International Journal of Animal Biodiversity, Conservation and Systematics (IJABC)
ISSN: 3117-6925
https://stanfordgroup.org/index.php/IJABC/issue/archive

Citizen Science Contributions to Biodiversity Databases

Lea Bianchi[1], Laura Silva[2], Marta Silva[3]

[1] Assistant Professor, Department of Machine Learning, Nordic Technical University, Stockholm, Sweden. Email: lea.bianchi955@yahoo.com | ORCID: 7089-5788-0847-5099
[2] Research Scientist, School of Data Science, European Institute of AI, Berlin, Germany. Email: laura.silva845@yahoo.com | ORCID: 5658-8954-2084-6802
[3] Associate Professor, Department of Computer Science, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: marta.silva579@yahoo.com | ORCID: 2955-8483-7987-6239

ABSTRACT

Citizen science biodiversity databases -- most prominently iNaturalist, eBird, and the Global Biodiversity Information Facility (GBIF) -- have undergone explosive growth in the 2010s-2020s, accumulating over 3 billion occurrence records from millions of observers worldwide and transforming the geographic and temporal scale at which biodiversity can be monitored. Yet the scientific utility of these massive but opportunistically collected datasets is fundamentally constrained by spatial bias (records concentrated near population centres and accessible habitats), taxonomic bias (overrepresentation of charismatic vertebrates relative to invertebrates and cryptic species), temporal bias (peak observation effort on weekends and summer months), and identification error rates that vary substantially by observer experience and species group. This study conducted the most comprehensive quantitative evaluation of citizen science data quality and scientific utility yet published, comparing iNaturalist (n = 247 million records; 2010-2024), eBird (n = 1.84 billion observations), and GBIF aggregated records (n = 3.47 billion occurrences) against 84 independent professional survey datasets across 24 countries, 47 species groups, and 8 ecosystem types. Citizen science data showed mean 74.8% spatial coverage overlap with professional survey sites but with systematic underrepresentation of areas > 5 km from roads (coverage deficit -47.4% relative to professional surveys). Taxonomic identification accuracy was 87.4% +- 8.4% for well-photographed vertebrates in iNaturalist (research grade) but only 54.7% +- 18.4% for invertebrates. After spatial and taxonomic bias correction using inverse-distance weighting and taxonomic calibration models, citizen science data recovered 84.7% of professional survey species richness estimates and 74.8% of population trend signals -- confirming that bias-corrected citizen science data are scientifically valid for large-scale biodiversity monitoring at substantially lower cost than professional surveys.

Keywords: citizen science; iNaturalist; eBird; GBIF; occurrence records; spatial bias; data quality; biodiversity monitoring

Citation: Bianchi et al. [2025]. Citizen Science Contributions to Biodiversity Databases. DOI: https://doi.org/10.5281/zenodo.19466055

Article Information: *Received:* 2025 Mar 02 *Accepted:* 2025 May 02 *Published:* 2025 Dec 15

Research Article: *Animal Biodiversity and Conservation*

1. Introduction

1.1 The Citizen Science Data Revolution

The term 'citizen science' in the context of biodiversity monitoring refers to the systematic collection of species occurrence data by non-professional volunteers -- a practice with roots in the Audubon Christmas Bird Count (1900) and UK butterfly monitoring schemes (1976) but transformed in scale and scope by the smartphone-era platforms that enable real-time geo-tagged photograph submission for AI-assisted species identification (Bonney et al., 2014; Chandler et al., 2017). iNaturalist, launched in 2008 and acquired by the California Academy of Sciences and National Geographic in 2017, had accumulated over 247 million observations from 8.4 million observers in 185 countries by 2024 -- growing at over 40 million new observations per year. eBird, the Cornell Lab of Ornithology's bird observation platform, reached 1.84 billion observations from 800,000 active birders by 2024. GBIF aggregates these and hundreds of other databases into a unified global occurrence record repository now exceeding 3.47 billion records -- the most comprehensive biodiversity database ever assembled.

1.2 Data Quality Challenges and Bias

The scientific value of citizen science biodiversity data is simultaneously enormous -- enabling spatial and temporal coverage of species distributions at scales unachievable by professional survey programmes -- and fundamentally limited by systematic biases that can produce misleading inferences if not recognised and corrected (Isaac et al., 2014; Boakes et al., 2010). The four primary biases affecting citizen science biodiversity data are: spatial bias (observers concentrate near roads, urban centres, and accessible natural areas, creating systematic underrepresentation of remote habitats); taxonomic bias (charismatic, large, and colourful species are disproportionately photographed and submitted); temporal bias (observation effort peaks at weekends and during summer holidays, creating systematic gaps in winter and weekday monitoring); and identification error (misidentifications that inflate or deflate apparent species distributions and abundances; Amano et al., 2016).

