Social media data reveal novel habitats for invasive species

Shawan Chowdhury*[1], Rochita Debnath[2], Niloy Hawladar[2], Scarlett R. Howard[1], Kathryn A. Hodgins[1], Bob B.M. Wong[1], Ivan Jarić[3, 4]

* corresponding author (shawan.chowdhury@monash.edu)

[1]School of Biological Sciences, Monash University, Clayton, Australia
[2]Department of Zoology, University of Dhaka, Bangladesh
[3]CNRS, AgroParisTech, Ecologie Société Evolution, Université Paris-Saclay, Gif-sur-Yvette, France
[4]Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, České Budějovice, Czech Republic

Keywords

Citizen science; data integration; Facebook; GBIF; iEcology; invasive species; online data; tropics

Acknowledgement

We thank the numerous volunteers for collecting data and sharing their records on GBIF and social media. IJ was supported by grant no. 23-07278S from the Czech Science Foundation. SRH acknowledges the ARC Discovery Early Career Researcher Award (DECRA): DE230101556 for funding support.

Author contributions

SC conceptualised the idea, everyone contributed to the idea; SC, RD, and NH collected the Facebook data; SC wrote the first draft of the paper, everyone contributed to the paper; SC did the analyses, everyone contributed to the analyses.

Conflict of interest

The authors declare no competing interests.

Data and code availability

The GBIF data is publicly available (GBIF, 2025). We uploaded the Facebook data in the supplementary section.

All the codes used in the analysis are publicly available in the GitHub repository (https://github.com/ShawanChowdhury/InvasiveSpecies_SocialMedia_Bd).

Abstract

Invasive alien species pose significant threats to biodiversity, yet their distributions remain poorly documented across much of the tropics. Using Bangladesh, a megapopulated tropical country, we combine species distribution data from Facebook and the Global Biodiversity Information Facility (GBIF) to evaluate how data integration improves invasive alien species distribution. Our compiled dataset contains 11,469 occurrence records for 65 species. Although Facebook contributed only 6% of the total records, it provided more data than GBIF for two-thirds of the species and served as the unique source of distribution data for 23 species. Incorporating Facebook data increased estimated range sizes for 44 species and expanded the spatial extent of species distributions by 14%. Facebook records also exhibited distinct environmental patterns, often in urban and human-impacted areas. Our study demonstrates that social media can help fill critical biodiversity data gaps in under-sampled regions, and should be integrated into invasive species monitoring and conservation planning frameworks.

Background

Global biodiversity faces an existential crisis due to various natural and human-induced threats (Pereira et al., 2012; Dirzo et al., 2014). Over 47,000 species are at risk of extinction, ~28% of those assessed by the International Union for Conservation of Nature (IUCN) Red List (IUCN, 2025). The Living Planet Index Report indicates a staggering 73% decrease in wildlife over the last 50 years (WWF, 2024). Climate change, agricultural intensification, and invasive alien species are major threats impacting these alarming biodiversity trends (Capinha et al., 2015; Seebens et al., 2017; Tilman et al., 2017; Bradshaw et al., 2021). Such losses are not just of ecological concern—they affect economies, food security, and human well-being (Cardinale et al., 2012). Ecologists and conservationists are actively working to improve the situation, leading to the emergence of global biodiversity targets (Green et al., 2019). The most recent of these, set by the Convention on Biological Diversity, is the Kunming-Montreal Global Biodiversity Framework (GBF), which aims to minimise the biodiversity decline by protecting 30% of the globe by 2030 (CBD, 2022).

Biological invasions are a major driver of global biodiversity decline, as non-native species can outcompete, prey on, or spread disease to native species (Simberloff et al., 2013). These invasions also disturb ecosystem functions and lead to significant economic and social costs worldwide (Early et al., 2016). GBF Target 6 aims to reduce the introduction of invasive alien species by half and minimise their impact (CBD, 2022). Achieving such targets requires detailed species distribution data, which are absent from the vast majority of the planet (Troudet et al., 2017). Our understanding of global biodiversity is limited primarily due to the lack of monitoring data from the tropics (Collen et al., 2008; Hortal et al., 2015). While systematic monitoring has been practised in the developed world for decades, it remains scarce in the tropics, even though tropical rainforests harbour over half the world's known biodiversity (Collen et al., 2008). The situation is gradually improving, however, thanks to the growing popularity of citizen science, where anyone can contribute to biodiversity recording (Callaghan et al., 2021; Mason et al., 2025). Over the past few decades, biodiversity observation data volume has increased significantly, primarily due to citizen science initiatives and technological advances (Pocock et al., 2018; Heberling et al., 2021). For instance, species occurrence data in the most extensive biodiversity repository, the Global Biodiversity Information Facility (GBIF), rose twelvefold from 2009 to 2021, largely due to the inclusion of citizen science records (Heberling et al., 2021). However, citizen science data come with limitations. In particular, many authors have highlighted the temporal and spatial bias in citizen science data: most observations are for birds and originate from the developed world, particularly Western Europe and North America (Amano et al., 2016; Bowler et al., 2025). This is because citizen science applications have struggled to gain traction in the tropics (Pocock et al., 2019) due to various barriers, such as limited access to technology, language and cultural barriers, and the lack of institutional support and data infrastructure (Danielsen et al., 2014).

