

Citizen-science and empirical field data yield limited agreement about avian diversity

A recent study by Zhang et al. (2026) reported that stronger solar expansion policies in China were associated with lower bird diversity estimated from citizen-science records. We assessed how well bird diversity derived from such records represented bird communities documented through field surveys. We compared records from the birdwatching database with our field surveys of forest birds in southern China, selecting citizen-science observations made within 5 km of each survey region. Across nine regions represented in both datasets, neither Shannon entropy nor coverage-standardized Hill diversity was significantly correlated between the two data sources, although the small number of paired regions limited the precision of the correlation estimates. The two lists of the ten most abundant species shared only three species, and an ordination based on Bray–Curtis dissimilarities clearly separated the assemblages recorded by the two sources. On a log–log scale, total abundance was positively and linearly related to record frequency in both datasets, but this relationship was significantly steeper in the field survey data. Our findings suggest that citizen-science and empirical survey data may not provide interchangeable descriptions of bird communities, and highlight the need to consider sampling locations, habitat coverage, and recording practices when citizen-science data are used to assess the ecological effects of environmental policies.

The rapid expansion of photovoltaic energy creates an urgent need to assess its ecological costs alongside its climate benefits. Recently, Zhang et al. (2026) reported that a one-standard-deviation increase in policy strength was associated with a 2.1% decline in bird Shannon diversity. The interpretation of this result has prompted methodological debate. Lai & Zhang (2026) questioned how statistical significance and model explanatory power were used to characterize the ecological importance of the association. Concerns have also been raised about attributing the observed association to physical photovoltaic development, the evidence supporting the proposed ecological mechanisms, and the scope of the biodiversity assessment (Jin, 2026; Yang et al., 2026).

The bird data used by Zhang et al. (2026) came from a widely used citizen-science database, the China Bird Report Online Database (www.birdreport.cn). Citizen science

provides biodiversity observations at a scale that would be difficult to achieve through conventional monitoring alone. However, drawing ecological conclusions from these records requires careful consideration of where observers go, which species they detect, and what they choose to report (Johnston et al., 2023). Regarding the abundance-based Shannon index, the relative frequencies of recorded individuals of different species directly determine the resulting value. These frequencies reflect not only the composition of the ecological community but also habitat coverage, observation effort, species detectability, and reporting practices. A change in the index may therefore reflect ecological change, a change in the observation process, or both. Although Zhang et al. (2026) included a covariate related to observation effort and conducted sensitivity analyses using different sampling thresholds, these adjustments alone do not establish that the recorded assemblages consistently represent the ecological communities of interest.

These considerations raise a complementary question: how representative is the diversity derived from citizen-science birdwatching records of the bird communities compared to empirical surveys by ornithologists? To address this question, we selected ten study regions in South China and compared citizen-science records from the same birdwatching data source used by Zhang et al. (2026) with field observations collected by our team. Notably, we examined a different spatial and temporal subset of citizen-science birdwatching records and did not replicate Zhang et al. (2026)'s nationwide dataset or reanalyse their association between policy strength and bird diversity. Our aim was to assess consistency between citizen-science and empirical survey data in three respects: (1) alpha diversity, using Shannon entropy and coverage-standardized Hill numbers; (2) species composition, by comparing the most abundant species and evaluating beta diversity using Bray–Curtis dissimilarities; and (3) recording patterns, by examining how the frequency of records varied with total abundance in each dataset.

We conducted our empirical survey in 10 nature reserves covering a broad latitudinal range (N18.7°–N33.6°) of forest ecosystems in China: Badagongshan, Dinghushan, Foping, Jianfengling, Maoershan, Nanling, Thousand Island Lake (Qiandaohu), Shennongjia, Wuyishan, and Yaoluoping (Figure 1). This gradient encompasses subtropical forests,

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ©2026 Editorial Office of Zoological Research: Diversity and Conservation, Kunming Institute of Zoology, Chinese Academy of Sciences

Received: 09 September 2026; Accepted: 18 September 2026; Online: 18 September 2026

Foundation items: This research was funded by the National Natural Science Foundation of China (W2431023, 32471628), Basic and Applied Basic Research Foundation of Guangdong Province (2025A1515011932), XJTLU Research Development Fund (RDF-22-02-023), XJTLU Postgraduate Research Scholarship (PGRSSP220901).

