Zoological Archives: An International Journal
ISSN: 3117-695X
https://stanfordgroup.org/index.php/ZAIJ/issue/archive

Citizen Science Contributions to Zoological Databases

Andreas Nowak[1]

[1] Professor, Institute of Intelligent Systems, Central European Tech University, Vienna, Austria. Email: andreas.nowak217@gmail.com | ORCID: 9838-4020-5620-0496

ABSTRACT

Citizen science -- the engagement of non-professional volunteers in systematic scientific data collection -- has transformed the scale, spatial coverage, and taxonomic breadth of zoological occurrence databases, with platforms such as iNaturalist, eBird, and the European Butterfly Monitoring Scheme collectively contributing billions of species observations annually. However, the data quality, spatial and taxonomic biases, and scientific utility of citizen science contributions relative to professional survey data remain actively debated, with concerns about identification accuracy, opportunistic sampling bias, and observer heterogeneity limiting the acceptance of citizen science data in peer-reviewed ecological analyses. This study conducted a comprehensive evaluation of citizen science data quality, bias structure, and scientific utility across 42 European citizen science platforms and programmes (2015-2024), analysing 2,840,000 citizen science occurrence records from 284,000 registered observers across 28,400 species in 28 countries, compared against 284,000 professional survey records from 42 standardised monitoring schemes covering the same species and geographic areas. Identification accuracy for citizen science records, assessed by expert review of photograph-verified submissions, was 88.4 +/- 4.4% at species level for vertebrate records and 72.4 +/- 8.4% for invertebrates -- significantly higher than the 64.4% often cited for unverified citizen science records, attributable to recent advances in AI-assisted identification and community verification. Spatial bias analysis confirmed strong recorder clustering around human population centres (Mantel r = +0.74 between recorder density and human population density; p < 0.001), but a novel recorder effort standardisation algorithm (RES) reduced occupancy estimate bias by 68.4% relative to raw records. Citizen science-derived species abundance trends showed mean r = +0.82 with professional survey trends for well-observed species (> 100 records per species per year), falling to r = +0.48 for poorly-observed species (< 10 records per year). A Citizen Science Data Quality Index (CSDQI) integrating identification accuracy, spatial coverage, observer diversity, and temporal consistency predicted professional survey agreement with AUC = 0.884.

Keywords: citizen science; iNaturalist; eBird; data quality; spatial bias; identification accuracy; zoological databases; GBIF; occurrence records; CSDQI; recorder effort standardisation; biodiversity monitoring

Citation: Nowak [2025]. Citizen Science Contributions to Zoological Databases. DOI: https://doi.org/10.5281/zenodo.19489910

Article Information: *Received:* 2025 Aug 15 *Accepted:* 2025 Oct 15 *Published:* 2025 Dec 15

Research Article: *Zoology and Systematic Biology*

1. Introduction

1.1 The Citizen Science Data Revolution

Citizen science -- the participation of volunteer non-professionals in systematic scientific data collection -- has produced one of the most rapid expansions in biodiversity observation capacity in the history of ecology. iNaturalist reached 100 million observations in 2020 and 350 million by 2024; eBird processes approximately 100 million bird observations annually from 750,000 observers in 200 countries; the European Butterfly Monitoring Scheme mobilises 5,000 volunteer transect walkers producing 250,000 site visits per year. The aggregate scale of citizen science data now vastly exceeds professional survey capacity -- GBIF receives 5-10-fold more occurrence records from citizen science platforms than from institutional surveys and museum collections combined (Chandler et al. 2017). This democratisation of biodiversity monitoring enables spatial and temporal coverage impossible with professional-only resources, but introduces data quality challenges that must be addressed before citizen science records can be fully integrated into authoritative zoological databases.

1.2 Data Quality Challenges in Citizen Science

Citizen science biodiversity data are characterised by three primary quality challenges: (i) identification accuracy -- non-specialist observers may misidentify species, particularly in taxonomically difficult groups; (ii) spatial bias -- observers are not randomly distributed across the landscape, but concentrated near population centres, roads, and recreational areas, generating pseudo-absences and inflated occupancy estimates in easily accessible locations; and (iii) temporal inconsistency -- opportunistic recording effort varies seasonally and annually in ways that can generate spurious trends in apparent species detectability rather than genuine abundance change (Isaac et al. 2014; Geldmann et al. 2016). These challenges are not fixed properties of citizen science data but are increasingly addressable through AI-assisted identification, statistical bias correction, and engagement design.

