Received: 28 November 2024 | Revised: 31 August 2025 | Accepted: 24 November 2025

DOI: 10.1111/cobi.70234

Conservation Biology

CONTRIBUTED PAPER

Harnessing social media data to track species range shifts

Shawan Chowdhury[1] | Niloy Hawladar[2] | Ripon C. Roy[2] | César Capinha[3,4] | Phillip Cassey[5] | Ricardo A. Correia[6,7,8] | Gideon Gywa Deme[8,9,10] | Moreno Di Marco[11] | Enrico Di Minin[6,7,12] | Ivan Jarić[13,14] | Richard J. Ladle[15] | Jonathan Lenoir[16] | Mohammad Momeny[6] | Jooel J. Rinne[6] | Uri Roll[17] | Aletta Bonn[18,19,20]

[1]School of Biological Sciences, Monash University, Melbourne, Victoria, Australia

[2]Department of Zoology, University of Dhaka, Dhaka, Bangladesh

[3]Centre of Geographical Studies, Institute of Geography and Spatial Planning, University of Lisbon, Lisbon, Portugal

[4]Associate Laboratory Terra, Lisbon, Portugal

[5]School of Biological Sciences, The University of Adelaide, Adelaide, South Australia, Australia

[6]Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

[7]Helsinki Institute of Sustainability Science, University of Helsinki, Helsinki, Finland

[8]Department of Biodiversity Sciences, University of Turku, Turku, Finland

[9]Turku Collegium for Science, Medicine, and Technology (TCSMT), University of Turku, Turku, Finland

[10]Department of Zoology, University of Jos, Jos, Nigeria

[11]Department of Biology and Biotechnologies Charles Darwin, Sapienza University of Rome, Rome, Italy

[12]Centre for Sustainability Transitions, Stellenbosch University, Stellenbosch, South Africa

[13]Université Paris-Saclay, CNRS, AgroParisTech, Ecologie Systématique Evolution, Gif-sur-Yvette, France

[14]Institute of Hydrobiology, Biology Centre of the Czech Academy of Sciences, České Budějovice, Czech Republic

[15]Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió, Brazil

[16]UMR CNRS 7058, Ecologie et Dynamique des Systèmes Anthropisés (EDYSAN), Université de Picardie Jules Verne, Amiens, France

[17]Mitrani Department of Desert Ecology, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion, Israel

[18]Department of Biodiversity and People, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

[19]Institute of Biodiversity, Ecology and Evolution, Friedrich Schiller University Jena, Jena, Germany

[20]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

Correspondence
Shawan Chowdhury, School of Biological Sciences, Monash University, 25 Rainforest Walk, Clayton VIC 3168, Australia.
Email: shawan.chowdhury@monash.edu

Article impact statement: There is the potential to harness social media data to track biodiversity redistribution in response to climate change.

Abstract

Biodiversity monitoring programs and citizen science data remain heavily biased toward the Global North. Especially in megadiverse countries with limited biodiversity records, incorporating social media data can help address existing data gaps. However, whether such data can significantly improve our understanding of range-shifting species is still unknown. We tested whether social media data improved our knowledge of the range dynamics of a rapid range-shifting butterfly, the tawny coster (*Acraea terpsicore*). We collated locality data from Flickr and Facebook and compared these with occurrence data from the Global Biodiversity Information Facility (GBIF). We used species distribution models (SDMs) and niche assessments, which we calibrated with data from GBIF alone and both sources combined

Funding information
German Research Foundation, Grant/Award Number: DFG-FZT 118 and 202548816; Australian Research Industry Laureate Fellow; Combatting Wildlife Crime and Preventing Environmental Harm, Grant/Award Number: IL230100175; British Ecological Society, Grant/Award Number: CE24/1043; European Union, Grant/Award Number: ERC, BIOBANG, 101171602; CNPq, Grant/Award Number: 447598/2025-2, 441125/2023-9, 306174/2025-1

(GBIF and social media data) to analyze range shift dynamics. Social media data increased occurrence records by 35%, and the proportion of social media data was higher in countries poorly represented in GBIF. In addition, we obtained new distributional information from well-represented countries (e.g., Australia and Malaysia). Over time, the SDMs calibrated with GBIF and social media data showed greater expansion rates than SDMs based solely on GBIF data. The niche assessments revealed that GBIF-only data failed to capture regions with relatively low maximum temperature, relatively low precipitation and high elevation. Our results highlight the potential of harnessing social media data to track rapid biodiversity redistribution in response to climate change.

KEYWORDS
biodiversity shortfall, citizen science, Facebook, Flickr, iEcology, invasive species, range shift, social media

1 | Introduction

The current era is marked by a pressing biodiversity crisis (Diaz et al., 2019; Dirzo et al., 2014; Pimm et al., 2014), driven by various biodiversity threats, such as agricultural intensification and expansion, habitat loss and fragmentation, biological invasions, and climate change (Butchart et al., 2010; Jaureguiberry et al., 2022; Joppa et al., 2016; Maxwell et al., 2016; Wilson, 1989). In response to these combined pressures on the environment, many species have already shifted (and are still shifting) their distributions (Chan et al., 2024; Chen et al., 2011; Di Marco et al., 2021; Lenoir & Svenning, 2015; Lenoir et al., 2020; Yackulic et al., 2011). Indeed, biodiversity redistribution in response to the consequences of human activities is now a global phenomenon (Lenoir & Svenning, 2015; Lenoir et al., 2020).

However, our understanding of this phenomenon is biased both geographically and taxonomically (Feeley et al., 2017; Hortal et al., 2015; Lenoir & Svenning, 2015). For instance, a systematic review of 258 peer-reviewed studies reporting species range shifts revealed that nearly 40% of the identified range shifts (12,415 species) were of flowering plants, and 22%, 12%, 4%, and 0.5% were of insects, birds, fishes, and mammals, respectively (Lenoir et al., 2020). Even within insects, there is a strong bias toward specific groups, such as orthoptera and dragonflies (5% of total documented range shifts; Lenoir et al., 2020). This taxonomic bias is compounded by a severe geographical bias; only a limited number of studies are from the tropics and the Global South (Lawlor et al., 2024; Lenoir et al., 2020; Parker et al., 2024; Chowdhury, 2023), including many megadiverse countries (i.e., countries with multiple biodiversity hotspots). Assessing species range shifts requires detailed species distribution data, typically unavailable for species from the Global South (Hortal et al., 2015; Hughes et al., 2021). This, in turn, biases our global overview of species undergoing climate-induced range shifts. To improve our understanding of species redistribution at a truly global scale, it is thus essential that we compile data from all sources to better test how species are responding to global change.

Rapid technological advances involving smartphones, digital cameras, and fast Internet have revolutionized data collection (Sheard et al., 2024; Van Klink et al., 2022). Moreover, anyone from anywhere in the world now has the potential to share their biodiversity observations on a variety of digital platforms (e.g., apps, web pages, and social media; Caley & Cassey, 2023; Chandler et al., 2017; Pocock et al., 2018; Toivonen et al., 2019). When appropriately harvested, such data can reduce knowledge gaps in biodiversity distribution and monitoring and help answer questions about how species respond to global changes (Jarić et al., 2020; Soriano-Redondo et al., 2024). This is especially relevant for countries where systematic biodiversity monitoring programs are uncommon and that are poorly represented in global biodiversity repositories (Chowdhury, Aich, et al., 2023; Marcenò et al., 2021; Mota et al., 2022). For example, Global Biodiversity Information Facility (GBIF) data are highly biased toward temperate countries, whereas very little is known about the distribution of tropical species (Troudet et al., 2017).

