HAL open science

Vegetation unit assignments: phytosociology experts and classification programs show similar performance but low convergence

Lise Maciejewski, Paulina Pinto, Stéphanie Wurpillot, Jacques Drapier, Serge Cadet, Serge Muller, Pierre Agou, Benoît Renaux, Jean-claude Gégout

To cite this version:

Lise Maciejewski, Paulina Pinto, Stéphanie Wurpillot, Jacques Drapier, Serge Cadet, et al.. Vegetation unit assignments: phytosociology experts and classification programs show similar performance but low convergence. Applied Vegetation Science, 2020, 23 (4), pp.698-709. 10.1111/avsc.12516 . hal-03048253

HAL Id: hal-03048253

https://hal.science/hal-03048253v1

Submitted on 3 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vegetation unit assignments: phytosociology experts and classification programs show similar performance but low convergence

Running title: comparisons of vegetation unit assignments

Lise Maciejewski[1,2], Paulina Elcira Pinto[1], Stéphanie Wurpillot[3], Jacques Drapier[4], Serge Cadet[5], Serge Muller[6], Pierre Agou[7], Benoît Renaux[8], Jean-Claude Gégout[1]

ORCID:

Lise Maciejewski https://orcid.org/0000-0002-9142-0507

Serge Muller https://orcid.org/0000-0002-4545-9427

Jean-Claude Gégout https://orcid.org/0000-0002-5760-9920

Authors' institutional affiliations

[1] Université de Lorraine, AgroParisTech, INRAE, UMR Silva, 14 rue Girardet F-54000 Nancy, France.

[2] OFB, MNHN, CNRS, UMS 2006 PatriNat, CP 41, 36 rue Geoffroy-Saint-Hilaire F-75005 Paris, France.

[3] IGN (Institut National de l'Information Géographique et Forestière), 73 Avenue de Paris F-94160 Saint-Mandé, France.

[4] IGN (Institut National de l'Information Géographique et Forestière), Direction interrégionale nord-est, 1 rue des Blanches Terres F-54390 Champigneulles, France.

[5] ONF (Office national des forêts), Bureau d'études territorial 13/84, 46 av. Cézanne F-13098 Aix-en-Provence, France.

[6] MNHN, CNRS, UPMC, EPHE, UMR 7205 ISYEB, F-75005 Paris, France.

[7] Biotope, Centre Bourgogne, 122-124 Faubourg Bannier F-45000 Orléans, France.

[8] Conservatoire botanique national du Massif Central, Le Bourg F-43230 Chavaniac-Lafayette, France.

Correspondence

Lise Maciejewski, Université de Lorraine, AgroParisTech, INRAE, UMR Silva, 14 rue Girardet F-54000 Nancy, France.

Email : lise.maciejewski@agroparistech.fr

Funding information

This study was supported by the French National Institute for Agricultural Research (Forest Grassland and Freshwater Ecology Department, EFPA) through the ONF-INRA Interface Grant "Station, distribution, croissance et choix des essences dans un contexte environnemental changeant", by the French National Research Agency (ANR) through the Laboratory of Excellence ARBRE (ANR- 12-LABXARBRE-01), and by the Regional Council of Lorraine through the project "Contribution à l'identification et à la cartographie à fine résolution des zones humides forestières à l'aide du caractère hygrophile des plantes".

ABSTRACT

Aims

Assigning vegetation plots to vegetation units is a key step in biodiversity management projects. Nevertheless, the process of plot assignment to types is usually non-standardized, and assignment consistency remains poorly explored. To date, the efficiency of automatic classification programs has been assessed by comparing them with a unique expert judgment. Therefore, we investigated the consistency of five phytosociology expert judgments, and the consistency of these judgements with those of automatic classification programs.