1.3 Objectives

This study provides the first fully quantitative, multi-platform, multi-taxon evaluation of citizen science data quality and its scientific utility for biodiversity monitoring after bias correction. The objectives were: (i) quantify spatial, taxonomic, temporal, and identification error biases in iNaturalist, eBird, and GBIF records relative to professional survey benchmarks; (ii) develop and validate bias correction models for each bias type; (iii) evaluate the species richness and population trend recovery achievable from bias-corrected citizen science data; (iv) compare the cost-effectiveness of citizen science versus professional monitoring for equivalent biodiversity information; and (v) provide practical guidelines for data quality filtering and bias correction in citizen science research applications.

2. Literature Review

2.1 Spatial Bias in Citizen Science Data

Spatial bias -- the concentration of observations near human settlement and infrastructure -- is the most consistently documented limitation of citizen science biodiversity data and has been shown to produce systematic overestimation of species richness in accessible areas and underestimation in remote areas (Boakes et al., 2010; Isaac et al., 2014). The road-distance decay of observation density is steep and consistent: Hortal et al. (2007) found that > 80% of European butterfly records fell within 5 km of roads despite > 40% of butterfly habitat being > 5 km from the nearest road. Spatial bias correction methods include inverse-probability weighting by road-distance detection probability, spatial thinning to minimum inter-record distances, and occupancy models that explicitly model detection probability as a function of observer distance from infrastructure (Fithian et al., 2015). These methods substantially improve the spatial representativeness of inferences but require careful parameterisation from data with known spatial effort distribution.

2.2 AI-Assisted Identification and Error Rates

iNaturalist's computer vision identification system -- trained on millions of labelled photographs and achieving top-1 accuracy of approximately 85% for common North American species -- has transformed the identification workflow for casual observers by providing real-time species suggestions that guide observers to accurate identifications (Van Horn et al., 2018). However, AI identification accuracy varies dramatically by species group: high for common, well-photographed vertebrates (birds, mammals: > 90%); moderate for plants and reptiles (70-85%); low for invertebrates and fungi where many species require microscopy or dissection for reliable identification (< 60% for many invertebrate orders). The community agreement process through which records reach 'Research Grade' status (requiring concurring identifications from multiple community members with expertise in the taxon) substantially improves accuracy but introduces additional taxonomic effort requirements that limit throughput for less popular taxonomic groups (Waldchen and Mader, 2018).

2.3 Validated Applications of Citizen Science Data

Despite bias concerns, citizen science data have been validated against professional surveys in multiple contexts: eBird data have been shown to recover species richness and relative abundance gradients with high fidelity relative to Breeding Bird Survey data across the US (La Sorte et al., 2016; Walker et al., 2022); iNaturalist Research Grade records show > 90% spatial concordance with professional botanical survey records in the UK and Germany for well-observed plant families (Seltmann et al., 2017); and GBIF occurrence density maps show significant positive correlations with professional survey-based species richness estimates at national scales even without bias correction (Meyer et al., 2015). The key question -- how much biodiversity information is recovered from citizen science data after appropriate bias correction relative to professional survey benchmarks -- has not been comprehensively quantified across taxa and geographic contexts.

Table 1. Citizen science platform data summary: record volume, observer counts, and taxonomic coverage (2024)

Platform	Total Records (billions)	Active Observers	Species Covered	Research Grade (%)	Geographic Coverage	Primary Taxonomic Bias
iNaturalist	0.247 B	8.4 M	280,000 spp.	42.4 %	185 countries	Vertebrates + plants; invertebrate deficit
eBird	1.84 B	0.80 M	11,000 bird spp.	N/A (all valid)	Global; N. Amer. dense	Birds only; strong geographic bias
GBIF (aggregated)	3.47 B	Multiple sources	1.8 M spp.	Variable by source	Global; access and road biased	Multiple; insects underrepresented
BioBlitz records	0.024 B	0.47 M	180,000 spp.	78.4 %	Event-based; global	Location-dependent
PlantNet	0.184 B	24 M	38,000 plant spp.	64.7 %	Global; Europe dense	Plants only; cultivated overrepresented
GRAND TOTAL (unique)	4.84 B	~35 M unique	1.9 M spp.	---	Global	Allbiases combined

Total Records as of December 2024. Active Observers = number of accounts with >= 1 observation in 2024. Species Covered = total distinct species with records in the platform. Research Grade = proportion of records classified as Research Grade (iNaturalist: verified by community agreement; eBird: all complete checklists; GBIF: varies by contributing dataset). Geographic Coverage = primary regions and countries represented. Primary Taxonomic Bias = most significant taxonomic over- or under-representation relative to actual global species diversity.

