Suitability of open digital species records for assessing biodiversity patterns in cities: a case study using avian records

Marufa Sultana * and Ilse Storch

Chair of Wildlife Ecology and Management, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

*Corresponding author. E-mail: marufa.sultana@wildlife.uni-freiburg.de, marufa_sultana@hotmail.com

Submitted: 22 October 2020; Received (in revised form): 3 March 2021. Accepted: 15 April 2021

Abstract

Openly available species observation records on various online platforms achieve good coverage in urban areas. Thus, such digital data could provide a basis for biodiversity assessments in cities. Here, we investigated suitability of open digital species occurrence data, compared with systematically field surveyed data, in Freiburg, Germany (a Western city) and Dhaka, Bangladesh (a global-South city). We focused on resident bird species richness as an indicator of local biodiversity. We collected avian records for urban areas from ‘ornitho.de’ in Freiburg and ‘gbif.org’ in Dhaka. Additionally, we conducted point count surveys at several urban locations in both cities. Using these records, we prepared three grid (cell size 250 m × 250 m) based datasets—open digital dataset (i.e. records compiled from well-surveyed grid cells), field surveyed dataset (i.e. records of systematic bird surveys) and combined dataset (i.e. digital data and field data combined). We compared the relationship of resident bird richness with different habitat factors by applying linear regression models, separately using each of the three datasets. We assessed suitability of data from online platforms by comparing the variables retained after model selection based on digital data versus field surveyed data. We found that field surveyed data and combined data did not alter general understanding of the key driving factors of bird richness patterns we obtained from open digital data. This held for both case examples, Freiburg vs. Dhaka, respectively. This suggests that open digital data from well-surveyed urban locations can provide a suitable basis to assess drivers of biodiversity patterns within cities.

Key words: GBIF, Ornitho.de, bird, Freiburg, Dhaka, urbanization

Introduction

Consistent data on occurrence and distribution of species underpin biodiversity conservation planning. A lack of such data leads to information gaps and affects the understanding of spatial and temporal biodiversity patterns (Soria-Auza and Kessler 2008; Amano et al. 2016). The enormous amounts of species observation records available from a variety of online platforms can be useful to minimize such gaps. Since the coverage of species occurrence reporting at online platforms is generally higher in urban areas than in rural sites (Tye et al. 2017; Speed et al. 2018; Petersen et al. 2021), readily available digital species records hold a huge potential to advance biodiversity research and conservation in cities. In fact, open digital data are increasingly advanced in urban settings, particularly to assess and monitor population trends (e.g. of birds and butterflies), to model species distribution, and to measure species specific responses to urbanization (Wang Wei et al. 2016; Callaghan et al. 2020). Nevertheless, digital species observation records commonly remain unused due to potential biases (Sousa-Baena et al. 2014), and representation of such data in urban ecological research is below expectation (Wang Wei et al. 2016). Besides, as indicated by a lack of data and reports in developed vs developing regions...
In this article, we compared species observation records from well-surveyed sites acquired from open digital sources, with species records collected using standard field survey protocols (Bibby et al. 2000; Gregory et al. 2004), for biodiversity pattern assessments in two cities—Dhaka in Bangladesh and Freiburg in Germany. We assessed the species richness of birds in Dhaka and Freiburg with high IC representing sufficient survey effort, where an additional observation/survey is likely unnecessary (Sousa-Baena et al. 2014). Presumably, species records at such sites are complete enough to identify local differences in species richness, and to assess explanatory factors for such patterns.

In this article, we compared species observation records from well-surveyed sites acquired from open digital sources, with species records collected using standard field survey protocols (Bibby et al. 2000; Gregory et al. 2004), for biodiversity pattern assessments in two cities—Dhaka in Bangladesh and Freiburg in Germany. We assessed the species richness of birds (i.e. as an indicator of biodiversity), since avian records are the most prominent in comparison to all other taxa within the currently available digital databases (Chandler et al. 2017; Freeman and Peterson 2019). Moreover, birds are one of the most studied taxon in urban ecology (Strohbach et al. 2009; Trimble and van Aarde 2012; Li et al. 2019). Our study focused on urban areas, since digitally available species observation records are more frequent and pronounced there than in rural sites; thus, online portals provide better spatial coverage in settled areas (Tye et al. 2017; Speed et al. 2018). Further, we chose cities on two continents, because sites in developed countries (e.g. in North America, Europe, Australia) tend to be better represented in online portals than sites in developing countries of the Global South (Amano et al. 2016; Chandler et al. 2017).

Our aim was to investigate whether avian species occurrence records from well-surveyed urban sites available on online portals can be suitable for assessing patterns of bird species richness in relation to habitat factors in urban areas, as commonly used in the context of monitoring and planning of urban biodiversity. We infer that digital data are suitable, if the predictors of species richness patterns derived from it are consistent with those from data collected using standard field methods. Specifically, we expected to find bird species richness to decline in relation to increasing imperviousness (Silva et al. 2015; Sultana et al. 2021), whereas high bird species richness associates with high proportions of green structures (Evans et al. 2009; Mayurga et al. 2020).

**Methods**

**Study area**

We conducted our investigation in two cities, Freiburg and Dhaka, to cover two alternate scenarios in availability of open digital data. For Freiburg, as is typical for Western world cities, plenty of digital species occurrence records are available from online platforms, whereas we considered Dhaka as a model of cities located in the global South where digital data collection retains a deficiency. For our intended study, systematic field surveyed data were easily accessible in both cities.

Freiburg is a medium-sized Western European city with a population of ~230 000, located in southern Germany (Statistisches Landesamt Baden-Württemberg 2018). Within the administrative limits (153 km²), this city boasts 47.1% green space of which ~32% is forest cover (Beatley 2000; Medearis and Daseking 2012). It is recognized as the model of green and sustainably cities globally (Buehler et al. 2011) and is a popular site for bird watching for local ornithologists.

In contrast, Dhaka is a megalcity located in Southern Asia with a population of ~14 million (www.citypopulation.de) and it is the capital of Bangladesh. We focused on the central city region of greater Dhaka which has a surface area of 306 km², of which ~117 km² are built-up areas (Khaleda et al. 2017). A large extent of farmland exists within the boundary of the city area (about 64 km²) and its surroundings (around 193 km² at 5 km buffered area) (GISAT 2011); farmland extent has decreased due to increasing urbanization in the last few decades (Dewan et al. 2012; Khaleda et al. 2017). Moreover, only 8% tree coverage remains within Dhaka, due to a rapid increase in built-up areas (Islam 2001; Byomkesh et al. 2012).

