence.info/species/index.html	Database
Birds of the World	https://birdsoftheworld.org/bow/home	Database
Copernicus	https://dataspace.copernicus.eu/	Database
National Ecological Observatory Network	https://data.neonscience.org/	Database
National Phenology Network	https://www.usanpn.org/	Database
Pan-European Phenology Project	Templ et al., 2018 http://www.pep725.eu/	Database
VertNet	https://www.vertnet.org/	Database
AVONET	Tobias et al., 2025 https://figshare.com/s/b990722d72a26b5bfead	Dataset
BirdBase	Şekercioğlu et al., 2025 https://doi.org/10.6084/m9.figshare.27051040	Dataset
BirdTree	Jetz et al., 2012 https://birdtree.org/	Dataset
EtlonTrait	Wilman et al., 2014 https://opentraits.org/datasets/elton-traits.html	Dataset
Finnish Moth Monitoring Scheme	https://demo.gbif.org/dataset/8fd3f044-b7f5-4ff0-9d78-9b1881f4942f	Dataset
MacroPopGen	Lawrence et al., 2019 https://doi.org/10.6084/m9.figshare.7207514.v1	Dataset
Open Tree of Life	https://tree.opentreeoflife.org/	Dataset
PanTHERIA	Jones et al., 2009 https://doi.org/10.6084/m9.figshare.c.3301274.v1	Dataset
Swedish and Finnish	https://www.gbif.org/fr/dataset/be77e203-486c-4651-91b9-8347968b728c https://www.gbif.org/dataset/181eab51-9399-4baa-a0df-8de01a3acf19	Dataset

Butterfly Monitoring Scheme		
TRY	Kattge et al., 2011 https://www.try-db.org/TryWeb/Home.php	Dataset
ESA (European Space Agency)	https://earth.esa.int/eogateway/catalog	Organization
NASA (Aeronautics and Space Administration)	https://data.nasa.gov/	Organization
Austrian Department of the Environment	https://www.dcceew.gov.au/environment/environmental-information-data/databases-applications/snes	Repository
Bird Life	https://datazone.birdlife.org/	Repository
BOLD	https://boldsystems.org/	Repository
DataOne	https://www.dataone.org/	Repository
Dryad	https://datadryad.org/	Repository
Finnish Biodiversity Info Facility	https://laji.fi/en	Repository
GBIF	https://www.gbif.org/	Repository
Norfolk Biodiversity Information Service	http://www.nbis.org.uk/	Repository
Predictor Variables Covered		
Climate and Economic Justice Screening Tool (CEJST)	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/B6ULET (archived due to website removal)	Data Tool
CHELSA	https://www.chelsa-climate.org/	Database
Copernicus Land Monitoring Service (CLMS)	https://land.copernicus.eu/en	Database
Highway Performance Monitoring System	https://www.fhwa.dot.gov/policyinformation/hpms.cfm	Database
History database of the Global	https://landuse.sites.uu.nl/datasets/	Database

Environment (HYDE)		
IPUMS	https://www.ipums.org/	Database
Landsat	https://landsat.gsfc.nasa.gov/	Database
Mapping Inequality database (historical HOLC maps)	Nelson et al., 2023 https://dsl.richmond.edu/panorama/redlining/	Database
Meteorological Satellite Program Operational Linescan System	https://www.ncei.noaa.gov/products/dmsp-operational-linescan-system	Database
MODIS	https://modis.gsfc.nasa.gov/	Database
National Land Cover Databases	https://www.usgs.gov/centers/eros/science/nlcd-product-suite	Database
WorldClim	https://www.worldclim.org/	Database
2010 Global Tree Cover Dataset	https://www.earthdata.nasa.gov/data/catalog/lpcloud-gfcc30tc-003	Dataset
Anthropogenic biomes or Anthromes	https://anthroecology.org/anthromes/	Dataset
Basin Climate Characterization Model (BCM)	https://www.usgs.gov/centers/california-water-science-center/science/basin-characterization-model-bcm#overview	Dataset
CalEnviroScreen	https://oehha.ca.gov/calenviroscreen	Dataset
ClimateNA	Wang et al., 2016 https://climatena.ca/	Dataset
CORINE Land Cover	https://land.copernicus.eu/en/products/corine-land-cover	Dataset
European Settlement Map	https://www.eea.europa.eu/en/datahub/datahubitem-view/42783031-34d5-4b38-b9ba-a0783723053c	Dataset
Global Man-made Impervious Surface (GMIS)	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-ulandsat-gmis-v1-1.0	Dataset
Global Roads Open Access Dataset	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-groads-v1-1.0	Dataset
Gridded Population of the World	https://www.earthdata.nasa.gov/data/catalog/sedac-ciesin-sedac-gpwv4-popdens-r11-4.11	Dataset