Citizen science data need not be confined to structured platforms like iNaturalist; individuals can contribute to biodiversity recording in various ways. One such option is online data sharing (Jarić et al., 2020, 2021; Correia et al., 2021; Caley & Cassey, 2023). With the increasing availability of smartphones and fast internet, along with the growing popularity of social media (e.g., Facebook, Instagram), anyone from anywhere in the world can take biodiversity photographs and share them on social platforms (Callaghan et al., 2022). If properly harvested, these data can address numerous ecological questions (Di Minin et al., 2015; Toivonen et al., 2019; Sbragaglia et al., 2022; Chowdhury et al., 2023a; Vardi et al., 2024; Baasanmunkh et al., 2025). For instance, social media data can potentially lessen the biodiversity knowledge gap (Marcenò et al., 2021; Chowdhury et al., 2023a), increase knowledge of threatened taxa (Rosa & Freitas, 2024), improve conservation assessments (Chowdhury et al., 2023b, 2024a), track the movement of highly range-shifting species to understand their movement dynamics better (Chowdhury et al., 2025), provide data on underrepresented species (ElQadi et al., 2017; Moore et al., 2024), track changes in species caused by seasonal or climatic change (Elquadi et al., 2023), map ecosystem services, promote conservation through marketing and education, and facilitate conservation communication (Di Minin et al., 2015). Such insights are especially valuable in countries that are poorly represented in global biodiversity repositories. While previous studies have used social media data to monitor charismatic species, their use to track invasive alien species distributions and spread remains largely underexplored (Jarić et al., 2021). However, such additional records from social media can be utilised to obtain new distributions for invasive alien species (e.g., Allain et al., 2019; Rothman et al., 2020; Šmejkal et al., 2024).

Here, we evaluate the utility of social media data as potential data sources to improve understanding of invasive alien species distributions. We focus on Bangladesh, a densely populated tropical country, which serves as a useful and representative case study where biodiversity data are poorly represented in global repositories. Similarly, we use Facebook as a social media platform due to its popularity in Bangladesh (Chowdhury et al., 2023a). By collating species occurrence records from Facebook and GBIF, we investigate how supplementary data from Facebook improves the grid-based distribution and the range of invasive alien species using the minimum convex polygon. Additionally, we use generalised linear models to explore the preferential bias in documenting species observations.

Data

We obtained the list of invasive alien species for Bangladesh from Mukul et al. (2020), which is the most comprehensive invasive alien species database for Bangladesh. We collated the species occurrence data using two approaches. First, we extracted the occurrence data from GBIF (https://www.gbif.org/; GBIF, 2025) using the rgbif package (Chamberlain et al., 2025) in R (R Core Team, 2022, version 4.2.2). GBIF contains data from thousands of citizen science applications, including iNaturalist (Heberling et al., 2021), so we did not use any other database.

We obtained the species occurrence data from Facebook following Chowdhury et al. (2024b). Initially, we created a list of relevant Facebook groups by searching individual taxa (plants, insects, molluscs, fishes, and birds), obtained from the invasive alien species list. Once we created the list (97 groups, Supplementary Table S1), we searched each group by species scientific name, common name, and local names. We compiled species common names and local names using IUCN Bangladesh (2015). We carefully checked each post to obtain the event date (day, month, year), location, life stage, and photographer's name. As Facebook does not provide specific geolocation information, we georeferenced the location (longitude and latitude) using Google Maps (https://www.google.com/maps).

We collected predictor variable data to compare the nature of the distributions between different data sources. Using the geodata R package (Hijmans et al., 2024), we extracted the annual mean temperature, rainfall, and elevation for Bangladesh from WorldClim (https://www.worldclim.org/). We downloaded the most recent human footprint index map (for 2020) from the Wildlife Conservation Society (WCS, 2005; https://wcshumanfootprint.org/data-access), and built areas from the Global Human Settlement Database (Pesaresi et al., 2024; https://human-settlement.emergency.copernicus.eu/download.php).

Data cleaning and preparation

We cleaned the Facebook and GBIF data using two different approaches. We excluded GBIF data if i) the location information was missing, ii) species names were mismatched, iii) records were outside the borders of Bangladesh, and iv) there were duplicate records. We excluded observations from Facebook if the i) photograph was not clear enough to identify up to the level of species, ii) location was coarser than 100 km^2, or iii) photograph was not from Bangladesh.

We compiled species occurrence data from both sources to create three groups: Facebook, GBIF, and Overall (data from Facebook and GBIF). We used Facebook and GBIF to compare the distribution of species occurrence records using the grid-based distribution. In contrast, we used GBIF and Overall datasets to compare the differences in area, which were calculated using the minimum convex polygon.

Using the terra R package (Hijmans, 2025), we reprojected, cropped, and masked all environmental predictor layers to a uniform spatial resolution of 1 km × 1 km, using the WGS84 coordinate reference system and the geographic extent of Bangladesh. We used the tidyverse R package (Wickham et al., 2019) for all data processing and the ggplot2 R package (Wickham, 2016) for all visualisations.

Grid-based distribution

To quantify the difference between GBIF and Facebook data, we converted species distribution into a 1 km × 1 km grid. We considered a species to be present in a grid if at least one occurrence record existed within that particular grid. We created separate presence maps for each data source - GBIF, Facebook and Overall. Some species were not represented in all data sources, so the number of maps depended on the number of sources from which we obtained their data.

Minimum convex polygon

We calculated a minimum convex polygon (MCP) surrounding the occurrence records (Joppa et al., 2016) for each invasive alien species in Bangladesh. We used the sf R package (Pebesma & Bivand, 2023) and created two MCPs for each species, using the GBIF and Overall datasets. For computational efficiency, we rasterised the MCP using the fasterize R package (Ross, 2024) and removed areas outside the borders of Bangladesh.