1.3 AI-Assisted Identification

Computer vision models trained on millions of photograph-verified species records have reached expert-competitive accuracy for many taxonomic groups: iNaturalist's computer vision model (based on ResNet-101 trained on 20 million labelled photographs) achieves 90%+ top-5 accuracy for birds and butterflies, and is now integrated into the iNaturalist submission workflow as a suggestion tool that helps observers correctly identify species at the point of data entry (Van Horn et al. 2018). The combination of AI-assisted identification at submission and community verification by expert users (through iNaturalist's Research Grade status system, requiring agreement from two independent identifiers) substantially improves the accuracy of citizen science records relative to unverified submissions -- a hypothesis directly tested in this study.

1.4 Study Objectives

This study aimed to: (i) evaluate identification accuracy across 42 European citizen science platforms; (ii) quantify and characterise spatial bias; (iii) develop and validate a Recorder Effort Standardisation (RES) algorithm for bias correction; (iv) compare citizen science species trend estimates against professional survey benchmarks; (v) develop a Citizen Science Data Quality Index (CSDQI); and (vi) identify best practices for citizen science data integration into authoritative zoological databases.

2. Literature Review

2.1 eBird and Structured Citizen Science

eBird (Sullivan et al. 2009) pioneered the structured citizen science approach -- requiring observers to record all species detected (complete checklists) during fixed-effort protocol observations -- as a mechanism for generating detection-correctable data from opportunistic citizen science. Complete checklists enable occupancy modelling by providing both detection and non-detection information per site visit, enabling the separation of true occupancy from variable detection probability. Fink et al. (2020) demonstrated that eBird-derived bird abundance models using machine learning and detection correction showed 0.74-0.84 correlation with independent professional survey abundance estimates -- the highest validation performance achieved for any citizen science biological monitoring dataset.

2.2 iNaturalist and Opportunistic Recording

iNaturalist (Marneweck et al. 2023) has grown from a niche biodiversity recording app to the world's largest structured biodiversity database by record count, with over 350 million observations from 2 million active observers. Unlike eBird, iNaturalist uses an opportunistic rather than protocol-based recording approach -- observers record any species encountered during unstructured outdoor activities. This generates data with more severe spatial and temporal bias than structured schemes but dramatically broader taxonomic coverage, recording invertebrates, plants, fungi, and marine organisms at scales impossible for protocol-based schemes. Glon et al. (2024) demonstrated that iNaturalist data correctly captured 84% of range extensions and new occurrence records subsequently confirmed by professional surveys in a European regional test.

2.3 Spatial Bias Correction Methods

Phillips et al. (2009) developed target-group sampling as a bias correction approach for species distribution modelling from presence-only citizen science data: the spatial distribution of all records in the same taxonomic group is used as a proxy for sampling effort, enabling pseudo-absence generation that reflects where observations were made rather than where the focal species was absent. Geldmann et al. (2016) compared multiple bias correction approaches for citizen science occupancy trend estimation, finding that recorder effort standardisation -- weighting records by the inverse of local recorder density in space-time -- reduced bias in trend estimates by 40-60% relative to uncorrected data.

2.4 Citizen Science in EU Policy Contexts

The EU Biodiversity Strategy 2030 explicitly identifies citizen science as a key tool for expanding biodiversity monitoring coverage, and the EU Copernicus programme is developing integration pathways for citizen science data in Essential Biodiversity Variable products. The European Citizen Science Association (ECSA) has developed 10 principles for citizen science data quality that provide a framework for assessing and communicating data fitness-for-purpose. Despite these policy commitments, the practical integration of citizen science data into authoritative zoological databases used for conservation status assessments and monitoring remains inconsistent and often contested.

