UNIVERSITÀ DI SIENA 1240

USiena

Observer-expertise bias in multiscale forest vegetation surveys: effects on species detection, richness and composition across spatial grains and forest contexts

This is the peer reviewed version of the following article:

Original:

Bacaro, G., Tordoni, E., Maccherini, S., Castello, M., Chiarucci, A. (2026). Observer-expertise bias in multiscale forest vegetation surveys: effects on species detection, richness and composition across spatial grains and forest contexts. PLANT DIVERSITY [10.1016/j.pld.2026.07.019].

Availability:

This version is availablehttp://hdl.handle.net/11365/1331654 since 2026-11-10T13:47:58Z

Published:

DOI: http://doi.org/10.1016/j.pld.2026.07.019

Terms of use:

Open Access

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. Works made available under a Creative Commons license can be used according to the terms and conditions of said license.
For all terms of use and more information see the publisher's website.

(Article begins on next page)

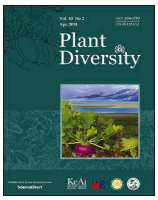
26 November 2026

Plant Diversity xxx (xxxx) xxx

Contents lists available at ScienceDirect

Plant Diversity

journal homepage: http://www.keaipublishing.com/en/journals/plant-diversity/
http://journal.kib.ac.cn



Research paper

Observer-expertise bias in multiscale forest vegetation surveys: effects on species detection, richness and composition across spatial grains and forest contexts

Giovanni Bacaro [a,*], Enrico Tordoni [b], Simona Maccherini [c], Miris Castello [a], Alessandro Chiarucci [b]

[a] Department of Life Sciences, University of Trieste, Via L. Giorgieri 10, 34127 Trieste, Italy
[b] Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, Via Irnerio 42, 40126 Bologna, Italy
[c] Department of Life Sciences, University of Siena, Via P.A. Mattioli 4, 53100 Siena, Italy

ARTICLE INFO

Article history:
Received 30 April 2026
Received in revised form
12 July 2026
Accepted 17 July 2026
Available online xxx

Keywords:
Forest understory
Observer-expertise bias
Nonsampling error
Permanent plots
Pseudoturnover
Vegetation monitoring

ABSTRACT

Observer-related uncertainty can affect vegetation surveys, but its magnitude and ecological consequences remain poorly quantified in forest understories. We assessed observer-expertise-related variation using independent presence–absence records collected by three observer categories differing in botanical experience, calibration history and familiarity with nested permanent plots in six Tuscan forest stands. The design combined three spatial grains (1, 10 and 100 m^2), replicated stands within forest type and repeated observations of identical sampling units. We quantified effects on pooled taxon records, stratum-specific floristic record richness, three-observer concordance, pairwise pseudoturnover and species composition. Differences among observer-expertise categories were measurable but strongly context dependent. Effects on record-based richness were statistically significant in only two of eighteen site × grain combinations, indicating limited richness effects rather than a general expert–novice gradient. By contrast, concordance and pseudoturnover revealed broader disagreement in species detection: agreement was highest in beech forests and lower in species-rich, structurally complex oak stands. Plot size altered disagreement, but the direction and magnitude of this effect depended on forest context. PERMANOVA showed that site and plot size explained most compositional variation, whereas observer-expertise-related compositional differences were restricted to oak forests and were primarily site dependent. Thus, observer-expertise variation did not override the main ecological gradients, but it affected field outputs, reduced concordance, inflated apparent turnover and altered some compositional comparisons. Because each expertise category was represented by a single observer team, these effects should be interpreted as observer-expertise-related variation rather than as bias among observers with comparable expertise.

© 2026 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Vegetation surveys remain a central empirical foundation of plant ecology, vegetation science and biodiversity monitoring. Permanent plots, resurveys and repeated community inventories are widely used to infer ecological change, compare habitats and evaluate management outcomes. These applications assume that the observation process is sufficiently consistent for differences among surveys to be interpreted mainly as ecological signal. Yet species lists and abundance estimates are produced by human observers, and therefore include an observational component that may affect the response variables most often analysed in vegetation science, including species richness, i.e. the number of species recorded in an assemblage (Chao and Chiu, 2016), species composition, turnover and cover estimates (Lepš and Hadincová, 1992; Archaux et al., 2006; Vittoz and Guisan, 2007; Milberg et al., 2008; Morrison, 2016, 2021; Morrison et al., 2023).

* Corresponding author.
E-mail address: gbacaro@units.it (G. Bacaro).
Peer review under the responsibility of Editorial Office of Plant Diversity.

https://doi.org/10.1016/j.pld.2026.07.019
2468-2659/© 2026 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article as: G. Bacaro, E. Tordoni, S. Maccherini *et al.*, Observer-expertise bias in multiscale forest vegetation surveys: effects on species detection, richness and composition across spatial grains and forest contexts, Plant Diversity, https://doi.org/10.1016/j.pld.2026.07.019

Observer-related error is now recognised as a persistent component of uncertainty in vegetation data rather than a marginal inconvenience (Ahrends et al., 2011). Morrison (2021) framed such errors within nonsampling error, distinguishing overlooking, misidentification and estimation error, while earlier and subsequent studies showed that disagreement among observers can affect species detection, abundance estimates, richness and apparent turnover across a range of habitats and protocols (Scott and Hallam, 2003; Vittoz and Guisan, 2007; Milberg et al., 2008; Morrison, 2016; Morrison et al., 2020, 2023, 2024; Seidling et al., 2020; Boch et al., 2022). This is particularly relevant for long-term monitoring, where unmeasured observer effects may be mistaken for ecological change when surveys are compared across years, teams or institutions (Kapfer et al., 2017; Verheyen et al., 2018; Douda et al., 2023). In the present study, however, we focus on one specific component of this broader problem: variation among observer teams differing in botanical expertise, previous calibration and familiarity with the permanent plots. We therefore use "observer bias" when referring to the general literature, but refer to our empirical contrasts as observer-expertise-related variation or observer-expertise effects.

One of the clearest manifestations of observer-related uncertainty is pseudoturnover, namely apparent species turnover generated by differences among observers rather than by true ecological change (Morrison, 2017). Observer error may also be substantial in focused population surveys, including surveys of rare plant species, further showing that plant records should not be assumed to be error-free simply because plants are sessile organisms (Morrison and Young, 2016). This is particularly problematic in resurveys and monitoring programmes, where species gains and losses are often interpreted as directional change. If observers differ in their ability to detect inconspicuous taxa, distinguish similar species or apply inclusion rules consistently, part of the observed turnover is artefactual. Richness and diversity metrics may appear more stable, because they compress species identities into summary values, yet they can still mask substantial disagreement in species detection among observers (Morrison, 2016; Morrison et al., 2023, 2025).

Observer-related error is also strongly context dependent. Species richness, census duration, habitat type, seasonality, location, plot geometry and scale can all influence survey completeness and observer performance (Archaux et al., 2006; Vittoz and Guisan, 2007; Kéry and Schmidt, 2008; Burg et al., 2015; Boch et al., 2022; Morrison et al., 2024). Observer bias therefore cannot be reduced to a single correction term, but is better understood as the outcome of interactions among observer attributes, vegetation structure, species detectability and the practical rules applied during field recording (Lin et al., 2020; Moreno-Fernández et al., 2024).

Spatial grain is a key component of this context dependence (Portier et al., 2022). Plot size changes both the ecological information available and the practical task faced by the observer. Small plots constrain the search area but emphasize fine-scale detection and plot-border decisions, whereas larger plots include more species and microhabitats but also increase search area, within-plot heterogeneity and the probability of inconsistent recording. The effect of plot size on observer disagreement is therefore unlikely to be monotonic or universal. In nested or multiscale sampling designs, this matters because the relationship between spatial grain and diversity may itself be affected by imperfect detection and observer-related variation (Archaux et al., 2006; Vittoz and Guisan, 2007; Seidling et al., 2020; Perret et al., 2023; Martínez-Villegas et al., 2024).