Location: mainland France

Methods

We used 273 vegetation plots distributed across France and covering the diversity of the temperate and mountainous forest ecosystems of Western Europe. We asked a representative panel of five French organizations with recognized expertise in phytosociology to assign each plot to vegetation units. We provided a phytosociological classification including 228 associations, 43 alliances and eight classes. The assignments were compared among experts using an agreement ratio. We then compared the assignments suggested by three automatic classification programs with the expert judgments.

Results

We observed small differences among the agreement ratios of the expert organizations; a given expert organization agreed with another one on association assignment one time in four on average, and one time in two on alliance assignment. The agreement ratios of the automatic classification programs were globally lower, but close to expert judgments.

Conclusions

The results support the current trend toward unifying the existing classifications and specifying the assignment rules by creating guiding tools, which will decrease inter-observer variation. As compared to a pool of phytosociology experts, programs perform similarly to individual experts in vegetation unit assignment, especially at the alliance level. Although programs still need to be improved, these results pave the way for the creation of habitat time series crucial for the monitoring and conservation of biodiversity.

Keywords: vegetation classification, habitat, forest ecosystem, phytosociology, Natura 2000, vegetation typology.

INTRODUCTION

Public policies for nature protection were initially developed for species conservation. In 1992, the European Directive 92/43/EEC (EC Council 2006), known as the Habitats Directive, also targeted ecosystem conservation by defining a list of habitats of Community interest. Thanks to a historical pan-European use of phytosociology, phytosociological units were chosen as a basis for the definition of the terrestrial habitats of Community interest of the Habitats Directive, also called 'Natura 2000 habitats' (European Commission DG Environment, 2013; Guarino et al., 2018). The Habitats Directive states that all European Union countries must protect and restore the threatened species and habitats of Community interest listed in the annexes. Each member state must contribute to the creation of the Natura 2000 network by designating sites as special areas of conservation where the appropriate management plans specifically designed for the sites and the targeted habitats are established (EC Council 2006).

In this context, decision-makers defined the vegetation types to be protected that stakeholders now need to inventory in the field to create maps of endangered ecosystems and actually implement management actions. Site managers must assign new vegetation observations to previously defined vegetation types in accordance with the way these types were originally defined. The classification of vegetation plots within vegetation units is traditionally performed manually on the basis of expert human knowledge, based on the observation of diagnostic plant species in the field. The assignment of surveys to vegetation units usually only relies on a single expert judgment, and this process can be slow, subjective and fraught with difficulties (see Willner 2011). It considerably varies among countries and individual researchers or managers because no formal assignment rules have ever been used until recently (Guarino et al., 2018). Experts rarely indicate how to determine whether a given new plot record fits into any of the published types. The absence of explicit assignment rules can lead to erroneous classifications and be at the origin of inappropriate management or legal restrictions (Cherrill, 2016), or of the failure to detect or monitor land-cover changes (Prosser and Wallace, 1999), or even of the destruction of rare ecosystems due to ignorance. Although the question of how to assign new sample plots to existing units is of importance, the issue has hardly ever been addressed (Guarino et al., 2018).

In the absence of any clear definition and application of assignment rules, it is difficult to estimate the consistency of the assignments of several phytosociology experts, and this in turn makes it difficult to compare two vegetation units over time and across space. Expert habitat identifications have hardly ever been compared, or only compared across a few vegetation units over a small area (Couvreur et al., 2015). Several studies have compared the ways experts mapped land-cover types, but were usually limited to small-sized sites (Eriksen et al., 2019;

Hearn et al., 2011; Stevens et al., 2004; Ullerud et al., 2018). These studies all conclude that the level of agreement among experts is usually quite low and dependent on the number of units in the provided repository, as the probability of two identical suggestions increases when the number of units decreases.

The absence of standardized assignment rules and the financial and time costs of manual assignment have led to the development of several automatic classification programs with explicit assignment rules (e.g. Kočí et al. 2003; Oliver et al. 2013; Tichý et al. 2014). Automatic classification programs assign large numbers of floristic surveys to vegetation units using standardized methods. Once the tool has been developed, the running cost and the running time are negligible as compared to the time and cost of seeking an assignment by an expert, and the program can be improved with the progress of knowledge. Moreover, the potential deviations of the programs can be detected and then taken into account.