Recent studies have demonstrated that biodiversity data posted on Facebook can be more representative than those available from the GBIF (Chowdhury, Aich, et al., 2023). Using a comprehensive set of biodiversity data in ecological research is key to assessing species redistribution in the context of global change. Combining social media data with GBIF data could substantially improve conservation assessments (Chowdhury, Fuller, et al., 2024), which can be useful to identify priority areas for protection (Chowdhury, Ahmed, et al., 2024). Nevertheless, biodiversity observation data from social media have rarely been used in conservation assessments (Di Minin et al., 2015; Jensen et al., 2019; Phakoago et al., 2024) or for tracking range-shifting species (Chowdhury et al., 2026; Sbragaglia et al., 2024).

As a proof of concept, we aimed to address this important knowledge gap by harvesting and harnessing social media data to evaluate the distribution pattern of an ecologically important butterfly that is currently undergoing a rapid range expansion in response to global changes: the tawny coster (*Acraea terpsicore*). We chose the tawny coster as a model species first because this species is currently experiencing a well-documented active range expansion driven by global change, making it ecologically significant for understanding real-time shifts in distribution (Chowdhury, Fuller et al., 2021). Second, as a highly visible and easily recognizable species, the tawny coster is often photographed and shared on public platforms, boosting its detectability through social media. Although this method might not be equally effective for more cryptic or less charismatic species, our study helps identify the conditions where social media data can be most valuable and highlights the opportunity to focus similar efforts on other highly visible species.

We compiled occurrence records of the tawny coster from social media (Facebook and Flickr) and combined these with GBIF records. We anticipated that social media data would add new locality information when combined with GBIF records, covering otherwise unrepresented areas. We hypothesized that the geographic range of this iconic species would differ from the known range extent (Chowdhury, Braby, et al., 2021) and that this would produce higher range expansion rates. Finally, we predicted that the additional social media data would identify novel combinations of the environmental niche space. Building on our findings, we provided recommendations on using social media data to answer general ecological questions related to species' biogeography and macroecology that can affect real change in conservation policies.

2 | Methods

2.1 | Tawny coster

The tawny coster butterfly has a well-documented geographic range area (Chowdhury, Braby, et al., 2021), and its charisma, like in many other butterfly species, attracts high public attention on social media. This butterfly is native to the Indian subcontinent (India, Bangladesh, and Sri Lanka) (Braby et al., 2014; Chowdhury, Braby, et al., 2021), and since the 1980s, it has rapidly expanded its range to other parts of South Asia (e.g., Bhutan, Nepal, and Pakistan) and Southeast Asia (e.g., Malaysia, Singapore, and Thailand), eventually entering Australia in April 2012 (Braby et al., 2014). It was first recorded in Australia near Darwin in the Northern Territory (Sanderson et al., 2012). In subsequent years, the species spread toward Western Australia (to Kimberly) and subsequently toward Queensland. Since its arrival in Australia, the tawny coster has expanded its range within the country at a rate of approximately 135 km/year, while remaining within its native climatic niche (Chowdhury, Braby, et al., 2021). The species can cover a wide range of habitats and can migrate long distances (Chowdhury, Fuller, et al., 2021), which has facilitated its expansion into new regions.

2.2 | Data

We collated occurrence records and locality data for tawny coster's sightings from 3 different sources: GBIF, Flickr, and Facebook. For all 3 data sources, we selected data from January 2005 to May 2024 to maintain a comparable sampling period. We used 2 approaches to remove potential duplicates in occurrence records: exact duplicates (i.e., rows containing the exact same values for all the columns) and a comparison of the distribution of occurrence data and occurrence records in the same 21.625-km^2 (4.65 × 4.65 km^2) grid cells (based only on longitude and latitude coordinates), which we used for niche modeling and the niche overlap analysis. When searching for distribution data on Flickr and Facebook, we used only the scientific name (*Acraea terpsicore*) and the English common name (tawny coster) for consistency with GBIF. Therefore, we did not use any local language names, which might have caused some geographic bias in our social media data collection, but this should only underestimate the power of social media data over GBIF data.

We downloaded GBIF data manually from the website (https://www.gbif.org/; GBIF, 2024). The GBIF portal is a collection of hundreds of citizen science applications, including iNaturalist (Heberling et al., 2021), so we did not look for citizen science data from other potential sources. In particular, 95% of GBIF data for this species come from iNaturalist. When downloading occurrence records from GBIF for the tawny coster, we kept only the presence data with coordinate uncertainty below 10 km. It should be noted that many GBIF occurrence records lack information on coordinate uncertainty, and our choice to ignore these records resulted in a reduced but more reliable sample.

We used a Python script, which relies on Flickr's application programming interface (API) (https://www.flickr.com/services/developer) and its keyword search, to collect all publicly available Flickr posts related to the tawny coster. We then removed duplicates and any posts not containing a geotag from the data. Using the URLs of the posts, we manually confirmed the identity of the tawny coster in the photographs.

For Facebook data, we followed the protocol developed by Chowdhury, Ahmed, et al. (2024). Specifically, the data extraction process was divided into 3 steps. First, we searched for butterfly groups through a combination of taxon and country names (Appendix S1). We collected all the known distributions (17 country names used as keywords) of the tawny coster from Chowdhury, Braby, et al. (2021). When searching for Facebook groups, we included 10 more countries from the surrounding area. With these 27 keywords corresponding to 27 countries, we identified 41 Facebook groups from 17 countries (Appendix S1). We reviewed each photograph and validated the species' information. From each photograph, we extracted date (day, month, year), location, and photographer's information. We excluded photographs if their quality was unsuitable for identification of the species, if a specific date and location were not provided, or if the location provided in the photographs was unspecified (>10 km uncertainty). Finally, using the locations obtained from Facebook posts, we searched the locality's name in Google Maps (https://www.google.com/maps) to georeference it and retrieve its coordinates (longitude and latitude).

During the initial data cleaning process, we removed all duplicate records (same coordinates) and kept only records between January 2005 and May 2024. Our compiled dataset included 6459 occurrence records (GBIF: 4206; Flickr: 325; Facebook: 1928). The Facebook and Flickr data are in Appendix S2, and the GBIF data are publicly available (GBIF, 2024).

We used the TerraClimate database (https://www.climatologylab.org/terraclimate.html; Abatzoglou et al., 2018) to obtain annual climatic predictor variables (from 2005–2023) at a 21.625-km^2 resolution. We used TerraClimate instead of commonly used climate databases like WorldClim or CHELSA, because it provided more recent variables that we used for temporal matching. The climatic data for 2024 were not available at the time we ran our analyses. We focused on 10 ecologically relevant bioclimatic variables available in TerraClimate to model the species' distribution, encompassing environmental factors that influence butterfly life cycles, particularly temperature, water availability (evapotranspiration, precipitation, and soil moisture), and environmental stress indicators, such as the drought index and wind speed. These included 10 bioclimatic variables: actual evapotranspiration, climate water deficit, potential evapotranspiration, precipitation (monthly total), soil moisture, maximum temperature, minimum temperature, wind speed, runoff (monthly total), and the Palmer Drought Severity Index. We further downloaded elevation data from the WorldClim (https://www.worldclim.org/data/worldclim21.html) database at the same spatial resolution of 21.625 km^2.