**Compilation of bird records**

For both Freiburg and Dhaka, we collected readily available bird occurrence records from digital platforms (‘ornitho.de’ in Freiburg; ‘gbif.org’ in Dhaka). Additionally, we surveyed birds in the field at pre-selected urban sites using standard protocols. We compiled three different datasets: (i) an open digital dataset comprising bird observation records from online platforms; (ii) a
field dataset comprising bird observation records from systematic bird surveys; (iii) a combined dataset of bird records from both open digital sources and field surveys.

Following a grid-based data aggregation approach, we mapped all bird observation records from each dataset in ArcMap (version 10.5.1). We used the spatial join tool to allocate each bird record to one of the grid cells, at 250 m × 250 m scale extent, across the urban area of the cities, and excluded any cells which were predominantly forest sites/agricultural areas (Figs 1 and 2).

**Open digital dataset comprising bird observation records from online platforms**

For Freiburg, digital bird records from a 5-year period, 2012–6, were collected from the ‘Ornitho.de’ database (www.ornitho.de). These consisted of 48 471 observation records of birds in Freiburg city. Each record comprised of a species occurrence at a geographic location along with the number of individuals observed. Ornitho.de is considered the main digital platform for avifaunal data in Germany, whose legal entity is DDA (Dachverband Deutscher Avifaunisten). This database is associated with different partner organizations, working groups and specialist groups, facilitating volunteer bird observation at the local level. Dhaka lacks a regional database platform comparable to Ornitho.de. Here, we considered digital bird records from GBIF: GBIF (www.gbif.org) is the largest global online platform of species occurrences records for different taxa. It contains an enormous amount of avian records, most of which come from eBird (www.ebird.org), the largest online platforms for bird observations. We examined all available species occurrence records for Dhaka in GBIF (gbif.org 2019). Aves occurrence observations. We examined all available species occurrence records for an extended period of 1995–2017. In total, we found only 1062 bird records for Dhaka city; each record comprised of a species incidence at a geographic location.

We considered only human observation records; any museum specimen records were removed. We checked and removed records with geospatial issues (e.g. if recorded coordinate, geodetic datum or coordinate precision appeared invalid). Additionally, we cleaned the datasets to remove any potential species identification errors in consultation with local ornithologists. We considered only the records which contained species-level identifications of birds.

All bird records were joined with the grid cells of the cities. For our analysis, we used only the records from selected grid cells, which were well-surveyed sites at the open digital source level. For this, we measured IC at each grid cell that contained bird observation records. We followed the formula, IC = \( \frac{S_{obs}}{S_{exp}} \) (Colwell and Coddington 1994; Sousa-Baena et al. 2014; Freeman and Peterson 2019). Here, \( S_{obs} \) is the total number of observed species per grid cell; \( S_{exp} \) is the expected true species richness, which we measured following Chao² estimates of species richness (Chao 1984; Colwell and Coddington, 1994), using the \( R \) package fossil (Vavrek 2011). In selection of well-surveyed grid cells, the used level of IC varied in the existing literature; on average, a value ≥0.50 was used to confirm well-surveyed areas (Sousa-Baena et al. 2014; Troia and McManamay 2016). Accordingly, our selection of the well-surveyed grid cells at the open digital dataset considered the threshold of IC ≥0.50 and frequency of species occurrence records ≥10 (Figs 1 and 2). This selection allowed us to remove sites with low sample size despite their high IC.

**Field dataset comprising bird observation records from systematic bird surveys**

We surveyed birds in Freiburg and Dhaka, following commonly used techniques of fixed radius point counts (Bibby et al. 2000; Gregory et al. 2004). Bird counts have been used to investigate avian communities along urbanization gradients by many other urban ecologists (Blair 1996; Clergeau et al. 1998; Sandström et al. 2006; Menon et al. 2014; van Heezik and Seddon 2017). Since cities include many restricted areas, traditional sampling designs for rural/urban areas were unsuitable for the urban environment (DeGraaf et al. 1991; van Heezik and Seddon 2017). We thus limited our surveys to easily accessible public sites.

In Freiburg, we conducted a field survey at 45 point locations during March–May in 2018. The locations were distributed at nine urban plots (each with ~1 km² area) across different densities of built-up areas.

In Dhaka, we conducted an extensive field survey during February–June in 2018. Dissimilar to Freiburg, Dhaka is a big city, but the number of bird records available from open digital platform was low. Therefore, we extended the duration of the survey to allow us to inspect a larger number of urban point locations. We surveyed 172 point locations, which were distributed at 17 urban plots (each with ~1 km² area) and along 10 line transects (1–8 km in length) across the city.

For our bird survey, we visited our preselected plots/line transect sites during four hours in the early morning and four hours before sunset, when birds’ activities are generally high. Within each study site, we chose our first survey point randomly; subsequent points were chosen by walking between 100 and 200 m following any accessible walking paths. Thus, all of our field points were at least 100 m apart from each other. At each point, all bird species occurrences were recorded by field ornithologists following a point count method for 10–15 min and using a 50 m radius.

**Combined dataset of bird records from open digital sources and systematic field surveys**

The data from selected well-surveyed grid cells at the open digital source level and the data of our field surveyed grid cells were merged.

In both Freiburg and Dhaka, open digital datasets and field datasets varied in terms of the data collection period; thus, the proportions of observed migratory birds were higher on online platforms owing to records being collected across the seasons. To reduce bias and for better comparison between the datasets, we considered only records of bird species thought to be resident at the city scale. For this, we identified the origin and migration status of birds in both cities using species fact sheets and range maps of BirdLife International (BirdLife International and Handbook of the Birds of the World 2018; BirdLife International 2020). In the case of Freiburg, this might include some species which are partly resident: in the Black Redstart, for instance, a few individuals are year-round residents, but most are migrants. However, surmising the migratory status of individual birds is beyond the scope of current paper. We thus strictly followed BirdLife range maps to identify and extract the records of resident bird species. For our analysis, we measured resident bird species richness (i.e. number of species) per grid cell in each of three grid-based datasets.
Predictor variables

Habitat factors
We extracted data for different predictor variables at the grid cells from available raster datasets and local urban atlas map datasets. To represent the urbanization effect on bird species richness, we looked at the percentage of impervious surface as a proxy for the extent of built-up areas (Brown de Colstoun et al. 2017). Besides, we considered the amount of high density urban fabric which refers to the total area (m²) with >50% soil sealing level (GISAT 2011; Ares 2012; EEA 2017). To represent the green structures’ effect, we looked at the vegetation fraction (%) (Broxton et al. 2014) and the total urban green area (m²) per grid cells. Urban green area included public gardens, parks, zoos,
suburban natural sites bordered by urban fabric, water bodies, and any green sites used for sports, leisure or recreational activities (GISAT 2011; Ares 2012; EEA 2017). In addition to this, we measured proximity (i.e. distance, m) from each grid cell’s centroid to the nearest edge of forest area in Freiburg, and to the nearest edge of agricultural area in Dhaka (GISAT 2011; EEA 2017). Among all habitat factors (Supplementary table 1), we performed a ‘Pearson correlation analysis’ separately within each dataset. We estimated the variance inflation factor (VIF) to check for correlation effects (i.e. correlation coefficient $>0.70$.

