Vegetation Ecology and Diversity 63 (2026) e189819 | DOI 10.3897/ved.189819

PENSOFT.

VEGETATION ECOLOGY AND DIVERSITY

Research Article

Training and intercalibration reduce observer-induced variability in forest vegetation surveys

Marco Cervellini[1*], Luciano Ludovico Maria De Benedictis[1*], Leonardo Salvatori[1], Stefano Chelli[1], Giandiego Campetella[1], Federico Selvi[2], Giovanni Iacopetti[2], Arianna Ferrara[3], Alessandro Chiarucci[3], Antonella Canini[4], Maura Francioni[1], Chiara Scalet[1], Roberto Canullo[1]

1 *Biodiversity and Ecosystem Management Unit, School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy*
2 *Department of Agriculture, Food, Environment and Forestry, University of Florence, PlantDive Lab, Florence, Italy*
3 *BIOME Lab, Department of Biological, Geological & Environmental Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy*
4 *Department of Biology, University of Rome "Tor Vergata", Rome, Italy*

Corresponding author: Leonardo Salvatori (leonardo.salvatori@unicam.it)

Academic editor: Irena Axmanová ♦ **Received** 25 February 2026 ♦ **Accepted** 24 April 2026 ♦ **Published** 13 May 2026

Abstract

The increasing availability of vegetation-plot data results from decades of field surveys conducted by numerous observers. This highlights observer error not as a marginal issue but as a structural component of uncertainty, with direct consequences for estimates of species turnover and temporal trends. Observer-related sampling bias can be investigated through variation in species richness, species composition and abundance. Specifically, pseudoturnover refers to changes in species composition caused by overlooking or misidentification of taxa between sampling events or among different observers surveying the same plot. Although the causes and implications of observer error have been widely discussed, it remains unclear whether observer-related pseudoturnover decreases within observer groups as a result of training. Using data from training and intercalibration sessions carried out in 2023 and 2025 within two forest monitoring programmes in Italy (the LIFE project ModerNEC and the Italian National Forest Inventory), we assessed whether targeted training and collective briefing reduce observer-induced pseudoturnover. We applied Bayesian multilevel models to estimate changes in inter-observer species richness variability and inter-observer dissimilarity. The former decreased across observers in both years, while the latter declined after training when using Jaccard and Euclidean distances in both years; Bray–Curtis dissimilarity decreased only in 2023 and increased in 2025. Overall, training and intercalibration are likely to reduce observer-induced pseudoturnover related to species presence, while variability in abundance estimation needs further study and remains a key challenge for future vegetation monitoring programmes.

Keywords

Intercalibration, observer error, pseudoturnover, training, vegetation resurvey

Introduction

There is a large and increasing availability of vegetation plot data, particularly in Europe, covering wide spatial extents as well as long-term temporal gradients, largely resulting from decades of field surveys conducted by hundreds of surveyors. These data are fundamental to understand vegetation patterns and the drivers shaping them, including global changes (Chytrý et al. 2016; Bruelheide et al. 2019; Knollová et al. 2024). However, even when

* These authors contributed equally to this work.

backed by botanical and field expertise, sampling relies on visual and inherently subjective judgements made by human observers, both in terms of species identification and estimation of relative abundances (Vittoz and Guisan 2007; Morrison 2016; Salvatori 2024). Owing to this subjectivity, vegetation surveys are prone to several forms of observer-related error, which can substantially affect the reliability and comparability of ecological data (Boch et al. 2022). For instance, following Kéry and Gregg (2003), detectability is defined as the probability of recording a species given its true presence in the plot. Imperfect detectability can generate false absences, leading to spurious estimates of local extinctions and vegetation change if the analyses do not explicitly account for this bias. Futschik et al. (2020) found that observer-related errors can affect our interpretation of the effect of climate change on alpine grassland, specifically, they found differences of 5 to 13% of species in the records conducted by different observers. Similarly, Verheyen et al. (2018) showed that a substantial proportion of the floristic changes attributed to eutrophication or climate are sensitive to detection errors in the herbaceous layer. Observer error is therefore a pervasive issue in vegetation science, particularly in long-term monitoring programmes and multi-observer surveys (Dengler et al. 2011; Bruelheide et al. 2019; Seidling et al. 2020; Morrison et al. 2024), affecting estimates of biodiversity, turnover, and trends. Although it is evident that observer error in vegetation surveys can influence the interpretation of our results, there is still limited awareness of this issue, and the strategies to reduce or control this source of error are rarely implemented (Kercher et al. 2003; Morrison 2021).

One way to quantify observer error is to compare observations with the true value of the variable of interest. However, in vegetation studies, true values are rarely known because a complete and exhaustive list of species in a given area is often unavailable (Di Biase et al. 2025). While accuracy—the closeness of an observation to the true value—is therefore difficult to assess (Salvatori 2024), precision—the degree to which repeated observations agree with each other—can be explicitly evaluated and provides a meaningful measure of observer performance. Observer-related sampling bias can be investigated by examining variation in species richness, species composition and species abundances. Morrison (2016) identified three main types of observer error in vegetation surveys: (i) overlooking error, when species present in the plot are not detected; (ii) misidentification error, when species are incorrectly identified; and (iii) visual estimation error, when species abundances (e.g., in terms of projected ground cover) are inaccurately assessed. Other sources of error are related to "attributes associated with observers" and could include errors (typos) made during data digitalisation due to mental or physical fatigue. When repeated surveys by the same observer yield different results, this is referred to as intra-observer error; when discrepancies arise among different observers or teams, it is referred to as inter-observer error. Focusing on species as observational units, several studies have shown that overlooking small, rare or inconspicuous species inflates apparent turnover among observers (Nilsson and Nilsson 1985; Morrison 2016). This issue, also called pseudoturnover, has been defined as the proportion of species either missed or misidentified by an observer or team between two sampling occasions or among different observers surveying the same plot. Vegetation surveys often include species abundances attributed to different vertical layers, e.g., in forest ecosystems. Vertical stratification is generally recorded more consistently for dominant tree layers, whereas shrub and herbaceous layers are more prone to both intra- and inter-observer variability (Kennedy and Addison 1987; Klimeš et al. 2001; Kéry and Gregg 2003; Archaux 2009). High variability among observers in estimating percentage cover has been widely documented, with systematic tendencies to under- or overestimate cover depending on size of individuals and plot density (Morrison et al. 2020).