Statistical analyses

We tested whether there was any difference between species occurrence records from Facebook and GBIF, using five environmental predictors (built areas, human footprint index, temperature, rainfall, and elevation) that impact species distribution. We fitted linear models using the broom R package (Robinson et al., 2025) to estimate standardised effect sizes for each environmental predictor. We standardised values (z-scored) within each environmental variable to allow comparison across variables with different units and magnitudes. For each taxonomic group (e.g., fishes, birds), we fitted separate linear models for each predictor using the formula where the response variable was the standardised environmental value and the categorical predictor was the data source, resulting in 25 comparisons in total. From each model, we extracted the estimated effect size (i.e., the mean difference between Facebook and GBIF records), standard error, and p-value. We considered differences statistically significant at $p < 0.05$. Effect sizes and 95% confidence intervals were visualised using coefficient plots stratified by taxonomic group. To account for multiple comparisons, we also applied the Benjamini–Hochberg false discovery rate correction and report adjusted p-values (p_adj) for transparency (Supplementary Table S2).

Distribution of spatial data

Our compiled data contained 11,469 occurrence records (GBIF: 10,745, Facebook: 710) for 65 species. Although Facebook data accounted for only 6% of the compiled data, there were substantial differences across taxa. For example, we obtained more records from Facebook for plants, molluscs, and fishes. Nearly 85% of the GBIF data (N = 9,687) were for a single bird species (the rock pigeon, *Columba livia*), and for many species, we obtained a higher number of occurrence records from Facebook (for 43 species (66%); Figure 1A). There were 34 species, which were present in either Facebook or GBIF (unique to the data source). For 23 species, we exclusively obtained occurrence data from Facebook (the majority being plants), while 11 species were sourced only from GBIF (spanning multiple taxa) (Figure 1A). Four species were entirely absent from GBIF but had over 20 occurrence records each in the Facebook dataset, including two fish species (*Pterygoplichthys multiradiatus*, *Pangasius sutchi*) and two plants (*Hyptis suaveolens*, *Triumfetta rhomboidei*).

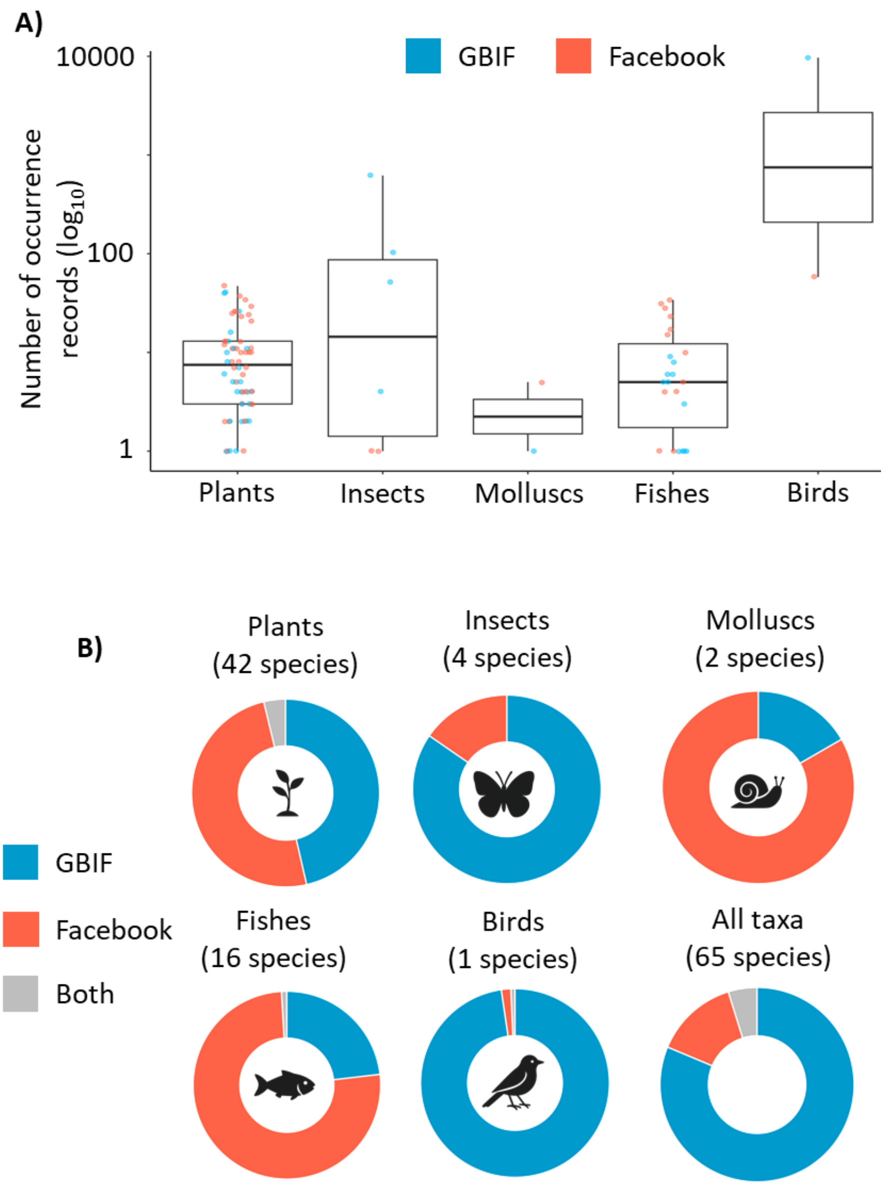

Figure 1. Taxonomic patterns in species occurrence data from GBIF and Facebook. A) The number of species occurrence records by taxa and data source, where each circle represent one species, and B) the proportion of 1 × 1 km grid cells containing species records from GBIF (blue), Facebook (red), or both sources (grey) for each taxonomic group. 'Both' (in B) indicates species were recorded by both sources within the same grid cell.

Our grid-based distribution analysis shows that the total number of distribution grids increased by 14% after incorporating Facebook data, while there was a 5% overlap in the grids (Figure 1B). Same as for occurrence records, we obtained more occurrence grid cells using Facebook data for plants, molluscs, and fishes (Figure 1B). For 43 species, we obtained a larger number of grid cells using Facebook data, whereas for the billygoat-weed (*Ageratum conyzoides*) we obtained an equal number of grid cells using GBIF and Facebook data.