Figure 2: Distribution of bird survey locations in urban areas of Dhaka city. (A) Selection of well-surveyed grid cells of bird observation from open digital data sources. The well-surveyed cells consist of a threshold of IC metric $>0.50$ and frequency of species occurrence records $>10$. (B) Grid cell-based aggregation of systematically collected field survey data on birds using the point count method. (C) Combined data from selected well-surveyed grid cells at the open digital source level and field surveyed grid cells. Note that the point symbols and legend symbols are enlarged here to improve visibility. The grid cell size used in the assessment was $250\times250$ m and projected coordinate system was WGS_1984_UTM_Zone_46N.
may induce multicollinearity); we considered VIF < 5 to confirm that collinearity was not an issue (Akinwande et al. 2015).

**Sampling factors**

Along with the habitat-related predictor variables, we assumed that frequency of observation records (FR), observed days (FoD), observers and surveyed point locations (Supplementary table 1) might have affected resident bird records per grid cell. In case of open digital data, we presumed additional sampling effects due to varying levels of observer experience (Kelling et al. 2015; Johnston et al. 2018) (Supplementary Table 1). We compared Pearson correlation coefficient of resident bird species richness with all these sampling-related factors in each dataset separately.

The sampling effect among the datasets varied, possibly owing to the variation in study effort. Hereafter, we identified the topmost correlated sampling factor of bird species richness within each dataset separately in both cities. We considered the identified sampling factor as a covariate of bird species richness during model assessment using the corresponding dataset. More precisely, in Freiburg, FoD was the identified sampling factor in the open digital dataset and in the combined dataset; FR was the sampling factor in the field surveyed dataset. In Dhaka, FR was the identified sampling factor of bird species richness in all the datasets.

**Model analysis**

Our response variable was number of resident bird species (i.e. species richness) in the grid cells (250 m x 250 m). To achieve normality, this number of resident bird species was log10-transformed in case of Dhaka.

For the datasets of Freiburg, we checked model fit (i.e. using a full model that combines all variables) by applying linear regression (lm) and a generalized linear regression model (glm—family = Poisson and negative binomial); the linear regression model showed a better fit with the lowest AIC. For the datasets of Dhaka, linear regression model fit was also appropriate. Hereafter, we applied linear regression models to explore and compare the best causal relationship of resident bird species richness with the urban habitat-related variables in both cities.

We built similar models, using the (i) open digital dataset, (ii) field surveyed dataset and (iii) combined dataset. We formed a global model that included all the variables of habitat factors; one model included only the variables related to urbanization effect; one model included only the variables related to green structures’ effect. Additionally, five simpler models were created for each of the habitat factors separately. All eight of these models included the identified sampling effect factor as a covariate of bird species richness.

To assess the best causal relationship of bird species richness with habitat factors, we compared the coefficient parameters among the models. We compared the entire set of models and

### Table 1: A comparison of regression models of observed bird species richness within urban areas in Freiburg city

| Model Structure | df | AICc | ΔAIC | wi |
|-----------------|----|------|------|----|
| **(A) Open digital data** | | | | |
| Global model | IS+HU+UG+PF+VG+FoD | 8 | 356.5 | 10.34 | 0.004 |
| Urban effect | IS+HU+FoD | 5 | 348.3 | 2.13 | 0.227 |
| Impervious surface | IS+FoD | 4 | 346.1 | 0 | 0.659 |
| High density urban fabric | ~HU+FoD | 4 | 353.1 | 7 | 0.02 |
| Green effect | ~UG+PF+VG+FoD | 6 | 356.2 | 10.11 | 0.004 |
| Vegetation fraction | ~VG+FoD | 4 | 351.8 | 5.65 | 0.039 |
| Urban green area | ~UG+FoD | 4 | 352.7 | 6.52 | 0.025 |
| Distance to forest area | ~PF+FoD | 4 | 353 | 6.82 | 0.022 |
| **(B) Field surveyed data** | | | | |
| Global model | IS+HU+UG+PF+VG+FR | 8 | 181.8 | 1.86 | 0.173 |
| Urban effect | IS+HU+FR | 5 | 181 | 1.09 | 0.255 |
| Impervious surface | IS+FR | 4 | 179.9 | 0 | 0.44 |
| High density urban fabric | ~HU+FR | 4 | 187.8 | 7.94 | 0.008 |
| Green effect | ~UG+PF+VG+FR | 6 | 186.4 | 6.46 | 0.017 |
| Vegetation fraction | ~VG+FR | 6 | 184.2 | 4.3 | 0.051 |
| Urban green area | ~UG+FR | 4 | 192.4 | 12.48 | 0.001 |
| Proximity to forest area | ~PF+FR | 4 | 184.1 | 4.17 | 0.055 |
| **(C) Combined data** | | | | |
| Global model | IS+HU+UG+PF+VG+FoD | 8 | 581 | 6.9 | 0.02 |
| Urban effect | IS+HU+FoD | 5 | 576.4 | 2.25 | 0.207 |
| Impervious surface | ~IS+PF+Od | 4 | 574.1 | 0 | 0.639 |
| High density urban fabric | ~HU+FoD | 4 | 581.6 | 7.41 | 0.016 |
| Green effect | ~UG+PF+VG+FoD | 6 | 582.7 | 8.54 | 0.009 |
| Vegetation fraction | ~VG+FoD | 4 | 578.7 | 4.54 | 0.066 |
| Urban green area | ~UG+FoD | 4 | 582.2 | 8.02 | 0.012 |
| Proximity to forest area | ~PF+FoD | 4 | 580.2 | 6.09 | 0.03 |

The contrast is shown using three separate datasets—open digital data, systematic field survey data and a combined dataset (i.e. field data and open digital data are joined). Results shown here are produced using resident bird species richness (i.e. number of bird species) as the response variable. Top ranked models with lowest AICc score and highest Akaika weight (wi) are shown in bold.

- **Predictor:** IS — Impervious surface (%), HU — High density urban fabric (m²), UG — Urban green area (m²), VG — Vegetation fraction (%), PF — Proximity (i.e. distance) to Forest area (m).
- **Covariate:** FR — Frequency of observation records, FoD — Frequency of observation days.
identified the top ranked models based on the smallest AICc score and the highest Akaike weight (\( w_i \)). Next, we inspected average model summaries, parameter coefficients at 95% confidence interval and the significance level of the variables in a set of models by controlling delta AICc (\( \Delta \)) < 2 (Burnham and Anderson 2004). Furthermore, we visualized the fitted relationship of bird species richness to the predictor variables among all three datasets. For model validation and normality check, we performed the Kolmogorov–Smirnov goodness-of-fit test and conducted a visual inspection of model residuals; we also checked for heteroskedasticity (Breusch–Pagan test).