In this context, Boch et al. (2022) reported that "it remains unclear whether observer-related pseudoturnover declines with time within a group of observers, thanks to their continuous training in species identification and their increasing experience in conducting vegetation records". Morrison (2016) and Morrison et al. (2024) confirmed that the use of multiple observers, additional training including active feedback approaches, and continual evaluation and calibration among observers are recommended strategies to reduce observer error in vegetation surveys.

Forest ecosystems represent one of the most frequently studied habitats in observer error research (Morrison 2016), highlighting the relevance of this issue for forest monitoring programmes. Improving our understanding of observer-related precision is particularly important for long-term monitoring schemes involving numerous observers and repeated surveys over time, such as the ICP Forests Level II network (Allegrini et al. 2009; Ferretti 2021) and the Italian National Forest Inventories (Gasparini et al. 2009; D'Amico et al. 2025).

Using data collected during training and intercalibration sessions carried out within these programmes in Italy, we aim to assess whether targeted training and collective briefing can reduce observer-induced pseudoturnover in forest vegetation surveys. Specifically, we estimate whether inter-observer variability decreases between pre- and post-training exercises in terms of species richness and species composition, the latter measured by three dissimilarity indices.

Materials and methods

Training and intercalibration approach

In 2023, within Action B4 "Training for monitoring operators" of the LIFE project ModerNEC (https://lifemodernec.eu/), four teams, each composed of two botanists, surveyed four 10 m × 10 m forest plots. Observers recorded species presence and percentage cover for herbaceous, shrub and tree layers. The same approach was adopted in

2025 within the action "Training and intercalibration for plant diversity surveyors" of the IFNI 2025 project (D'Amico et al. 2025), involving 15 individual botanists surveying eight 10 m × 10 m forest plots. All plots were set up using wooden stakes and tape at two sites within the Marsiliana State Reserve (https://rgpbio.it/riserva/marsiliana/ - Tuscany, Italy) by personnel of the Carabinieri Forestali (Figure 1). Forest sites are characterised by a dominance of *Fraxinus ornus* with presence of *Arbutus unedo*, *Phillyrea latifolia*, *Quercus cerris* and *Quercus pubescens*.

In 2023, the project protocol prescribed the use of the 7-degree Braun-Blanquet cover-abundance scale (Braun-Blanquet 1964; Canullo et al. 2012) while in 2025 species cover was assessed using direct percentage estimates. For the statistical analysis of 2023 data, the Braun-Blanquet values were transformed into the respective median value of each category (i.e., 5 = 87.5%; 4 = 62.5%; 3 = 37.5%; 2 = 15%; 1 = 3%; + = 0.5%; r = 0.01%; Canullo et al. 2020).

Training and intercalibration activities were structured over three consecutive days and included two survey exercises with the following data entry phase, and a collective briefing session between the surveys. In 2025, a list of species occurring in an area that includes all the plots, produced by a control team, was supplied to the observers. While this was not performed in 2023, control team and observers still assembled a field-based consensus list before the first survey exercise. For both projects, observers were selected on a geographical basis (i.e., Italian administrative regions) and playing with a common species list is a tool to simulate their familiarity with the real field flora. Prior to the surveys, an on-field operations manual was distributed to all observers. The manual described the training survey protocol in detail, outlining the sequence of actions and specific procedures required for plant diversity surveys. The first survey exercise ("pre-training", EX1) was conducted on the first day, with the observers cycling through the plots. On the second day, data entry was followed by a briefing session, during which observers discussed difficulties and self-perceived errors, including taxonomic issues. The second survey exercise ("post-training", EX2) consisted of re-surveying the same plots and was carried out on the third day. During the briefing session, observers participated in a "species familiarization" exercise (with plant specimens collected outside the plots) aimed at identifying taxa considered critical for correct identification (Canullo et al. 2012). Diagnostic comparisons were performed to clarify key morphological traits, taxonomic ambiguities and differences between morphologically similar species. A group exercise for cover estimation was performed before EX1, showing percentage cover charts of squares with different proportions of area (e.g. 0.5%, 1%, 5%, 10%, etc.), and then on-field before EX2 to harmonise the visual estimation of 1 m^2 of vegetation cover, using a commonly occurring species, *Ruscus aculeatus*. Graphical outputs summarising species richness, abundances, and herb-layer cover recorded during EX1 by each observer in each plot