2.3 | Data preparation

We analyzed the range-shift dynamic of the tawny coster by splitting the 2005–2024 period into 5 intervals of 4 years each (2005–2008; 2009–2012; 2013–2016; 2017–2020; and 2021–2024) and assigned the occurrence and climatic data, available at a yearly resolution, according to each of those 5 periods. For occurrence data, initially we held GBIF and social media datasets separate. For each year interval and dataset, we kept a single occurrence record per grid cell (21.625 km^2). Finally, we merged the GBIF and social media datasets to create the combined dataset, so we had 2 datasets for the subsequent analyses: the GBIF-only dataset versus the combined dataset.

For each of the 11 environmental variables, we cropped the layers to the study extent (minimum longitude 60.875, maximum longitude 158.9583, minimum latitude −54.75, maximum latitude 53.54167) and calculated the mean conditions for all bioclimatic variables except elevation and for the 5 time intervals. Given that the bioclimatic data for 2024 were not published yet, for the last period (2021–2024) we computed mean climatic conditions over 3 years instead of 4 years. We checked for multicollinearity among the 9 bioclimatic variables and elevation and removed highly correlated ones ($|r| > 0.7$) (Brun et al., 2020). Consequently, we removed 4 bioclimatic variables and kept the 7 remaining variables for the final analysis: climate water deficit, precipitation (monthly total), soil moisture, maximum temperature, wind speed, the Palmer Drought Severity Index, and elevation.

2.4 | Habitat suitability maps

To obtain habitat suitability maps for the tawny coster, we fitted MaxEnt species distribution models (SDMs) (Elith et al., 2010; Phillips et al., 2017) in R (R Core Team, 2024) using the ENMeval package (Kass et al., 2021). We ran the model twice for each of the 5 time intervals, once for the GBIF-only dataset and a second time with the combined dataset.

We fit SDMs with 7 predictor variables at a 21.625-km^2 resolution (2.5 arc minutes). To control for potential biases in the citizen science observations, for background points, we followed the target-group pseudoabsences approach (Phillips et al., 2009). Specifically, we used the occurrence records of a similar range-shifting butterfly species, *Papilio demoleus* (9834 points; GBIF, 2025), as the pseudoabsence points. For all 5 intervals and for both the model relying on GBIF-data only and the model relying on the combined dataset, we used the same set of pseudoabsences across the entire study extent (Appendix S3). We did that to avoid the possibility that the background selection strategy will affect model outputs when comparing model performances over time and between data sources (GBIF-only vs. combined data). Before fitting the model, we removed duplicate values in each raster cell and created a 500-km buffer around the spatial records. We cropped the environmental variables to the buffered region to limit model overfitting. We assigned the records to grid cells and then randomly assigned grid cells to particular folds (validation and training bins; Kass et al., 2021). We used the checkerboard2 evaluation method (with the presence and background points), which can handle overinflation of model performance, at least from biased sampling. This evaluation method partitions geospatial records and background points into evaluation bins to reduce spatial autocorrelation between points in the testing and training bins.

To improve MaxEnt's modeling performance, we performed a calibration procedure by fitting the model under different combinations of parameters and hyperparameters. Specifically, we fitted the model under 6 feature class combinations (L, LQ, H, LQH, LQHP, and LQHPT, where L is linear, Q is quadratic, H is hinge, P is product, and T is threshold) and 8 different regularization multipliers (0.5–4 at 0.5 intervals). Although the feature class allows MaxEnt to develop composite models to ensure a good fit to the data, regularization multiplier values control model overfitting (Kass et al., 2021).

Overall, there were 48 models (6 feature classes × 8 regularization multipliers) for each data group (GBIF-only vs. combined data) for each time interval. We used the raster package in R (Hijmans, 2023) to choose the best model with the lowest corrected Akaike information criterion (AICc) (model evaluation metric) (Kass et al., 2021). We used the 10% omission rate threshold value (Pearson et al., 2007) and transformed the suitability map into binary classes based on the threshold value (1 for presence with suitability value ≥ threshold value and 0 for absence with suitability value < threshold value). We calculated the centroid position of these binary maps and used the geosphere (Hijmans, 2022) package in R to calculate the range expansion rate.

2.5 | Niche assessment

We used the ecospat R package (Broennimann et al., 2023) to evaluate whether the use of additional data from social media led to significant differences in the realized niche space occupied by the species. We quantified the niche overlap between the GBIF-only and combined dataset, for each time interval separately. We used the same 7 environmental variables that were used for fitting SDMs. We extracted the environmental data corresponding to the occurrence records and ran the principal component analysis (PCA) to reduce dimensionality. We transformed the first 2 components (PC1 and PC2) into density by kernel smoothers. We quantified niche overlap with Schoener's *D*. This metric varies between 0%, if there is no overlap between niches estimated from the 2 datasets, and 100%, if there is complete overlap (Broennimann et al., 2012).

To test whether the niches of GBIF-only and combined data were statistically equivalent, we conducted a niche equivalency test by randomly permuting the identities of the occurrences across datasets ($n = 100$ permutations). Afterward, we conducted niche similarity tests (also with 100 permutations) in both directions, comparing the observed overlap to a null distribution of overlaps generated by randomly shifting one niche within the background environment of the other. We separately ran both tests for each time interval and assessed the significance by the proportion of simulated overlaps that were equal to or greater than the observed value following Broennimann et al. (2012).

3 | Results

Our cleaned and compiled dataset included 6459 occurrence records, of which 65% (4206) were from GBIF and 35% (2253) were from social media (Flickr, 5%; Facebook, 30%). We noticed marked differences between data sources when we compared the number of occurrence records across countries (Figure 1). For most countries (13 out of 17), the number of occurrence records substantially increased (>25%) after including social media data (e.g., data increased from 10 to 224 for Bangladesh and from 262 to 468 for Malaysia) (Figure 1a,b). The distribution of the tawny coster is known from 17 countries (Chowdhury, Braby, et al., 2021), and we obtained a higher percentage of data than GBIF from social media for 5 of these countries (Figure 1a,b; Appendix S2). For example, we obtained 3096 occurrence records from India, of which 64% (1968) were from GBIF and 36% (1128) were from social media (Figure 1a). The percentages of species occurrence records from social media were higher for countries with a lower number of total occurrence records, but >10% in all cases (range 10–44%) (Figure 1b).

Over time, there were substantial differences in the temporal distribution of occurrence records (Figure 1c). Although the initial period, from 2005 to 2007, had a larger percentage of occurrence records from GBIF, subsequent years (2008–2018) were characterized by a higher percentage of occurrence records from social media, except for 2013. Following a substantial decline during 2017–2022, the proportion of social media data stabilized recently (Figure 1c).

The addition of social media data in our SDMs contributed substantially to the identification of potentially new suitable areas for the tawny coster. Although the mean model performance, based on the area under the curve (AUC) statistic, was higher when using combined data, the AUC value was very similar across time intervals (Appendix S4). The total surface area predicted to be suitable for the tawny coster was always larger (0.9–4.6 million km^2) when combining social media data with GBIF data than when relying on only GBIF data (Figure 2). New suitable areas identified with the addition of social media data were mostly distributed in South Asia (especially toward higher elevations in the Indo-Himalaya region) during 2005–2008, whereas new suitable areas identified with the addition of social media data during 2017–2020 were distributed throughout the entire region and especially toward higher latitudes (Figure 2; Appendix S5). When we combined all the suitability maps over the 5 intervals, we found that predictions from the models relying on GBIF-only data missed many areas at higher latitudes and at mid to high elevations (Figure 2g). In terms of range expansion, the combined data captured a larger expansion area, except in the third interval (2013–2016). However, it increased by 201–513 km in the other intervals (Figure 2h).