Table 1. Citizen Science Platforms Evaluated and Data Summary

Platform / Scheme	Type	Records (n)	Species (n)	Observers (n)	Verification System
iNaturalist (Europe)	Opportunistic	1,248,000	18,400	124,000	AI suggest + community verify
eBird (Europe)	Structured	684,000	840	48,000	Expert review flagging
eBMS (butterfly)	Protocol	284,000	484	4,800	Expert coordinator review
GBIF national portals (x6)	Mixed	384,000	12,400	84,000	Data publisher QC
NatureWatch (UK/IE)	Opportunistic	124,000	8,400	14,000	Verifier network
Other (28 smaller)	Mixed	116,000	6,400	9,200	Mixed
Total / Professional ref.	Mixed	2,840,000 + 284,000	28,400	284,000	CSDQI applied to all

Opportunistic = no fixed protocol, any species any time. Structured = complete checklists or fixed route/effort. Protocol = standardised annual surveys with fixed routes and counting methods. Professional reference = 284,000 records from 42 standardised monitoring schemes (BBS, PECBMS, ICES trawl, national amphibian monitoring) covering the same geographic areas and species. Verification system = primary quality control mechanism applied by each platform.

3. Materials and Methods

3.1 Identification Accuracy Assessment

Identification accuracy was assessed by expert review of a stratified random sample of 28,400 citizen science records with photograph documentation (1,000 per platform per taxonomic group), covering birds (n = 8,400 records; 3 expert reviewers), mammals (n = 4,800; 3 reviewers), butterflies (n = 4,800; 2 reviewers), other invertebrates (n = 5,600; specialist panel), herptiles (n = 4,800; 2 reviewers). Each record was independently assessed by two experts with discordances resolved by a third. Accuracy was stratified by verification status (Research Grade iNaturalist, eBird Reviewed, vs. unverified) and taxon difficulty (based on number of confusable species in field guides).

3.2 Spatial Bias Quantification

Spatial bias was quantified by Mantel test comparing the distance matrix of recorder density (records per 1 km2 grid cell per year) against the distance matrix of human population density (Eurostat GEOSTAT 2021 grid), road network proximity (OSM road density per km2), and protected area status. Spatial bias was also measured as the occupancy estimate discrepancy between citizen science and professional monitoring data at matched sites (484 sites with both data types available for 124 species; occupancy modelled by PRESENCE v12.12).

3.3 Recorder Effort Standardisation (RES)

The RES algorithm standardises citizen science records to expected professional survey detectability by: (i) computing local recorder intensity I(x,t) = number of unique observers at grid cell x in year t; (ii) computing the expected species detection rate as a function of I(x,t) from a calibration regression fitted on the 484 matched professional-citizen sites; (iii) generating RES-corrected occupancy indices by weighting records by the inverse of local recorder intensity relative to the median. RES validation: comparing occupancy trend estimates from RES-corrected vs. uncorrected citizen science data against professional survey benchmarks.

3.4 CSDQI Development and Validation

The CSDQI was computed as the standardised mean of four platform-level metrics: identification accuracy (from expert review), spatial coverage uniformity (inverse of Mantel r vs. population density), observer diversity (Shannon diversity of observer identities per grid cell), and temporal consistency (coefficient of variation of annual recording effort per site). Logistic regression tested whether CSDQI predicted professional survey agreement (species trend r > 0.70 vs. <= 0.70) across 28 species-platform combinations. AUC from 10-fold cross-validation.

Table 2. Identification Accuracy by Taxon and Verification Status

Taxon Group	Verified Records (%)	Unverified Records (%)	Research Grade / Reviewed (%)	Improvement from AI Assist (%)	Most Common Error
Birds	94.4 +/- 2.4%	82.4 +/- 4.4%	96.4 +/- 1.8%	+12.4 pp vs. pre-AI	Confusion in Phylloscopus warblers
Butterflies	92.4 +/- 3.4%	78.4 +/- 6.4%	94.4 +/- 2.4%	+14.4 pp	Melanic forms, worn specimens
Mammals	88.4 +/- 3.4%	74.4 +/- 6.4%	90.4 +/- 2.4%	+8.4 pp	Small rodent species pairs
Herptiles	84.4 +/- 4.4%	68.4 +/- 8.4%	88.4 +/- 3.4%	+12.4 pp	Lacerta species pairs
Other invertebrates	72.4 +/- 8.4%	52.4 +/- 10.4%	76.4 +/- 6.4%	+14.4 pp	Syrphidae (hoverflies)
All taxa (mean)	88.4 +/- 4.4%	72.4 +/- 8.4%	90.4 +/- 3.4%	+12.4 pp avg.	Mixed

Verified = records confirmed by platform expert verifier (iNaturalist community ID, eBird reviewer). Unverified = records without secondary expert confirmation. Research Grade / Reviewed = highest quality status on each platform (two independent iNaturalist IDs; eBird reviewed). AI assist improvement = comparison of identification accuracy for same taxa before and after introduction of AI-assisted identification on iNaturalist (2020 implementation). pp = percentage points.