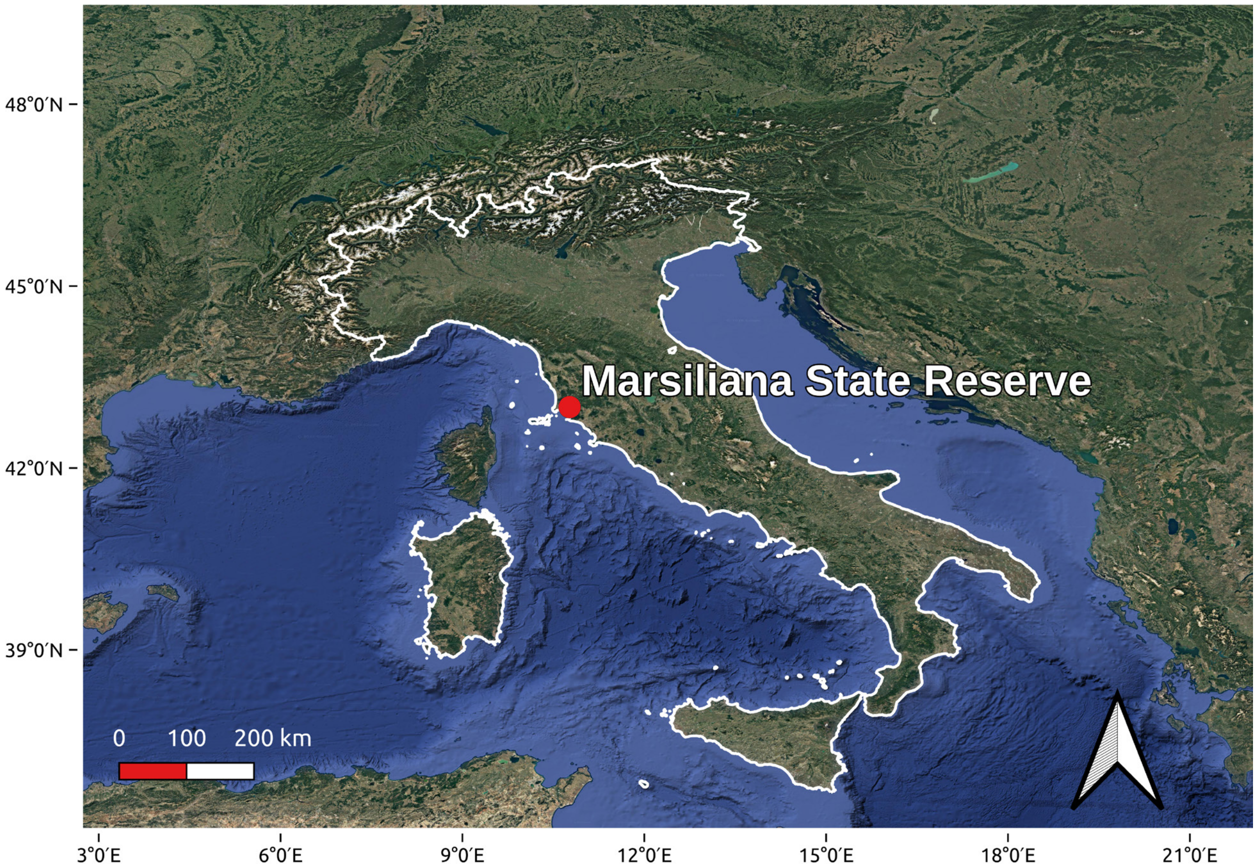

Figure 1. Location of the Marsiliana State Reserve, with the Italian national border as white outline. Base map data 2026 © Google.

were also displayed and then discussed during the briefing, allowing observers to reflect on differences in visual estimation (Suppl. material 1). The briefing concluded with a collective discussion on the interpretation and application of the field manual.

Statistical analyses

We used Bayesian multilevel models fitted through the 'brms' package (Bürkner 2017) to estimate the change in precision in assessing species richness and the change in inter-observer dissimilarity between the two exercises. Bayesian models fit through Markov Chain Monte Carlo (MCMC) allow for complex model structures and a wide choice of likelihoods, flexibility in the specification of group-level terms, easy computation of probability statements about any model parameter, and generation of many kinds of predictions.

Since the two years represent different training experiences, different structure of the teams, and different vegetation, they are analysed separately.

All analyses and data handling have been performed with R, version 4.5.3 (R Core Team 2026). The PDF notebook in the code repository reports the specific version of each package and other system information.

Inter-observer variability in species richness

In order to measure the effect of training on observer-induced variability in the estimation of species richness, we assessed whether the standard deviation between observers decreased after training. Species richness was modeled as a function of the population-level effects of exercise (EX1 and EX2) and plots (SU), including a group-level effect for observers estimated separately for each exercise. Since the focus was on observer variation, not reliable inference about plot-level richness, SU was chosen as a population-level effect reflecting the differences in species richness among plots. Weakly informative priors were used, together with prior predictive simulations, to constrain outcomes and effects within known bounds (Gelman et al. 2013, 2017). Convergence was verified using numerical and visual diagnostics (Vehtari et al. 2021). Inter-observer variability was estimated through the posterior standard deviation of the observer group effect and contrasted between EX2 and EX1 through their difference.

Inter-observer variability in species composition

Since most of the plant biodiversity in temperate forest ecosystems is found in the herbaceous layer (Gilliam 2007), the average dissimilarity of specific composition between observers in EX1 and EX2 (β-diversity) was assessed using data from this layer (see also Archaux 2009). For each exercise × SU combination, species × observer matrices were created. Percentage abundances were transformed using log(1 + cover) to attenuate the influence of species with high cover (Legendre and Legendre 2012). In cases where an observer reported the same species more than once, the first occurrence was retained. For the 2025 dataset, one observer had not carried out the survey in one of the eight SUs before the training; to avoid unbalanced comparisons, this observer was also excluded in the post-training phase for that specific SU. β-diversity calculation was performed pairwise between observers within each exercise × SU combination. Three dissimilarity indices were calculated for each matrix: Jaccard (based on presence/absence), Bray–Curtis (based on abundances), and Euclidean distance between abundance vectors (abundances and double-zeros are taken into account); the latter was transformed into a dissimilarity value bounded between 0 and 1 dividing it by the largest overall value, so that it can be treated similarly to the other two dissimilarities. While the Euclidean dissimilarity is considered unsuitable as a measure of resemblance between plots because of the spurious effect of double-zeros (Legendre and Legendre 2012; Ricotta 2021), it is useful to compare observers in reference to a fixed vegetation object, because an absence reported by two observers represents a meaningful similarity between them.

Dissimilarity was estimated as a function of a population-level effect of exercise and group-level effects of SU and observer, allowing systematic differences in dissimilarity among SUs and observers. Since the dissimilarities take values bounded between 0 and 1, with values near the bounds likely, the beta likelihood was used, with the logit link for the mean and the identity link for the sample size parameter. The beta distribution does not allow exact 0s and 1s, so we subtracted 10^{-6} from the 1s resulting from scaling the Euclidean distance into this range.