The efficiency of automatic classification programs has usually been assessed based on one expert judgment considered as a "true assignment" (e.g. Černá and Chytrý 2005; van Tongeren et al. 2008; De Cáceres et al. 2009), which could be merely regarded as a validation. Moreover, the absence of explicit assignment rules has led to significant inter-expert variation (Eriksen et al., 2019; Hearn et al., 2011; Stevens et al., 2004; Ullerud et al., 2018). Consequently, the choice of the reference is of crucial importance, and could be contested. Besides, with a unique expert as a reference, the only possibility is to reach the expert level, it is impossible to outperform it. Finally, to date we do not know whether the differences between assignment by an automatic classification program and an expert judgment are comparable to the differences among several assignments by phytosociology experts.

We explored the consistency of several phytosociology expert assignments, and checked the ability of automatic classification programs to reproduce the assignments of floristic surveys to vegetation units by several human experts. To highlight a current reality, we used a representative panel of expert organizations with recognized skills in phytosociology and three automatic classification programs to perform a large-scale study encompassing a wide range of forest ecosystems.

METHODS

We quantified inter-observer variation among five expert organizations with recognized skills in phytosociology. We did not compare each assignment with a reference considered as "the truth", but kept each of the five expert judgments as one sole reference, and used identical assignments among expert organizations to compute an index of consistency that we called 'agreement ratio' among experts. Then, we compared the assignments suggested by three automatic classification programs with the assignments of the expert organizations to check

whether the automatic classification programs could perform as well as or even outperform the expert organizations by being more consensual. We used identical assignments to calculate an 'agreement ratio' to investigate if their agreement ratios were lower than, similar to or higher than those obtained by the human experts.

Vegetation plots

Two hundred and seventy-three surveys were carried out between May and September 2013 by six teams of the National Forest Inventory (NFI) of France across the forests of mainland France (Fig. 1) (see Hervé 2016 for a presentation of the French NFI method). Each plot had been randomly located on the forest map before the fieldwork phase, which implied that the field agents did not re-locate the plot before beginning the survey. The Mediterranean area as defined by the NFI was excluded from the study. Vascular species found in the understory layer and terricolous bryophytes were recorded across each circular 700 m² plot. All herbaceous and shrub species, and tree species below 7.5 cm in diameter at 1.30 m high were included. The cover of each species was visually assessed using the Braun-Blanquet approach (Braun-Blanquet et al., 1932), after completion of the survey (Pinto et al., 2016). Tree species above 7.5 cm in diameter were also included, but their cover was assessed visually using 10 % percentage classes. The mean species richness of the floristic surveys was 32.9 species (STD=15.6).

Selection of organizations with recognized expertise in phytosociology

We selected a panel of five French organizations from the private and public sectors with recognized expertise to represent the diversity of backgrounds and practices in phytosociology, namely the National Botanical Conservatory of the Massif Central (Bot. Cons.); Biotope, an engineering office (Eng. Off.); the botanist's network of the National Forestry Office of France (For. Manag.); the National Museum of Natural History of France (Museum); and the National Forest Inventory (For. Invent.) of France.

The National Botanical Conservatory of the Massif Central is a public institution that aims to develop knowledge and conservation of the flora, plant communities and natural habitats. At a local scale, their ecologists are involved in or consulted about all local vegetation studies in the Massif Central (in central France), and in particular they are in charge of developing and unifying the classification of the plant communities of the Massif Central. At a national scale, they are in charge of classifying four phytosociological classes of the French temperate forests.

Biotope is a French private engineering office specialized in ecological studies, created in 1993. It gathers experts in phytosociology across the French territory; one of them is a national expert in phytosociology.