Across all 5 consecutive time intervals, the model combining GBIF data with social media data captured a broader niche space than the model relying on GBIF-only data (Figure 3). Although the niche overlap between the model relying on GBIF-only data and the model combining social media data with GBIF data was fairly large, reaching 70%, 74%, and 62% in 2005–2008, 2009–2012, and 2021–2024, respectively, the overlap was much smaller during the other 2 periods (37% and 29% for 2013–2016 and 2017–2020, respectively) (calculated using Schoener's *D*).

The niche equivalency tests revealed that the observed overlap was not significantly different from randomized overlaps ($p > 0.05$) during 2005–2008, 2013–2016, 2017–2020, and 2021–2024. However, during 2009–2012, the equivalency test was marginally significant ($p = 0.04$), suggesting a potential deviation from strict niche equivalency. In contrast, niche similarity tests identified consistent niche conservatism across years. In every time interval, the observed overlap was significantly greater than expected by chance ($p \leq 0.01$ in at least one direction for each time interval; Appendix S6), suggesting that niches were more similar than expected under a null model of random overlap, even when equivalency was not supported.

The PC1 and PC2 explained 59–63% of the total variance. The PCA identified precipitation (monthly total), maximum temperature, and elevation to be the most important environmental variables determining the differences in the covered environmental niche space between the model relying on GBIF-only data and the model combining both GBIF and social media data. Across time intervals, the GBIF-only data failed to capture regions with lower maximum temperature, lower precipitation (monthly total), and higher elevation.

4 | Discussion

Social media data can help reduce the global biodiversity data shortfall (Chowdhury, Aich, et al., 2023; Chowdhury, Fuller et

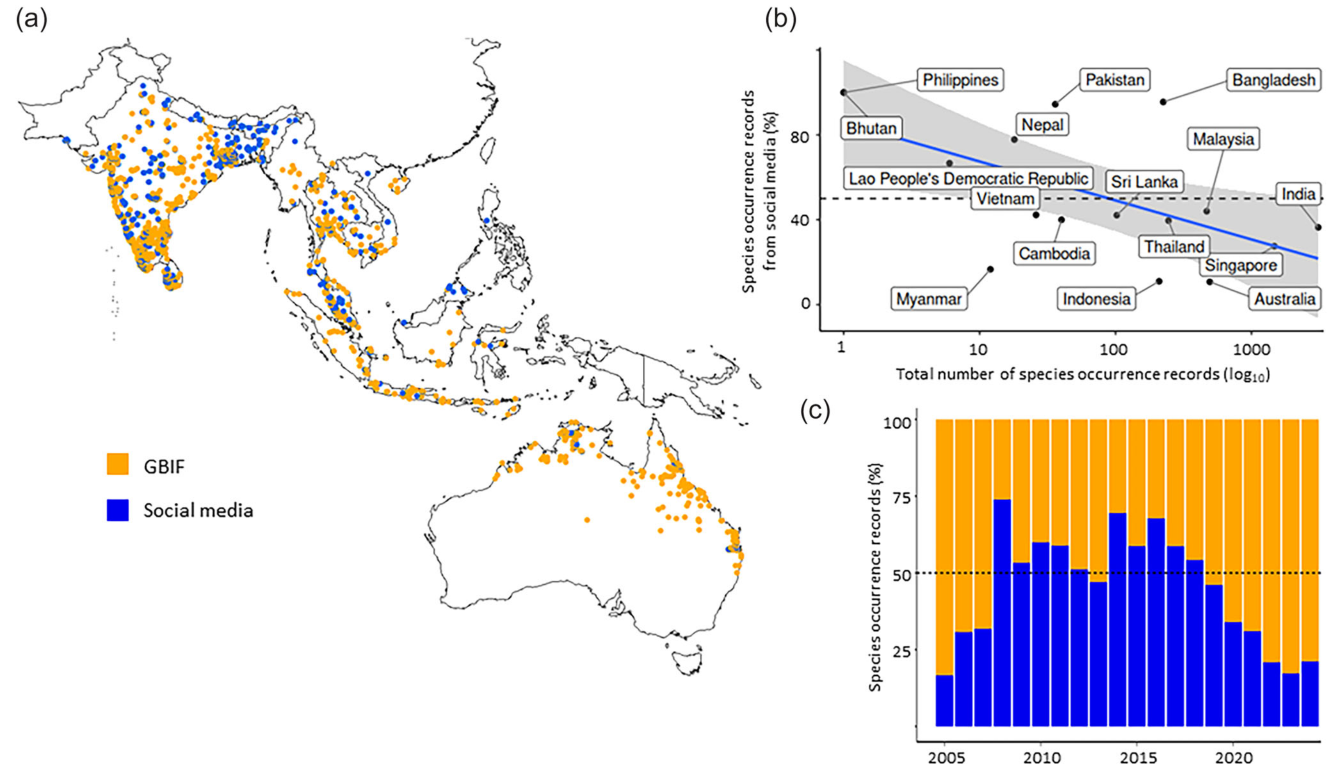

FIGURE 1 Distribution of occurrence records of tawny coster (*Acraea terpsicore*) by source (records from the Global Biodiversity Information Facility [GBIF] and Facebook and Flickr): (a) entire known distribution (each point is one occurrence record), (b) association between the total number of occurrence records and the percentages of occurrence records obtained from social media, and (c) temporal trends in occurrence records by data source (GBIF vs. social media).

al., 2023; Di Minin et al., 2015; Pollock et al., 2025) and improve understanding of biodiversity (re)distribution for conservation purposes (Chowdhury, Fuller, et al., 2024). Yet, such data have rarely been used in large-scale studies (Di Minin et al., 2015). Using standardized protocols to scrape data from social media, we harvested more occurrence data for modeling the potential redistribution of a rapidly expanding species than would typically be used in traditional SDM studies that rely on GBIF data only. We demonstrated that social media data can help identify new portions of a species range: occurrence records increased by 35% (4206 to 6459). In recent years, the percentages of social media data decreased slightly, which could be explained by the COVID pandemic and lockdowns (Chowdhury, Ahmed, et al., 2024). During this period, people were often limited in their movement and travels, resulting in less biodiversity records from the field posted on social media. When we combined occurrence records from social media with those from GBIF and after fitting SDMs to project habitat suitability maps, the suitable area of potential occupancy increased in all 5 studied time intervals, and the total amount of additional areas increased over time compared with a model relying on GBIF data only.

Systematic biodiversity monitoring programs and (semistructured) citizen science provide important biodiversity data sources for scientists, conservation biologists, legislators, and managers (Mesaglio & Callaghan, 2021). We identified the potential key role social media data can have in improving basic understanding of species' redistribution and spread, even in well-surveyed (e.g., developed) countries. For example, biodiversity data from Australia are well represented in global biodiversity repositories like GBIF, compared with other countries, but we still obtained many new localities from social media that represented uncharted conditions, from a GBIF perspective, in the climatic space of the tawny coster butterfly. The total number of occurrence records increased by 12% (440 to 493) and the suitable area of potential occupancy increased by 9% (1.64 million km^2 to 1.79 million km^2). Such gains were much more pronounced for megadiverse countries of the Global South, such as Bangladesh, where the total number of occurrence records retrieved from social media was 22.4 times higher than GBIF. By harnessing social media data, we managed to identify many new localities, representing areas otherwise overlooked. These localities are mostly distributed at higher latitudes and at higher elevations, chiefly representing climate conditions from colder environments with lower maximum temperature and lower precipitation. We also showed that social media data helped to capture a broader niche space exploited by the tawny coster.