Regression coefficients from generalised linear models cannot usually be interpreted directly, representing a non-linear change in the outcome through the link function (Rohrer and Arel-Bundock 2025), thus the results of these models are presented through posterior predictive distributions of the expected dissimilarity in the two exercises, and as percent lift of the exercise effect on the outcome scale—(EX2 − EX1) / EX1 × 100.

Results

Species richness showed no consistent change between EX1 and EX2 in 2023 (0.92 species, 95% posterior quantile interval: -4.42–5.79) and 2025 (-1.65 species, 95% posterior quantile interval: -2.87 – -0.41), while the precision of species richness estimates increased across observers in both years. Inter-observer variability was higher before training, as indicated by the posterior distribution of the EX2 - EX1 contrast (Figure 2), with a probability of a decrease in inter-observer variability of 70% in 2025 and 67% in 2023.

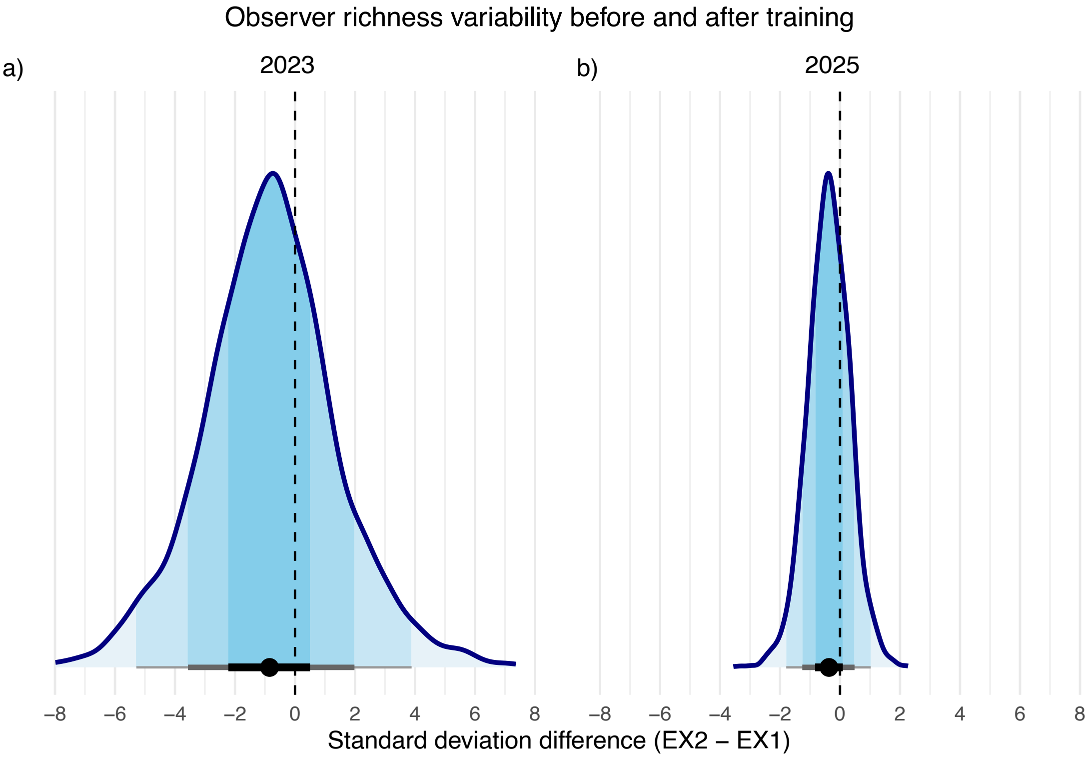

Figure 2. Distribution of the difference in posterior draws for the standard deviation of the observer group-level effect. The colour bands and the intervals represent 50%, 80%, 95% widths, the point represents the median.

Dissimilarity analyses revealed a general reduction in inter-observer dissimilarity after training when using Jaccard index and Euclidean distance in both years. In contrast, Bray–Curtis dissimilarity decreased only in 2023, whereas there is evidence of an increase in 2025 (Figure 3). The percentage reduction in dissimilarity ranged between 10 and 30%, with the exception of Bray–Curtis in 2025 (Figure 4).

Discussion

Training and intercalibration substantially reduced inter-observer variability in species richness and dissimilarity in the 2023 and 2025 exercises. In our case study, the briefing session seems to have produced a similar effect on variability reduction in measuring species richness in both 2023 (4 teams with 2 observers each) and 2025 (16 individual observers) exercises.

The reduction in Jaccard-based dissimilarity confirms the briefing session enhanced agreement in species detection, highlighting its effectiveness in reducing overlooking and misidentification errors. Importantly, a similar reduction in variability was observed in 2023, even without a reference species list, indicating that the interactive briefing session per se was sufficient to promote species familiarization and observer alignment (Vittoz and Guisan 2007).

In contrast, the absence of a consistent reduction in Bray–Curtis dissimilarity indicates that individual differences in abundance estimation persisted despite training, particularly in the 2025 campaign, which involved a larger number of single observers. This pattern was observed even though the 2025 protocol prescribed the use of direct percentage cover estimates, generally considered more robust than Braun-Blanquet classes (Dengler and Dembicz 2023). This indicates that species cover estimation remains a critical source of observer variability, especially for the most abundant species (e.g. *Ruscus aculeatus*), and may require more specific calibration tools. While plants with low cover are associated with greater observational variability when focusing on species detection (Klimeš et al. 2001; Archaux 2009), in this case it is the cover of abundant species that has the largest impact and potential variation (Vittoz and Guisan 2007). Particular attention must therefore be paid to training observers to arrive at a consensus on abundance estimates by organizing group exercises in which the same species is assessed and discussed simultaneously by all. This exercise can be extended and applied to morphological groups of species (e.g., graminoids such as *Carex* sp.).