All model analyses were conducted with R 3.6.1 (R Core Team 2019) in RStudio 1.2.1335 (RStudio Team 2019). The used packages were fossil (Vavrek 2011), vegan (Oksanen et al. 2019), MASS (Venables and Ripley 2002), MuMIn (Barton 2013), PerformanceAnalytics (Peterson et al. 2018) and effects (Fox 2003).

**Results**

Open digital data revealed similar key predictor variables of urban bird species richness as resulted from field surveys. This held true not only for the Western city of Freiburg with its rich coverage in digital data platforms, but also for the poorly covered global Southern city of Dhaka.

**Freiburg**

In Freiburg, 630 urban grid cells accounted for 21 150 resident bird observation records obtained from open digital sources. Among these, 48 urban grid cells were selected as well-surveyed, and included observation records of 70 bird species. We covered 42 urban grid cells during field surveys and documented 32.8% (23 resident bird species) of the species recorded in open digital source. Combining all records, we found that the most frequently observed species were Turdus merula (Eurasian Blackbird), Fassér domesticus (House Sparrow), Parus major (Great Tit), Fringilla coelebs (Common Chaffinch) and Parus caeruleus (Eurasian Blue Tit), a pattern which largely held true across all datasets (Fig. 3).

**Model results**

The model analysis of bird species richness with predictors, using all three datasets, showed that highest weight and lowest AICc was achieved by the model which included the variable of impervious surface (Table 1). The selected models pointed to the negative effect of proportion of impervious surface on bird species richness. The explained variation (adjusted \( R^2 \)) in the selected models using open digital data was 45%, using field surveyed data it was 29% and using combined data it was 77%.

The averaged model coefficients using separate datasets similarly indicated impervious surface as a significant predictor of bird species richness (Table 2). The negative effect of the variable ‘proximity (i.e. distance) to forest area’ was retained as significant for bird species richness only in the model average summary using field surveyed data.

The fitted value of bird species richness did not indicate any alteration in the linear relationship of observed richness to the predictor variable ‘impervious surface’ among the datasets (Fig. 4). Using the combined dataset, the plot of fitted values of bird species richness showed some grouping effect indicating different sets of data. However, the trends in the linear relationship of bird species richness to the predictor variables using separated datasets was similar. Outside of these, the effect of sampling factors was stronger in open digital data than in field-collected data; inclusion of topmost correlated sampling factor ‘frequency of observation days (FoD)’ as a covariate of bird species richness improved the model performance (i.e. for the selected model with FoD, AIC = 345.2 and for the similar model without FoD, AIC = 372.5). There was no issue of normality in the selected models. We examined any heteroskedasticity by adjusting variance covariance Heteroskedasticity Consistent (vcovHC) which showed no inconsistency in the standard error and significance level of the variables.

**Table 2: A comparison of averaged coefficient parameters of variables in a set of models at AICc < 2 in the case of Freiburg city**

|                      | Estimate | 95% Confidence intervals | Std.Error | Pr(>|t|) |
|----------------------|----------|--------------------------|-----------|---------|
|                      |          | 2.50%                    | 97.50%    |         |
| (A) Open digital data |          |                          |           |         |
| (Intercept)          | 26.63    | 24.19                    | 29.06     | 1.21    | <2e−16*** |
| Impervious surface   | −3.42    | −5.93                    | −0.91     | 1.25    | 0.00865**  |
| Frequency of observation days | 7.66 | 5.15                     | 10.17     | 1.25    | 0.0000019*** |
| (B) Field surveyed data |       |                          |           |         |
| (Intercept)          | 6.69     | 6.10                     | 7.28      | 0.29    | <2e−16***  |
| Impervious surface   | −1.11    | −1.83                    | −0.39     | 0.36    | 0.00259**  |
| High density urban fabric | −0.33 | −0.99                    | 0.33      | 0.33    | 0.33       |
| Urban green areas    | 0.19     | 0.45                     | 0.84      | 0.32    | 0.56       |
| Distance to forest area | −1.24 | −2.32                    | −0.17     | 0.53    | 0.02371*   |
| Vegetation fraction  | −0.79    | −2.16                    | 0.58      | 0.68    | 0.26       |
| Frequency of observation records | 0.48 | 0.13                     | 1.10      | 0.30    | 0.12       |
| (C) Combined data    |          |                          |           |         |
| (Intercept)          | 2.57     | 2.49                     | 2.65      | 0.04    | <2e−16***  |
| Impervious surface   | −0.16    | −0.24                    | −0.08     | 0.04    | 0.000234*** |
| Frequency of observation days | 0.70 | 0.62                     | 0.79      | 0.04    | <2e−16***  |

The contrast is shown using three separate datasets—open digital data, systematic field survey data, and a combined dataset (i.e. field data and open digital data are joined). The results show the average coefficient estimate, the 95% confidence intervals of the estimate, standard error (Std.Error) and the significance level of the variables. Significance codes, 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.
Dhaka

In Dhaka, 45 urban grid cells (<1% of the total city grid cells) contained 947 resident bird observation records at the open digital data level. Among these, 12 grid cells were selected as well-surveyed (Fig. 2), and included records of 67 bird species. Our field surveys covered 128 grid cells, where we documented 73% (49 species) of the species reported in digital data. Combining all observation records, we found that the most frequently observed species were *Passer domesticus* (House Sparrow), *Corvus splendens* (House Crow), *Acridotheres tristis* (Common Myna), *Milvus migrans* (Black Kite) and *Gracupica contra* (Asian Pied Starling), a pattern which largely held true across the datasets (Fig. 5).

Model results

The model assessment using all three datasets separately revealed comparable results, showing that the lowest AICc and highest weight was achieved by the model which included the variable vegetation fraction (Table 3). The summary of the top ranked models pointed to the positive effect of vegetation fraction for bird species richness in Dhaka. The explained variation in the selected model using open digital data was 69%, using field survey data it was 89% and using combined dataset it was 91%.

The averaged model coefficients using separate datasets indicated vegetation fraction was a significant predictor of bird species richness in field surveyed data and combined data (Table 4). However, the significance of this predictor variable using the digital dataset differed from the results obtained using the field survey and combined datasets (Table 4). This variation in the significance level of the predictor might have resulted from the small sample size of well-surveyed locations.

Residual plots, the test of normality and heteroscedasticity in the selected top ranked models did not reveal any issues with
Figure 5: The 10 most common bird species in Dhaka in the 3 datasets. The vertical bar represents total frequency of occurrences.