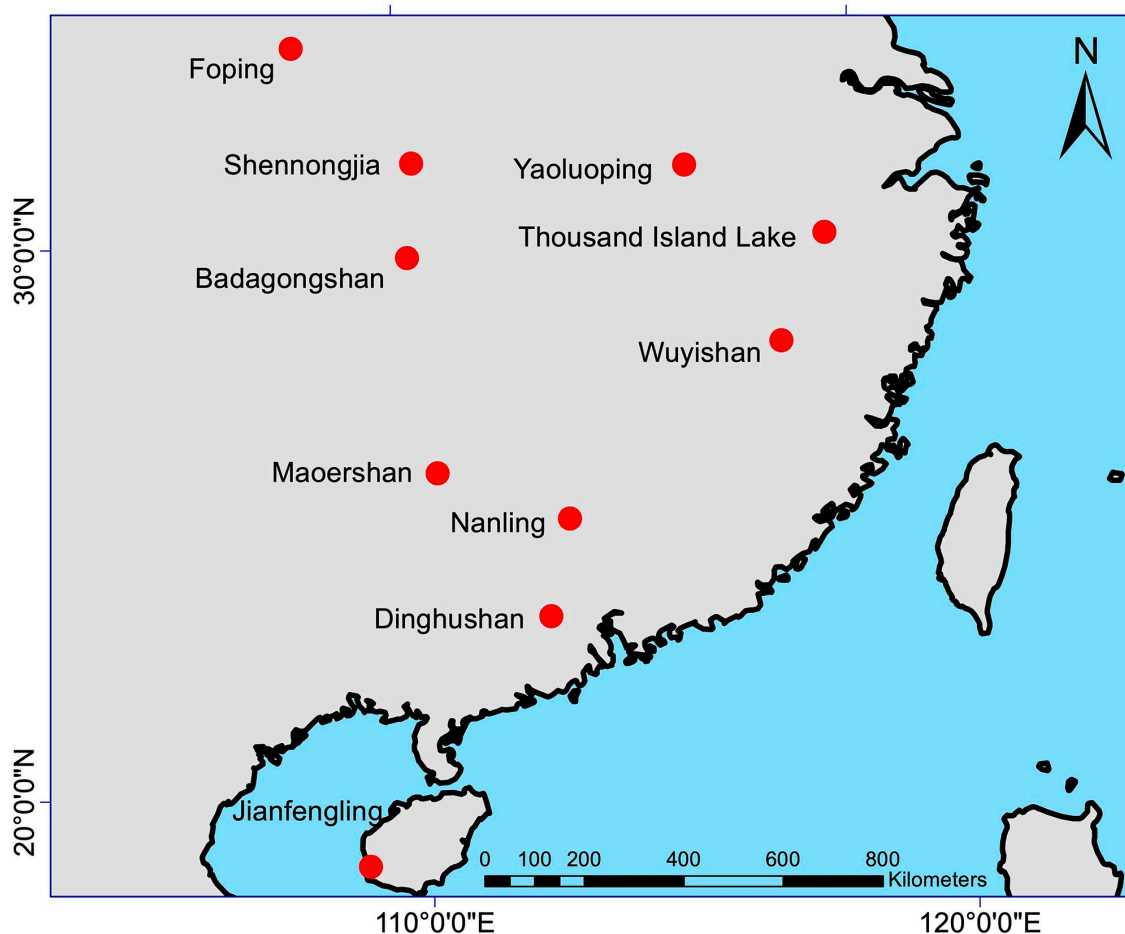


Figure 1 Geographic locations of the regions where we conducted empirical bird surveys

extending into tropical and temperate zones at its southern and northern ends, respectively. At each reserve, we established three 2-km line transects, separated by at least 1 km to ensure spatial independence; all transects were located in well-protected secondary forests with minimal human disturbance. We recorded all forest birds within 20 m of the transect line. We conducted surveys in both the breeding (May–August) and non-breeding (November–February) seasons during 2024 and 2025, and to reduce climatic variation the survey route progressed from south to north during the breeding season and reversed (north to south) for the non-breeding season. Each reserve was surveyed twice per season (two replicates during the breeding season and two during the non-breeding season). We emphasize that this survey method was designed to detect most common forest species and has its own biases regarding how representative it is of bird communities; for example, we did not conduct stratified or full grid sampling, or try to integrate detectability differences between species by occupancy modeling or distance adjustment.