Changes in invasive alien species distribution

We obtained MCPs for 54 species. Using the Overall data (combined GBIF and Facebook data), the range size increased for 44 species, compared to using only GBIF data, while there was no change for 10 species (Figure 2). Among species for which the range increased using the Overall data, the range size varied substantially across taxa (range 15-94,913 km^2, median = 24,811 km^2). For example, the range size increased by 15 km^2 for the invasive aquatic weed, the giant salvinia (*Salvinia molesta*), and 94,913 km^2 for the shrub, the diamond burbark (*Triumfetta rhomboidea*). For 15 species, the range polygon increased by > 10,000 km^2. Compared to the GBIF data, the mean range size increased dramatically across taxa using the Overall data, except for birds. The range size increased between 4,801 km^2 (insects) and 31,291 km^2 (fishes) (Figure 2).

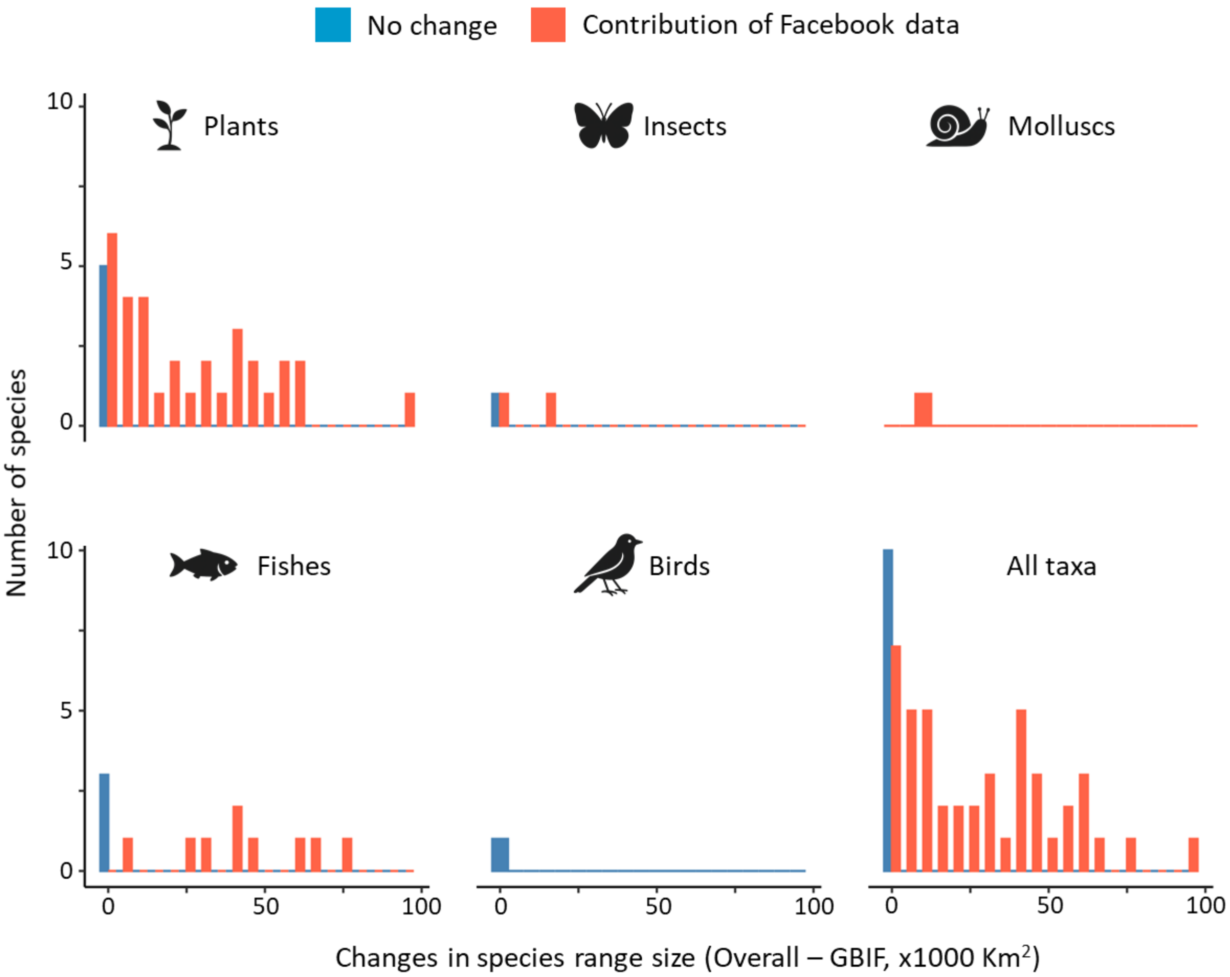

Figure 2. The contribution of species occurrence data from Facebook to expand invasive alien species range estimates. Histograms show changes in estimated range size (minimum convex polygons) across taxa when Facebook records were added to GBIF data. A value of 0

indicates no change in range size, which could result either from a lack of Facebook records for that species or from Facebook records being fully nested within existing GBIF ranges.

Patterns of data collection

Our linear models identified statistically significant differences ($p < 0.05$) between GBIF and Facebook data for 15 out of 25 predictor–taxon combinations (Figure 3; Supplementary Table S2). However, after correcting for multiple comparisons using the Benjamini–Hochberg false discovery rate, only 11 comparisons remained significant ($p_{adj} < 0.05$), all of which involved plants, insects, or molluscs. The strongest divergences were observed for plants and insects, where Facebook records were associated with significantly higher temperature, human footprint, and built area values. Among molluscs, only rainfall showed a significant difference, with lower values in Facebook records. In contrast, birds and fishes exhibited minimal differences between sources; for birds, only built areas were marginally significant ($p = 0.015$), and for fishes, elevation differed ($p = 0.038$), but neither remained significant after correction.