Observer expertise is especially relevant to the present study. More experienced botanists are expected to produce more complete and internally consistent species lists than less experienced observers, but expertise is not a single attribute. It combines taxonomic competence, field experience, familiarity with local floras, previous exposure to the same plots and calibration with other surveyors. Studies comparing observers with different levels of experience or training show that expertise can influence survey completeness and species detection, but that its effect depends on habitat complexity, survey protocol and field conditions (Archaux et al., 2006; Burg et al., 2015; Morrison, 2016, 2021). At the same time, differences among experienced observers may remain substantial, indicating that expertise reduces but does not eliminate observer-related error (Lepš and Hadincová, 1992; Vittoz and Guisan, 2007; Milberg et al., 2008; Boch et al., 2022; Morrison et al., 2024). For this reason, the relevant question is not whether experts are "correct" and others are "biased", but how variation associated with expertise, calibration and site familiarity interacts with vegetation context and spatial grain.

These issues are particularly consequential in forest understory surveys, where small-scale heterogeneity, seasonal detectability, variable light environments, litter cover and patchy species distributions can make species detection difficult (Kirby et al., 1986). More generally, the same problem is relevant whenever recorded species lists are used as evidence of biodiversity status, because non-detection, unrecorded species pools and uncertainty in the completeness of biodiversity knowledge may affect ecological interpretation and environmental assessment (Bacaro, 2025; Bacaro and Fonda, 2026). In such settings, observer disagreement may arise from taxonomic uncertainty, unequal familiarity with the local flora, different search strategies and inconsistent plot-border decisions. Explicit observer-error studies are therefore useful because they help separate ecological variation from variation introduced by the observation process itself (Archaux et al., 2006; Seidling et al., 2020; Douda et al., 2023).

The present study addresses this problem using a field design that combines repeated independent observations, nested spatial grains and contrasting forest contexts. The same permanent plots were surveyed by observer teams representing different levels of botanical expertise, calibration history and prior plot familiarity in six forest stands, with vegetation recorded at 1, 10 and 100 m^2. This design allows us to ask three questions. First, how strongly do observer-expertise categories affect recorded species richness and agreement in species detection? Second, how does pseudoturnover associated with observer-expertise categories vary with plot size and forest context? Third, to what extent does observer-expertise-related variation alter multivariate inference on community composition relative to the larger ecological effects of site and forest type?

We expected observer-expertise-related variation to be detectable, structured and context dependent, but not to dominate the main ecological gradients in the dataset. Specifically, we expected disagreement to increase in richer and more heterogeneous stands, where species diversity and structural complexity create more opportunities for omission and inconsistent detection. We also expected plot size to influence disagreement in a non-linear way, because increasing grain both integrates more of the local flora and increases the complexity of the survey task. Finally, we expected teams with greater botanical experience, previous calibration and site familiarity to produce more convergent records, although not as a uniform main effect across all forest contexts and spatial grains.

2. Materials and methods

2.1. Study system and permanent plot network

This study used data from the Tuscan forest monitoring network MON.I.TO. (MONitoraggio Intensivo foreste Toscane), a

long-term programme established in 1995 to support repeated ecological assessment of representative forest stands across the region (Chiarucci et al., 2001). We analysed six long-term sampling sites distributed across Tuscany and spanning the main regional forest types along a broad ecological gradient from Mediterranean coastal stands to montane Apennine forests. Two sites were located in holm oak forests dominated by *Quercus ilex* L. (Cala Violina, CV; Colognole, CG), two in Turkey oak forests dominated by *Quercus cerris* L. (Amiata, AM; Ulignano, UG), and two in beech forests dominated by *Fagus sylvatica* L. (Foresta del Teso, FT; Vallombrosa, VO) (Fig. 1). The network was originally designed to support repeated biodiversity measurements across standardised permanent plots and multiple spatial grains, making it well suited to the analysis of observer-related variation in species detection and composition (Chiarucci et al., 2001; Chiarucci and Bonini, 2005).

At each site, vegetation was surveyed in ten permanently marked sampling plots. Each plot comprised three nested subplots (1, 10 and 100 m^2), yielding a replicated multiscale design within each forest stand (Fig. 1c). Although this design was developed before later formalised multiscale vegetation protocols, it is consistent with their general logic for nested richness sampling (Chiarucci et al., 2001; Dengler, 2009). The dataset therefore combines three features that are critical for quantifying observer-expertise-related variation: repeated observations of the same permanent plots, replicated sites within forest type, and explicit grain dependence in the observation process.

2.2. Observer categories and survey protocol

Field surveys were conducted in spring and summer 2008. Each site was surveyed independently by three single-observer teams representing contrasting levels of botanical experience and prior familiarity with the permanent plots. To avoid an *ex post* classification of ability based on the results themselves, observer teams were assigned *a priori* to three experience categories. The expert team (E) consisted of botanists with strong taxonomic competence and previous direct experience with the monitored sites and their permanent plots. The trained team (T) consisted of botanists who had surveyed the same plots in the previous year, following direct calibration with the expert team. The inexpert team (I) consisted of botanically competent early-career researchers with general field experience but no prior direct experience with the permanent plots before the 2008 campaign. This classification follows the general logic adopted in later work on observer performance in vegetation surveys, in which field experience, site familiarity and previous calibration are treated as distinct but complementary components of observer expertise (Burg et al., 2015). Because each expertise category was represented by a single observer team, our design does not allow observer identity to be separated from expertise category, prior calibration or site familiarity. The differences quantified below should therefore be interpreted as variation among predefined observer-expertise categories, rather than as a full estimate of observer bias among surveyors with comparable levels of expertise.

The three observer teams surveyed the same plots independently on the same day or within a narrow sampling window at each site, so that differences among teams could not be attributed to temporal changes in vegetation state. A common protocol was agreed upon before fieldwork, and teams were not allowed to inspect each other's records during sampling. Observers did not receive complete species lists from previous years, so each inventory had to be reconstructed solely from direct field observation. Teams approached plots independently and from different positions to minimise visual cueing and path dependence, and trampling within plots was reduced as much as possible, especially