The citizen-science bird data were derived from the China Bird Reporting Center (www.birdreport.cn). To improve temporal comparability, we restricted the citizen-science data to the same months covered by our empirical surveys. To improve spatial comparability while retaining sufficient data for analysis, we further filtered the citizen-science records and retained only those with reported coordinates within 5 km of our field-survey transects (which could be outside the reserve). Applying a narrower buffer substantially reduced data availability. For example, a 2-km buffer retained citizen-

science records from only three study regions, which was insufficient for meaningful paired correlation analyses. We therefore used the 5-km buffer and interpreted the resulting citizen-science data as representing landscape-scale assemblages surrounding the survey transects, not the forest communities sampled directly along the transects. The citizen-science dataset comprised 573 checklists submitted by 274 observers across 106 observation sites in nine reserves (no citizen-science records from Yaoluoping met the selection criteria), spanning 134 survey days. Observation durations per checklist ranged from <1 day to 5 days, consistent with the ~5-day field stay per reserve in our survey.

For both data sources, scientific names were harmonized to the IOC World Bird List v.14.1, before analysis. Overall, the field survey dataset contained 16 824 recorded individuals, 245 species, and 4 117 species-level records (i.e., one entry for a species and its recorded number of individuals) across the 10 regions. After selection, the citizen-science dataset contained 23 211 recorded individuals, 344 species, and 10 953 species-level records across nine regions. Sampling effort was uneven among reserves for the citizen-science dataset, with one reserve (Jianfengling) accounting for 70% of all checklists. Citizen-science observer expertise and habitat inclusion were also unstandardized, so these data represent regional-scale assemblages rather than the forest communities sampled along the transects.

For each region and data source, we pooled all available observations by summing the abundance records of each species across sampling dates. This regional aggregation increased the number of observations contributing to each

assemblage and thereby improved the stability of the diversity estimates. We calculated Shannon entropy (H') for the citizen-science and empirical survey datasets. Because raw diversity might be sensitive to unequal sample completeness, we additionally calculated coverage-standardized Hill diversity at 95% sample coverage, and considered $q=0$ (species richness), $q=1$ (the exponential of Shannon entropy) and $q=2$ (the inverse Simpson concentration) (Chao & Jost, 2012). Pearson linear correlation and Spearman rank correlations were used to quantify the cross-source agreement. In addition, to assess seasonal variation, we repeated all comparisons separately for each region and calendar month. Observations from the same month were pooled across years, and only combinations of regions and months represented in both datasets were retained, which eventually resulted in 24 paired samples. This analysis broadly parallels the monthly approach of Zhang et al. (2026), although they treated each county, year and month as a separate observation.

To compare community composition, we first checked the top ten most dominant species to see how many species were in common between the two datasets. We then calculated Bray–Curtis dissimilarities for each study region of both data sources, with count data transformed into relative abundance before analysis. We visualized the dissimilarity matrix using two-dimensional non-metric multidimensional scaling (NMDS), with the ordination dimensionality fixed at $k=2$. Multiple random starting configurations were used, with a maximum of 100 random starts, to reduce the risk of retaining a local optimum. The final two-dimensional solution had a stress value of 0.13, below the commonly used heuristic threshold of 0.2 for interpreting broad patterns in an NMDS ordination (Clarke et al., 2014).

To compare recording patterns, we calculated record frequency as the number of species-level records for each species within each dataset, with each record corresponding to one entry for a species and its reported number of individuals. Records were counted across all retained regions and dates within each data source. Total reported abundance was calculated by summing the numbers of individuals across these records. We fitted a linear model with $\log_{10}$ -transformed total reported abundance as the response variable and $\log_{10}$ -transformed record frequency, data source, and their interaction as explanatory variables.

The alpha-diversity correlations were based on the nine regions represented in both datasets. For the pooled dominant-species analysis, NMDS ordination, and abundance–record frequency analysis, we retained all available observations, including empirical survey data from Yaoluoping, for which no corresponding citizen-science data were available. These analyses therefore included nine citizen-science regions and ten empirical survey regions.

All analyses were conducted in R (R Core Team 2026). Package iNEXT v.3.0.2 (Hsieh et al., 2016) was used to calculate the coverage-standardized Hill diversity; package vegan v.2.7.5 (Oksanen et al., 2018) was used to calculate the Bray–Curtis dissimilarities and NMDS; package reshape2 v.1.4.5 (Wickham, 2007) and ggplot2 v.4.0.3 (Wickham, 2009) were used in data organization and visualization.

Across the nine regions represented in both datasets, correlations between citizen-science and empirical estimates of bird diversity were not statistically significant for Shannon entropy (Figure 2A) or coverage-standardized Hill diversity at $q=0$ (Figure 2B), $q=1$ (Figure 2C), or $q=2$ (Figure 2D).