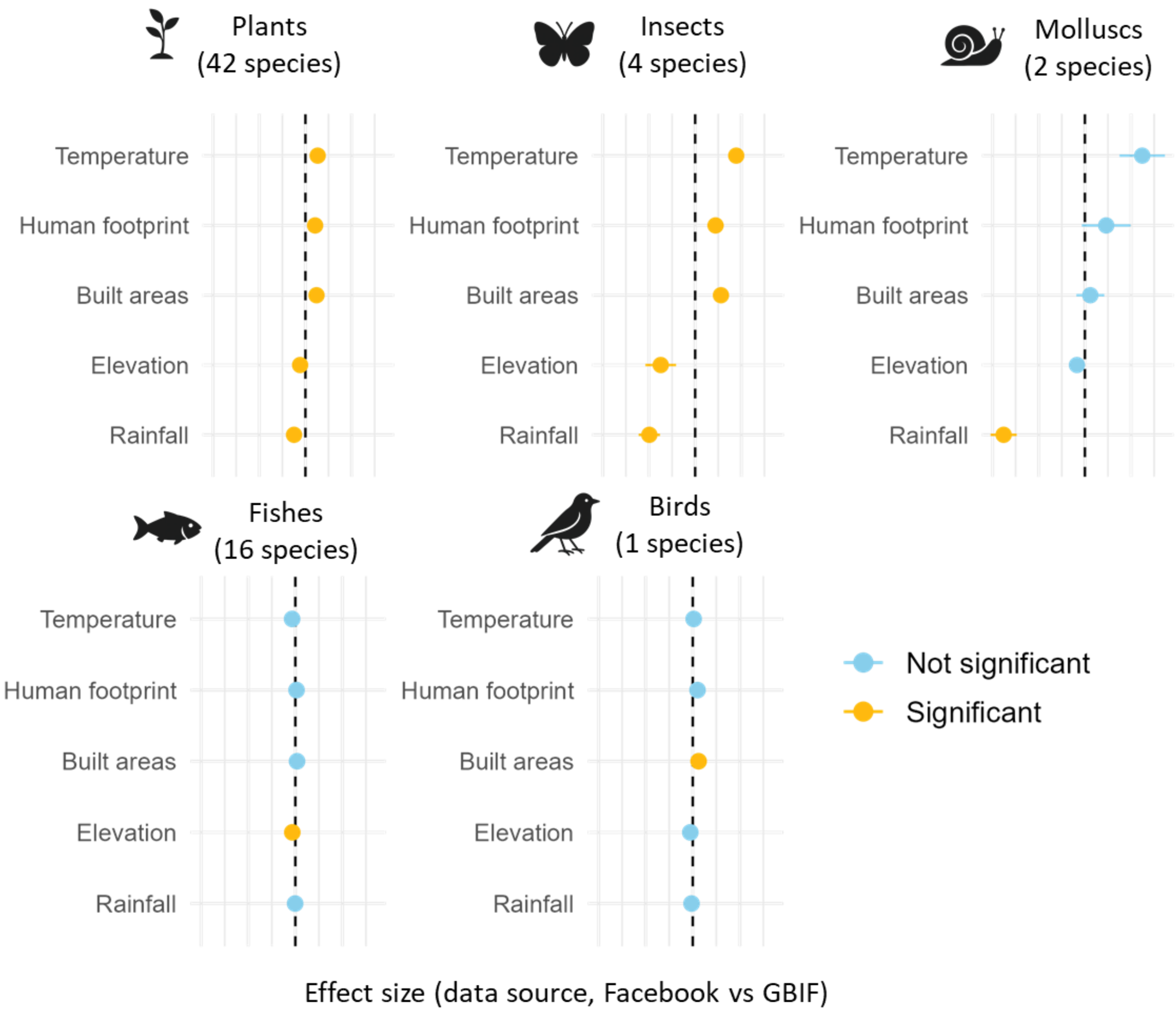

Figure 3. Effects of environmental variables on species occurrence across taxa and data sources.

Facebook can reduce critical distribution gaps

Invasive alien species are well-known for their severe negative impacts on ecosystems, economies, and human well-being (Mollot et al., 2017; Diagne et al., 2021). If not properly monitored, they can disrupt entire ecosystems (Peller & Altermatt, 2024). The GBF aims to halve the impact of invasive alien species by 2030. Achieving this goal requires detailed, up-to-date distribution data for each species—data that remain scarce across much of the tropics. Using Bangladesh as a case study, we demonstrate that incorporating species distribution data from Facebook can reduce this knowledge shortfall. Although Facebook contributed a relatively small proportion of the total records (6%), it provided more occurrence data than GBIF for two-thirds of the species analysed. In general, Facebook data provided a more diverse picture: for 66% of the 65 invasive alien species assessed, more occurrence records were available from Facebook than from GBIF. Notably, for 23 species, Facebook was the only source of distribution data. Incorporating Facebook data also expanded the species range size by two-thirds of the assessed species. These expanded ranges could reflect more realistic distribution patterns of invasive alien species that GBIF failed to capture due to limited sampling effort or reporting biases. Our results highlight the untapped potential of social media in improving invasive alien species monitoring, particularly in underrepresented regions. While we only focused on Facebook, future efforts could benefit from integrating data across multiple social platforms (e.g., Instagram, TikTok) to enhance biodiversity monitoring further.