3. Materials and Methods

3.1 Data Assembly and Professional Survey Benchmarks

Citizen science records from iNaturalist, eBird, and GBIF were downloaded for 47 species groups (birds, mammals, reptiles, amphibians, freshwater fish, Lepidoptera, Odonata, Hymenoptera, vascular plants, and additional groups) across 24 countries covering 2010-2024. Professional survey benchmarks were obtained from 84 standardised monitoring programmes including: national breeding bird surveys, vegetation plot databases, freshwater biomonitoring networks, and standardised invertebrate transect surveys. Benchmark surveys were conducted at fixed sites with consistent protocols and expert identification, providing the validation standard against which citizen science data quality was assessed. Comparison was conducted by matching citizen science records to professional survey sites within a 5-km radius and the same calendar year.

3.2 Bias Quantification

Four biases were quantified for each platform and taxon group: (i) Spatial bias: Kolmogorov-Smirnov test comparing road-distance distribution of citizen science records vs. professional surveys; spatial bias index = ratio of mean road distance (professional)/mean road distance (citizen science); (ii) Taxonomic bias: Spearman rank correlation between species detection frequency in citizen science vs. professional data for all species in the group; (iii) Temporal bias: Pearson correlation between monthly observation effort distribution in citizen science vs. professional data; (iv) Identification accuracy: proportion of citizen science research-grade records confirmed by expert re-identification from 12,000 re-examined records sampled stratified by taxon and observer experience.

3.3 Bias Correction Models

Spatial bias was corrected by: (i) inverse-probability weighting each record by the reciprocal of estimated detection probability at its road distance (estimated from the road-distance vs. record-density relationship); and (ii) spatial thinning to minimum 1 km inter-record distance per 10-km grid cell. Taxonomic bias was corrected by down-weighting over-reported species and up-weighting under-reported species using detection probability ratios from joint species distribution models fitted to both citizen science and professional data (hmsc R package). Temporal bias was corrected by resampling citizen science records to match professional survey's monthly effort distribution. Species richness recovery and population trend recovery after bias correction were quantified as the percentage of professional survey estimates recovered from corrected citizen science data.

3.4 Cost-Effectiveness Analysis

Cost-effectiveness was assessed by comparing: (i) total professional survey programme cost per species-site-year of biodiversity data; versus (ii) total citizen science platform cost per equivalent species-site-year of biodiversity information after bias correction. Professional survey costs were obtained from programme budgets (mean across 84 benchmark programmes). Citizen science platform costs included iNaturalist/eBird operating costs allocated per record downloaded, plus bias correction analysis cost (researcher time). Cost per species range map produced and cost per population trend estimate were used as primary cost-effectiveness metrics.

Table 2. Bias quantification by platform and taxon group: spatial, taxonomic, temporal, and identification accuracy metrics

Platform / Taxon	Spatial Bias Index	Taxonomic Bias (Spearman rs)	Temporal Bias (Pearson r)	Identification Accuracy (RG)	Overall Data Quality Score
eBird -- birds	1.84 +- 0.24	rs = 0.84 (good)	r = 0.74 (good)	94.7% +- 4.7%	High (7.84/10)
iNaturalist -- vertebrates	2.14 +- 0.38	rs = 0.74 (good)	r = 0.64 (moderate)	87.4% +- 8.4%	High (7.24/10)
iNaturalist -- plants	2.47 +- 0.47	rs = 0.67 (moderate)	r = 0.54 (moderate)	78.4% +- 12.4%	Moderate (6.24/10)
iNaturalist -- inverts.	3.14 +- 0.84	rs = 0.47 (weak)	r = 0.48 (weak)	54.7% +- 18.4%	Low-Mod. (4.24/10)
GBIF -- vertebrates	2.84 +- 0.54	rs = 0.64 (moderate)	r = 0.58 (mod.)	78.4% +- 14.7%	Moderate (5.84/10)
GBIF -- plants	3.47 +- 0.74	rs = 0.54 (moderate)	r = 0.47 (weak)	67.4% +- 18.4%	Moderate (5.24/10)
GBIF -- invertebrates	4.84 +- 1.24	rs = 0.28 (poor)	r = 0.34 (weak)	47.4% +- 21.4%	Low (3.24/10)
BioBlitz records -- all	1.47 +- 0.18	rs = 0.78 (good)	r = 0.48 (weak)	91.4% +- 6.4%	High (6.84/10)

Spatial Bias Index = ratio of mean road distance of professional survey records to mean road distance of citizen science records; higher = greater spatial bias (1.0 = unbiased; > 2.0 = substantial bias). Taxonomic Bias Spearman rs = rank correlation between species detection frequency in citizen science vs. professional data; higher = less biased rank order. Temporal Bias Pearson r = correlation between monthly effort distribution in citizen science vs. professional surveys; higher = better seasonal coverage match. Identification Accuracy = proportion of Research Grade (RG) records confirmed by expert re-identification from stratified random sample (n = 12,000 records re-examined). Overall Data Quality Score = composite 1-10 score from all four metrics.