Pearson's correlation coefficients ranged from -0.04 to 0.26 , and Spearman's rank correlation coefficients ranged from 0.12 to 0.42 , with all P values greater than 0.1 . Comparisons conducted separately for each calendar month produced the same overall result, with none of the correlations statistically significant (Supplementary Figure S1). Pearson's correlations ranged from 0.03 to 0.35 and Spearman's correlations from 0.16 to 0.36 (all $P>0.05$).

The lists of the ten most abundant species shared three species: Huet's fulvetta (*Alcippe hueti*), chestnut bulbul (*Hemixos castanonotus*), and scarlet minivet (*Pericrocotus speciosus*) (Figure 3A, B). The puff-throated bulbul (*Alophoixus pallidus*) was the most abundant species in the citizen-science data, accounting for 4.6% of all recorded individuals, but was not among the ten most abundant species in the empirical survey data. Huet's fulvetta (*A. hueti*) dominated the empirical survey data (13.7%), but ranked fourth in the citizen-science data (3.2%). NMDS ordination showed partial separation between citizen-science and empirical survey assemblages, with some overlap between the two groups in terms of the 95% distribution of sampling locations (stress=0.13; Figure 3C).

Record abundance (log scale) increased linearly with frequency (log scale) in both datasets (Figure 3D). The full model showed a significant interaction between the data source and the abundance–frequency relationship (full model $R^2=0.9$; $P=0.001$ for the interaction), indicating that the relationship was steeper in the empirical survey data. Separate regressions for each data source showed $\beta=1.08\pm0.018$ ($R^2=0.91$) and 1.18 ± 0.027 ($R^2=0.87$) for citizen-science and empirical survey data, respectively.

Our analysis suggests that citizen-science and empirical survey data provided different descriptions of bird communities across the study regions. The lack of agreement in Shannon and Hill diversity, using both pooled regional estimates and monthly estimates, indicates that the differences were not specific to a single diversity measure. The three species shared between the lists of dominant species and the partial overlap in the NMDS ordination show that the datasets captured some common features of bird communities. Nonetheless, there is still a clear separation in the ordination plot, indicating different assemblages between the two datasets. An observation-based diversity index should therefore be viewed as a product of both the ecological community and the process used to generate the observations (Kelling et al., 2019). However, the absence of statistically significant correlations should be interpreted cautiously, given the small number of paired regions.

Some of this disagreement likely occurred because the citizen-science observations covered a broader or different range of habitats than the empirical field observations. Although this difference limits direct numerical comparisons, it also highlights the central validation issue: county-level estimates like those used by Zhang et al. (2026) are affected when the distribution of observations across habitats changes, even if the communities within individual habitats remain stable. Ecological claims should therefore clearly identify the habitat domain and target population that the response variable is intended to represent (Johnston et al., 2023).

We found a positive linear relationship between total abundance and record frequency in both datasets, but the slope was significantly steeper in the empirical survey data. At the same record frequency, the fitted relationship predicted

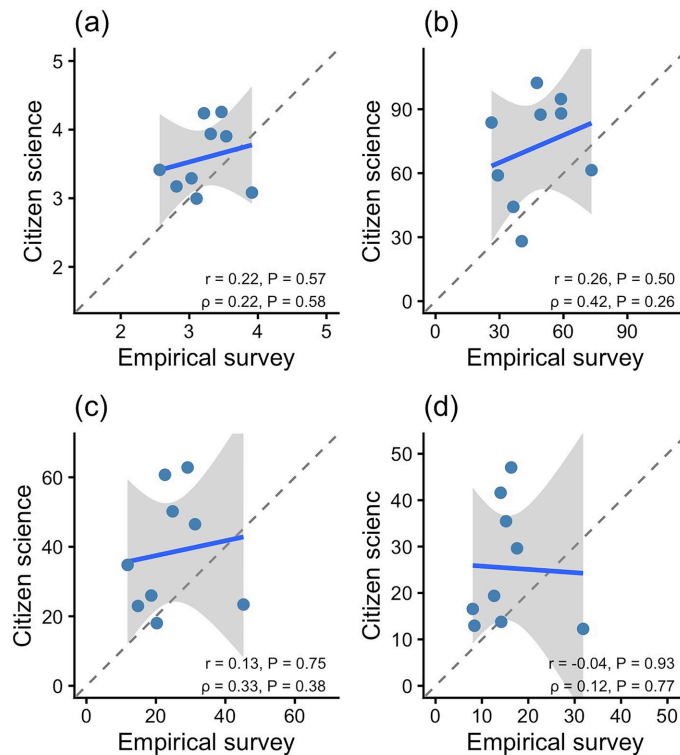


Figure 2 Cross-source agreement in region-level avian alpha diversity

(a) Shannon entropy. (b–d) Coverage-standardized Hill diversity at 95% sample coverage for $q=0$, $q=1$, and $q=2$, respectively. Points are study regions; lines refer to linear regression fits with 95% Confidence Interval (CI); dashed lines indicate 1:1 agreement; r and ρ refer to the Pearson and Spearman correlation coefficients, and respective P values.