We used 2 social media channels—Facebook and Flickr—to harvest more occurrence data for the tawny coster. We demonstrate the utility of social media data for range-shifting, highly visible, and easily detectable species. This approach may not be suitable for cryptic, small, or taxonomically challenging taxa (e.g., many moths, beetles, or fungi), which are far less likely to appear on public platforms and would require disproportion-

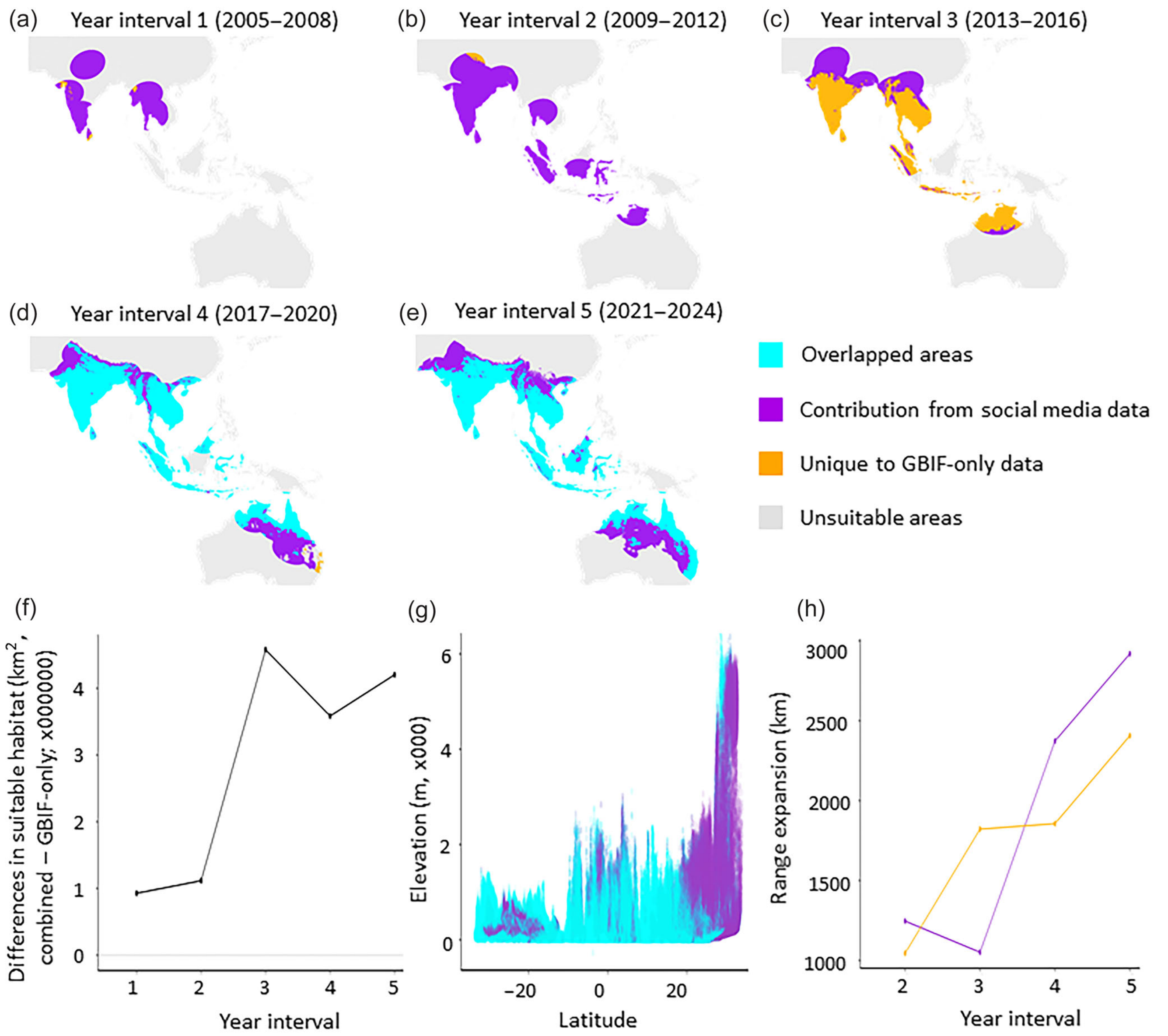

FIGURE 2 Range shift dynamics of tawny coster (*Acraea terpsicore*) from 2005 to 2024: (a–e) habitat suitability in 5-year intervals (purple, areas where model predictions changed after adding social media data to the Global Biodiversity Information Facility data), (f) differences in the surface area obtained from binarized model predictions, (g) latitudinal and elevational distribution of the suitable areas (for all year intervals), and (h) differences in estimated range expansion.

ate effort to detect. Although the additional data substantially improved our understanding of the range dynamics, we faced several obstacles. First, we used an automated procedure to scrape data from Flickr (following Hausmann et al. [2018]), but we had to manually extract data from Facebook (following Chowdhury, Ahmed, et al. [2024]), which was time-consuming compared with an approach that relies solely on artificial intelligence (AI). In the future, it should be possible to develop an automated approach to extract species' occurrences from Facebook, which would save a substantial amount of time (Chowdhury, Ahmed, et al., 2024; Correia et al., 2021; Jarić et al., 2020). For example, Castro et al. (2024) showed that the success rate of AI models in extracting information from unstructured text is quite high, making them valuable tools for managing ecological data efficiently.

Second, we faced 2 major data problems when using Flickr: photographs with no location data and photographs erroneously flagged as the tawny coster. Because of that, we could use only 5% of the data we initially scraped from Flickr. It is important to improve the Flickr data extraction process by checking individual photographs and validating whether or not they represent the target species. Finally, photographs shared on social media might not be the species the photographers assume they are (Gorleri et al., 2023). To address this problem, having someone with taxonomic expertise to validate the content of photographs is essential.