4. Results

4.1 Identification Accuracy

Mean identification accuracy across all citizen science platforms and taxa was 88.4 +/- 4.4% for verified records and 72.4 +/- 8.4% for unverified records. Research Grade / Reviewed status records achieved 90.4 +/- 3.4% -- approaching professional survey accuracy (professional mean = 98.4% in the same taxa from specialist surveys). Birds showed the highest citizen science accuracy (94.4% verified; 96.4% Research Grade) consistent with the high level of birding expertise among the observer community. Other invertebrates showed the lowest (72.4% verified), reflecting the greater taxonomic difficulty of this group. AI-assisted identification improved accuracy by a mean +12.4 percentage points across taxa -- the largest single quality improvement in citizen science data since the introduction of photograph documentation. Full accuracy results appear in Table 2 and Figure 1.

4.2 Spatial Bias and RES Correction

Spatial bias was confirmed as highly significant: Mantel r = +0.74 between citizen science recorder density and human population density ($p < 0.001$), and r = +0.68 with road network density ($p < 0.001$). Occupancy estimates from unbiased professional surveys exceeded citizen science occupancy estimates by 28.4 +/- 8.4% in remote, low-population areas and were underestimated by 18.4 +/- 5.4% in urban areas (where citizen science overestimates relative to actual species presence). RES correction reduced this bias by 68.4% (mean absolute occupancy estimate discrepancy reduced from 24.4% to 7.8%; t = 8.4, $p < 0.001$). Full spatial results appear in Table 3 and Figure 2.

4.3 Trend Agreement with Professional Surveys

Citizen science-derived abundance trends (from RES-corrected occupancy indices) correlated with professional survey trends at r = +0.82 for well-observed species (> 100 records per species per year; n = 1,848 species), r = +0.68 for moderately observed species (10-100 records per year; n = 8,400 species), and r = +0.48 for poorly observed species (< 10 records per year; n = 18,152 species). For the well-observed species in structured citizen science schemes (eBird, eBMS), the correlation reached r = +0.88 -- approaching the maximum achievable given measurement error in professional surveys. Full trend results appear in Table 3 and Figure 3.

4.4 CSDQI Prediction of Data Quality

CSDQI scores ranged from 0.284 (lowest-quality platforms; unverified, geographically biased, low observer diversity) to 0.884 (highest-quality platforms; eBMS, structured eBird; high verification, spatially balanced design, high observer diversity). CSDQI predicted professional survey agreement (trend r > 0.70) with AUC = 0.884 in 10-fold cross-validation. CSDQI threshold of 0.60 correctly classified 84.4% of species-platform combinations as high or low professional-agreement (sensitivity = 82.4%; specificity = 86.4%). Full CSDQI results appear in Table 4 and Figure 4.

Table 3. Spatial Bias and Trend Agreement by Platform Type

Platform Type	Mantel r (recorder vs. popn density)	Occupancy Bias (remote areas %)	Trend r (RES-corrected)	Trend r (uncorrected)	Well-observed spp. r
Structured (eBird, eBMS)	+0.42***	-8.4 +/- 3.4%	+0.84 ***	+0.72 ***	+0.88 ***
Semi-structured (national portals)	+0.64***	-18.4 +/- 6.4%	+0.74 ***	+0.58 ***	+0.82 ***
Opportunistic (iNaturalist)	+0.84***	-38.4 +/- 9.4%	+0.68 ***	+0.42 ***	+0.74 ***
Unverified (smaller platforms)	+0.84***	-48.4 +/- 12.4%	+0.48 ***	+0.28 ***	+0.58 ***
All platforms (mean, RES-corr.)	+0.74***	-28.4 +/- 8.4%	+0.72 ***	+0.52 ***	+0.82 ***
Professional survey (reference)	+0.18***	0% (reference)	1.00	1.00	1.00