fewer individuals in the citizen-science data, indicating a lower average number of individuals per record. This pattern may reflect differences in how individuals are counted and reported. Because the field surveys followed a standardized protocol, they provided a more controlled approximation of relative abundance within the surveyed habitats. This result suggests that individual birds may be reported less completely in the citizen-science data, particularly for frequently recorded species, although differences in the bird assemblages encountered and the exact sampling locations could also contribute to this pattern.

Overall, across nine study regions in South China, we found that citizen-science and empirical survey bird data produced weakly concordant alpha-diversity estimates and different community composition. We acknowledge that with only nine paired regions, statistical power to detect moderate correlations was limited. In addition, spatial matching was also imperfect because citizen-science records within 5 km represented broader landscape assemblages, whereas transects sampled forest communities. Future validation should include more paired regions and stratified sampling across habitats. Nonetheless, as the empirical surveys were conducted by ornithologists along fixed transects using a standardized protocol, they provided a more controlled reference for relative abundance within the surveyed habitats. The lack of agreement means that the two data sources are not interchangeable. Therefore, when citizen-science data are used to assess the ecological effects of environmental policies, as in Zhang et al. (2026), differences in sampling locations and habitat coverage should be considered, and the representativeness of the observations should be evaluated. If diversity estimates are influenced by data reporting practices, an association with policy strength cannot be assumed to

represent changes in the underlying bird communities, unless these observation processes are adequately controlled or the estimates are independently validated.

DATA AVAILABILITY

All data and code used in the analyses are available in the GitHub repository: https://github.com/YiZou01/CSEM_bird.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.Z. conceived the idea. J.W. and Q.Z. collected the field survey data. Y.Z. and J.W. performed the analyses and wrote the initial draft, with substantial contributions from E.G. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We are grateful to the authorities of the nature reserves for their permission and support. We also thank Jiangshan Lai for helpful discussions that contributed to the development of this work.

Yi Zou^{1,*,#}, Jia-Hao Wu^{1,2,#}, Qiang Zhang³, Eben Goodale¹

¹ Centre of Excellence for One Health and Environmental Sustainability, Academy of Life and Natural Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China

² Institute of Infection, Veterinary and Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, Merseyside L3 5TR, UK

³ Guangdong Key Laboratory of Animal Conservation and Resource Utilization, Institute of Zoology, Guangdong Academy of Sciences, Guangzhou, Guangdong 510275, China