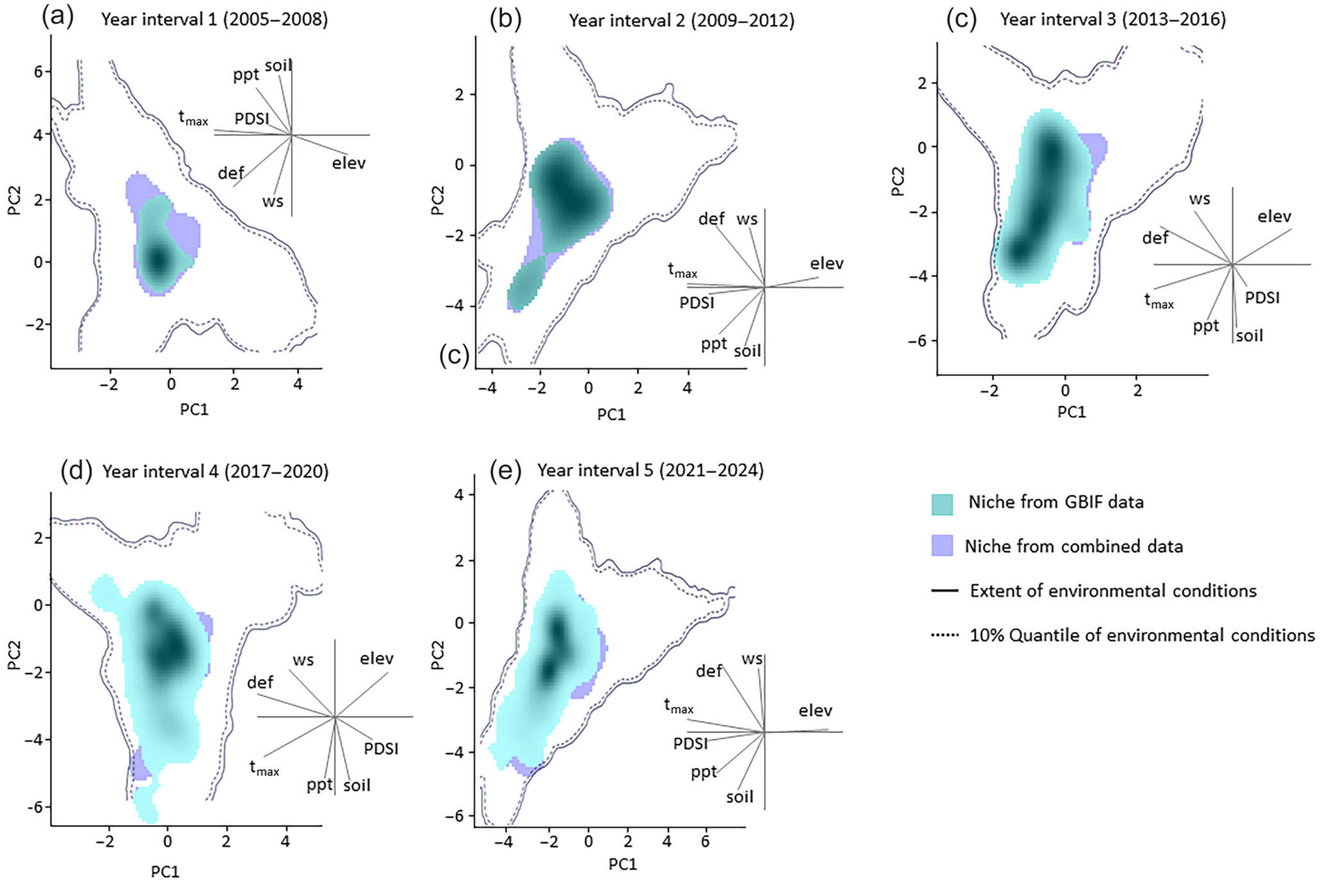

FIGURE 3 Differences in identified niche space of the tawny coster (*Acraea terpsicore*) after adding occurrence records from social media at different year intervals. Inset correlation plots show the importance and direction of impact of each predictor variable (def, climate water deficit; ppt, precipitation monthly total; soil, soil moisture; t_{max}, maximum temperature; ws, wind speed; PDSI, Palmer Drought Severity Index; elev, elevation).

It is important to think strategically to get the maximum value from social media data. We used Facebook and Flickr as social media platforms, as these are generally popular in countries where our target species occur. Including other popular platforms for particular countries (e.g., Weibo in China or possibly Instagram in other countries) could provide many more new records. We recommend that future studies assess data quality performance across several social media channels. We also did not compare how the results might vary by individually comparing Facebook and Flickr with GBIF, which will also be a good avenue for future research. Furthermore, we had to remove many Facebook records due to data quality problems (e.g., the locations were unspecified and photographs were unclear). To solve these problems, group moderators are needed and should maintain strict rules about sharing biodiversity observations.

The Kunming–Montreal Global Biodiversity Framework (CBD, 2022) aims to ensure that the best data are available for conservation assessments (Target 21) and to protect 30% of Earth by 2030 (Target 3). By comparing data distribution between the most comprehensive global biodiversity repository (GBIF) and social media, we showed that species occurrence data shared on social media can improve scientific knowledge on species distributions, even in countries well represented in global biodiversity repositories like GBIF or iNaturalist. Due to environmental changes, range-shifting species (including invasive species) are expanding rapidly, and social media data can be especially powerful in such situations as they allow for almost real-time monitoring. Moreover, social media occurrence data are often available across large geographic scales and international borders, which is not always possible when relying on GBIF data solely. This makes social media data especially useful to set up early warning systems of species colonization (Soriano-Redondo et al., 2024).

Although our study focuses on the tawny coster butterfly, these findings are transferable to other visible taxa with high recognition and photographic potential. By improving the temporal and spatial resolution of occurrence data, our approach contributes directly to real-time conservation assessment, such as early detection of invasive species, identifying climate refugia, or refining species threat assessments. To translate this approach into conservation practice, researchers and practitioners can use image recognition tools to mine social media platforms for real-time sightings, engage with nature photography communities to crowdsource occurrence data, and feed validated records into biodiversity databases like GBIF or iNaturalist for monitoring range shifts and informing rapid response actions.

In addition to other platforms that are currently difficult to access (e.g., Instagram), there is also incidental (or secondary) biodiversity data (e.g., posted photographs of flowers that, by chance, have a butterfly on them; Pernat et al., 2024). These data might come into play with improvements in automated species recognition tools. Such approaches can help better understand and track ongoing species' movements, novel biotic interactions, and future biological invasions (Capinha et al., 2024; Cardoso et al., 2024). The current conservation literature is highly biased, chiefly stemming from North America and western Europe (Dawson et al., 2024; Di Marco et al., 2017), because of significant and long-standing human capacity limitations (e.g., due to limited resources) in the tropics. Our findings suggest that combining data from multiple sources can contribute to answering key ecological questions, especially for countries with limited biodiversity observation records currently registered in global biodiversity repositories.

AUTHOR CONTRIBUTIONS

Shawan Chowdhury conceptualized the idea. Shawan Chowdhury, Niloy Hawladar, Ripon C. Roy, Jooel J. Rinne, and Mohammad Momeny collected the data. Shawan Chowdhury developed the method; all authors contributed to the method and the analyses. Shawan Chowdhury wrote the paper; all authors contributed to the writing.

ACKNOWLEDGMENTS

We are grateful to the numerous volunteers for data collection and for sharing their records on GBIF and social media. A.B. gratefully acknowledge the support of the German Centre for Integrative Biodiversity Research (iDiv) and the sMon project funded by the German Research Foundation (through grant nos. DFG-FZT 118 and 202548816). P.C. is an Australian Research Industry Laureate Fellow (grant title Combatting Wildlife Crime and Preventing Environmental Harm, IL230100175). R. A. C. acknowledges support from the Research Council of Finland (grant agreement no. 348352) and the KONE Foundation (grant agreement no. 202101976). GGD acknowledges the support from the British Ecological Society (through grant no. CE24/1043). EDM was funded by the European Union (ERC, BIOBANG, 101171602) and the KONE Foundation (project 202309134). RJL was supported by CNPq (grants 447598/2025-2, 441125/2023-9, 306174/2025-1).

DATA AVAILABILITY STATEMENT

The GBIF (2024) data (https://doi.org/10.15468/dl.wq7282) are publicly available. Flickr and Facebook data are in Appendix S2. The R scripts are available in the GitHub repository https://github.com/ShawanChowdhury/SocialMedia_RangeChange_TC.

ORCID

Shawan Chowdhury https://orcid.org/0000-0003-2936-5786
Gideon Gywa Deme https://orcid.org/0000-0002-0537-6707
Moreno Di Marco https://orcid.org/0000-0002-8902-4193
Enrico Di Minin https://orcid.org/0000-0002-5562-318X
Richard J. Ladle https://orcid.org/0000-0003-3200-3946
Mohammad Momeny https://orcid.org/0000-0002-8512-8307
Uri Roll https://orcid.org/0000-0002-5418-1164
Aletta Bonn https://orcid.org/0000-0002-8345-4600

REFERENCES

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

Braby, M. F., Bertelsmeier, C., Sanderson, C., & Thistleton, B. M. (2014). Spatial distribution and range expansion of the tawny Coster butterfly, Acraea terpsicore (Linnaeus, 1758) (Lepidoptera: Nymphalidae), in South-East Asia and Australia. *Insect Conservation and Diversity*, *7*(2), 132–143.