Facebook posts reflect distinct environmental patterns

The data from GBIF indicated a significant bias towards birds, with one species, the rock pigeon, making up 85% of all records. This taxonomic bias corresponds with earlier studies on citizen science contributions, reflecting a predominance of birds and taxa from the Global North (Troudet et al., 2017). In contrast, Facebook exhibited a more equitable distribution of records across various taxa, especially for fishes, molluscs, and plants, including species not represented in GBIF. To further investigate the preferential bias among different data source, we compared Facebook and GBIF data using multiple environmental predictors. Facebook posts were primarily from areas with higher temperatures, greater human impacts, and more urban environments, particularly for plants and insects. This observation highlights that social media can yield valuable insights into species distributions in human-altered environments, which are frequently overlooked in conventional biodiversity monitoring. Such representation is crucial for tracking invasive alien species, considering that many tend to flourish in disrupted or urban settings (McKinney, 2006). By documenting these occurrences, social media could provide early warnings or new insights into the proliferation and spread of invasive alien species.

Although we only focused on Bangladesh, our results highlight the value of the approach to other tropical countries with limited biodiversity monitoring. However, it should be noted that the popularity of social media data may vary depending on the region, digital access, and taxonomic expertise, among other factors. Future research should explore how social media can be leveraged across different regions and platforms, and how data quality can be improved through community engagement, automated species recognition, and verification tools.

Limitations and ethical considerations in using social media data

Due to the lack of an automated extraction process, harvesting species locality data from Facebook is time-consuming (Chowdhury et al., 2024b). On average, it took about 25 minutes to complete the entire process (all locality records) for each species (from data extraction to geo-referencing). While time-intensive, this approach yielded valuable information, particularly for tropical species that are poorly represented in conventional databases. Given this potential, there is a pressing need to develop automated or semi-automated tools to extract and standardise biodiversity data from social media platforms (Sheard et al., 2024).

Like any other online data, when extracting distribution data from Facebook, we should always strictly maintain standard data privacy concerns (Di Minin et al., 2021). Only posts from public groups or pages should be used unless private group administrators give explicit permission. Furthermore, we should be careful when sharing such data publicly. Observer names should be anonymised, and detailed location data should be generalised where necessary to prevent unintended consequences, such as poaching of wildlife or habitat disturbance (Di Minin et al., 2021). Responsible data handling practices are essential to ensure that efforts to close biodiversity knowledge gaps do not introduce new risks for people or biodiversity.

In addition to privacy concerns, social media data may be prone to issues such as species misidentification, lack of metadata (e.g., date or observer effort), and duplicated records. Addressing these challenges requires careful curation, cross-validation with expert datasets, and, where possible, the integration of machine learning tools for automated image verification (van Klink et al., 2022; Sheard et al., 2024).

Conclusion

Detailed species distribution data are crucial for effective conservation planning and achieving global biodiversity goals, including those specified in the GBF. Using various approaches, we show that social media platforms such as Facebook can greatly improve the spatial and taxonomic coverage of invasive alien species data, especially in areas with limited data, like the tropics. Incorporating social media data with those from global biodiversity databases, such as GBIF, can help bridge the key knowledge gaps, highlight poorly-known species, and improve the known distribution of invasive alien species, many of which present immediate threats to ecosystems and local livelihoods. Conservation agencies and biodiversity data platforms should consider integrating social media records into biodiversity monitoring systems to harness this potential fully. By investing in data collection tools, validation processes, and community involvement in high-biodiversity yet under-monitored regions, the resolution and effectiveness of invasive alien species surveillance could see significant enhancement. Including these unconventional data sources in national and global biodiversity strategies can render monitoring more inclusive, adaptable, and impactful, particularly in the most needed areas.

References

Allain, S. J. (2019). Mining Flickr: a method for expanding the known distribution of invasive species. *Herpetological Bulletin*, *148*, 11-4.

Amano, T., Lamming, J. D., & Sutherland, W. J. (2016). Spatial gaps in global biodiversity information and the role of citizen science. *Bioscience*, *66*(5), 393-400.

Baasanmunkh, S., Oyuntsetseg, B., Tsegmed, Z., Undruul, A., Munkhtulga, D., Urgamal, M., ... & Choi, H. J. (2025). iNaturalist projects represent a valuable resource for aggregating plant observations and engaging society: A case study of the Flora of Mongolia project. *Plants, People, Planet* (in press).

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J., & Pocock, M. J. (2025). Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews*, *100*(1), 50-67.

Bradshaw, C. J., Hoskins, A. J., Haubrock, P. J., Cuthbert, R. N., Diagne, C., Leroy, B., ... & Courchamp, F. (2021). Detailed assessment of the reported economic costs of invasive species in Australia. *NeoBiota*, 67, 511-550.

Caley, P., & Cassey, P. (2023). Do we need to mine social media data to detect exotic vertebrate-pest introductions? *Wildlife Research*, *50*(11), 869-875.

Callaghan, C. T., Poore, A. G., Mesaglio, T., Moles, A. T., Nakagawa, S., Roberts, C., ... & Cornwell, W. K. (2021). Three frontiers for the future of biodiversity research using citizen science data. *BioScience*, *71*(1), 55-63.

Callaghan, C. T., Mesaglio, T., Ascher, J. S., Brooks, T. M., Cabras, A. A., Chandler, M., ... & Young, A. N. (2022). The benefits of contributing to the citizen science platform iNaturalist as an identifier. *PLoS Biology*, *20*(11), e3001843.

Capinha, C., Essl, F., Seebens, H., Moser, D., & Pereira, H. M. (2015). The dispersal of alien species redefines biogeography in the Anthropocene. *Science*, *348*(6240), 1248-1251.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., ... & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59-67.