Table 3: A comparison of regression models of observed bird species richness within urban areas in Dhaka city

| Model structure                  | df | AICc | ΔAIC | wi |
|----------------------------------|----|------|------|----|
| (A) Open digital data            |    |      |      |    |
| Global model                     | 8  | 25.5 | 34.68| 0  |
| Urban effect                     | 5  | 3.1  | 6.06 | 0.02|
| Impervious surface               | 4  | 0.0  | 0.81 | 0.2 |
| High density urban fabric        | 4  | 8.7  | 0.48 | 0.2 |
| Green effect                     | 6  | 0.9  | 8.22 | 0.01|
| Urban green area                 | 4  | 6.2  | 2.91 | 0.1 |
| Proximity to agricultural area   | 4  | 7.4  | 1.73 | 0.1 |
| (B) Field surveyed data          |    |      |      |    |
| Global model                     | 8  | 293.6| 6.07 | 0.03|
| Urban effect                     | 5  | 288.7| 10.9 | 0.003|
| Impervious surface               | 4  | 290.6| 9.07 | 0.01|
| High density urban fabric        | 4  | 288.9| 10.8 | 0.003|
| Green effect                     | 6  | 298.1| 1.6  | 0.3 |
| Urban green area                 | 4  | 290.2| 9.5  | 0.01|
| Proximity to agricultural area   | 4  | 286.8| 12.9 | 0.001|
| (C) Combined data                |    |      |      |    |
| Global model                     | 8  | 297.7| 4.9  | 0.05|
| Urban effect                     | 5  | 292.7| 9.9  | 0.004|
| Impervious surface               | 4  | 294.3| 8.3  | 0.01|
| High density urban fabric        | 4  | 292.4| 10.2 | 0.003|
| Green effect                     | 6  | 302.1| 0    | 0.52|
| Urban green area                 | 4  | 293.2| 9.4  | 0.005|
| Proximity to agricultural area   | 4  | 289.3| 13.23| 0.001|

The contrast is shown using three separate datasets—open digital data, systematic field survey data, and a combined dataset (i.e. field data and open digital data are joined). Results shown here were produced using resident bird species richness (i.e. log10-transformed for normality) as the response variable. Top ranked models with lowest AICc score and highest Akaike weight (\(w_i\)) are shown in bold.

Covariate: FR = Frequency of observation records.
This article focused on the urban context, the features of which are broadly represented by increasing levels of built-up areas and human settlement. We explored bird records for well-surveyed urban sites (i.e., locations with sufficient species documentation) available on digital platforms, using an IC metric (Sousa-Baena et al. 2014; Troia and McManamay 2016). We found that the top ranked relationship of bird species richness to the key predictor variables retained using data of such well-surveyed sites were not different from that using data collected through standard field survey. This indicates that if a sufficient amount of data of well-surveyed sites can be obtained from online platforms for a city, those data can be suitable for investigating driving factors of bird species richness patterns across its urban areas. Therefore, if field data are not available, biodiversity monitoring and planning may be initiated based on data from online platforms alone, until systematic field data become available. Note, however, that this notion does not necessarily hold for taxa and/or other indicators of biodiversity, such as species diversity or abundance of threatened species, which our study did not address.

In our study, the distribution of well-surveyed urban sites indicated gaps in the spatial coverage of bird observation records within both cities. In Freiburg, even though the city has more than 45 K digital bird records at Ornitho.de, not all locations are well surveyed; Fig. 1 shows poorly surveyed grid cells. The case of Dhaka is shown in Fig. 2, where the majority of the grid cells reflect the fact that GBIF holds less than 1.5 K avian records in this city. In general, spatial coverage of bird observation records is higher in urbanized areas than in forest/agricultural sites in both cities. Several earlier studies also indicated spatial and temporal gaps within open digital species databases (i.e. for single species or multiple taxonomic groups) at the scale of continents, countries and regions (Soberón et al. 2007; Mora et al. 2008; Nakamura and Soberón 2009; Sousa-Baena et al. 2014). Digital data usually covers a portion of the entire spectrum of environmental gradients in any geographic region (Hortal et al. 2007). However, species reporting by citizens was found to be more frequent in human-settled areas (Tye et al. 2017) and well-surveyed sites were found clustered around major cities (Freeman and Peterson 2019). Despite such bias, existing studies showed that web-based data collected by citizens can provide reliable predictions of species–habitat relationships (Tye et al. 2017).

### Table 4: A comparison of average coefficient parameters of variables in a set of models at ΔAIC < 2 in the case of Dhaka city

| Variable                              | Estimate | 95% Confidence intervals | Std.Error | Pr(>|t|) |
|---------------------------------------|----------|--------------------------|-----------|---------|
|                                       |          | 2.50%                    | 97.50%    |         |
| A) Open digital data                  |          |                          |           |         |
| (Intercept)                           | 1.27     | 1.20                     | 1.34      | 0.03    | <2e−16*** |
| Impervious surface                    | −0.05    | −0.13                    | 0.03      | 0.03    | 0.200281  |
| High density urban fabric             | −0.05    | −0.13                    | 0.02      | 0.03    | 0.17      |
| Vegetation fraction                   | 0.06     | −0.02                    | 0.13      | 0.03    | 0.14      |
| Urban green area                      | 0.06     | −0.02                    | 0.13      | 0.03    | 0.15      |
| Proximity to agricultural area        | −0.04    | −0.13                    | 0.04      | 0.04    | 0.31      |
| Frequency of observation records      | 0.14     | 0.06                     | 0.22      | 0.04    | 0.000466 *** |
| B) Field surveyed data                |          |                          |           |         |
| (Intercept)                           | 0.83     | 0.82                     | 0.84      | 0.01    | <2e−16*** |
| Vegetation fraction                   | 0.02     | 0.01                     | 0.04      | 0.01    | 0.000388 *** |
| Urban green area                      | 0.01     | −0.01                    | 0.02      | 0.01    | 0.31      |
| Proximity to agricultural area        | 0.01     | 0.00                     | 0.02      | 0.01    | 0.19      |
| Frequency of observation records      | 0.21     | 0.20                     | 0.22      | 0.01    | <2e−16*** |
| C) Combined data                      |          |                          |           |         |
| (Intercept)                           | 0.87     | 0.86                     | 0.88      | 0.01    | <2e−16*** |
| Vegetation fraction                   | 0.03     | 0.01                     | 0.04      | 0.01    | 0.000418 *** |
| Urban green area                      | 0.01     | −0.01                    | 0.02      | 0.01    | 0.24      |
| Proximity to agricultural area        | 0.01     | 0.00                     | 0.02      | 0.01    | 0.13      |
| Frequency of observation records      | 0.24     | 0.23                     | 0.25      | 0.01    | <2e−16*** |

The contrast is shown using three separate datasets—open accessed digital data, systematic field survey data, and a combined dataset (i.e. field data and open digital data are joined). The results show the average coefficient estimate, the 95% confidence intervals of the estimate, standard error (Std.Error) and the significance level of the variables.