Broennimann, O., Di Cola, V., & Guisan, A. (2023). *ecospat: Spatial ecology miscellaneous methods*. R package version 4.0.0. https://CRAN.R-project.org/package=ecospat

Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N. E., Graham, C. H., & Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, *21*(4), 481–497.

Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., & Zimmermann, N. E. (2020). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, *47*(1), 130–142.

Butchart, S. H., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., … Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*, *328*(5982), 1164–1168.

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.

Capinha, C., Ceia-Hasse, A., de-Miguel, S., Vila-Viçosa, C., Porto, M., Jarić, I., Tiago, P., Fernandez, N., Valdez, J., McCallum, I., & Pereira, H. M. (2024). Using citizen science data for predicting the timing of ecological phenomena across regions. *Bioscience*, *74*(6), 383–392.

Cardoso, A. S., Malta-Pinto, E., Tabik, S., August, T., Roy, H. E., Correia, R., Vicene, J. R., & Vaz, A. S. (2024). Can citizen science and social media images support the detection of new invasion sites? A deep learning test case with *Cortaderia selloana*. *Ecological Informatics*, *81*, Article 102602.

Castro, A., Pinto, J., Reino, L., Pipek, P., & Capinha, C. (2024). Large language models overcome the challenges of unstructured text data in ecology. *Ecological Informatics*, *82*, Article 102742.

Chan, W. P., Lenoir, J., Mai, G. S., Kuo, H. C., Chen, I. C., & Shen, S. F. (2024). Climate velocities and species tracking in global mountain regions. *Nature*, *629*(8010), 114–120.

Chandler, M., See, L., Copas, K., Bonde, A. M., López, B. C., Danielsen, F., Legind, J. K., Masinde, S., Miller-Rushing, A. J., Newman, G., Rosemartin, A., & Turak, E. (2017). Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation*, *213*, 280–294.

Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, *333*(6045), 1024–1026.

Chowdhury, S. (2023). Threatened species could be more vulnerable to climate change in tropical countries. *Science of the Total Environment*, *858*, Article 159989.

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. (2024). A protocol for harvesting biodiversity data from Facebook. *Conservation Biology*, *38*(4), Article e14257.

Chowdhury, S., Aich, U., Rokonuzzaman, M., Alam, S., Das, P., Siddika, A., Ahmed, S., Labi, M. M., Di Marco, M., Fuller, R. A., & Callaghan, C. T. (2023). Increasing biodiversity knowledge through social media: A case study from tropical Bangladesh. *Bioscience*, *73*(6), 453–459.

Chowdhury, S., Braby, M. F., Fuller, R. A., & Zalucki, M. P. (2021). Coasting along to a wider range: Niche conservatism in the recent range expansion of the Tawny Coster, Acraea terpsicore (Lepidoptera: Nymphalidae). *Diversity and Distributions*, *27*(3), 402–415.

Chowdhury, S., Debnath, R., Hawladar, N., Howard, S. R., Hodgins, K. A., Wong, B. B. M., & Jarić, I. (2026). Social media data reveal novel habitats for invasive species. *Biological Conservation*, *314*, Article 111668.

Chowdhury, S., Fuller, R. A., Ahmed, S., Alam, S., Callaghan, C. T., Das, P., Correia, R., Di Marco, M., Di Minin, E., Jarić, I., Labi, M. M., Ladle, R., Rokonuzzaman, M., Roll, U., Sbragaglia, V., Siddika, A., & Bonn, A. (2024). Using social media records to inform conservation planning. *Conservation Biology*, *38*(1), Article e14161.

Chowdhury, S., Fuller, R. A., Dingle, H., Chapman, J. W., & Zalucki, M. P. (2021). Migration in butterflies: A global overview. *Biological Reviews*, *96*(4), 1462–1483.

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

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

Correia, R. A., Ladle, R., Jarić, I., Malhado, A. C., Mittermeier, J. C., Roll, U., Soriano-Redondo, A., Veríssimo, D., Fink, C., Hausmann, A., Guedes-Santos, J., Vardi, R., & Di Minin, E. (2021). Digital data sources and methods for conservation culturomics. *Conservation Biology*, *35*(2), 398–411.

Dawson, N. M., Coolsaet, B., Bhardwaj, A., Brown, D., Lliso, B., Loos, J., Mannocci, L., Martin, A., Oliva, M., Pascual, U., Sherpa, P., & Worsdell, T. (2024). Reviewing the science on 50 years of conservation: Knowledge production biases and lessons for practice. *Ambio*, *53*, 1395–1413. https://doi.org/10.1007/s13280-024-02049-w

Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., Balvanera, P., Brauman, K. A., Butchart, S. H. M., Chan, K. M. A., Garibaldi, L. A., Ichii, K., Liu, J., Subramanian, S. M., Midgley, G. F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., … Zayas, C. N. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, *366*(6471), Article eaax3100.

Di Marco, M., Chapman, S., Althor, G., Kearney, S., Besancon, C., Butt, N., Maina, J. M., Possingham, H. P., von Bieberstein, K. R., Venter, O., & Watson, J. E. (2017). Changing trends and persisting biases in three decades of conservation science. *Global Ecology and Conservation*, *10*, 32–42.

Di Marco, M., Pacifici, M., Maiorano, L., & Rondinini, C. (2021). Drivers of change in the realised climatic niche of terrestrial mammals. *Ecography*, *44*(8), 1180–1190.

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

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

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*(4), 330–342.

Feeley, K. J., Stroud, J. T., & Perez, T. M. (2017). Most 'global' reviews of species' responses to climate change are not truly global. *Diversity and Distributions*, *23*(3), 231–234.

Global Biodiversity Information Facility (GBIF). (2024). *Occurrence*. https://doi.org/10.15468/dl.wq7282

Global Biodiversity Information Facility (GBIF). (2025). *Occurrence Download*. https://doi.org/10.15468/dl.925tdp

Gorleri, F. C., Jordan, E. A., Roesler, I., Monteleone, D., & Areta, J. I. (2023). Using photographic records to quantify accuracy of bird identifications in citizen science data. *Ibis*, *165*(2), 458–471.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters*, *11*(1), Article e12343.

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 of the United States of America*, *118*(6), Article e2018093118.

Hijmans, R. (2022). *geosphere: Spherical trigonometry*. R package version 1.5-18. https://CRAN.R-project.org/package=geosphere

Hijmans, R. (2023). *raster: Geographic data analysis and modeling*. R package version 3.6-26. https://CRAN.R-project.org/package=raster

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.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, *44*(9), 1259–1269.

Jarić, I., Correia, R. A., Brook, B. W., Buettel, J. C., Courchamp, F., Di Minin, E., Firth, J. A., Gaston, K. J., Jepson, P., Kalinkat, G., Ladle, R., Soriano-Redondo, A., Souza, A. T., & Roll, U. (2020). iEcology: Harnessing large online resources to generate ecological insights. *Trends in Ecology & Evolution*, *35*(7), 630–639.

Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D. E., Coscieme, L., Golden, A. S., Guerra, C. A., Jacob, U., Takahashi, Y., Settele, J., Díaz, S., Molnár, Z., & Purvis, A. (2022). The direct drivers of recent global anthropogenic biodiversity loss. *Science Advances*, *8*(45), Article eabm9982.