Convention on Biological Diversity (CBD). (2022). *Kunming–Montreal Global biodiversity framework*. Draft decision submitted by the President CBD/COP/15/L.25, 18 December 2022. https://www.cbd.int/conferences/2021-2022/cop-15/documents

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., Ram, K. (2025). _rgbif: Interface to the Global Biodiversity Information Facility API_. R package version 3.8.1, <https://CRAN.R-project.org/package=rgbif>.

Chowdhury, S., Hawladar, N., Roy, R. C., Capinha, C., Cassey, P., Correia, R. A., ... & Bonn, A. (2025). Harnessing social media data to track species range shifts. *EcoEvoRxiv*. https://doi.org/10.32942/X2R63N.

Chowdhury, S., Fuller, R. A., Ahmed, S., Alam, S., Callaghan, C. T., Das, P., ... & Bonn, A. (2024a). Using social media records to inform conservation planning. *Conservation Biology*, *38*(1), e14161.

Chowdhury, S., Ahmed, S., Alam, S., Callaghan, C. T., Das, P., Di Marco, M., Di Minin, E., Jarić, I., Labi, M. M., Rokonuzzaman, M., Roll, U., Sbragaglia, V., Siddika, A., & Bonn, A. (2024b). A protocol for harvesting biodiversity data from Facebook. *Conservation Biology*, 38, e14257.

Chowdhury, S., Aich, U., Rokonuzzaman, M., Alam, S., Das, P., Siddika, A., ... & Callaghan, C. T. (2023a). Increasing biodiversity knowledge through social media: A case study from tropical Bangladesh. *BioScience*, *73*(6), 453-459.

Chowdhury, S., Fuller, R. A., Rokonuzzaman, M., Alam, S., Das, P., Siddika, A., ... & Hanson, J. O. (2023b). Insights from citizen science reveal priority areas for conserving biodiversity in Bangladesh. *One Earth*, *6*(10), 1315-1325.

Collen, B., Ram, M., Zamin, T., & McRae, L. (2008). The tropical biodiversity data gap: addressing disparity in global monitoring. *Tropical Conservation Science*, *1*(2), 75-88.

Correia, R. A., Ladle, R., Jarić, I., Malhado, A. C., Mittermeier, J. C., Roll, U., ... & Di Minin, E. (2021). Digital data sources and methods for conservation culturomics. *Conservation Biology*, *35*(2), 398-411.

Danielsen, F., Jensen, P. M., Burgess, N. D., Altamirano, R., Alviola, P. A., Andrianandrasana, H., ... & Young, R. (2014). A multicountry assessment of tropical resource monitoring by local communities. *BioScience*, *64*(3), 236-251.

Di Minin, E., Fink, C., Hausmann, A., Kremer, J., & Kulkarni, R. (2021). How to address data privacy concerns when using social media data in conservation science. *Conservation Biology*, *35*(2), 437-446.

Di Minin, E., Tenkanen, H., & Toivonen, T. (2015). Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, *3*, 63.

Diagne, C., Leroy, B., Vaissière, A. C., Gozlan, R. E., Roiz, D., Jarić, I., ... & Courchamp, F. (2021). High and rising economic costs of biological invasions worldwide. *Nature*, *592*(7855), 571-576.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, *345*(6195), 401-406.

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., ... & Tatem, A. J. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications*, *7*(1), 12485.

ElQadi, M. M., Dyer, A. G., Vlasveld, C., & Dorin, A. (2023). The spatiotemporal signature of cherry blossom flowering across Japan revealed via analysis of social network site images. *Flora*, *304*, 152311.

ElQadi, M. M., Dorin, A., Dyer, A., Burd, M., Bukovac, Z., & Shrestha, M. (2017). Mapping species distributions with social media geo-tagged images: Case studies of bees and flowering plants in Australia. *Ecological Informatics*, *39*, 23-31.

GBIF. (2025). GBIF Occurrence Download https://www.gbif.org/occurrence/download/0002603-250525065834625. Accessed from R via rgbif (https://github.com/ropensci/rgbif) on 2025-05-27.

Green, E. J., Buchanan, G. M., Butchart, S. H., Chandler, G. M., Burgess, N. D., Hill, S. L., & Gregory, R. D. (2019). Relating characteristics of global biodiversity targets to reported progress. *Conservation Biology*, *33*(6), 1360-1369.

Heberling, J. M., Miller, J. T., Noesgaard, D., Weingart, S. B., & Schigel, D. (2021). Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences*, *118*(6), e2018093118.

Hijmans R (2025). _terra: Spatial Data Analysis_. R package version 1.8-50, <https://CRAN.R-project.org/package=terra>.

Hijmans RJ, Barbosa M, Ghosh A, Mandel A (2024). _geodata: Download Geographic Data_. R package version 0.6-2, <https://CRAN.R-project.org/package=geodata>.

Hortal, J., De Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 523-549.

IUCN 2025. The IUCN Red List of Threatened Species. Version 2025-1. <https://www.iucnredlist.org>

IUCN Bangladesh. (2015). *Red list of Bangladesh: A brief on assessment result 2015*. International Union for Conservation of Nature, Bangladesh Country Office.

Jarić, I., Correia, R. A., Brook, B. W., Buettel, J. C., Courchamp, F., Di Minin, E., ... & Roll, U. (2020). iEcology: harnessing large online resources to generate ecological insights. *Trends in Ecology & Evolution*, *35*(7), 630-639.

Jarić, I., Bellard, C., Correia, R. A., Courchamp, F., Douda, K., Essl, F., ... & Roll, U. (2021). Invasion culturomics and iEcology. *Conservation Biology*, *35*(2), 447-451.

Joppa, L. N., O'Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., ... & Burgess, N. D. (2016). Filling in biodiversity threat gaps. *Science*, *352*(6284), 416-418.