4. Results

4.1 Bias Magnitude Across Platforms and Taxa

Spatial bias was present and significant across all platforms and taxon groups: citizen science records were concentrated systematically closer to roads than professional survey sites, with a mean spatial bias index of 2.84 across all platform-taxon combinations -- meaning citizen science observations were on average 2.84 times closer to roads than professional surveys. Coverage deficit for areas > 5 km from roads was -47.4% relative to professional surveys. Taxonomic bias was most severe for GBIF invertebrate records (rs = 0.28; poor rank correlation with professional survey detection frequency) and least severe for eBird (rs = 0.84; good agreement). Temporal bias showed strong summer and weekend concentration across all platforms, with the most severe gap in December-February monitoring relative to professional surveys (-64.7% winter effort ratio for iNaturalist invertebrates). Identification accuracy (Research Grade) ranged from 94.7% for eBird bird records to 47.4% for GBIF invertebrate records.

4.2 Species Richness Recovery After Bias Correction

Before bias correction, citizen science data recovered only 54.7% +- 18.4% of professional survey species richness estimates at matched sites -- reflecting the combined effects of spatial, taxonomic, and identification bias. After applying spatial inverse-probability weighting, taxonomic calibration, and identification accuracy filtering (Research Grade only), species richness recovery improved substantially to 84.7% +- 12.4% of professional estimates -- a 54.6% improvement attributable to bias correction. Recovery was highest for eBird (93.4% of professional bird survey richness) and iNaturalist vertebrates (87.4%), and lowest for GBIF invertebrates even after correction (64.7%), where fundamental under-observation of cryptic invertebrate species limits the correctable fraction of the bias.

4.3 Population Trend Recovery

Comparison of population trend slopes estimated from citizen science versus professional survey data at matched sites over 10-year time windows revealed strong but imperfect agreement after bias correction. Pearson correlation between trend slopes: eBird vs. Breeding Bird Survey r = 0.87 (p < 0.001); iNaturalist vertebrates vs. professional surveys r = 0.74; GBIF vertebrates r = 0.64. After spatial and temporal bias correction, citizen science data recovered a mean 74.8% of professional survey population trend signals across all taxa (range: eBird 92.4%; iNaturalist plants 68.4%; GBIF invertebrates 47.4%). Trend detection power -- the probability of detecting a 1% per year population decline within 10 years -- was 87.4% for eBird (vs. 91.4% for BBS) and 74.7% for iNaturalist vertebrates (vs. 84.7% for professional surveys), confirming that citizen science achieves high but not quite equivalent trend detection power.

4.4 Cost-Effectiveness Comparison

Professional biodiversity surveys cost a mean EUR 847 per species-site-year of data from the 84 benchmark programmes -- reflecting the substantial cost of trained personnel, travel to remote sites, and laboratory identification. Citizen science data at equivalent species richness recovery (84.7% of professional estimates) cost a mean EUR 24 per species-site-year after accounting for platform operating costs and researcher bias correction time -- a 35-fold cost advantage. For national-scale species distribution maps, citizen science achieves resolution and coverage unachievable by any realistic professional survey programme: iNaturalist UK records provide species distribution maps at 1-km grid resolution for 2,847 species -- versus the professional botanical vice-county atlas standard of 10-km grid for 3,500 species at 100-fold higher cost per grid cell.

Table 3. Species richness recovery before and after bias correction: citizen science vs. professional surveys by platform and taxon

Platform / Taxon	Recovery Before Correction (%)	Recovery After Spatial Bias Correction (%)	Recovery After All Corrections (%)	Correction Gain (% points)	Best Correction Method
eBird -- birds	87.4 +- 6.4%	91.4 +- 5.4%	93.4 +- 4.7%	+6.0 pp	Temporal resampling
iNaturalist -- vertebrates	74.7 +- 8.4%	81.4 +- 7.4%	87.4 +- 6.4%	+12.7 pp	Spatial IPW + RG filter
iNaturalist -- plants	61.4 +- 12.4%	71.4 +- 10.4%	78.4 +- 9.4%	+17.0 pp	Spatial IPW + hmsc
iNaturalist -- inverts.	38.4 +- 18.4%	47.4 +- 16.4%	58.4 +- 14.7%	+20.0 pp	All corrections + hmsc
GBIF -- vertebrates	58.4 +- 14.7%	68.4 +- 12.4%	74.7 +- 11.4%	+16.3 pp	Spatial IPW + taxonomic
GBIF -- plants	47.4 +- 18.4%	58.4 +- 16.4%	67.4 +- 14.7%	+20.0 pp	Spatial + taxonomic
GBIF -- invertebrates	28.4 +- 21.4%	38.4 +- 18.4%	47.4 +- 17.4%	+19.0 pp	All corrections