The observed decrease in Euclidean dissimilarity further supports a general convergence among observers, especially in terms of shared detections and non-detections. It has to be stressed again that this distance considers “double zeros” in the same way as “double presences”. Usually, this feature does not allow the use of this index for community comparisons (Orlóci paradox, see Ricotta 2021), but in the context of measuring the differences

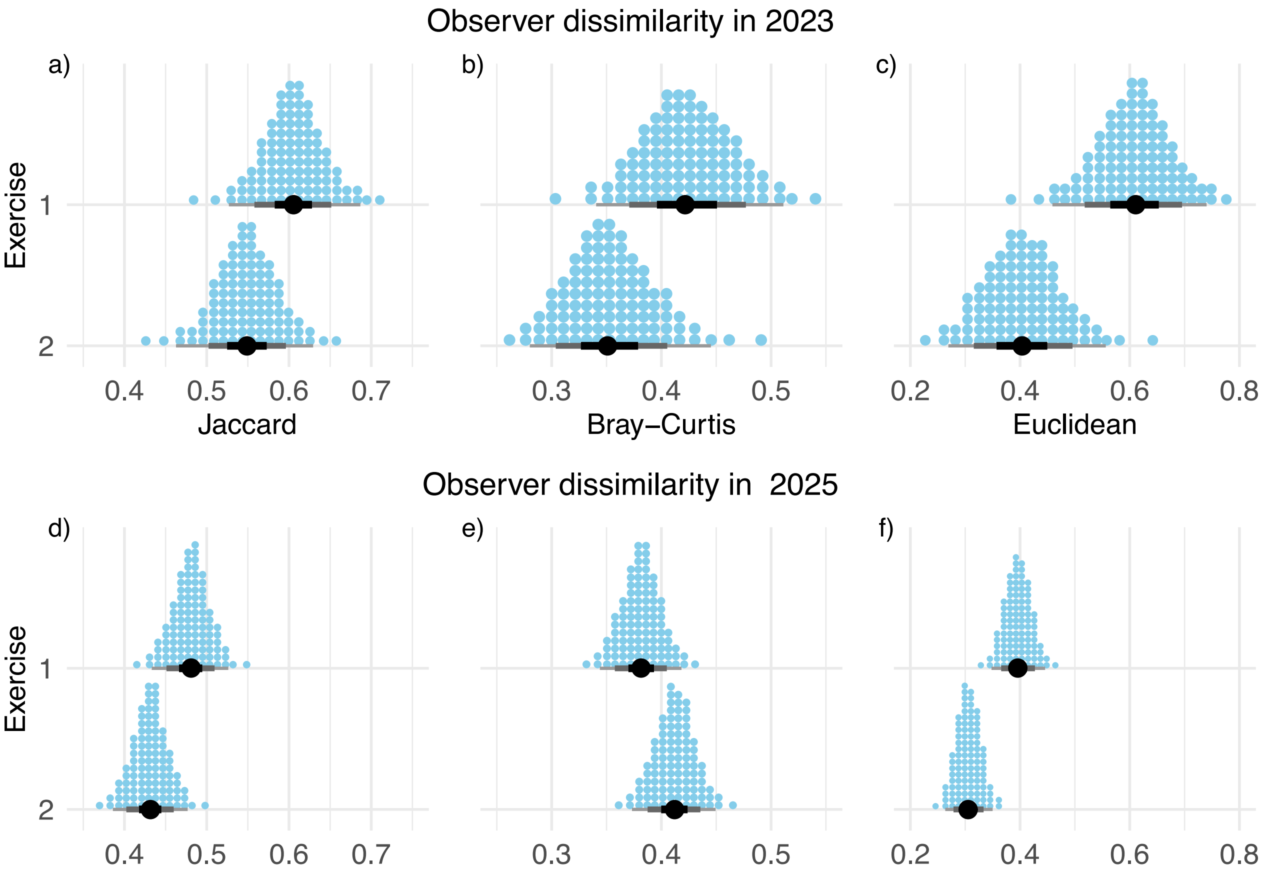

Figure 3. Quantile dot-plot of the expectation of the posterior predictions of beta diversity in the two exercises. The blue dots represent 100 evenly spaced quantiles of the distribution. The intervals represent 50%, 80%, 95% widths, the point represents the median.

among observers in detecting species (or in not detecting the same species) it seemed to be effective in mitigating the discrepancies related to plant cover found with the Bray–Curtis dissimilarity. The different patterns observed among dissimilarity indices illustrate how methodological choices may always influence conclusions, in this case about observer agreement, even when their differences are well known and chosen intentionally to answer different questions.

Overall, these results demonstrate that training and intercalibration are effective in reducing observer-induced pseudoturnover related to species presence, while variability in abundance estimation remains an important challenge for future vegetation monitoring programmes. The approach presented in this study addresses the needs of quality assurance for monitoring in forest ecosystems and complies with the requirements and criteria of the ICP Forests Monitoring Programme (Ferretti et al. 2021, section 4.2), but it can be adapted and tailored to the specific type of vegetation being surveyed.

Focusing on key elements of vegetation surveys, we provide some general advice. In programmes covering a wide geographic extent, project planning would assign observers on a geographical basis (e.g., Italian administrative regions). Playing with a common species list is a tool to simulate familiarity with the flora, which is otherwise specific to the survey exercises, providing a baseline to assess remaining biases in terms of species identification and overlooking. Often, the survey protocol requires subdividing the vegetation into vertical layers; this is a potential cause of large discrepancies in cover estimation, therefore the training should focus on correct application of the protocol and all possible edge cases or misinterpretations that could occur. Sampling units differ in both size and structure (e.g., nested or transect) depending on the ecosystem considered, whether it is characterized by denser vegetation with few dominant species (e.g., grassland) or by sparsely distributed species (e.g., shrubland); the training phase should reflect as much as possible all those sampling conditions.