Marcenò, C., Padullés Cubino, J., Chytrý, M., Genduso, E., Salemi, D., La Rosa, A., Gristina, A. S., Agrillo, E., Bonari, G., Giusso Del Galdo, G., Ilardi, V., Landucci, F., & Guarino, R. (2021). Facebook groups as citizen science tools for plant species monitoring. *Journal of Applied Ecology*, *58*, 2018–2028.

Mason, B. M., Mesaglio, T., Heitmann, J. B., Chandler, M., Chowdhury, S., Gorta, S. B. Z, Grattarcola, F., Groom, Q., Hitchcock, C., Hoskins, L., Lowe, S. K., Marquis, M., Pernat, N., Shirey, V., Baasanmunkh, S., & Callaghan, C. T. iNaturalist is Shaping the Future of Biodiversity Research. BioScience (in press).

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. *Biological Conservation*, *127*(3), 247-260.

Mollot, Grégory, J. H. Pantel, and T. N. Romanuk (2017). The effects of invasive species on the decline in species richness: a global meta-analysis. *Advances in Ecological Research* 56, 61-83.

Moore, R. A., Symonds, M. R., & Howard, S. R. (2024). Leveraging social media and community science data for environmental niche models: A case study with native Australian bees. *Ecological Informatics*, *84*, 102857.

Mukul, S. A., Khan, M. A. S. A., & Uddin, M. B. (2020). Identifying threats from invasive alien species in Bangladesh. *Global Ecology and Conservation*, *23*, e01196.

Pebesma, E., & Bivand, R. (2023). Spatial Data Science: With Applications in R. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016.

Peller, T., & Altermatt, F. (2024). Invasive species drive cross-ecosystem effects worldwide. *Nature Ecology & Evolution*, *8*(6), 1087-1097.

Pereira, H. M., Navarro, L. M., & Martins, I. S. (2012). Global biodiversity change: the bad, the good, and the unknown. *Annual Review of Environment and Resources*, *37*(1), 25-50.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J. H., ... & Kemper, T. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. *International Journal of Digital Earth*, *17*(1), 2390454.

Pocock, M. J., Roy, H. E., August, T., Kuria, A., Barasa, F., Bett, J., ... & Trevelyan, R. (2019). Developing the global potential of citizen science: Assessing opportunities that benefit people, society and the environment in East Africa. *Journal of Applied Ecology*, *56*(2), 274-281.

Pocock, M. J., Chandler, M., Bonney, R., Thornhill, I., Albin, A., August, T., ... & Danielsen, F. (2018). A vision for global biodiversity monitoring with citizen science. *Advances in Ecological Research* 59, 169-223.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robinson D, Hayes A, Couch S (2025). _broom: Convert Statistical Objects into Tidy Tibbles_. R package version 1.0.8, https://CRAN.R-project.org/package=broom.

Rosa, A. H., & Freitas, A. V. (2024). The role of citizens in conservation science: a case study with threatened Brazilian butterflies. *Journal of Insect Conservation*, *28*(6), 1149-1160.

Ross N (2024). _fasterize: Fast Polygon to Raster Conversion_. R package version 1.1.0, <https://CRAN.R-project.org/package=fasterize>.

Rothman, S. B., Gayer, K., & Stern, N. (2020). A long-distance traveler: the peacock rockskipper Istiblennius meleagris (Valenciennes, 1836) on the Mediterranean intertidal reefs. *Biological Invasions*, *22*(8), 2401-2408.

Sbragaglia, V., Espasandín, L., Coco, S., Felici, A., Correia, R. A., Coll, M., & Arlinghaus, R. (2022). Recreational angling and spearfishing on social media: insights on harvesting patterns, social engagement and sentiments related to the distributional range shift of a marine invasive species. *Reviews in Fish Biology and Fisheries*, *32*(2), 687-700.

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., ... & Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*(1), 14435.

Sheard, J. K., Adriaens, T., Bowler, D. E., Büermann, A., Callaghan, C. T., Camprasse, E. C., ... & Bonn, A. (2024). Emerging technologies in citizen science and potential for insect monitoring. *Philosophical Transactions of the Royal Society B*, *379*(1904), 20230106.

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., ... & Vilà, M. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*, 28(1), 58-66.

Šmejkal, M., Dočkal, O., Thomas, K., Verma, C. R., Kumkar, P., & Kalous, L. (2024). First record of highly invasive Chinese sleeper Perccottus glenii Dybowski, 1877 (Perciformes: Odontobutidae) in the Elbe River Basin, Czechia. *Aquatic Ecology*, *58*(1), 125-130.

Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., & Packer, C. (2017). Future threats to biodiversity and pathways to their prevention. *Nature*, *546*(7656), 73-81.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., ... & Di Minin, E. (2019). Social media data for conservation science: A methodological overview. *Biological Conservation*, *233*, 298-315.

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports*, *7*(1), 9132.

Van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., ... & Bowler, D. E. (2022). Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution*, *37*(10), 872-885.

Vardi, R., Soriano-Redondo, A., Gutiérrez, J. S., Dylewski, Ł., Jagiello, Z., Mikula, P., ... & Sbragaglia, V. (2024). Leveraging social media and other online data to study animal behavior. *PLoS Biology*, *22*(8), e3002793.

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." _Journal of Open Source Software_, *4*(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wildlife Conservation Society - WCS, and Center for International Earth Science Information Network - CIESIN - Columbia University. 2005-12-31. Last of the Wild Project, Version 2, 2005 (LWP-2): Global Human Footprint Dataset (Geographic). Version 2.00. Palisades, NY. Archived by National Aeronautics and Space Administration, U.S. Government, NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4M61H5F. https://doi.org/10.7927/H4M61H5F.

WWF. (2024). Living Planet Report 2024 – A System in Peril. WWF, Gland, Switzerland.