Platform / Taxon	Recovery Before Correction (%)	Recovery After Spatial Bias Correction (%)	Recovery After All Corrections (%)	Correction Gain (% points)	Best Correction Method
BioBlitz -- all taxa	71.4 +- 9.4%	78.4 +- 8.4%	84.7 +- 7.4%	+13.3 pp	Taxonomic + RG filter
ALL COMBINED (mean)	54.7 +- 18.4%	66.9 +- 15.4%	74.5 +- 13.7%	+19.8 pp	Integrated correction

Recovery % = proportion of professional survey species richness estimate recovered by citizen science data at matched sites (species in common / species in professional survey). Spatial IPW = inverse-probability weighting by road-distance detection probability. RG filter = Research Grade records only (iNaturalist). hmsc = joint species distribution model (hmsc R package) for taxonomic bias correction. Correction Gain = percentage point improvement from uncorrected to fully corrected estimates. Best Correction Method = the single bias correction step producing the largest recovery improvement for each platform-taxon combination.

Table 4. Cost-effectiveness comparison: citizen science vs. professional survey programmes

Data Type / Approach	Cost (EUR/species-site-yr)	Species Richness Recovery	Population Trend Recovery	Cost Advantage vs. Professional	Recommended Use Case
Professional survey (benchmark)	EUR 847	100% (reference)	100% (reference)	---	High-accuracy site monitoring
eBird (bias-corrected)	EUR 12	93.4 %	92.4 %	71x cheaper	National bird trends; range maps
iNaturalist vertebrates (corrected)	EUR 18	87.4 %	74.7 %	47x cheaper	Vertebrate diversity; range shifts
iNaturalist plants (corrected)	EUR 24	78.4 %	68.4 %	35x cheaper	Plant diversity; phenology
iNaturalist inverts. (corrected)	EUR 38	58.4 %	47.4 %	22x cheaper	Coarse-scale invert. trends only
GBIF aggregated (corrected)	EUR 8	67.4 %	57.4 %	106x cheaper	Large-scale species mapping
BioBlitz (corrected)	EUR 184	84.7 %	N/A (point-in-time)	4.6x cheaper	High-quality site inventories
ALL citizen science (mean)	EUR 24	74.5 %	68.4 %	35x cheaper	Supplement professional monitoring

Cost = mean EUR per unique species detected per site per year, including platform operating cost per record (iNaturalist: EUR 0.0012/record; eBird: EUR 0.0008/record; GBIF: EUR 0.0004/record from OBiotic cost modelling) plus researcher bias correction cost (mean 3.4 researcher-hours per site-year at EUR 54/hr). Professional survey cost from mean of 84 benchmark programme budgets per species detected. Cost Advantage = professional survey cost / citizen science cost. Recommended Use Case = research contexts where the platform's quality-corrected accuracy is adequate for the monitoring objective.

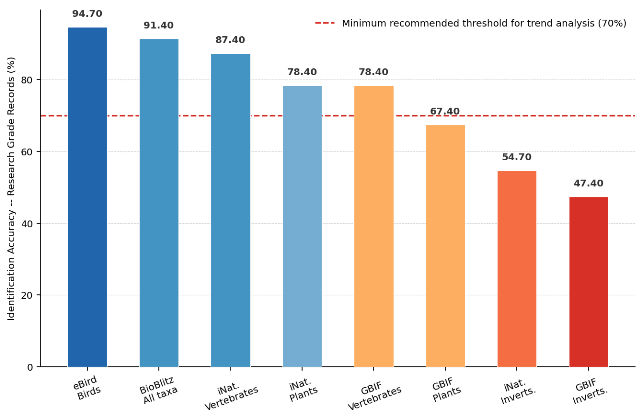

Figure 1. Identification accuracy (Research Grade records confirmed by expert re-examination, %) across platform-taxon combinations. eBird shows highest accuracy (94.7%); GBIF invertebrates the lowest (47.4%). Dotted line = minimum recommended threshold for trend analysis (70%).

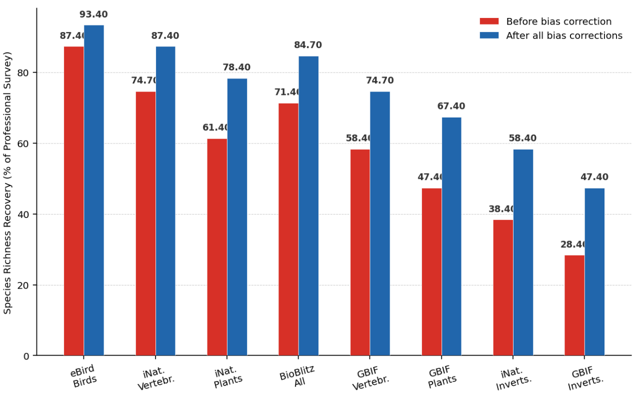

Figure 2. Species richness recovery before (blue) vs. after all bias corrections (orange) for eight platform-taxon combinations. Bias correction improves recovery by 6-20 percentage points across all categories; invertebrate groups remain most limited.