Jensen, T. J., Auliya, M., Burgess, N. D., Aust, P. W., Pertoldi, C., & Strand, J. (2019). Exploring the international trade in African snakes not listed on CITES: Highlighting the role of the internet and social media. *Biodiversity and Conservation*, *28*, 1–19.

Joppa, L. N., O'Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., Watson, J. E. M., Butchart, S. H. M., Virah-Sawmy, M., Halpern, B. S., Ahmed, S. E., Balmford, A., Sutherland, W. J., Harfoot, M., Hilton-Taylor, C., Foden, W., Di Minin, E., Pagad, S., Genovesi, P., … Burgess, N. D. (2016). Filling in biodiversity threat gaps. *Science*, *352*(6284), 416–418.

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla-Buitrago, G. E., Boria, R. A., Soley-Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602–1608.

Lawlor, J. A., Comte, L., Grenouillet, G., Lenoir, J., Baecher, J. A., Bandara, R. M. W. J., Bertrand, R., Chen, I-C., Diamond, S. E., Lancaster, L. T., Moore, N., Murienne, J., Oliveira, B. F., Pecl, G. T., Pinsky, M. L., Rolland, J., Rubenstein, M., Scheffers, B. R., Thompson, L. M., … Sunday, J. (2024). Mechanisms, detection and impacts of species redistributions under climate change. *Nature Reviews Earth & Environment*, *5*, 351–368. https://www.nature.com/articles/s43017-024-00527-z

Lenoir, J., Bertrand, R., Comte, L., Bourgeaud, L., Hattab, T., Murienne, J., & Grenouillet, G. (2020). Species better track climate warming in the oceans than on land. *Nature Ecology & Evolution*, *4*(8), 1044–1059.

Lenoir, J., & Svenning, J. C. (2015). Climate-related range shifts—A global multidimensional synthesis and new research directions. *Ecography*, *38*(1), 15–28.

Marcenò, C., Padullés Cubino, J., Chytrý, M., Genduso, E., Salemi, D., La Rosa, A., Gristina, A. S., Agrillo, E., Bonari, G., del Galdo, G. G., Ilardi, V., Landucci, F., & Guarino, R. (2021). Facebook groups as citizen science tools for plant species monitoring. *Journal of Applied Ecology*, *58*(10), 2018–2028.

Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. M. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, *536*(7615), 143–145.

Mesaglio, T., & Callaghan, C. T. (2021). An overview of the history, current contributions and future outlook of iNaturalist in Australia. *Wildlife Research*, *48*(4), 289–303.

Mota, L. L., Boddington, S. J., Brown Jr, K. S., Callaghan, C. J., Carter, G., Carter, W., Dantas, S. M., Dolibaina, D. R., Garwood, K., Hoyer, R. C., Robbins, R. K., Soh, A., Willmott, K. R., & Freitas, A. V. (2022). The butterflies of Cristalino Lodge, in the Brazilian southern Amazonia: An updated species list with a significant contribution from citizen science. *Biota Neotropica*, *22*(3), Article e20221367.

Parker, E. J., Weiskopf, S. R., Oliver, R. Y., Rubenstein, M. A., & Jetz, W. (2024). Insufficient and biased representation of species geographic responses to climate change. *Global Change Biology*, *30*, Article e17408.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. (2007). Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography*, *34*(1), 102–117.

Pernat, N., Canavan, S., Golivets, M., Hillaert, J., Itescu, Y., Jarić, I., Mann, H. M. R., Pipek, P., Preda, C., Richardson, D. M., Teixeira, H., Vaz, A. S., & Groom, Q. (2024). Overcoming biodiversity blindness: Secondary data in primary citizen science observations. *Ecological Solutions and Evidence*, *5*, Article e12295.

Phakoago, M. V., Maloney, S. K., Kamerman, P. R., Meyer, L. C., Weyer, N. M., & Fuller, A. (2024). Social media as a tool to understand the distribution and ecology of elusive mammals. *Journal of Mammalogy*, *105*(1), 206–214.

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, *40*(7), 887–893.

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, *19*(1), 181–197.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187), Article 1246752.

Pocock, M. J., Chandler, M., Bonney, R., Thornhill, I., Albin, A., August, T., Bachman, S., Brown, P. M. J., Cunha, D. G. F., Grez, A., Jackson, C., Peters, M., Rabarijaon, N. R., Roy, H. E., Zaviero, T., & Danielsen, F. (2018). A vision for global biodiversity monitoring with citizen science. *Advances in Ecological Research*, *59*, 169–223.

Pollock, L. J., Kitzes, J., Beery, S., Gaynor, K. M., Jarzyna, M. A., Mac Aodha, O., Meyer, B., Rolnick, D., Taylor, G. W., Tuia, D., & Berger-Wolf, T. (2025). Harnessing artificial intelligence to fill global shortfalls in biodiversity knowledge. *Nature Reviews Biodiversity*, *1*, 166–182.

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sanderson, C., Braby, M. F., Thistleton, B. M., & Neal, M. (2012). First record of the Tawny Coster butterfly Acraea terpsicore (Linnaeus, 1758) from Australia. *Myrmecia. News Bulletin of the Australian Entomological Society*, *48*, 39–41.

Sbragaglia, V., Espasandín, L., Jarić, I., Vardi, R., Ramírez, F., & Coll, M. (2024). Tracking ongoing transboundary marine distributional range shifts in the digital era. *Marine Ecology Progress Series*, *728*, 103–114.

Sheard, J. K., Adriaens, T., Bowler, D. E., Büermann, A., Callaghan, C. T., Camprasse, E. C. M., Chowdhury, S., Engel, T., Finch, E. A., von Gönner, J., Hsing, P.-Y., Mikula, P., Oh, R. Y. R., Peters, B., Phartyal, S. S., Pocock, M. J. O., Wäldchen, J., & Bonn, A. (2024). Emerging technologies in citizen science and potential for insect monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*, Article 20230106.

Soriano-Redondo, A., Correia, R. A., Barve, V., Brooks, T. M., Butchart, S. H., Jarić, I., Kulkarni, R., Ladle, R. J., Vaz, A. S., & Di Minin, E. (2024). Harnessing online digital data in biodiversity monitoring. *PLoS Biology*, *22*(2), Article e3002497.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., & 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), 932. https://doi.org/10.1038/s41598-017-09084-6

van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., Høye, T. T., Jongejans, E., Menz, M. H. M., Miraldo, A., Roslin, T., Roy, H. E., Ruczyński, I., Schigel, D., Schäffler, L., Sheard, J. K., Svenningsen, C., Tschan, G. F., Wäldchen, J., … Bowler, D. E. (2022). Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution*, *37*, 872–885.

Wilson, E. O. (1989). Threats to biodiversity. *Scientific American*, *261*(3), 108–117.

Yackulic, C. B., Sanderson, E. W., & Uriarte, M. (2011). Anthropogenic and environmental drivers of modern range loss in large mammals. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(10), 4024–4029.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chowdhury, S., Hawladar, N., Roy, R. C., Capinha, C., Cassey, P., Correia, R. A., Deme, G. G., Di Marco, M., Di Minin, E., Jarić, I., Ladle, R. J., Lenoir, J., Momeny, M., Rinne, J. J., Roll, U., & Bonn, A. (2026). Harnessing social media data to track species range shifts. *Conservation Biology*, e70234. https://doi.org/10.1111/cobi.70234