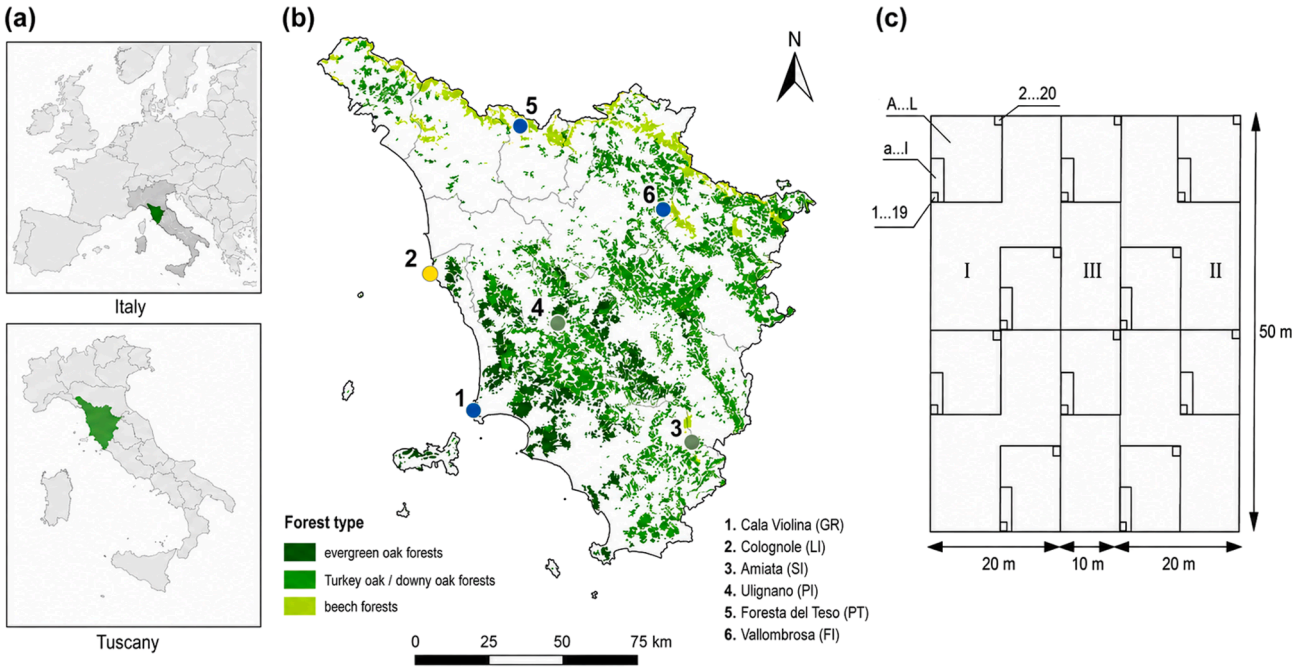


Fig. 1. Geographic location of the study area, distribution of the permanent forest monitoring sites, and nested multiscale sampling design. **(a)** Location of Tuscany within Europe and Italy. **(b)** Distribution of the six MON.I.TO. permanent monitoring sites across Tuscany, shown against the main forest types represented in the study: evergreen oak forests, Turkey oak/downy oak forests, and beech forests. Site numbers indicate: 1, Cala Violina (GR); 2, Colognole (LI); 3, Amiata (SI); 4, Ulignano (PI); 5, Foresta del Teso (PT); and 6, Vallombrosa (FI). **(c)** Nested sampling design adopted within each permanent plot for multiscale vegetation surveys. Each site consisted of a 2500 m^2 permanent plot (50 × 50 m), including two 1000 m^2 subplots (I and II), one 500 m^2 subplot (III), twelve 10 × 10 m subplots (A–L), twelve 2 × 5 m subplots (a–l), and twenty 1 × 1 m subplots (1–20), allowing species detection and community patterns to be assessed across multiple spatial grains.

at the smallest grain. No fixed time limit was imposed for plot completion; each team ended the survey only when it considered the inventory complete. No samples were taken from the plots. Doubtful taxa were checked only after they had been independently recorded in the field notebook, so that taxonomic support did not affect the original decision to include or exclude a species from the census. The final analytical dataset comprised 540 observer-specific plot-level inventories, corresponding to six sites × ten permanent plots per site × three spatial grains (1, 10 and 100 m^2) × three observer-expertise categories. Thus, the design included 180 plot × grain units, each independently surveyed by the three observer teams. Within each forest type, this corresponded to 60 observer-specific inventories per spatial grain, given by two sites × ten plots per site × three observer teams.

2.3. Vegetation recording and data structure

For each plot, all detected vascular plant species were recorded. Species occurrence was available at all three grains, whereas visual cover estimates were recorded only for the 100 m^2 plots using a continuous percentage scale. This reflects the original MON.I.TO. field protocol, in which species occurrence was the common currency across scales, while quantitative cover was restricted to the largest grain to maintain feasible field effort and comparability among repeated surveys (Chiarucci et al., 2001; Chiarucci and Bonini, 2005).

During field recording, taxa were distinguished as rooted species or covering-only species. The latter category included species whose above-ground canopy intercepted the vertical projection of the plot but whose rooting system lay outside the plot border. This distinction is relevant in small forest plots, where the treatment of plot-border species can materially affect richness estimates, agreement among observers and apparent turnover. The original field notebooks and archived analytical matrices preserve both the full species lists and, in separate files, the split between rooted and covering-only records. In the analyses presented here, observer-expertise-related variation was assessed using complete survey records, whereas the rooted/covering distinction was retained only as a diagnostic aid for interpreting disagreement in species detection.

After fieldwork, species names were harmonised across notebooks according to the nomenclature adopted for the whole dataset, which followed Pignatti (1982). Orthographic variants, duplicate spellings and equivalent names were standardised before analysis to ensure that the study captured observer-related differences in field detection rather than artefacts arising from inconsistent nomenclature.

2.4. Response variables

Observer-expertise-related variation was quantified through complementary univariate and multivariate summaries derived from the original field notebooks and the archived plot-by-species matrices. First, for each site × plot-size combination, we calculated the pooled set of taxa recorded by the three observer teams. Although this does not represent “true richness” in a strict detection-theory sense, it provides a pragmatic upper bound for direct comparison with observer-specific field records. Second, we calculated observer-specific stratum-specific floristic record richness for each plot. This was defined as the number of vascular-plant field records entered by each observer team within a given plot and spatial grain, retaining the original field distinction between rooted and covering-only records. Consequently, this variable reflects the number of field records generated by the protocol rather than a fully deduplicated count of taxonomic species whenever the same taxon was recorded in more than one stratum or record category. These plot-level values were summarised by site, spatial grain and observer-expertise category as means and standard deviations across the ten permanent plots. This record-based response is appropriate for evaluating observer-expertise effects on the actual field output produced by the survey protocol, including differences arising from taxonomic detection, omission and inclusion decisions at plot boundaries. Third, we quantified pairwise pseudoturnover between observer teams. Pseudoturnover expresses the proportion of apparent compositional change that arises solely because two observers produce different species lists for the same plot, even though no ecological change has occurred. For each observer pair and plot, pseudoturnover (PT) was calculated following Nilsson and Nilsson (1983):

$$PT = ((b + c) / (2a + b + c)) \times 100$$

where *a* is the number of species shared by the two observers, and *b* and *c* are the species recorded exclusively by the first and second observer, respectively. PT equals 0 when two observers produce identical species lists and approaches 100 as the inventories become completely different. PT was calculated for all three observer pairs within each site and grain: expert versus trained (E–T), expert versus inexpert (E–I), and trained versus inexpert (T–I). Fourth, we summarised concordance in species detection directly from the observer notebooks. Three-observer concordance was used as a descriptive complement to pseudoturnover and was expressed as the percentage of records jointly detected by all three observer-expertise teams within each site × plot-size combination. This index was calculated for the complete set of field records and, separately, for covering-only records. Pairwise agreement among observer categories was also summarised descriptively to support interpretation of observer-pair differences. Finally, to test whether observer-expertise category affected not only richness but also the multivariate structure of the vegetation records, we assembled species composition matrices from the original plot data and analysed them separately for each forest type.

2.5. Statistical analyses

2.5.1. Univariate analyses of species richness and pseudoturnover

Univariate analyses followed the factorial structure of the sampling design. For observer-specific stratum-specific floristic record richness, we tested whether observer-expertise category affected the number of field records produced by each survey within each site × plot-size combination. Because this response variable is a count, observer-expertise effects were analysed using Poisson generalized linear mixed models. Models were fitted separately for each site × plot-size combination, with observer-expertise category as a fixed effect and plot identity as a random intercept to account for repeated observations of the same permanent plots by the three observer-expertise teams. Model overdispersion was checked, and the Poisson specification was retained where no relevant overdispersion was detected. Observer-expertise effects were assessed using likelihood-ratio χ^2 tests with df = 2. When the observer-expertise effect was significant, post-hoc pairwise contrasts among expert, trained and inexpert teams were used to identify the observer categories responsible for the difference. The full count-model results and significant post-hoc contrasts are reported in Table S1.

Pseudoturnover was analysed separately because it is a continuous percentage response calculated for observer pairs rather than a count recorded by individual observer teams. Plot-level PT values were computed for each observer pair, spatial grain and site. Observer-pair effects were then tested within each

site × plot-size combination using linear mixed models, with observer-pair category as a fixed effect and plot identity as a random intercept to account for repeated observer-pair comparisons on the same permanent plots. Observer-pair effects were assessed using likelihood-ratio χ^2 tests with df = 2. When significant effects were detected, post-hoc pairwise contrasts among the three observer-pair categories were used to identify the contrasts responsible for the difference. Full model results and significant post-hoc contrasts are reported in Table S2.

As a descriptive complement to these analyses, we calculated sample-based species accumulation curves separately for each forest type × plot-size × observer-expertise category combination. These curves were used to visualise how accumulated taxon richness changed with sampling effort and to compare the relative behaviour of expert, trained and inexpert observer teams across ecological contexts and spatial grains. Rarefaction analyses were carried out in R using the *Rarefy* package (Thouverai et al., 2026a, 2026b).