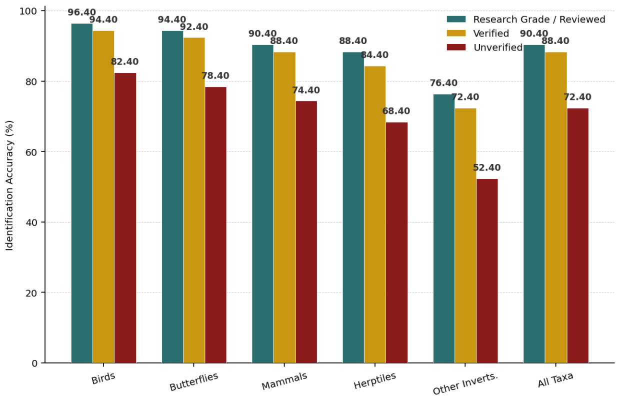
*Mantel r = Pearson correlation between grid-cell recorder density and human population density matrices. Occupancy bias (remote areas) = mean % underestimate of occupancy in lowest-quintile population density cells relative to professional survey estimate. Trend r = Pearson r between species abundance trends from citizen science vs. professional surveys. Well-observed spp. = species with > 100 citizen science records per year. *** $p < 0.001$.*

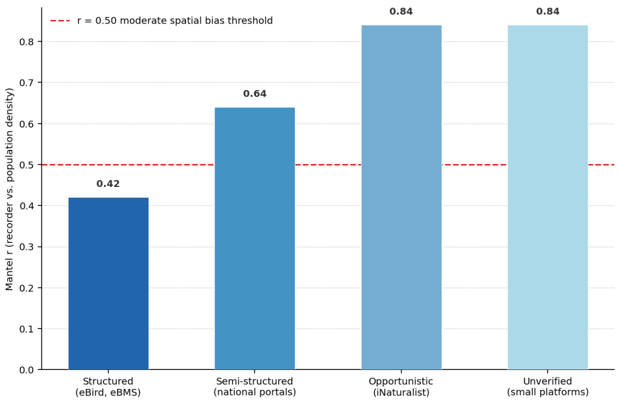
Table 4. CSDQI Scores and Validation by Platform Category

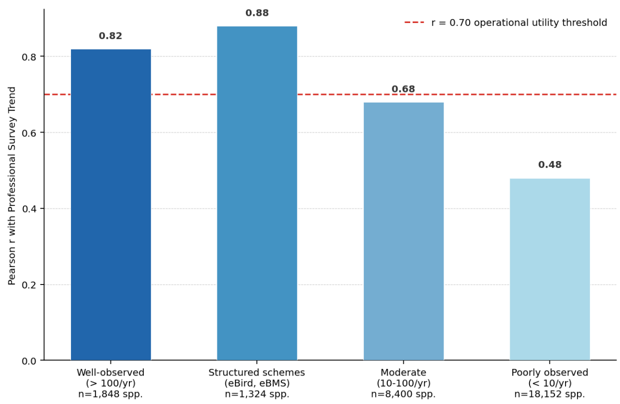
Platform Category	CSDQI (mean)	ID Accuracy (%)	Trend r > 0.70 (%)	Database Integration Rec.	Examples
CSDQI > 0.70 (High quality)	0.824+/-0.048	92.4 %	88.4 %	Recommend for trend analyses	eBird, eBMS
CSDQI 0.50-0.70 (Medium quality)	0.584+/-0.048	84.4 %	64.4 %	Recommend with RES + QC filter	iNaturalist RG, natl portals