Conclusions

Our results provide evidence that targeted training and intercalibration are effective tools for reducing observer-induced pseudoturnover related to species presence in vegetation surveys, even when observer heterogeneity increases.

Monitoring programmes and resurvey projects should include training and intercalibration exercises as a way to establish a common ground not only in terms of a standard procedure but also of its interpretation, guaranteeing

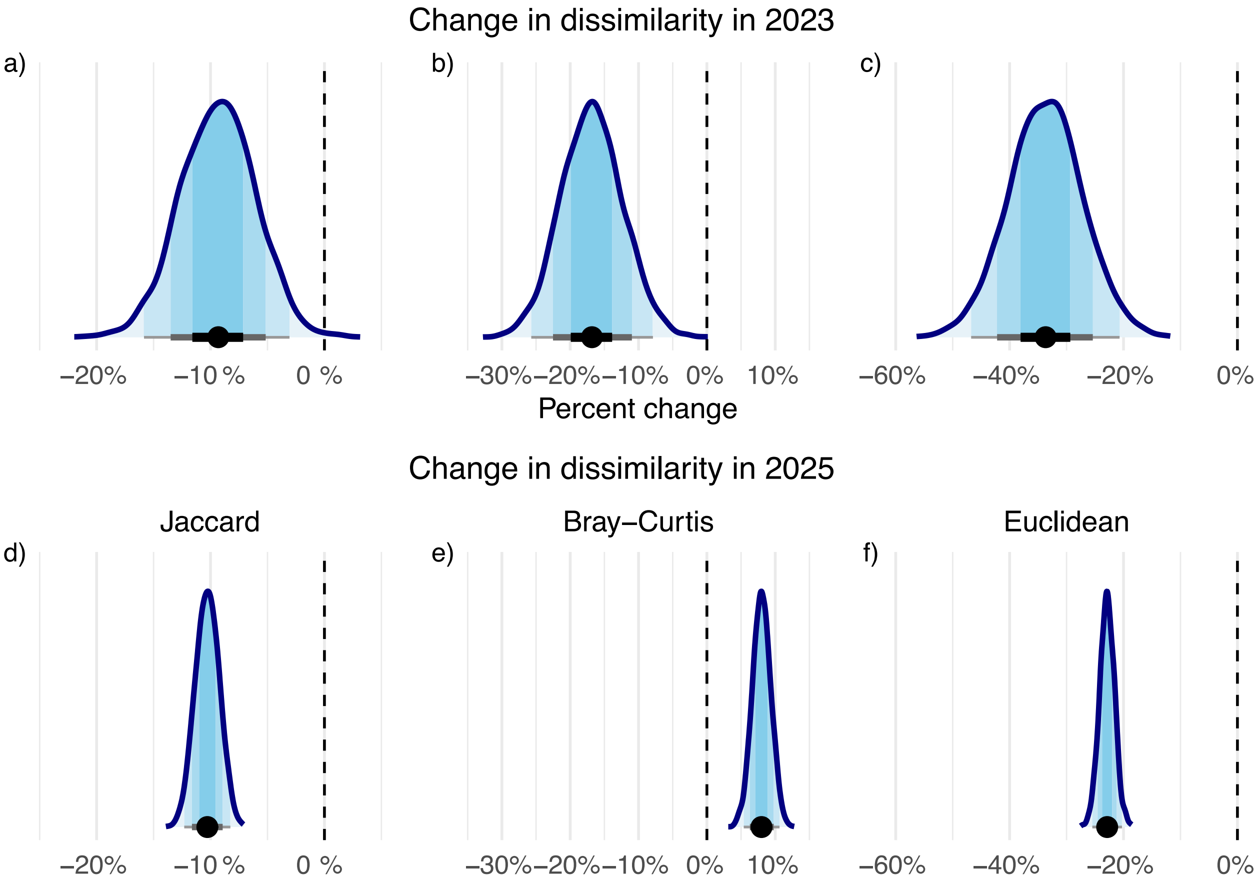

Figure 4. Percent lift of posterior predictions, representing the percent change in dissimilarity in exercise 2 relative to the exercise 1 baseline: (EX2 – EX1) / EX1 × 100. The colour bands and the intervals represent 50%, 80%, 95% widths, the point represents the median.

the achievement of programmes' goals and the expected relative team performances. This is in line with Morrison et al. (2021), who indicated that measures to avoid nonsampling error "can be taken, with or without the use of multiple observers at each site". These are effective to intercalibrate observers able to adhere to the same standard (e.g. on-field manual and reference species list), regardless of whether they remain the same throughout the survey period or whether they change.

Briefing sessions are a group exercise where observers share their survey experiences allowing them to avoid systematic errors (e.g., protocol misinterpretations) and harmonize judgement (despite personal professional background and knowledge). They consistently improve agreement in species detection, by reducing overlooking and misidentification errors and enhancing the reliability of species richness and presence-absence metrics. In contrast, observer variability in species cover estimation remains largely unresolved, despite the use of direct percentage cover values, indicating that abundance assessment is inherently more subjective and sensitive to observer-specific attributes. These findings highlight that standard training protocols are sufficient to improve detection-related data quality, but not to fully control variability in abundance estimates. Future large-scale vegetation monitoring programmes should therefore complement training with dedicated calibration tools for cover estimation to ensure robust assessments of vegetation change. In the absence of training on plant species cover, we suggest using binary presence/absence data for scientific purposes.

Acknowledgements

The authors thank the "Follonica Carabinieri Biodiversity Department" (https://rgpbio.it/reparto/follonica/) staff assigned to the "Marsiliana" State Nature Reserve and all the botanists who participated in the training and intercalibration exercises (LIFE project ModerNEC on-field manual https://lifemodernec.eu/documenti/Raccolta_manuali_operazio_i_di_campagna_Rete_NEC_Italia.pdf; IFNI 2025 project on-field manual https://zenodo.org/records/17022354).