The National Forestry Office is the public institution in charge of managing the French public forests (8.5% of the French mainland territory), from ecological engineering to timber sales. Their best regional naturalists are gathered in a national network composed of tens of botanists and phytosociology experts involved in or consulted about all local vegetation studies in public forests, to which they bring their regional and practical knowledge.

The National Museum of Natural History is a reference public institution in France. It carries out a wide range of missions, including basic and applied research, the conservation and expansion of its collections, education, expertise, and dissemination of knowledge. The scientific Head of the National Herbarium, who was involved in this study, is a doctor in phytosociology.

The National Forest Inventory of France is a public institution that has carried out a permanent inventory of forest resources across the country since the 1960's. They have been recording ecological and floristic data in addition to dendrometric information for more than 25 years. At a national level, the data are checked and validated by nationally recognized ecologists specialized in botany and phytosociology.

<u>Assignment of vegetation plots to vegetation units</u>

Several studies have shown a lack of exhaustiveness of vegetation inventories, as well as inter-observer variations (e.g. Archaux et al. 2006; Vittoz and Guisan 2007; Milberg et al. 2008; Morrison 2016). Variations in the list of species observed in the field can be a source of differences among experts' assignments to vegetation units. Therefore, we sent the same vegetation plots to all the expert organizations, asked them to assign a phytosociological association to each plot, and compared their assignments to vegetation units on an *ex-post* basis. Thus, the potential differences among the assignments were bound to be attributed to variability among the expert judgments and not to the potential variability in the lists of the species observed in the field.

Each of the five expert organizations received the 273 vegetation plots consisting of the list of species and their cover, and additional information such as the date of the survey and plot location variables (city, "département", region, altitude, slope, exposure). We provided a phytosociological repository of the forests of the French mainland territory in temperate and mountainous areas (Mediterranean area excluded), consisting of a list of 228 associations, 43 alliances, 21 orders, and eight classes, and their full citations. The repository included four lowland classes: *Querco roboris-Fagetea sylvaticae and Quercetea ilicis* in the well-drained soils, and *Alnetea glutinosae* and *Salicetea purpureae* in moist conditions. It also included four highland classes: two classes of pine forests (*Erico carneae-Pinetea sylvestris* and *Pino sylvestris-Juniperetea sabinae*), one class of spruce forest (*Vaccinio myrtilli-Piceetea abietis*), and one

treeline class (*Loiseleurio procumbentis-Vaccinietea microphylli).* This typology was based on the units published at the association and alliance levels (Bardat et al., 2004; Bioret et al., 2014; Gégout et al., 2009). We asked the organizations to assign a phytosociological association to each vegetation plot based on this information, and the suited phytosociological alliance and class in accordance with the hierarchical nested system we provided. But it was also possible for them to answer that they could not or did not want to assign one if they thought that it was not feasible.

For the National Botanical Conservatory of the Massif Central and the engineering office Biotope, the assignments of the vegetation plots to vegetation units were directly performed on an *ex-post* basis by the national phytosociology experts of each organization. For the National Museum of Natural History, they were performed by the Head of the National Herbarium. For the National Forestry Office, they were performed by the regional phytosociology experts and then gathered and checked by the national phytosociology expert. For the National Forest Inventory, they were performed in the field by the field agents when the vegetation surveys were carried out using vegetation keys when available (118 vegetation plots), and the remaining assignments were performed by the local phytosociology experts on an ex-post basis. All the assignments were then gathered and checked by the national phytosociology expert of the National Forest Inventory. Thus, we collected five assignments to associations (and the suited alliances) for each of the 273 vegetation plots. Each expert organization also reported the time taken to fulfill the mission.