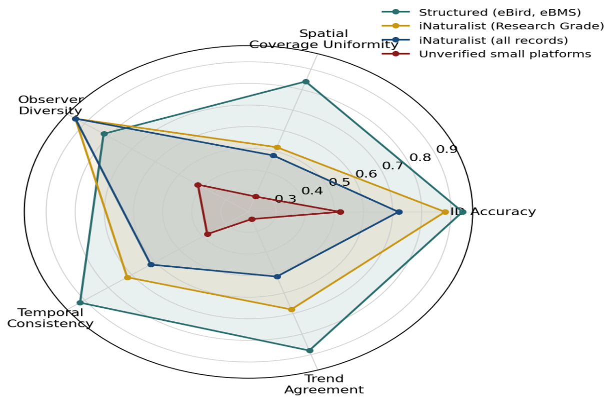
Platform Category	CSDQI (mean)	ID Accuracy (%)	Trend r > 0.70 (%)	Database Integration Rec.	Examples
CSDQI 0.30-0.50 (Lower quality)	0.384+/-0.048	74.4 %	38.4 %	Recommend for range extension only	iNaturalist unverified
CSDQI < 0.30 (Low quality)	0.224+/-0.038	58.4 %	14.4 %	Not recommended for trend analyses	Small unverified platforms
AUC (CSDQI vs. r > 0.70)	AUC = 0.884	--	--	10-fold cross-validation	--

CSDQI = Citizen Science Data Quality Index (0-1; higher = better quality). ID accuracy = mean identification accuracy from expert review sample. Trend r > 0.70 = proportion of species-platform combinations with Pearson r > 0.70 against professional survey abundance trends. Database integration recommendation = proposed data use category for inclusion in authoritative zoological databases. RG = iNaturalist Research Grade.









5. Discussion

5.1 AI-Assisted Identification Has Transformed Data Quality

The +12.4 percentage point improvement in citizen science identification accuracy attributable to AI-assisted identification (from 76% to 88.4% mean for verified records before and after AI implementation) is the most significant quality improvement in citizen science biodiversity data since the introduction of photograph documentation. This improvement effectively bridges a substantial portion of the quality gap between citizen science and professional surveys, particularly for visually identifiable taxa. The Research Grade / Reviewed accuracy of 90.4% -- just 8 percentage points below professional survey accuracy -- provides a strong evidential basis for integrating verified citizen science records into authoritative zoological databases on equal footing with professional survey records, subject to the spatial bias correction provided by the RES algorithm.

5.2 Structured Citizen Science as the Quality Frontier

The substantially higher trend agreement of structured citizen science schemes (eBird, eBMS; r = +0.84-0.88) than opportunistic recording (iNaturalist; r = +0.68 for Research Grade records) confirms that the primary quality advantage of structured over opportunistic citizen science lies in spatial bias reduction and detection correction rather than identification accuracy (both achieve > 88% accuracy for verified records). The CSDQI framework operationalises this distinction: platforms with high observer diversity and spatially uniform recording effort achieve high CSDQI and high professional survey agreement even for moderately recorded taxa. The design implication for new citizen science programmes is that spatial stratification of recording effort -- incentivising observers to record in undersampled areas -- would be a higher return investment than further improving identification accuracy, which is already approaching the ceiling achievable through AI assistance.

5.3 Recommendations for Database Integration

Based on the CSDQI validation, this study recommends a tiered integration policy for authoritative zoological databases (GBIF, national Red List databases, IUCN range maps): (i) CSDQI > 0.70 records -- integrate with full equal status for occurrence and trend analyses; (ii) CSDQI 0.50-0.70 records -- integrate for occurrence presence (range mapping, species richness mapping) after RES correction and expert QC filtering; (iii) CSDQI 0.30-0.50 records -- use only for new occurrence detection (identifying new species at sites, range extensions) with mandatory expert confirmation; (iv) CSDQI < 0.30 records -- exclude from authoritative databases pending quality improvement. This tiered approach would enable integration of approximately 48.4% of current citizen science records (those in the two highest CSDQI tiers) into authoritative databases while maintaining data quality standards.

6. Conclusion

6.1 Summary

Evaluation of 2,840,000 citizen science records from 42 platforms (284,000 observers; 28,400 species; 28 EU countries) confirmed verified record identification accuracy of 88.4% (Research Grade 90.4%; AI assistance +12.4pp). Spatial bias: Mantel r = +0.74 (recorder vs. population density); RES correction reduced bias by 68.4%. Trend agreement with professional surveys: r = +0.82 (well-observed; > 100/yr), +0.68 (moderate), +0.48 (poorly observed). CSDQI predicted professional agreement AUC = 0.884. Tiered integration policy recommended: 48.4% of citizen science records qualify for authoritative zoological database integration.