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Artificial Intelligence (AI) use

Regarding the use of AI in the preparation of this manuscript, the authors declare the following: English proofing for the first draft, then thoroughly revised.

Funding

The New Italian Forest Inventory Project (IFNI2025) was carried out with funding from the Carabinieri Command of Forestry, Environmental and Agri-Food Unit (CUFAA).

Author contributions

Marco Cervellini: Investigation, Conceptualization, Data curation, Software, Visualization, Methodology, Writing – original draft, Writing – review and editing; Luciano Ludovico Maria De Benedictis: Conceptualization, Data curation, Software, Formal analysis, Visualization, Methodology, Writing – original draft, Writing – review and editing; Leonardo Salvatori: Investigation, Writing – review and editing; Stefano Chelli: Investigation, Writing – review and editing; Giandiego Campetella: Investigation, Supervision, Writing – review and editing; Federico Selvi: Investigation, Writing – review and editing; Giovanni Iacopetti: Investigation, Writing – review and editing; Arianna Ferrara: Writing – review and editing; Alessandro Chiarucci: Funding acquisition, Writing – review and editing; Antonella Canini: Funding acquisition, Writing – review and editing; Maura Francioni: Writing – review and editing; Chiara Scalet: Writing – review and editing; Roberto Canullo: Funding acquisition, Investigation, Supervision, Writing – review and editing.

Author ORCIDs

Marco Cervellini https://orcid.org/0000-0002-0853-2330
Luciano Ludovico Maria De Benedictis https://orcid.org/0009-0001-5014-4526
Leonardo Salvatori https://orcid.org/0009-0004-2522-9204
Stefano Chelli https://orcid.org/0000-0001-7184-8242
Giandiego Campetella https://orcid.org/0000-0001-6126-522X
Federico Selvi https://orcid.org/0000-0002-3820-125X
Giovanni Iacopetti https://orcid.org/0000-0002-1472-4435
Arianna Ferrara https://orcid.org/0000-0002-8178-3910
Alessandro Chiarucci https://orcid.org/0000-0003-1160-235X
Antonella Canini https://orcid.org/0000-0003-1132-8899
Maura Francioni https://orcid.org/0009-0005-6839-0548
Chiara Scalet https://orcid.org/0009-0001-3353-5795
Roberto Canullo https://orcid.org/0000-0002-9913-6981

Data availability

The data used in this study is available at https://doi.org/10.5281/zenodo.18710554. The code and notebook reporting the analyses are available at https://doi.org/10.5281/zenodo.19677911.

References

Allegrini M-C, Canullo R, Campetella G (2009) ICP-Forests (International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests): Quality Assurance procedure in plant diversity monitoring. Journal of Environmental Monitoring 11: 782. https://doi.org/10.1039/b818170p

Archaux F (2009) Could we obtain better estimates of plot species richness from multiple-observer plant censuses? Journal of Vegetation Science 20: 603–611. https://doi.org/10.1111/j.1654-1103.2009.01079.x

Boch S, Küchler H, Küchler M, Bedolla A, Ecker KT, … Bergamini A (2022) Observer-driven pseudoturnover in vegetation monitoring is context-dependent but does not affect ecological inference. Applied Vegetation Science 25: e12669. https://doi.org/10.1111/avsc.12669

Braun-Blanquet J (1964) Pflanzensoziologie. Springer Vienna, Vienna, XIV, 866 pp. https://doi.org/10.1007/978-3-7091-8110-2

Bruelheide H, Dengler J, Jiménez-Alfaro B, Purschke O, Hennekens SM, … Zverev A (2019) sPlot – A new tool for global vegetation analyses. Chiarucci A (Ed.). Journal of Vegetation Science 30: 161–186. https://doi.org/10.1111/jvs.12710

Bürkner P-C (2017) brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80: 1–28. https://doi.org/10.18637/jss.v080.i01

Canullo R, Allegrini M-C, Campetella G (2012) Reference field manual for vegetation surveys on the Conecofor LII network, italy (national programme of forest ecosystems control - Unece, Icp Forests). vol. 48. https://www.scienzadellavegetazione.it/wp-content/uploads/2023/07/NUMERO-48.pdf

Canullo R, Starlinger F, Granke O, Fischer R, Aamlid D, Dupouey J (2020) MANUAL on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. In: UNECE ICP Forests Programme Coordinating Centre (Ed.) Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. ICP Forests Expert Panel on Biodiversity and Ground Vegetation, 14.

Chytrý M, Hennekens SM, Jiménez-Alfaro B, Knollová I, Dengler J, … Yamalov S (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. Pärtel M (Ed.). Applied Vegetation Science 19: 173–180. https://doi.org/10.1111/avsc.12191

D'Amico G, Chirici G, Campetella G, Canullo R, Cervellini M, … Corona P (2025) National Forest Inventory in Italy: new perspectives for forest monitoring. Annals of Forest Science 82: 35. https://doi.org/10.1186/s13595-025-01303-9

Dengler J, Dembicz I (2023) Should we estimate plant cover in percent or on ordinal scales? Vegetation Classification and Survey 4: 131–138. https://doi.org/10.3897/VCS.98379

Dengler J, Jansen F, Glöckler F, Peet RK, De Cáceres M, … Spencer N (2011) The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. Journal of Vegetation Science 22: 582–597. https://doi.org/10.1111/j.1654-1103.2011.01265.x

Di Biase RM, Fattorini L, Marcelli A (2025) A design-based view of species richness estimation in environmental surveys. Journal of the Royal Statistical Society Series C: Applied Statistics 74: 884–903. https://doi.org/10.1093/jrsssc/qlaf005

Ferretti M (2021) New appetite for the monitoring of European forests. Annals of Forest Science 78: 94. https://doi.org/10.1007/s13595-021-01112-w