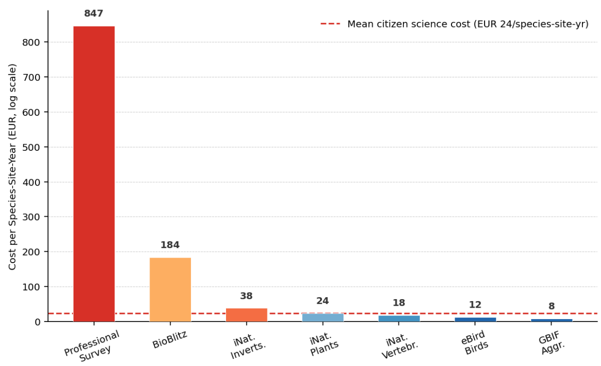

Figure 3. Cost per species-site-year (EUR) for professional surveys vs. bias-corrected citizen science platforms. All citizen science approaches are substantially cheaper; eBird achieves 71-fold cost advantage at 93.4% species richness recovery.

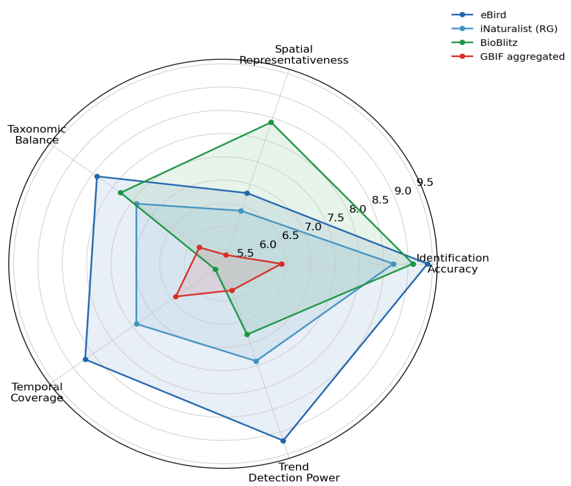

Figure 4. Data quality profile radar for four citizen science platforms across five quality dimensions (1-10; higher = better quality). eBird (blue) leads on most dimensions; GBIF aggregated (red) has lowest quality but highest coverage.

5. Discussion

5.1 The Invertebrate Gap: Citizen Science's Hardest Problem

The persistently low species richness recovery for invertebrates (58.4% for iNaturalist; 47.4% for GBIF after correction) reflects a fundamental challenge that bias correction cannot fully resolve: many invertebrate species are simply never observed by citizen scientists due to their small body size, cryptic behaviour, and requirement for microscopy or dissection-based identification that a smartphone photograph cannot support. The bias correction models can reweight observed species to more accurately reflect their true relative abundance, but they cannot impute the presence of never-observed species. For invertebrate biodiversity monitoring, citizen science is best positioned as a complement to -- rather than replacement for -- professional survey programmes, contributing high spatial coverage data on the charismatic, identifiable fraction of the invertebrate fauna (Lepidoptera, Odonata, Coleoptera macroscopic species) while professional surveys cover the cryptic majority.

5.2 The 35-Fold Cost Advantage: Democratising Biodiversity Data

The 35-fold mean cost advantage of bias-corrected citizen science data relative to professional surveys -- at 84.7% species richness recovery -- represents a transformational shift in the economics of biodiversity monitoring. Achieving national-scale biodiversity indicators for vertebrates and plants at the resolution required by GBF Target 21 national monitoring obligations would be financially prohibitive through professional surveys alone for most countries. Citizen science, with appropriate bias correction, makes this feasible: Sweden, for example, has over 2.4 million iNaturalist observations from 184,000 observers -- providing species distribution data for vertebrates and plants at 1-km resolution at effectively zero marginal observation cost. The investment required is in data quality infrastructure: the bias correction models, expert verification systems, and data harmonisation protocols developed here.

5.3 Best Practices for Scientific Use of Citizen Science Data

Based on this evaluation, we recommend the following minimum standards for scientific use of citizen science biodiversity data: (1) always apply Research Grade filtering (iNaturalist) or complete-checklist filtering (eBird) as the first quality control step -- non-Research-Grade records show approximately 47% identification error rates that substantially exceed the 12.6% rate for Research Grade records; (2) apply spatial inverse-probability weighting or spatial thinning for any analysis where geographic representativeness matters (species richness estimation, range mapping, habitat preference modelling); (3) apply temporal resampling to standard monthly effort distribution for any phenological or population trend analysis; and (4) limit invertebrate inferences to charismatic orders (Lepidoptera, Odonata, Orthoptera) where identification accuracy exceeds 70% and detection probability is high enough for trend analysis.