Significance codes: 0; ‘***’ 0.001; ‘**’ 0.01; ‘*’ 0.05.
In an urban context, our results showed that the key driving factor of bird species richness through model assessment using open digital datasets, field datasets and combined datasets were similar in both cities. Our study case in Freiburg was a model example where plenty of species observation records and well-surveyed sites were readily attainable from open digital sources for reliable analysis. The additional systematic field survey data and a combined dataset did not alter the general understanding of the key driving factors of bird species richness patterns we obtained using open digital data. The general trends in top ranked relationship of bird species richness to urban habitat factors, using different datasets, were consistent with earlier studies indicating that high imperviousness usually causes resident bird decline (Blair 1996; Garaffa et al. 2009; Silva et al. 2015; Sultana et al. 2021).

The study case in Dhaka added another model example, where the highest ranked relationship of bird species richness through model assessments with different datasets was also similar. It highlighted the benefit of vegetation surface on bird species richness within urban environments (Chace and Walsh 2006; Donnelly and Marzluff 2006; Evans et al. 2009). However, one issue with the digital data in Dhaka concerned the low number of well-surveyed sampling sites represented in the online sources. A robust ecological assessment of bird species richness can be ambiguous with such insufficient amount of data. In this case, the additional data we collected through field surveys, albeit for a single season, enhanced the spatial coverage of bird records within the city. A combined dataset provided sufficient sampling sites and boosted a reliable assessment by showing higher explaining power.

Caution should be taken in interpreting our results. We investigated the consistency of avian observation records available from online platforms for assessing drivers of bird species richness patterns in an urban context only. Similar studies in forest/rural sites may be achievable if species reporting on online platforms is extensive (Tye et al. 2017; Speed et al. 2018). It is also important to note the influence of the spatial resolution at the grid cells. In both cities, we selected small-sized grid cells to compile geo-referenced species records to minimize the loss of site-specific information for the observed records (Tye et al. 2017). Large grid cells (>1 km²) for assessment might be suitable for investigations focusing on large geographic regions (Hortal et al. 2007; Soberón et al. 2007; Troia and McManamay 2016).

In addition to spatial bias, sampling-related issues can hinder the suitability of digital data in species richness prediction across any environmental gradient (Hortal et al. 2007; Zhang et al. 2019). Sampling issues can include variation in survey effort by observers (i.e. cause observer-related bias) and undersampling of certain species groups (i.e. cause taxonomic bias). It can further include higher survey attempts in easily accessible areas that are often close to the observer’s residency, home institutes, near roads, human settlement, and in and around protected areas (Meyer et al. 2015; Tye et al. 2017; Freeman and Peterson 2019; Petersen et al. 2021). Thus, in a model assessment of bird species richness patterns in relation to different environmental predictors, an account for any such sampling efforts as covariates can be helpful to reduce associated bias within digital data (Warton et al. 2013; Tye et al. 2017). Such a modelling approach can help avoid the uncertainty in deriving diversity metrics from species level predictions, and sidestep the bias in prediction of richness patterns in relation to environmental predictors outside known communities (Zhang et al. 2019).

We observed that the generalist species with higher frequencies were similarly documented in both open digital data and field data (Figs 3 and 5). However, the number of all observed species was higher in digital data compared to the data we collected during field surveys in both cities. The higher number of species (i.e. including many less frequently observed species) in digital sources was largely due to the fact that digital data was collected over multiple years. We considered only resident birds’ occurrences for the purpose of better comparison between different datasets. For conservation purposes, it would, however, be important to consider both resident and migratory species (breeding/non-breeding) in any locality. Additionally, it is recommended to cross-check digitally available species records with taxa-specific local experts and city bird atlas data (if available), to confirm the presence of unusual species and to account for any misidentifications. During data compilation, one should also keep abreast of taxonomic changes and species synonyms to avoid inclusion of duplicate records.

In urban settings, species-specific conservation is complex due to shortage of suitable habitat patches and various factors.

---

**Figure 6:** Linear fit of resident bird species richness (i.e. log10 transformed for normality) in relation to vegetation fraction in the top-ranked models in Dhaka. The black line shows the trends in the fitted relationship of bird species richness to vegetation fraction (%). Shaded grey area indicates the marginal effect at a 95% confidence interval.
associated with human well-being (such as, provision of equal access to green spaces for better health benefits) (Dearborn and Kark 2010). Conservation efforts in human-developed areas can thus focus on overall species richness and prioritize steps to increase the presence of, e.g. rare and threatened species as much as possible. To collect reliable species richness data under any taxonomic class, a long-term consistent field survey would be desirable as the standard approach. However, this is time-consuming and relies on the availability of resources and expert investigators. Conversely, the number of online-based platforms (e.g. wildenachbarn.de, citynaturechallenge.org) to document species occurrences in cities has increased noticeably in the last decade, suggesting that digital data will continue to become increasingly available in the future. These species observation records, reported mostly by amateur citizens, can be a good fit in urban ecology since they are easily and rapidly accessible (van Heezik and Seddon 2017). We thus emphasize that our investigation on open digital data in Freiburg and Dhaka provides useful examples of the suitability of open digital data for assessing drivers of urban biodiversity patterns, and thus, its applicability in biodiversity monitoring and planning in cities.

Data availability

Digital species occurrence records of birds which were used for this assessment can be requested at ‘Ornitho.de’ for Freiburg, and can be accessed at GBIF (doi.org/10.15468/dl.beqzow) for Dhaka. Further species data and analysis script of this article can be found at the link: https://osf.io/yjmxq/?view_only=3ec41ace94cf4d96808f3733bf3be80

Authors’ contributions

M.S. and I.S. designed the study. M.S. prepared the data, conducted the analysis and wrote the first draft of the manuscript. I.S. guided throughout the study, contributed during writing and revising the manuscript.

Supplementary data

Supplementary data are available at JUECOL online.

Acknowledgements

We acknowledge the support of all the students and field assistants who contributed during field surveys in Dhaka and Freiburg. Special thanks to M. Niamul Naser and Mayeen Uddin for their assistance in organizing field surveys in Dhaka. We thank Joao Manuel Cordeiro Vale Pereira and Gernot Segelbacher and two anonymous reviewers for their suggestions on earlier drafts of the manuscript. We also thank Ottavio Jani for English proofreading of the final manuscript. M.S. was supported by LGFG (Landesgraduiertenförderungsgesetz) funding from International Graduate Academy (IGA) of the University of Freiburg.

Conflict of interest statement. None declared.

References

Akinwande, M. O., Dikko, H. G., and Samson, A. (2015) ‘Variance Inflation Factor: As a Condition for Inclusion of Suppressor Variable in Regression Analysis’, Open Journal of Statistics, 5: 754–67. Amano, T., Lamming, J. D. L., and Sutherland, W. J. (2016) ‘Spatial Gaps in Global Biodiversity Information and the Role of Citizen Science’, BioScience, 66: 393–400.