2.5.2. Multivariate analyses of species composition

To assess whether observer-related differences extended beyond richness to the full composition of the recorded communities, we performed permutational multivariate analysis of variance (PERMANOVA; Anderson, 2001) on presence/absence plot-by-species matrices assembled from the field records using PRIMER v6 with the PERMANOVA+ add-on (Clarke and Gorley, 2006; Anderson et al., 2008).

Because compositional differences among the three dominant forest types were expected to be much larger than the within-type observer effect, and because each forest type was represented by two sites, multivariate analyses were conducted separately for beech, Turkey oak and holm oak forests. For each forest type, Jaccard dissimilarities were calculated among vegetation records. Analyses were based on the species matrices archived for the original study and structured according to three fixed factors: site (ID), plot size (CA) and observer category (TIP_R), plus their interactions. Significance was assessed by permutation using 999 permutations. Pairwise comparisons and permutation tests were run within the PERMANOVA+ routine implemented in PRIMER v6. In addition to testing the significance of individual terms, we quantified the relative contribution of each model component to the total explained variation in order to compare variance structures among forest types. This framework partitions multivariate variation in species composition into fractions attributable to site context, spatial grain, observer category, and their joint effects, thereby distinguishing the ecological signal of habitat and scale from the observational signal introduced by different survey teams. When the Site × Observer interaction was significant, pairwise permutational post hoc comparisons were carried out within each site to identify which observer contrasts contributed to the interaction.

3. Results

3.1. Pooled richness and stratum-specific floristic record richness across sites and spatial grains

Pooled taxon richness showed a clear ecological structure, with strong differences among sites and a consistent increase with plot size (Table 1). Across spatial grains, Turkey oak forests produced the largest number of records, beech forests the smallest, and holm oak forests intermediate values, although differences between sites within the same forest type were also evident. The same broad pattern was observed for observer-specific stratum-specific floristic record richness (Table 2), indicating that site context and spatial grain were the dominant sources of variation in the field records produced by the survey protocol.

Table 1
Pooled species richness at the three spatial grains in the six permanent forest sites. Values represent the union of all taxa recorded by the three observer teams within each site × plot-size combination. Site codes: CV, Cala Violina; CG, Colognole; AM, Amiata; UG, Ulignano; FT, Foresta del Teso; VO, Vallombrosa. Forest types: HOL, Holm oak forests; TUR, Turkey oak forests; BEE, Beech forests.

Site	Type	1 m^2	10 m^2	100 m^2
CV	HOL	16	19	23
CG	HOL	25	29	43
AM	TUR	56	67	85
UG	TUR	51	61	72
FT	BEE	12	12	25
VO	BEE	8	9	14

Full univariate test results for the record-based richness response are reported in Table S1, while Table 2 reports the corresponding site- and grain-specific means. Differences among observer-expertise categories were detectable, but they were spatially restricted rather than pervasive across the dataset. Significant observer-expertise effects on record-based richness were detected only in two site × grain combinations. In Colognole at 1 m^2, the trained team recorded fewer field records than the expert team, while the inexpert team showed an intermediate value that did not differ significantly from either group. In Ulignano at 100 m^2, the expert team recorded more field records than both the trained and inexpert teams. In Colognole, the same descriptive tendency for lower values in the trained team was also observed at 10 and 100 m^2, but these contrasts did not reach conventional significance in the count-based models. No significant observer-expertise effect on record-based richness was detected in the remaining site × grain combinations. Thus, observer-expertise-related differences in field-record output were statistically

Table 2
Mean stratum-specific floristic record richness (±SD) recorded by the three observer-expertise categories at the three spatial grains in the six permanent forest sites. Values are means across the ten permanent plots within each site × plot-size combination.

Site	Type	Plot size	Expert (E)	Trained (T)	Inexpert (I)
CV	HOL	1 m^2	6.7 ± 1.76a	5.1 ± 1.59a	6.3 ± 1.25a
CV	HOL	10 m^2	10.9 ± 2.02a	9.2 ± 2.44a	10.2 ± 2.61a
CV	HOL	100 m^2	17.2 ± 3.55a	15.3 ± 3.23a	17.2 ± 2.25a
CG	HOL	1 m^2	11.7 ± 2.00a	8.0 ± 1.69b	10.5 ± 2.06 ab
CG	HOL	10 m^2	16.1 ± 2.80a	13.6 ± 2.63a	17.0 ± 2.70a
CG	HOL	100 m^2	26.4 ± 3.89a	22.0 ± 3.29a	26.9 ± 3.31a
AM	TUR	1 m^2	12.8 ± 3.67a	12.6 ± 2.67a	13.9 ± 2.96a
AM	TUR	10 m^2	25.0 ± 5.16a	25.7 ± 6.27a	27.7 ± 5.37a
AM	TUR	100 m^2	47.3 ± 7.93a	49.8 ± 4.30a	50.6 ± 3.47a
UG	TUR	1 m^2	16.7 ± 3.30a	15.3 ± 2.21a	15.4 ± 3.65a
UG	TUR	10 m^2	30.8 ± 4.34a	29.1 ± 2.92a	29.2 ± 3.67a
UG	TUR	100 m^2	55.5 ± 5.03a	47.7 ± 4.05b	46.6 ± 3.97b
FT	BEE	1 m^2	3.2 ± 1.03a	3.1 ± 0.56a	3.3 ± 0.82a
FT	BEE	10 m^2	5.6 ± 1.42a	6.0 ± 1.63a	5.9 ± 1.59a
FT	BEE	100 m^2	10.4 ± 2.31a	10.2 ± 2.65a	10.5 ± 1.95a
VO	BEE	1 m^2	3.9 ± 1.10a	4.2 ± 1.35a	4.6 ± 1.34a
VO	BEE	10 m^2	6.4 ± 0.90a	6.4 ± 1.57a	6.7 ± 1.41a
VO	BEE	100 m^2	8.3 ± 0.94a	8.4 ± 1.42a	8.4 ± 1.77a

The response variable represents the number of vascular-plant field records entered by each observer team, retaining the original distinction between rooted and covering-only records; it is therefore a record-based measure of plot-level floristic richness rather than a fully deduplicated count of taxonomic species. Different lowercase letters indicate significant post-hoc differences among observer-expertise categories within the same site × plot-size combination, based on the count-model contrasts reported in Table S1. Full univariate test results are reported in Table S1. Site codes: CV, Cala Violina; CG, Colognole; AM, Amiata; UG, Ulignano; FT, Foresta del Teso; VO, Vallombrosa. Forest types: HOL, holm oak forests; TUR, Turkey oak forests; BEE, beech forests.

limited, strongly context dependent, and did not follow a simple or uniform expert-versus-novice gradient.

3.2. Three-observer concordance and pairwise pseudoturnover

Full univariate test results for pseudoturnover are reported in Table S2, while Table 3 provides the corresponding descriptive means and standard deviations for each observer pair. Three-observer concordance in species detection and pairwise pseudoturnover revealed stronger observer-expertise-related differences than record-based richness alone (Fig. 2; Table 3). Concordance generally increased with plot size, while pseudoturnover tended to decline from 1 to 100 m^2, but these patterns were not uniform across forest contexts. Pairwise agreement among observer categories showed the same broad site- and grain-dependent structure, with higher agreement in the beech forests and more variable agreement in the oak-dominated stands (Fig. 3).

Agreement was highest and pseudoturnover lowest in the beech forests, especially at the largest grain, whereas the oak-dominated forests showed lower agreement and more persistent observer-related turnover. This indicates that larger plots generally improved shared detection, but did not fully remove disagreement in richer and structurally more complex stands.

Pseudoturnover was particularly informative because it showed that similar richness values could still correspond to different species lists. At the smallest grain, all forest types showed appreciable apparent turnover among observer teams, reflecting the sensitivity of fine-scale plots to omissions, local patchiness and plot-border decisions. With increasing grain, pseudoturnover decreased in most cases, but some oak-dominated stands retained non-negligible observer-related turnover even at 100 m^2. Agreement and pseudoturnover therefore support the same main interpretation: observer-expertise-related variation was strongest where vegetation complexity made species detection and inclusion decisions more difficult, and weakest in the simpler beech stands. When records were restricted to covering-only species, concordance patterns became much less regular than in the full field records (Fig. 2b). These values were not used as the main basis for statistical interpretation, but they support the conclusion that disagreement increases when field decisions depend on ambiguous plot-boundary inclusion.