*Corresponding author, E-mail: yi.zou@xjtlu.edu.cn

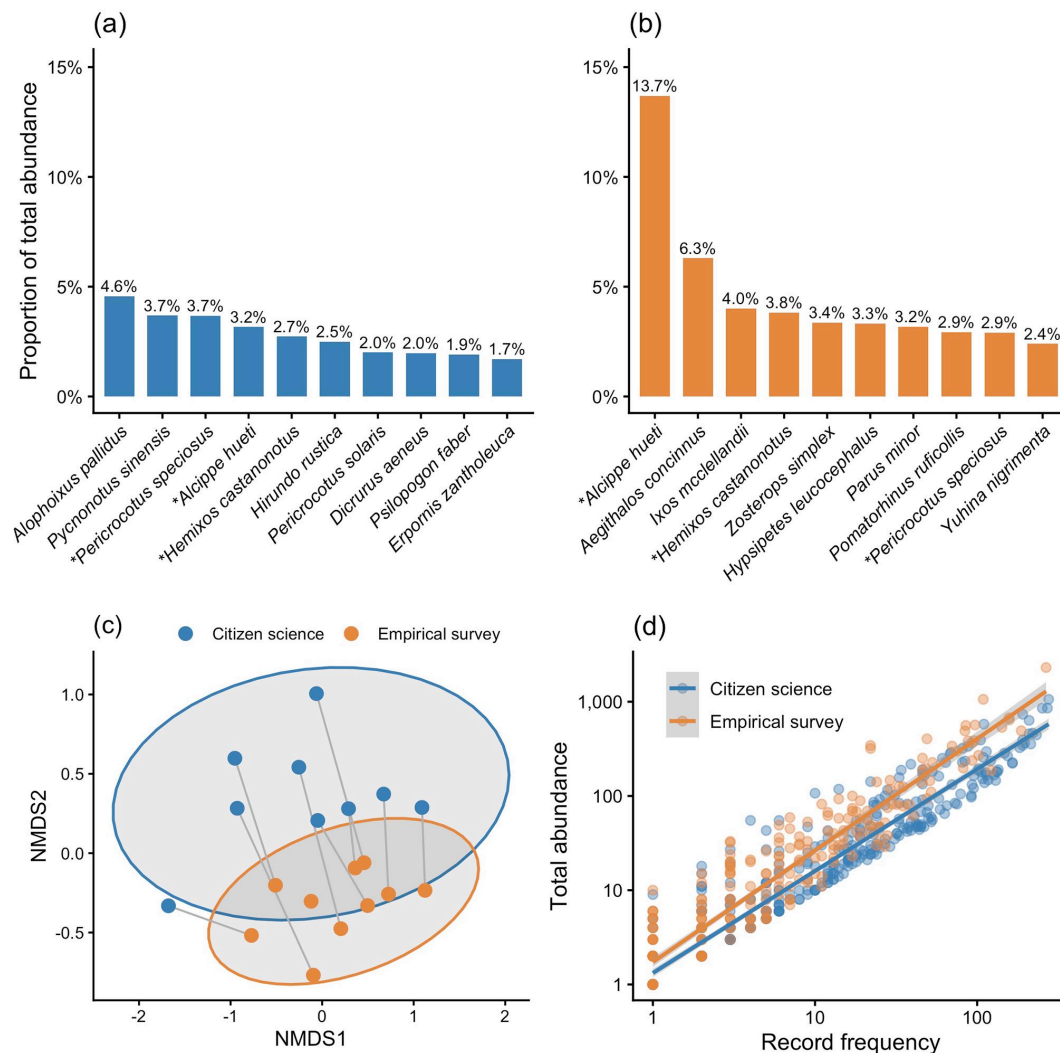


Figure 3 The ten most dominant species in citizen-science (a) and empirical survey (b), ranked by their proportion of total individuals; stars refer to species that appear in both data sources. (c) NMDS of Bray–Curtis dissimilarities (two dimensions; stress=0.13); points represent study region coloured by data source, grey segments join the two sources within a study region; ellipses represent the 95% distribution of sampling locations estimated using a multivariate t distribution. (d) Relationship between record abundance and frequency for each species in the two data sources; both axes are $\log_{10}$ transformed; lines show source-specific linear regression with 95% CIs.

REFERENCES

- Chao A, Jost L. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, **93**: 2533–2547.
- Clarke KR, Gorley RN, Somerfield PJ, et al. 2014. Change in marine communities: an approach to statistical analysis and interpretation. 3 edn. PRIMER-E, Plymouth.
- Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, **7**: 1451–1456.
- Jin H. 2026. Causal attribution and ecological interpretation of solar-policy effects. *Science eLetters*.
- Johnston A, Matechou E, Dennis EB. 2023. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, **14**: 103–116.
- Kelling S, Johnston A, Bonn A, et al. 2019. Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity. *Bioscience*, **69**: 170–179.
- Lai J, Zhang S. 2026. Statistical significance should not be conflated with explanatory importance: a comment on 'China's solar expansion policy reduces bird diversity'. *Statistics Innovation*, **3**: e013.
- Oksanen J, Blanchet FG, Friendly M, et al. 2018. vegan: Community Ecology Package. R package version 2.5-6. Available at: <http://CRAN.R-project.org/package=vegan>.
- R Core Team. 2026. R: A language and environment for statistical computing. Version 4.6.0. R Foundation for Statistical Computing, Vienna, Austria.
- Wickham H. 2007. Reshaping Data with the reshape Package. *Journal of Statistical Software*, **21**(12): 1–20.
- Wickham H. 2009. ggplot2: elegant graphics for data analysis. Springer Publishing Company, Incorporated.
- Yang Y, Zhang J, Xian Y, et al. 2026. Solar Expansion Needs System-Wide Biodiversity Accounting. *Science eLetters*.
- Zhang H, Zhang A, Wu K, et al. 2026. China's solar expansion policy reduces bird diversity. *Science*, **393**: 831–836.