Comparing the assignments: the agreement ratio

Choosing the reference

Comparisons among experts using one expert assignment (Couvreur et al., 2015; Stevens et al., 2004) or a collegial assignment as the reference (Eriksen et al., 2019) can be found in the literature. In these cases, all suggestions are compared with sole references considered as “the true assignments”. As the expert organizations solicited for our study had similar recognized expertise in phytosociology, we had no objective reason to consider one of them to be better or more trustworthy than the others. But to compare the assignments among them, we needed to define a reference, which in this study represented the most consensual assignment; this assignment was not the “truth”, but a close approximation of the convention to be followed when assigning new vegetation plots to vegetation units. We did not specifically define the most consensual assignment for each survey. But we investigated inter-observer variation and kept all five expert judgments as the reference, compared the assignments among one another, and investigated the number of identical assignments to compute an index of consistency that we called ‘agreement ratio’ (eq. 1). For any given expert, the higher the number of identical

assignments shared with each of the other four experts, the more consensual they were, and the higher their agreement ratio.

Calculation of the agreement ratio

For each plot, we compared the assignment of one expert organization with the other four. When the suggestions were identical, we scored one point. When they were different, we scored zero point. Then, for each vegetation plot and target expert organization, when a suggestion was similar with the other four, it scored four points. When the suggestion was different than the other four, it scored zero point. We repeated the comparison for the 273 vegetation plots and all five expert organizations. We summed up all the points scored by the identical suggestions and divided that number by the number of comparisons to obtain the agreement ratio:

$$Agreement\ ratio = \frac{Number\ of\ identical\ assignments}{Number\ of\ comparisons\ (Nc)} \quad (1)$$

Let *i* be the number of expert organizations and *n* the number of plots, the number of comparisons (Nc) was:

$$Nc = n * (i - 1) \quad (2)$$

When two expert organizations made no suggestion about a same vegetation plot, we did not consider it as an identical assignment and scored zero point.

The agreement ratio associated with an expert organization could theoretically reach 1 if the suggestions were all identical (they agreed in 100 % of cases) and the agreement ratio could reach 0 if the suggestions were all different (they all disagreed with one another). Nevertheless, the agreement ratios of the experts were interdependent. The value for a given expert could not be 1 if the other experts were not consensual themselves, because one cannot simultaneously agree with two other peers that do not agree with each other. Moreover, if one expert disagreed with the others, the agreement ratio of all experts was impacted. Therefore, we calculated the highest possible agreement ratio for each expert according to the suggestions of the other four.

Comparison of the expert judgments

We also made pairwise comparisons of each pair of expert organizations to potentially highlight patterns or similar behavior among experts by calculating the agreement ratios of the pairwise comparisons of the experts. When comparing two experts, the agreement ratio was similar to the simple matching coefficient, an index frequently used notably in genetics studies (Olden, Joy, and Death, 2006; Stiles, Lemme, Sondur, Morshidi, and Manshardt, 1993). We also tested the impact of the inclusion of an "inexperienced expert" on the agreement ratios of the five

expert organizations by adding a 6th reference with random assignments to the expert panel (see Supporting information, Appendix S1 for details).

For each vegetation plot, we calculated the number of experts who agreed on association and alliance assignment, and a typicality index to distinguish atypical vegetation plots that were more difficult to classify (Gégout and Coudun, 2012). A high typicality index highlights that plots are close to the core characteristics of the plant unit (Gégout and Coudun, 2012). Then, we used an analysis of variance to check the influence of species richness on the number of experts who agreed *per* plot to test whether too many or too few species could have destabilized the experts and impacted consistency among them. We also checked the correlation between the typicality of the surveys approximated by the typicality index and the number of experts who agreed.

Lastly, we investigated on which type of alliance the expert organizations tended to agree or disagree. Each time an alliance was cited, we collected the number of experts who agreed and calculated a mean value *per* alliance and reported it in a figure.