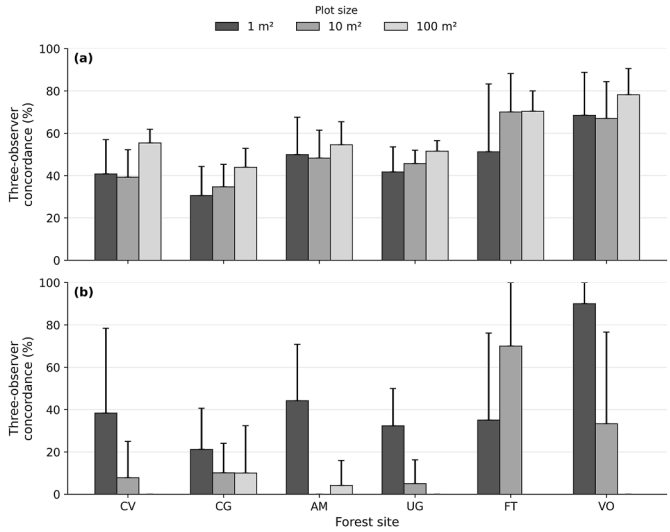


Fig. 2. Three-observer concordance in field records across forest sites and spatial grains. Concordance is expressed as the percentage of records jointly detected by the three observer-expertise teams within each site × plot-size combination. **(a)** Concordance based on the complete set of field records. **(b)** Concordance based only on covering records, i.e. taxa whose above-ground parts intercepted the plot but whose rooting system was outside the plot boundary. Bars show mean concordance values across plots, and upper error bars indicate the corresponding standard deviation. Site codes are: CV, Cala Violina; CG, Colognole; AM, Amiata; UG, Ulignano; FT, Foresta del Teso; and VO, Vallombrosa.

Table 3
Pseudoturnover (%) between observer pairs at the three spatial grains in the six permanent forest sites, expressed as mean ± SD across the ten permanent plots within each site × plot-size combination.

Plot size	Site	Type	E–T	T–I	E–I
1 m^2	CV	HOL	13.58 ± 8.95	23.90 ± 15.29	19.87 ± 13.42
1 m^2	CG	HOL	17.22 ± 12.96	26.99 ± 24.00	16.70 ± 11.10
1 m^2	AM	TUR	22.69 ± 16.38	20.13 ± 14.07	17.35 ± 11.41
1 m^2	UG	TUR	12.33 ± 20.25	23.79 ± 32.84	9.00 ± 11.74
1 m^2	FT	BEE	14.35 ± 4.58	11.96 ± 4.99	19.31 ± 9.13
1 m^2	VO	BEE	18.90 ± 17.90	15.01 ± 14.25	17.07 ± 16.94
10 m^2	CV	HOL	12.08 ± 7.77	11.61 ± 8.88	10.01 ± 6.59
10 m^2	CG	HOL	10.32 ± 7.19	11.46 ± 11.80	10.58 ± 6.48
10 m^2	AM	TUR	19.19 ± 9.30	18.98 ± 5.68	16.19 ± 7.19
10 m^2	UG	TUR	10.55 ± 5.77	11.88 ± 4.00	16.26 ± 4.59
10 m^2	FT	BEE	6.64 ± 7.33	9.77 ± 9.24	7.30 ± 6.90
10 m^2	VO	BEE	7.21 ± 8.69	7.91 ± 7.33	8.77 ± 6.86
100 m^2	CV	HOL	10.68 ± 6.66	12.06 ± 6.66	12.06 ± 5.57
100 m^2	CG	HOL	9.80 ± 4.92	10.07 ± 3.20	12.47 ± 4.38
100 m^2	AM	TUR	14.84 ± 5.70	14.88 ± 6.65	11.91 ± 4.60
100 m^2	UG	TUR	10.69 ± 3.65	9.19 ± 3.14	13.36 ± 3.90
100 m^2	FT	BEE	11.73 ± 4.86	13.76 ± 7.58	10.42 ± 4.44
100 m^2	VO	BEE	4.31 ± 2.92	8.17 ± 5.92	7.37 ± 4.41

Observer pairs are E–T, expert versus trained; T–I, trained versus inexpert; and E–I, expert versus inexpert. The table is intended as a descriptive summary of the magnitude of apparent turnover among observer pairs; statistical tests for pseudoturnover are reported separately in Table S2. Site codes: CV, Cala Violina; CG, Colognole; AM, Amiata; UG, Ulignano; FT, Foresta del Teso; VO, Vallombrosa. Forest types: HOL, holm oak forests; TUR, Turkey oak forests; BEE, beech forests.

3.3. Observer-expertise effects on species composition

PERMANOVA analyses showed that floristic composition was dominated by ecological variation among sites, while observer-expertise-related effects were restricted to oak-dominated forests and were primarily site-dependent rather than scale-dependent (Tables 4 and 5; Fig. 4). In all three forest types, Site and Plot size were significant sources of multivariate variation, and the Site × Plot size interaction was also significant, confirming that the main organisation of the data remained ecological rather than observational.

In beech forests (FT and VO), observer-expertise category was not significant as a main effect, nor were the Site × Observer and Plot size × Observer interactions significant. By contrast, Site was strongly significant, as were Plot size and Site × Plot size (Table 4). The variance partitioning shown in Fig. 4 was consistent with this pattern: Site explained 25.9% of the total variation, Plot size 22.0%, and Site × Plot size 13.7%, whereas observer-related components were negligible; residual variation accounted for 38.5%. In this subset, compositional structure was almost entirely controlled by ecological differences between sites and by variation across spatial grains.

Turkey oak forests (AM and UG) showed a more context-dependent pattern. Observer category was not significant as a main effect, but the Site × Observer interaction was significant. Site, Plot size and Site × Plot size were also all significant, whereas Plot size × Observer and the three-way interaction were not significant (Table 4). Fig. 4 shows that Site accounted for the largest share of variation (27.2%), followed by Plot size (18.1%) and Site × Plot size (12.9%), whereas observer-related fractions were smaller, with Site × Observer explaining 5.4% and Observer alone 2.6%; residual variation accounted for 33.8% of the total. Pairwise

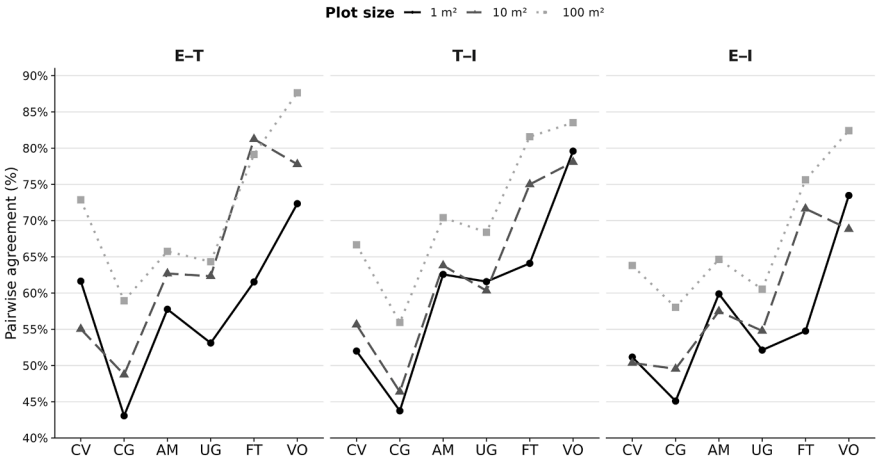


Fig. 3. Pairwise agreement in species detection among observer categories across the three nested plot sizes (1, 10 and 100 m^2) in the six forest sites. Values are shown separately for expert versus trained (E–T), trained versus inexpert (T–I) and expert versus inexpert (E–I) comparisons. The figure highlights how observer agreement varies jointly with forest context, plot size and observer pair.

Table 4
PERMANOVA results for species composition in the three forest types analysed separately using Jaccard dissimilarity. The design included observer category (TIP_R), site (ID) and plot size (CA), with all factors treated as fixed.