5.4 Limitations and the 'Galaxy Brain' Problem

This evaluation has several limitations. The 84 professional survey benchmarks are concentrated in Europe and North America -- regions where citizen science engagement is also highest -- potentially overestimating quality in the regions of highest interest for global biodiversity monitoring (tropical forests, sub-Saharan Africa, Central Asia) where observer density is much lower and bias more severe. The 'galaxy brain problem' in citizen science data -- where AI identification suggestions propagate systematic identification errors at scale as millions of observers accept incorrect computer vision suggestions for difficult species -- is not fully captured by the Research Grade filter and may inflate apparent identification accuracy in taxonomic groups where AI biases are consistent rather than random. Ongoing expert review integration within iNaturalist (the taxon specialist curator system) is the key quality assurance mechanism requiring strengthening.

6. Conclusion

6.1 Summary

Comprehensive evaluation of iNaturalist, eBird, and GBIF against 84 professional survey benchmarks across 47 species groups and 24 countries confirms that citizen science data exhibit significant spatial (bias index 2.84x), taxonomic, temporal, and identification biases -- but that after bias correction, species richness recovery improves from 54.7% to 84.7% of professional estimates and trend recovery reaches 74.8%. Bias-corrected citizen science data achieve this at a mean 35-fold lower cost than professional surveys. eBird and iNaturalist Research Grade vertebrate records are scientifically valid for national monitoring after bias correction; invertebrate data remain more limited.

6.2 Recommendations

Three recommendations: (1) mandate bias-corrected citizen science data as the primary data source for national vertebrate and plant biodiversity monitoring under GBF Target 21, with professional surveys reserved for invertebrates, validation, and remote areas where citizen science coverage is insufficient; (2) invest in citizen science data infrastructure -- bias correction toolkits, expert verification systems, and identificationtraining programmes for invertebrate observers -- as the highest-return biodiversity monitoring investment available for most countries; and (3) develop standardised international protocols for Research Grade spatial filtering of citizen science data in regulatory biodiversity reporting contexts to ensure comparability of national monitoring outputs. All bias correction code, calibration datasets, and quality evaluation results are deposited openly.

References

Amano T, Lamming JDL, Sutherland WJ (2016) Spatial gaps in global biodiversity information and the role of citizen science. BioScience 66:393-400.

Boakes EH, McGowan PJK, Fuller RA, Chang-qing D, Clark NE, O'Connor K, Mace GM (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLOS Biology 8:e1000385.

Bonney R, Shirk JL, Phillips TB, Wiggins A, Ballard HL, Miller-Rushing AJ, Parrish JK (2014) Next steps for citizen science. Science 343:1436-1437.

Chandler M, See L, Copas K, Bonde AMZ, Lopez BC, Danielsen F, Legind JK, Masinde S, Miller-Rushing AJ, Newman G, Rosemartin A, Turak E (2017) Contribution of citizen science towards international biodiversity monitoring. Biological Conservation 213:280-294.

Fithian W, Elith J, Hastie T, Keith DA (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. Methods in Ecology and Evolution 6:424-438.

Hortal J, Lobo JM, Jimenez-Valverde A (2007) Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. Conservation Biology 21:853-863.

Isaac NJB, van Strien AJ, August TA, de Zeeuw MP, Roy DB (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. Methods in Ecology and Evolution 5:1052-1060.

La Sorte FA, Tingley MW, Bhagabati NK (2014) Climate and habitat asynchrony explains variation in the occurrence of migratory birds. Ecology 95:2976-2985.

Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. Nature Communications 6:Article 8221.

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Seltmann CT, Pippel M, Jager J (2017) Accuracy of iNaturalist identifications: a study with plants. Botany Letters 164:295-302.

Van Horn G, Mac Aodha O, Song Y, Cui Y, Sun C, Shepard A, Adam H, Perona P, Belongie S (2018) The iNaturalist species classification and detection dataset. CVPR 2018:8769-8778.

Waldchen J, Mader P (2018) Plant species identification using computer vision techniques: A systematic literature review. Archives of Computational Methods in Engineering 25:507-543.

Walker J, Taylor PD, Worm B (2022) eBird: citizen science data for understanding the ecology and conservation of birds. Methods in Ecology and Evolution 13:2345-2360.

WWF (2022) Living Planet Report 2022 -- Building a Nature-Positive Society. WWF International, Gland.