Ferretti M, König N, Granke O, Nicolas M (2021) Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. In: Thünen Institute of Forest Ecosystems, Eberswalde, Germany. https://www.icp-forests.net/fileadmin/icp_forests/Dateien/Manual_Versions/2020-22/ICP_Manual_part03_2021_QAQC_version_2021-1.pdf

Futschik A, Winkler M, Steinbauer K, Lamprecht A, Rumpf SB, … Pauli H (2020) Disentangling observer error and climate change effects in long-term monitoring of alpine plant species composition and cover. Bartha S (Ed.). Journal of Vegetation Science 31: 14–25. https://doi.org/10.1111/jvs.12822

Gasparini P, Bertani R, De Natale F, Di Cosmo L, Pompei E (2009) Quality control procedures in the Italian national forest inventory. Journal of Environmental Monitoring 11: 761. https://doi.org/10.1039/b818164k

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian Data Analysis. 3rd ed. Chapman and Hall/CRC, New York, 675 pp. https://doi.org/10.1201/b16018

Gelman A, Simpson D, Betancourt M (2017) The prior can often only be understood in the context of the likelihood. Entropy 19: 555. https://doi.org/10.3390/e19100555

Gilliam FS (2007) The ecological significance of the herbaceous layer in temperate forest ecosystems. BioScience 57: 845–858. https://doi.org/10.1641/B571007

Kennedy KA, Addison PA (1987) Some considerations for the use of visual estimates of plant cover in biomonitoring. The Journal of Ecology 75: 151. https://doi.org/10.2307/2260541

Kercher SM, Frieswyk CB, Zedler JB (2003) Effects of sampling teams and estimation methods on the assessment of plant cover. Journal of Vegetation Science 14: 899–906. https://doi.org/10.1111/j.1654-1103.2003.tb02223.x

Kéry M, Gregg KB (2003) Effects of life-state on detectability in a demographic study of the terrestrial orchid *Cleistes bifaria*. Journal of Ecology 91: 265–273. https://doi.org/10.1046/j.1365-2745.2003.00759.x

Klimeš L, Dančak M, Hájek M, Jongepierová I, Kučera T (2001) Scale-dependent biases in species counts in a grassland. Journal of Vegetation Science 12: 699–704. https://doi.org/10.2307/3236910

Knollová I, Chytrý M, Bruelheide H, Dullinger S, Jandt U, … Essl F (2024) ReSurveyEurope: A database of resurveyed vegetation plots in Europe. Journal of Vegetation Science 35: e13235. https://doi.org/10.1111/jvs.13235

Legendre P, Legendre L (2012) Numerical Ecology. 3rd Edition, Volume 24, 1006 pp. https://shop.elsevier.com/books/numerical-ecology/legendre/978-0-444-53868-0 [February 19, 2026]

Morrison LW (2016) Observer error in vegetation surveys: a review. Journal of Plant Ecology 9: 367–379. https://doi.org/10.1093/jpe/rtv077

Morrison LW (2021) Nonsampling error in vegetation surveys: understanding error types and recommendations for reducing their occurrence. Plant Ecology 222: 577–586. https://doi.org/10.1007/s11258-021-01125-5

Morrison LW, Leis SA, DeBacker MD (2020) Interobserver error in grassland vegetation surveys: sources and implications. Journal of Plant Ecology 13: 641–648. https://doi.org/10.1093/jpe/rtaa051

Morrison LW, Leis SA, Short MF, DeBacker MD (2024) A spatiotemporal comparison of interobserver error in vegetation sampling. Journal of Vegetation Science 35: e13286. https://doi.org/10.1111/jvs.13286

Nilsson IN, Nilsson SG (1985) Experimental estimates of census efficiency and pseudoturnover on islands: Error trend and between-observer variation when recording vascular plants. Journal of Ecology 73: 65–70. https://doi.org/10.2307/2259768

R Core Team (2026) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ricotta C (2021) From the euclidean distance to compositional dissimilarity: What is gained and what is lost. Acta Oecologica 111: 103732. https://doi.org/10.1016/j.actao.2021.103732

Rohrer JM, Arel-Bundock V (2025) Models as prediction machines: How to convert confusing coefficients into clear quantities. PsyArXiv, preprint. https://doi.org/10.31234/osf.io/g4s2a_v2

Salvatori L (2024) Testing observers of plant compositional diversity in forest monitoring. https://doi.org/10.5281/zenodo.18315130

Seidling W, Hamberg L, Máliš F, Salemaa M, Kutnar L, … Canullo R (2020) Comparing observer performance in vegetation records by efficiency graphs derived from rarefaction curves. Ecological Indicators 109: 105790. https://doi.org/10.1016/j.ecolind.2019.105790

Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C (2021) Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). Bayesian Analysis 16: 667–718. https://doi.org/10.1214/20-BA1221

Verheyen K, Bažány M, Chećko E, Chudomelová M, Closset-Kopp D, … Baeten L (2018) Observer and relocation errors matter in resurveys of historical vegetation plots. Rocchini D (Ed.). Journal of Vegetation Science 29: 812–823. https://doi.org/10.1111/jvs.12673

Vittoz P, Guisan A (2007) How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. Journal of Vegetation Science 18: 413–422. https://doi.org/10.1111/j.1654-1103.2007.tb02553.x

Supplementary material 1

Examples of the graphical outputs used during the briefing session

Authors: Marco Cervellini, Luciano Ludovico Maria De Benedictis, Leonardo Salvatori, Stefano Chelli, Giandiego Campetella, Federico Selvi, Giovanni Iacopetti, Arianna Ferrara, Alessandro Chiarucci, Antonella Canini, Maura Francioni, Chiara Scalet, Roberto Canullo

Data type: pdf

Explanation note: **figure S1:** Plot displaying species richness for each observer and sampling unit; **figure S2:** Plot displaying the cover values for three species (minimum, median and maximum average cover) in a sampling unit; **figure S3:** Plot displaying herb layer cover for each observer and sampling unit.

Link: https://doi.org/10.3897/ved.189819.suppl1