Forest type	Source	df	SS	MS	Pseudo-F	P(perm)	perms
BEE	Observer	2	1440.7	720.36	0.58846	0.871	996
	Site	**1**	**50951.0**	**50951.0**	**41.622**	**0.001**	**999**
	Plot size	**2**	**50385.0**	**25192.0**	**20.580**	**0.001**	**999**
	Site × Plot size	**2**	**11687.0**	**5843.7**	**4.7738**	**0.001**	**998**
	Site × Observer	2	786.4	393.2	0.32121	0.987	999
	Plot size × Observer	4	768.63	192.16	0.15697	1.000	999
	Site × Plot size × Observer	4	436.76	109.19	0.089197	1.000	999
	Residuals	162	198310.0	1224.1			
	Total	179	314770.0				
TUR	Observer	2	4536.1	2268.1	1.3651	0.064	996
	Site	**1**	**98380.0**	**98380.0**	**59.214**	**0.001**	**998**
	Plot size	**2**	**60216.0**	**30108.0**	**18.122**	**0.001**	**999**
	Site × Plot size	**2**	**17713.0**	**8856.6**	**5.3306**	**0.001**	**997**
	Site × Observer	**2**	**5869.0**	**2934.5**	**1.7662**	**0.004**	**999**
	Plot size × Observer	4	1365.0	341.26	0.20540	1.000	995
	Site × Plot size × Observer	4	1637.0	409.26	0.24633	1.000	998
	Residuals	162	269150.0	1661.4			
	Total	179	458870.0				
HOL	**Observer**	**2**	**4432.2**	**2216.1**	**1.8402**	**0.006**	**998**
	Site	**1**	**87916.0**	**87916.0**	**73.003**	**0.001**	**998**
	Plot size	**2**	**37454.0**	**18727.0**	**15.550**	**0.001**	**997**
	Site × Plot size	**2**	**12728.0**	**6364.1**	**5.2846**	**0.001**	**997**
	Site × Observer	**2**	**4600.3**	**2300.1**	**1.9100**	**0.005**	**999**
	Plot size × Observer	4	2807.3	701.81	0.58276	0.981	997
	Site × Plot size × Observer	4	1296.1	324.03	0.26907	1.000	997
	Residuals	162	195090.0	1204.3			
	Total	179	346330.0				

Significance was assessed by permutation. Significant terms are in bold. Forest types: HOL, Holm oak forests; TUR, Turkey oak forests; BEE, Beech forests.

Table 5
Significant pairwise permutational post hoc comparisons for the Site × Observer interaction (ID × TIP_R) in Turkey oak and holm oak forests. Only significant contrasts are shown.

Type	Site	Contrast	t	P(perm)	perms
TUR	UG	Expert vs Inexpert	1.8591	0.001	998
TUR	UG	Trained vs Inexpert	1.5259	0.005	999
HOL	CG	Expert vs Inexpert	1.9027	0.001	999
HOL	CG	Trained vs Inexpert	1.6273	0.012	999

post hoc comparisons showed that observer-related differences were restricted to UG, where both Expert vs Inexpert and Trained vs Inexpert contrasts were significant, whereas no observer contrast was significant in AM (Table 5).

In holm oak forests (CV and CG), observer effects were stronger. Observer category was significant as a main effect, and the Site × Observer interaction was also significant. Site, Plot size and Site × Plot size were all significant, whereas Plot size × Observer and the three-way interaction were not significant (Table 4). Fig. 4 shows a comparable but slightly more observer-sensitive variance structure than in the Turkey oak forests: Site explained 29.3% of the total variation, Plot size 16.1%, Site × Plot size 12.4%, Site × Observer 5.7%, and Observer alone 3.9%, while residual variation accounted for 32.7%. Pairwise post hoc comparisons showed that observer-related differences were restricted to CG, where Expert vs Inexpert and Trained vs Inexpert contrasts were significant, while no observer contrast was significant in CV (Table 5).

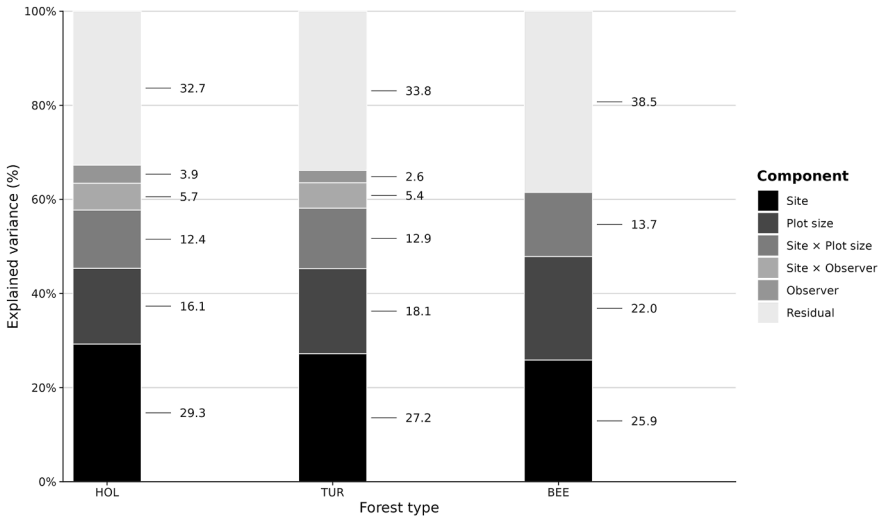


Fig. 4. Relative contribution (%) of the PERMANOVA variance components to total compositional variation in the three forest types, based on Jaccard dissimilarity. Each bar sums to 100% and shows the proportion of variation explained by observer category, site, plot size, their interactions, and residual variation. Forest types: HOL, holm oak forests; TUR, Turkey oak forests; BEE, beech forests.

The multivariate structure of the dataset therefore remained primarily ecological, but in the oak-dominated forests a measurable observer-expertise-related component emerged through site-specific divergence rather than through a uniform effect across plot sizes.

4. Discussion

Observer-expertise-related variation in these forest vegetation surveys was measurable, ecologically structured and methodologically relevant, but it did not obscure the dominant ecological signal in the data. Across the full design, differences among predefined observer-expertise categories affected species richness, agreement in species detection and pseudoturnover, whereas community composition remained primarily driven by site context and spatial grain. The results therefore reject both extreme interpretations: vegetation records are neither so dependent on observer-expertise category that ecological inference becomes unreliable, nor so robust that observer-related variation can be treated as negligible background noise. Instead, observer-expertise effects emerge as a structured component of the observation process, affecting different levels of inference to varying degrees.

4.1. Record-based richness across sites and spatial grains

A central finding is that observer-expertise-related variation was strongly context dependent. It was weakest in the beech forests and stronger in both oak-dominated forest types, namely the holm oak (*Quercus ilex*) and Turkey oak (*Q. cerris*) forests, with the clearest effects occurring in the more compositionally heterogeneous stands.

This pattern is consistent with the broader literature on observer error in vegetation surveys, while also indicating that the component isolated here is specifically linked to differences among expertise, calibration and site-familiarity categories. Earlier studies showed that disagreement among observers increases with census difficulty, species richness and sampling complexity (Archaux et al., 2006; Vittoz and Guisan, 2007; Milberg et al., 2008), while more recent work has confirmed that observer-driven pseudoturnover varies markedly across habitats and survey settings rather than following a uniform pattern (Seidling et al., 2020; Boch et al., 2022; Morrison et al., 2024). Our data fit this interpretation closely. The beech stands, with comparatively low richness and a simpler understory structure (Maccherini et al., 2019; Ottaviani et al., 2019), showed the highest agreement and the lowest observer-related turnover, especially at the largest grain. By contrast, the oak stands combined higher richness, greater structural heterogeneity and a broader range of potentially ambiguous field decisions, conditions likely to increase omission and disagreement.