Ares, R. (2012) Mapping Guide for a European Urban Atlas. Accessed on 29 July 2019. http://ec.europa.eu/regional_policy/sources/tender/pdf/2012066/annexe2.pdf.

Bartoš, K. (2013) MuMIn: Multi-Model Inference, R Package Version 1.9.13. http://CRAN.Rproject.org/package=MuMIn.

Beatley, T. (2000) Green Urbanism: Learning from European Cities. Washington, D.C.: Island Press.

Bibby, C. J. et al. (2000) Bird Census Techniques. Elsevier. Academic press, London.

BirdLife International and Handbook of the Birds of the World (2018) Bird species distribution maps of the world. Version 2018.1. Available at http://datazone.birdlife.org/species/requestdis.

BirdLife International (2020) IUCN Red List for Birds. Accessed on 6 April 2020. http://www.birdlife.org.

Blair, R. B. (1996) ‘Land Use and Avian Species Diversity along an Urban Gradient’, Ecological Applications, 6: 506–19.

Brown de Colstoun, E. C. et al. (2017) Global Man-Made Impervious Surface (GMIS) Dataset from Landsat. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Broxton, P. D. et al. (2014) ‘A MODIS-Based 1 km Maximum Green Vegetation Fraction Dataset’, Journal of Applied Meteorology and Climatology, 53: 1996–2004.

Buehler, R. et al. (2011) ‘How Germany Became Europe’s Green Leader: A Look at Four Decades of Sustainable Policymaking’, The Solutions Journal, 2: 5: 51–63.

Burnham, K. P. and Anderson, D. R. (2004) ‘Multimodel Inference: Understanding AIC and BIC in Model Selection’, Sociological Methods & Research, 33: 261–304.

Byomkesh, T., Nakagoshi, N., and Dewan, A. M. (2012) ‘Urbanization and Green Space Dynamics in Greater Dhaka’, Landscape and Ecological Engineering, 8: 45–58.

Callaghan, C. T. et al. (2020) ‘A Continental Measure of Urbanness Predicts Avian Response to Local Urbanization’, Ecography, 43: 528–38.

Chace, J. F. and Walsh, J. J. (2006) ‘Urban Effects on Native Avifauna: A Review’, Landscape and Urban Planning, 74: 46–69.

Chandler, M. et al. (2017) ‘Contribution of Citizen Science towards International Biodiversity Monitoring’, Biological Conservation, 213: 280–94.

Chao, A. (1984) ‘Nonparametric Estimation of the Number of Classes in a Population’, Scan H Statist, 11: 265–70.

Clergeau, P. et al. (1998) ‘Bird Abundance and Diversity along an Urban-Rural Gradient: A Comparative Study between Two Cities on Different Continents’, The Condor, 100: 413–25.

Colwell, R. K. and Coddington, J. A. (1994) ‘Estimating Terrestrial Biodiversity through Extrapolation’, Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 345: 101–18.

Dearborn, D. C. and Kark, S. (2010) ‘Motivations for Conserving Urban Biodiversity’, Conservation Biology, 24: 432–40.

DeGraaf, R. M., Geis, A. D., and Healy, P. A. (1991) ‘Bird Population and Habitat Surveys in Urban Areas’, Landscape and Urban Planning, 21: 181–8.

Dewan, A. M., Yamaguchi, Y., and Rahman, M. Z. (2012) ‘Dynamics of Land Use/Cover Changes and the Analysis of Landscape Fragmentation in Dhaka Metropolitan’, GeoJournal, 77: 315–30.

Donnelly, R. and Marzluff, J. M. (2006) ‘Relative Importance of Habitat Quantity, Structure, and Spatial Pattern to Birds in Urbanising Environments’, Urban Ecosystems, 9: 99–117.
Suitability of open digital species records for urban ecology

Drooge, S. (2007) ‘Just Because You Paid Them Doesn’t Mean Their Data Are Better’, in C. McEver et al. (eds) Citizen Science Toolkit Conference, pp 13–26. Cornell Laboratory of Ornithology, New York.

EEA (2017) Urban Atlas 2012 Dataset. Produced by European Environment Agency. Accessed on 20 April 2017. http://land.copernicus.eu/local/urban-atlas/urban-atlas-2012/view.

Evans, K. L., Newson, S. E., and Gaston, K. J. (2009) ‘Habitat Influences on Urban Avian Assemblages’, IBIS, 151: 19–39.

Fox, J. (2003) ‘Effect Displays in R for Generalised Linear Models’, Journal of Statistical Software, 8: 1–27.

Freeman, B. and Peterson, A. T. (2019) ‘Completeness of Digital Accessible Knowledge of the Birds of Western Africa: Priorities for Survey’, The Condor, 121: 1–10.

Garaffa, P. I., Filloy, J., and Bellocq, M. I. (2009) ‘Bird Community Responses along Urban-Rural Gradients: Does the Size of the Urbanised Area Matter?’ Landscape and Urban Planning, 90: 33–41.

GBIF.org (2019) All Species Occurrence Records of Dhaka, Duration 1990-2017. GBIF Occurrence Download. Accessed on 13 May 2019. https://doi.org/10.15468/db.beqzow.

GISAT (2011) Map dataset: EOW_GISAT_Dhaka_Status_2011, ‘Land cover/land use status - 2011’ in Dhaka, Bangladesh, EOWORLD Project. Credits: Gisat s.r.o. for ESA/World Bank.

Gregory, R. D., Gibbons, D. W., and Donald, P. F. (2004) ‘Bird Census and Survey Techniques’, in W. J. Sutherland, I. C. Woiwod (eds) Bird Ecology and Conservation: A Handbook of Techniques, pp 17–56. Oxford: University Press, Oxford.

Hortal, J. et al. (2015) ‘Seven Shortfalls That Beset Large-Scale Knowledge of Biodiversity’, Annual Review of Ecology, Evolution, and Systematics, 46: 523–49.

——— (2021), Lobo, J. M., and Jiménez-Valverde, A. (2007) ‘Limitations of Biodiversity Databases: Case Study on Seed-Plant Diversity in Tenerife, Canary Islands’, Conservation Biology, 21: 853–63.

Islam, N. (2001) Urbanization, Urban Planning and Development, and Urban Governance: A Reader for Studies. Dhaka: Center for Urban Studies (CUS).

Johnston, A. et al. (2018) ‘Estimates of Observer Expertise Improve Species Distributions from Citizen Science Data’, Methods in Ecology and Evolution, 9: 88–97.

Kelling, S. et al. (2015) ‘Can Observation Skills of Citizen Scientists Be Estimated with Species Accumulation Curves?’ PLoS One, 10: e0139600.