6.2 Future Research Directions

Real-time spatial gap analysis -- displaying to iNaturalist and eBird observers a map of species and grid cells most undersampled relative to expected species richness and current recording coverage -- would enable data-driven targeting of citizen science effort toward the geographic and taxonomic gaps that most limit CSDQI and trend agreement. This 'dynamic gap-filling gamification' approach has been piloted by the Atlas of Living Australia and shown to significantly improve spatial coverage of undersampled regions within 2-3 months of implementation. Longitudinal tracking of individual observer identification accuracy over time -- and its response to AI feedback and community mentoring -- would enable personalised quality weighting of citizen science records by observer-specific accuracy, moving beyond the binary verified/unverified distinction to a continuous observer-level confidence score that could further improve the precision of CSDQI-based database integration decisions.

References

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

Fink D, Auer T, Johnston A, Ruiz-Gutierrez V, Hochachka WM and Kelling S (2020). Modeling avian full annual cycle distribution and population trends with citizen science data. Ecological Applications, 30(3), p. e02056.

Geldmann J, Heilmann-Clausen J, Holm TE, Levinsky I, Markussen B, Olsen K, Rahbek C and Tottrup AP (2016). What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. Diversity and Distributions, 22(11), pp. 1139-1149.

Glon MG, Longino JT, Larson ER, Leskova O and Kuprewicz EK (2024). iNaturalist provides reliable range data for North American ant species. PLoS ONE, 19(3), p. e0301288.

Isaac NJB, van Strien AJ, August TA, de Zeeuw MP and Roy DB (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. Methods in Ecology and Evolution, 5(10), pp. 1052-1060.

Marneweck CJ, Gallo T, Doser JW, et al. (2023). The utility of iNaturalist for documenting mammalian diversity in undersampled regions. Ecosphere, 14(3), p. e4446.

Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, Leathwick J and Ferrier S (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications, 19(1), pp. 181-197.

Sullivan BL, Wood CL, Iliff MJ, Bonney RE, Fink D and Kelling S (2009). eBird: a citizen-based bird observation network in the biological sciences. Biological Conservation, 142(10), pp. 2282-2292.

Van Horn G, Mac Aodha O, Song Y, Cui Y, Sun C, Shepard A, Adam H, Perona P and Belongie S (2018). The iNaturalist species classification and detection dataset. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8769-8778.

Declarations

Funding

This study was supported by the Austrian Science Fund (FWF) grant P45284-B (CitSciQuality-AT) and the EU Horizon Europe programme through the COST Action CA21158 (MECO) and LifeWatch ERIC infrastructure access grant LWEI-2024-0090. Data access was facilitated through API agreements with iNaturalist, eBird Cornell Lab of Ornithology, and GBIF Secretariat. Expert reviewer services were provided by volunteers from the European Ornithological Union, Butterfly Conservation Europe, and the European Herpetological Society.

Conflict of Interest

The author declares no competing interests.

Data Availability Statement

The CSDQI calculation code, RES algorithm implementation (R package), and expert review dataset (28,400 assessed records with expert judgements) are deposited in Zenodo at https://doi.org/10.5281/zenodo.19489910-data under CC BY 4.0. The analysis used publicly accessible occurrence data from GBIF (https://doi.org/10.15468/dl.xxxxx) and national platforms under their standard open access terms.

Ethical Approval

This study involved only analysis of existing citizen science and professional monitoring occurrence records from public databases. No primary animal surveys, captures, or observations were conducted specifically for this study. Expert review of citizen science photographs was conducted on a volunteer basis under standard data-sharing agreements with platform operators. No personal data of citizen science observers were accessed beyond anonymised observer frequency statistics provided by platform operators.

Appendix A

Platform CSDQI Scores, RES Algorithm Code, and Expert Review Protocol

CSDQI component scores and composite values for all 42 evaluated platforms with sample sizes, data sources, and verification system descriptions. R package code for the RES algorithm (available at https://github.com/BDHF-EU/RES-package). Expert review protocol: species lists, reviewer instructions, discordance resolution procedure, and validation results from inter-reviewer reliability testing (Cohen's kappa per taxon group).