The record-based richness results refine this picture. Observer-expertise effects on stratum-specific floristic record richness were statistically limited to a small number of site × grain combinations, rather than being pervasive across the dataset. Significant differences were detected in Colognole at 1 m^2, where the trained team produced fewer field records than the expert team, and in Ulignano at 100 m^2, where the expert team produced more field records than both the trained and inexpert teams. In Colognole, the same descriptive tendency for lower values in the trained team was also visible at larger grains, but these contrasts did not reach conventional significance in the count-based models. This more conservative result reinforces the interpretation that observer-expertise effects on record-based richness were real but strongly context dependent, and should not be read as a uniform expert-versus-novice gradient.

4.2. Concordance, pseudoturnover and observer disagreement

Plot size also shaped observer disagreement, but its effect depended on forest context rather than following a simple linear pattern. Larger plots can reduce the influence of individual missed records by integrating more species from the regional species pool, yet they also increase search area, structural heterogeneity and the number of marginal or low-frequency species that may be inconsistently recorded. This trade-off explains why increasing grain generally reduced observer disagreement in simpler stands but did not fully eliminate observer-related turnover in richer, more heterogeneous forests, such as both oak forest types. Thus, the effect of scale is not purely geometric; rather, it should be interpreted as emerging from the interaction between plot size and vegetation complexity, rather than from plot size alone.

The contrast between record-based richness and composition-based results is particularly informative. In several cases, differences in the number of field records among observer-expertise categories were limited or absent, while pseudoturnover and multivariate analyses still showed that different observers did not necessarily record the same taxa. Record-based richness alone is therefore an incomplete descriptor of observer-expertise-related variation. Two observers may produce similar numbers of records while differing in species identity, especially when disagreement concerns rare, patchy, seasonal, taxonomically difficult or plot-border taxa. For monitoring purposes, this distinction is crucial: record-based richness may remain stable while apparent turnover is inflated, and such a pattern can easily be misinterpreted as ecological change if observer effects are not measured explicitly. Pseudoturnover is particularly informative because it captures the observational component of apparent compositional change more directly than record-based summaries do.

The pseudoturnover results directly support this interpretation. PT was highest and most variable at the smallest grain, then generally declined with increasing plot size, but often remained appreciable in oak-dominated forests even at 100 m^2. Repeated surveys of the same plots can therefore generate non-trivial apparent turnover even in the absence of real ecological change, particularly in species-rich stands and in fine-grained sampling. Our design quantified disagreement among different observer-expertise teams, but it did not measure within-observer repeatability across repeated surveys. We therefore cannot determine whether similar discrepancies would also arise when the same observer repeats the survey under challenging field conditions. For long-term vegetation monitoring, the key implication is broader than observer replacement alone: apparent gains and losses should be interpreted cautiously whenever survey repeatability has not been explicitly quantified. Such discrepancies may arise from changes in observer identity, but also from incomplete calibration, inconsistent application of inclusion rules, plot-border ambiguity, taxonomic uncertainty or variation in field decisions across repeated surveys. This interpretation is consistent with broader concerns about observer effects and other forms of sampling error in resurveys and historical plot comparisons (Kapfer et al., 2017; Verheyen et al., 2018; Douda et al., 2023).

The distinction between rooted and covering-only species, although not central to the main statistical analyses, helps explain part of the observed disagreement. Covering-only records were much less regular than full field records, especially where canopy interception and plot-boundary decisions became more frequent. This suggests that observer-expertise-related variation is not purely taxonomic. A non-negligible component is procedural and depends on how observers apply field rules at plot edges or in ambiguous situations. In small or nested plots, such decisions can materially influence apparent richness and turnover. Some observer-related variation is probably irreducible, but part of the variation observed among expertise categories can likely be reduced through stricter standardisation of field procedures, especially in relation to border rules, treatment of uncertain taxa and operational criteria for ending a census.

4.3. Observer-expertise effects on species composition

The multivariate analyses further corroborate this message, showing that observer-expertise-related variation did not overwhelm the dataset's main ecological structure. Site identity and plot size remained the dominant sources of compositional variation, whereas observer-expertise-related compositional effects emerged only in specific oak-dominated contexts. This indicates that observer-expertise-related variation was not a uniform property of the survey protocol, but depended on the interaction between observer-expertise category and local vegetation conditions. In simpler beech stands, species detection and recognition were apparently more repeatable, while richer and structurally more complex oak stands created greater opportunities for omission, inconsistent inclusion and taxonomic uncertainty. The compositional signal therefore supports a site-dependent interpretation of observer-expertise-related variation, consistent with the view that vegetation complexity and field ambiguity are major drivers of observer-related uncertainty in plant monitoring (Bacaro et al., 2009; Boch et al., 2022).

4.4. Implications for vegetation monitoring

These results support treating observer-related uncertainty, including variation associated with observer expertise, calibration and site familiarity, as a measurable component of vegetation data. Plants do not move, but plant detection remains imperfect, and the resulting record reflects both vegetation structure and human perception. The observer is therefore part of the measurement process, and making that component explicit strengthens ecological inference, especially when repeated plot data are used to detect subtle changes over time or across sites.

5. Conclusions

Observer-expertise-related variation in forest vegetation surveys was measurable and ecologically structured, but it did not override the dominant ecological signal in the data. Its effects were most evident in stratum-specific floristic record richness, agreement in species detection and apparent turnover, whereas species composition remained driven mainly by site and plot size. These effects were not uniform across the dataset: they were weak in beech forests and stronger in oak forests, where higher vegetation complexity increased the scope for omission, inconsistent detection and plot-boundary ambiguity. In multivariate terms, observer-expertise effects were primarily site-dependent rather than scale-dependent, indicating that they emerged from the interaction between observer-expertise category and local vegetation context rather than as a uniform effect across plot sizes. The broader implication is that observer expertise, calibration history and familiarity with permanent plots should be recorded as primary metadata in vegetation monitoring programmes and considered when interpreting repeated surveys. This is particularly important because pseudoturnover associated with changes in these observer-expertise attributes can be mistaken for ecological change when the observational component is ignored. Calibration procedures should therefore cover the main habitat types and plot sizes included in a monitoring programme, especially where vegetation is species-rich, structurally heterogeneous or sampled using fine-grained nested plots. Finally, record-based richness alone should not be used as the sole diagnostic of data quality, because similar numbers of field records may conceal substantial differences in recorded species identity; robust monitoring therefore requires both richness-based and composition-based indicators of repeatability.

CRediT authorship contribution statement

Giovanni Bacaro: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review and editing, Visualization, Supervision. Enrico Tordoni: Methodology, Formal analysis, Data curation, Writing – review and editing, Visualization. Simona Maccherini: Formal

analysis, Writing – review and editing. Miris Castello: Investigation, Resources, Writing – review and editing. Alessandro Chiarucci: Conceptualization, Methodology, Investigation, Resources, Writing – review and editing, Supervision.

Data availability

The raw data supporting this study are supplied as CSV files with the submission for peer review and will be deposited in a public repository upon acceptance. Repository details will be provided in the final published article with: The data and R code used to reproduce the analyses are available in the public GitHub repository: https://github.com/gbacaro/observer_bias.

Funding

No external funding was received for this study.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Elisa Baragatti, Elisa Santi, Sara Landi and Valerio Amici for their contribution to field data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2026.07.019.

References

Ahrends, A., Rahbek, C., Bulling, M.T., et al., 2011. Conservation and the botanist effect. Biol. Conserv. 144, 131–140. https://doi.org/10.1016/j.biocon.2010.08.008.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x.

Anderson, M.J., Gorley, R.N., Clarke, K.R., 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E, Plymouth.

Archaux, F., Gosselin, F., Bergès, L., et al., 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. J. Veg. Sci. 17, 299–306. https://doi.org/10.1111/j.1654-1103.2006.tb02449.x.

Bacaro, G., Baragatti, E., Chiarucci, A., 2009. Using taxonomic data to assess and monitor biodiversity: are the tribes still fighting? J. Environ. Monit. 11, 798–801. https://doi.org/10.1039/B818171N.

Bacaro, G., 2025. Shades of grey: a continuum of biodiversity understanding from dark to bright diversity. Ecol. Divers. 2, 10012. https://doi.org/10.70322/ecoldivers.2025.10012.