Khaleda, S., Mowla, Q. A., and Murayama, Y. (2017) ‘Dhaka Metropolitan Area’, in Y. Murayama, C. Kamusoko, A. Yamashita and R. Estoque (eds) Urban Development in Asia and Africa, pp 195–215. The Urban Book Series. Singapore: Springer.

Li, E. et al. (2019) ‘An Urban Biodiversity Assessment Framework That Combines an Urban Habitat Classification Scheme and Citizen Science Data’, Frontiers in Ecology and Evolution, 7: 277.

Mayorga, I., Bichier, P., and Philippot, S. M. (2020) ‘Local and Landscape Drivers of Bird Abundance, Species Richness, and Trait Composition in Urban Agroecosystems’, Urban Ecosystems, 23: 495–505.

Medearis, D. and Daseking, W. (2012) ‘Freiburg, Germany’s Eco-Capital’, in T. Beatley (ed.) Green Cities of Europe, pp 65–82. Washington: Island Press.

Menon, M., Devi, P., and Rangaswamy, M. M. (2014) ‘Avifaunal Richness and Abundance along an Urban Rural Gradient with Emphasis on Vegetative and Anthropogenic Attributes in Tiruchirappalli, India’, Landscape Research, 41(1):131–148, doi: 10.1080/01426397.2014.910294.

Meyer, C. et al. (2015) ‘Global Priorities for an Effective Information Basis of Biodiversity Distributions’, Nature Communication, 6: 8221.

Mora, C., Tittensor, D. P., and Myers, R. A. (2008) ‘The Completeness of Taxonomic Inventories for Describing the Global Diversity and Distribution of Marine Fishes’, Proceedings of the Royal Society B: Biological Sciences, 275: 149–55.

Nakamura, M. and Soberón, J. (2009) ‘Use of Approximate Inference in an Index of Completeness of Biological Inventories’, Conservation Biology, 23: 469–74.

Oksanen, J. et al. (2019) ‘Package “Vegan”’ Title Community Ecology Package’, Version 2.5-5. Community Ecology Package, 2: 1–297. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

Peterson, A. T., Soberón, J., and Krichaftka, L. (2015) ‘A Global Perspective on Decadal Challenges and Priorities in Biodiversity Informatics’, BMC Ecology, 15: 15.

Petersen, T. K. et al. (2021) ‘Species Data for Understanding Biodiversity Dynamics: What, Where and When of Species Occurrence Data Collection’, Ecological Solutions and Evidence, 2: e12048.

Peterson, B. G. et al. (2018) Package ‘PerformanceAnalytics’. R Team Cooperation.

R Core Team (2019) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

RStudio Team (2019) RStudio: Integrated development for R. Boston, MA: RStudio Inc. Boston.

Sandström, U. G., Angelstam, P., and Mikusiński, G. (2006) ‘Ecological Diversity of Birds in Relation to the Structure of Urban Green Space’, Landscape and Urban Planning, 77: 39–53.

Silva, C. P. et al. (2015) ‘Bird Richness and Abundance in Response to Urban Form in a Latin American City: Valdivia, Chile as a Case Study’, PLoS One, 10:1–16, doi: 10.1371/journal.pone.0138120.

Soberón, J. et al. (2007) ‘Assessing Completeness of Biodiversity Databases at Different Spatial Scales’, Ecography, 30: 152–60.

Soria-Auza, R. W. and Kessler, M. (2008) ‘The Influence of Sampling Intensity on the Perception of the Spatial Distribution of Tropical Diversity and Endemism: A Case Study of Ferns from Bolivia’, Diversity and Distributions, 14: 123–30.

Sousa-Baena, M. S., Garcia, L. C., and Peterson, A. T. (2014) ‘Completeness of Digital Accessible Knowledge of the Plants of Brazil and Priorities for Survey and Inventory’, Diversity and Distributions, 20: 369–81.

Speed, J. D. et al. (2018) ‘Contrasting Spatial, Temporal and Environmental Patterns in Observation and Specimen Based Species Occurrence Data’, PLoS One, 13: e0196417.

Statistisches Landesamt Baden-Württemberg – Bevölkerung nach Nationalität und Geschlecht am 31. Dezember 2018 (CSV-Datei) (Hilfe dazu). Accessed on 7 May 2020, https://www.statistik-bw.de/.

Strohbach, M. W., Haase, D., and Kabisch, N. (2009) ‘Birds and the City: Urban Biodiversity, Land Use, and Socioeconomics’, Ecology Society, 14: 31–45.

Sullivan, B. L. et al. (2014) ‘The eBird Enterprise: An Integrated Approach to Development and Application of Citizen Science’, Biological Conservation, 169: 31–40.
Sultana, M., Corlatti, L., and Storch, I. (2021) ‘The Interaction of Imperviousness and Heterogeneity Drives Bird Richness Patterns in South Asian Cities’, Urban Ecosystems, 24: 335–44.

Trimble, M. J. and van Aarde, R. J. (2012) ‘Geographical and Taxonomic Biases in Research on Biodiversity in Human-Modified Landscapes’, Ecosphere, 3: art119.

Troia, M. J. and McManamay, R. A. (2016) ‘Filling in the GAPS: Evaluating Completeness and Coverage of Open-Access Biodiversity Databases in the United States’, Ecology and Evolution, 6: 4654–69.

Tye, C. A. et al. (2017) ‘Evaluating Citizen vs. professional Data for Modelling Distributions of a Rare Squirrel’, Journal of Applied Ecology, 54: 628–37.

van Heezik, Y. and Seddon, P. J. (2017) ‘Counting Birds in Urban Areas: A Review of Methods for the Estimation of Abundance’, in E. Murgui and M. Hedblom (eds) Ecology and Conservation of Birds in Urban Environments. pp 185-207. Cham: Springer. 10.1007/978-3-319-43314-1_10.

Vavrek, M. J. (2011) ‘Fossil: Palaeoecological and Palaeogeographical Analysis Tools’, Palaeontologia Electronica, 14: 1T. R package version 0.4.0.

Venables, W. N. and Ripley, B. D. (2002) Modern Applied Statistics with S, 4th edn. New York: Springer.

Wang Wei, J., Lee, B.-H., and Bing Wen, L. (2016) ‘Citizen Science and the Urban Ecology of Birds and Butterflies — A Systematic Review’, PLoS One, 11: e0156425.

Warton, D. I., Renner, I. W., and Ramp, D. (2013) ‘Model-Based Control of Observer Bias for the Analysis of Presence-Only Data in Ecology’, PLoS One, 8: e79168.

Zhang, C. et al. (2019) ‘How to Predict Biodiversity in Space? An Evaluation of Modelling Approaches in Marine Ecosystems’, Diversity and Distributions, 25: 1697–708.