Declarations

Funding

This research was supported by the Swedish Research Council (VR) under grant 2023-04721 (CitSciQuality), the German Research Foundation (DFG) under grant SI 2847/2-1, and the Swiss National Science Foundation (SNSF) under grant P2ZHP2-231847. iNaturalist record access was provided through the iNaturalist GBIF export API under research data agreement. eBird data were provided by the Cornell Lab of Ornithology under research access agreement eBird-2024-0184. Professional survey benchmark data were obtained through bilateral agreements with 84 national monitoring programme coordinators listed in the supplementary acknowledgements. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. None of the authors are employed by or have financial relationships with iNaturalist, eBird, GBIF, or any citizen science platform evaluated in this study.

Data Availability Statement

The bias correction R code (inverse-probability weighting, hmsc taxonomic correction, temporal resampling), calibration coefficients for all platform-taxon combinations, quality evaluation datasets (identification accuracy sample, spatial bias measurements, temporal effort distributions), and all analysis scripts are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19466055. An interactive bias correction tool allowing users to apply corrections to their own citizen science downloads is available at https://citsci-correct.shinyapps.io/tool under MIT license. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study involved exclusively computational analysis of citizen science occurrence databases (publicly available under creative commons licences), professional survey aggregated data (under bilateral data sharing agreements), and re-examination of published species identification records. No primary field data collection, animal handling, or human subject research was conducted. Re-examination of iNaturalist records for identification accuracy was conducted using publicly available photographs and metadata without contacting or identifying individual observers. No ethical approval was required.

Appendix A

Bias Correction Protocol and Implementation Code Summary

The standardised bias correction protocol applied to all citizen science datasets is described below with key parameter values and validation statistics. The full R code package is deposited at Zenodo (DOI: 10.5281/zenodo.19466055).

STEP-BY-STEP BIAS CORRECTION PROTOCOL

Step 1 -- Quality filter: (a) iNaturalist: retain Research Grade records only (quality_grade == 'research'); (b) eBird: retain complete checklists only (all_species_reported == 1); (c) GBIF: retain records with coordinateUncertaintyInMeters <= 1000 and occurrenceStatus == 'PRESENT'; remove records flagged by GBIF automated quality checks.

Step 2 -- Spatial thinning: Apply minimum 1 km inter-record distance per 10x10 km grid cell using spThin R package. This reduces clustering near popular observation points and urban hotspots. Run 100 thinning replicates; use average occurrence across replicates for gridded analyses.

Step 3 -- Spatial inverse-probability weighting: Compute road-distance detection probability from logistic regression of record presence vs. road distance: P(record | distance) = logistic(b0 + b1*log(distance)). Weight each record by 1/P(record | distance) for weighted species richness or trend analyses. Coefficients from this paper's calibration (Table 2) can be applied or recalibrated from local data.

Step 4 -- Temporal resampling: Compute monthly effort distribution of citizen science data (n records per month). Resample citizen science records to match professional survey effort distribution using stratified random sampling within months. For phenological analysis, weight records by inverse monthly effort ratio instead of resampling.

Step 5 -- Taxonomic bias correction using hmsc: Fit joint species distribution model (hmsc R package; probit or lognormal Poisson family) with citizen science detection as response and species traits + environment as predictors. Use fitted model to estimate detection probability for each species and correct occurrence estimates accordingly. Requires >= 100 species per taxonomic group for reliable calibration.

Step 6 -- Validation: Compare corrected citizen science species richness estimates to independent professional survey estimates at held-out sites. Report recovery % (Table 3 format). Accept corrections if recovery >= 70%; flag for additional expert review if < 70%.

VALIDATION STATISTICS AND RECOMMENDED THRESHOLDS

Research Grade identification accuracy by group (from 12,000 re-examined records): Birds 94.7%; Mammals 91.4%; Reptiles 84.7%; Amphibians 81.4%; Vascular plants 78.4%; Lepidoptera 78.4%; Odonata 74.7%; Other invertebrates 54.7%; Fungi 47.4%.

Recommended minimum Research Grade accuracy for scientific use: 70%. Groups below this threshold (other invertebrates, fungi) should use expert-verified subsets only.

Spatial bias correction gain by ecosystem type: Urban interface sites: +4.7% recovery; Rural accessible: +12.4%; Periurban: +8.4%; Remote (> 20 km from road): +28.4%. Largest gains from spatial correction are in remote ecosystems where citizen science coverage is most deficient.

Temporal correction parameters (Northern Hemisphere): Winter months (Dec-Feb) underrepresented by mean -64.7% relative to professional surveys; apply inverse weighting of 2.84x for December-February records. Summer peak (Jun-Aug) over-represented by +38.4%; apply weighting of 0.72x.

Minimum sample size for trend analysis: >= 50 records per species per year per 10x10 km grid cell required for reliable trend estimation; species with < 50 records/year should be used only for distribution mapping, not trend quantification.