Bacaro, G., Fonda, F., 2026. Toward the full operationalization of the dark diversity concept in environmental impact assessments: a call to revise international EIA standards. Conserv. Lett. 19, e70034. https://doi.org/10.1111/con4.70034.

Boch, S., Küchler, H., Küchler, M., et al., 2022. Observer-driven pseudoturnover in vegetation monitoring is context-dependent but does not affect ecological inference. Appl. Veg. Sci. 25, e12669. https://doi.org/10.1111/avsc.12669.

Burg, S., Rixen, C., Stöckli, V., et al., 2015. Observation bias and its causes in botanical surveys on high-alpine summits. J. Veg. Sci. 26, 191–200. https://doi.org/10.1111/jvs.12211.

Chao, A., Chiu, C.-H., 2016. Species richness: estimation and comparison. In: Balakrishnan, N., Colton, T., Everitt, B., et al. (Eds.), Wiley StatsRef: Statistics Reference Online. Wiley. https://doi.org/10.1002/9781118445112.stat03432.pub2.

Chiarucci, A., Bonini, I., 2005. Quantitative floristics as a tool for the assessment of plant diversity in Tuscan forests. For. Ecol. Manag. 212, 160–170. https://doi.org/10.1016/j.foreco.2005.03.041.

Chiarucci, A., De Dominicis, V., Wilson, J.B., 2001. Structure and floristic diversity in permanent monitoring plots in forest ecosystems of Tuscany. For. Ecol. Manag. 141, 201–210. https://doi.org/10.1016/S0378-1127(00)00329-7.

Clarke, K.R., Gorley, R.N., 2006. PRIMER v6: User manual/tutorial. PRIMER-E, Plymouth.

Dengler, J., 2009. A flexible multi-scale approach for standardised recording of plant species richness patterns. Ecol. Indic. 9, 1169–1178. https://doi.org/10.1016/j.ecolind.2009.02.002.

Douda, J., Doudová, J., Holeštová, A., et al., 2023. Historical sampling error: a neglected factor in long-term biodiversity change research. Biol. Conserv. 286, 110317. https://doi.org/10.1016/j.biocon.2023.110317.

Kapfer, J., Hédl, R., Jurasinski, G., et al., 2017. Resurveying historical vegetation data – opportunities and challenges. Appl. Veg. Sci. 20, 164–171. https://doi.org/10.1111/avsc.12269.

Kéry, M., Schmidt, B., 2008. Imperfect detection and its consequences for monitoring for conservation. Community Ecol. 9, 207–216. https://doi.org/10.1556/ComEc.9.2008.2.10.

Kirby, K.J., Bines, T., Burn, A., et al., 1986. Seasonal and observer differences in vascular plant records from British woodlands. J. Ecol. 74, 123–131. https://doi.org/10.2307/2260353.

Lepš, J., Hadincová, V., 1992. How reliable are our vegetation analyses? J. Veg. Sci. 3, 119–124. https://doi.org/10.2307/3236006.

Lin, H.-T., Lam, T.Y., von Gadow, K., et al., 2020. Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. For. Ecol. Manag. 457, 117658. https://doi.org/10.1016/j.foreco.2019.117658.

Maccherini, S., Bacaro, G., Giovannetti, G., et al., 2019. Analysing methodological issues in short-term monitoring of rare European beech forests restoration. Plant Biosyst. 153, 60–67. https://doi.org/10.1080/11263504.2018.1454529.

Martínez-Villegas, J.A., Pisanty, I., Martorell, C., et al., 2024. Importance of accounting for imperfect detection of plants in the estimation of population growth rates. Oikos 2024, e10708. https://doi.org/10.1111/oik.10708.

Milberg, P., Bergstedt, J., Fridman, J., et al., 2008. Observer bias and random variation in vegetation monitoring data. J. Veg. Sci. 19, 633–644. https://doi.org/10.3170/2008-8-18423.

Moreno-Fernández, D., Cañellas, I., Hernández, L., et al., 2024. Nested plot designs used in forest inventory do not accurately capture tree species richness in Southwestern European forests. Ann. For. Sci. 81, 20. https://doi.org/10.1186/s13595-024-01237-8.

Morrison, L.W., 2016. Observer error in vegetation surveys: a review. J. Plant Ecol. 9, 367–379. https://doi.org/10.1093/jpe/rtv077.

Morrison, L.W., Young, C.C., 2016. Observer error in sampling a rare plant population. Plant Ecol. Divers. 9, 289–297. https://doi.org/10.1080/17550874.2016.1220989.

Morrison, L.W., 2017. Insular plant turnover across a 22-year interval: a critical retrospective of the roles of pseudoturnover and cryptoturnover. J. Biogeogr. 44, 1007–1017. https://doi.org/10.1111/jbi.12866.

Morrison, L.W., Bingham, S.N., Young, C.C., 2020. Inter-observer error in wetland vegetation surveys. Wetlands 40, 249–258. https://doi.org/10.1007/s13157-019-01173-8.

Morrison, L.W., 2021. Nonsampling error in vegetation surveys: understanding error types and recommendations for reducing their occurrence. Plant Ecol. 222, 1–10. https://doi.org/10.1007/s11258-021-01125-5.

Morrison, L.W., Leis, S.A., DeBacker, M.D., 2023. Observer error in grassland vegetation surveys: effects on species diversity metrics and species–abundance relationships. J. Plant Ecol. 16, rtad002. https://doi.org/10.1093/jpe/rtad002.

Morrison, L.W., Leis, S.A., Short, M.F., et al., 2024. A spatiotemporal comparison of interobserver error in vegetation sampling. J. Veg. Sci. 35, e13286. https://doi.org/10.1111/jvs.13286.

Morrison, L.W., Leis, S.A., DeBacker, M.D., 2025. Metrics for quantifying and comparing observer error across vegetation studies. J. Plant Ecol. 18, rtaf106. https://doi.org/10.1093/jpe/rtaf106.

Nilsson, S.G., Nilsson, I.N., 1983. Are estimated species turnover rates on islands largely sampling errors? Am. Nat. 121, 595–597.

Ottaviani, G., Götzenberger, L., Bacaro, G., et al., 2019. A multifaceted approach for beech forest conservation: environmental drivers of understory plant diversity. Flora 256, 85–91. https://doi.org/10.1016/j.flora.2019.05.006.

Perret, J., Besnard, A., Charpentier, A., et al., 2023. Plants stand still but hide: imperfect and heterogeneous detection is the rule when counting plants. J. Ecol. 111, 1483–1496. https://doi.org/10.1111/1365-2745.14110.

Pignatti, S., 1982. Flora d'Italia. Edagricole, Bologna.

Portier, J., Zellweger, F., Zell, J., et al., 2022. Plot size matters: toward comparable species richness estimates across plot-based inventories. Ecol. Evol. 12, e8965. https://doi.org/10.1002/ece3.8965.

Scott, W.A., Hallam, C.J., 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecol. 165, 101–115. https://doi.org/10.1023/A:1021441331839.

Seidling, W., Hamberg, L., Máliš, F., et al., 2020. Comparing observer performance in vegetation records by efficiency graphs derived from rarefaction curves. Ecol. Indic. 109, 105790. https://doi.org/10.1016/j.ecolind.2019.105790.

Thouverai, E., Pavoine, S., Tordoni, E., et al., 2026a. Rarefaction curves with the Rarefy package. In: Rocchini, D. (Ed.), R Coding for Ecology. Springer, Nature Switzerland, Cham, pp. 143–158. https://doi.org/10.1007/978-3-031-99665-8_6.

Thouverai, E., Pavoine, S., Tordoni, E., et al., 2026b. Rarefy: an R package for the calculation of taxonomic, functional and phylogenetic sample-based spatially explicit rarefaction curves. Ecol. Inform. 96, 103803. https://doi.org/10.1016/j.ecoinf.2026.103803.

Verheyen, K., Bažány, M., Chećko, E., et al., 2018. Observer and relocation errors matter in resurveys of historical vegetation plots. J. Veg. Sci. 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. J. Veg. Sci. 18, 413–422. https://doi.org/10.1111/j.1654-1103.2007.tb02553.x.