Human Footprint Index	https://wcshumanfootprint.org/data-access	Dataset
Human Population Density	Fang & Jawitz, 2018 https://doi.org/10.6084/m9.figshare.c.3890191	Dataset
NASA NEX-DCP30	https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-dcp30	Dataset
NASA NEX-GDDP	https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp	Dataset
National Park Service Noise Pollution	https://irma.nps.gov/DataStore/Reference/Profile/2217356	Dataset
Ocean Health Index Cumulative Human Impacts	https://oceanhealthindex.org/resources/data/	Dataset
OpenDataSoft	https://public.opendatasoft.com/explore/dataset/geonames-all-cities-with-a-population-1000/table/?disjunctive.cou_name_en&sort=name	Dataset
OpenPopGrid	Murdock et al., 2015 https://openpopgrid.geodata.soton.ac.uk/	Dataset
TerraClimate	Abatzoglou et al., 2018 https://www.climatologylab.org/terraclimate.html	Dataset
US National Transportation Noise Map	https://maps.dot.gov/BTS/NationalTransportationNoiseMap/	Dataset
USGS Hydrography Dataset	https://www.usgs.gov/3d-hydrography-program/access-3dhp-data-products	Dataset
VIIRS Nighttime Light	https://eogdata.mines.edu/products/vnl/	Dataset
World Urban Areas 2012	Kelso & Patterson. 2012 https://purl.stanford.edu/yk247bg4748	Dataset
PRISM	https://prism.oregonstate.edu/	Datasets
Australian Bureau of Statistics	https://www.abs.gov.au/	Organization
Environment Canda	https://climate.weather.gc.ca/	Organization
Statistics Canada	https://www.statcan.gc.ca/en/start	Organization
United States Census Bureau	https://data.census.gov/	Organization
Emerging and Underutilized Repositories/Datasets		

Genomic Observatories Metadatabase (GeOMe)	Deck et al., 2017 https://geome-db.org/	Database
Google Street View	https://developers.google.com/maps/documentation/streetview/get-api-key	Database
iDigBio	https://www.idigbio.org/	Database
Movebank	Kays et al., 2021 https://www.movebank.org/cms/movebank-main	Database
The Global Registry of Scientific Collections	https://scientific-collections.gbif.org/	Database
sPlotOpen	Sabatini et al., 2021 https://idata.idiv.de/ddm/Data/ShowData/3474?version=55	Dataset
TetraDENSITY 2.0	Santini et al., 2024 https://andrewzamp.github.io/TetraDENSITY/	Dataset
ArcGIS Living Atlas of the World	https://livingatlas.arcgis.com/en/home/	Repository
Arctos	Cicero et al., 2024 https://arctosdb.org/	Repository
Global Genome Biodiversity Network	Droege et al., 2016 https://www.ggbn.org/ggbn_portal/	Repository
MorphoSource	https://www.morphosource.org/	Repository
National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov/	Repository
Other helpful resources		
Opportunistic Database of Biodiversity Databases	https://earthskysea.org/biodiversity-databases/	Database
Urban Eco-Evo Toolkit, Trait Mapping Database	Savage et al., 2024 https://www.urbanevoecotools.org/	Database