Automatic classification programs

Automatic classification programs developed with supervised methods were chosen to assign vegetation plots to the same classification of vegetation used by the expert organizations. Automatic classification programs consist of two primary components: the "inference engine" and the "knowledge base". While the inference engine is a general algorithm, the knowledge base provides information needed for the automatic classification program to function. The knowledge base can be a set of plots assigned a priori (extensive definition of vegetation types) or assignment rules derived from a description of the vegetation units (intensive definition) (Tichý et al., 2019). Using intensive definitions can lead to non-assignment of part of the vegetation plots to vegetation units. Therefore, we defined the knowledge base using a set of 9,827 phytosociological surveys classified in the phytosociological system by experts (the same repository as the one provided to the expert organizations) and considered to be typical. These surveys were extracted from the phytosociological literature and were used to characterize the French forest associations of temperate and mountainous areas (Mediterranean areas excluded) floristically and ecologically. Then, we used individual species rather than groups of species to define fidelity and constancy to the associations, because using groups of species implies choosing a minimum number of species that need to be present to assign the survey, so that surveys can also be non-assigned. Lastly, we chose two well-known fidelity indexes and one similarity index, namely (i) IndVal (Dufrêne and Legendre, 1997), (ii) the fidelity index ϕ (the Phi coefficient) (Chytrý et al., 2002), and (iii) the frequency-positive fidelity index (FPFI) based on the frequency and the fidelity index of the species (Tichý, 2005). We used the 9,827

phytosociological surveys to calculate the Indval index, the fidelity index ϕ, and the frequency index of 1,648 species in 228 associations. We used those fidelity indexes and a similarity index to build three automatic classification programs (hereafter called 'the Phi-program', 'the IndVal-program' and 'the FPFI-program') to be able to assign an association to each of the 273 surveys.

The algorithms of the Phi- and Indval-programs were based on the model of an automatic classification program developed in 2012 (Gégout and Coudun, 2012): the automatic classification program tested the assignment of a given floristic survey to each of the 228 associations by calculating the average value of the fidelity indexes (the fidelity index for the Phi-program, the Indval index for the Indval-program) of the species present in the survey. The program assigned the survey to the association whose average fidelity index was the highest.

For the FPFI-program, the frequency-positive fidelity index (FPFI) of each floristic survey to each of the 228 associations was calculated by computing the frequency index and the fidelity index of the species present in the survey. The program assigned the survey to the association whose FPFI was the highest (Tichý, 2005).

We ran the three automatic classification programs which each assigned one association, and in turn the suited alliance and class to each of the 273 floristic surveys.

<u>Agreement ratio of the automatic classification programs</u>

To compare the assignments by the automatic classification programs with the assignments by the reference expert organizations, the process was the same as described before. For each survey, we compared the vegetation units assigned by each automatic classification program with the five assignments by the expert organizations. When the suggestions were identical, we scored one point. When they were different, we scored zero point. We repeated the comparison for the 273 surveys. We summed up all the points scored by the identical suggestions and divided that number by the number of comparisons to obtain the agreement ratio (eq. 1) of each program with the five reference expert organizations taken. A lower agreement ratio than the ones of the expert organizations would have meant that the automatic classification program agreed with the phytosociology experts less often than one expert with the other four, leading to a poor ability of the automatic program to reproduce expert judgments. A higher agreement ratio would have meant that an automatic classification program could agree with several experts more often than one expert with several others, suggesting that an automatic classification program could provide a more consensual assignment.

When calculating an agreement ratio for an automatic classification program, the number of comparisons was higher than when calculating an agreement ratio for the five reference expert

organizations. The same number of comparisons could have been reached by comparing the suggestions of the automatic classification programs with only four expert organizations, calculating an agreement ratio, repeating this four more times removing a different expert organization each time, and calculating the mean agreement ratio. But this approach would have provided exactly the same results. Furthermore, we calculated the highest possible agreement ratio for each program to compare them with the highest possible agreement ratio of each expert, to study whether their peaks were equivalent.

Random assignment

To estimate the role of chance in the convergence of an automatic classification program and the five expert organizations, we compared their assignments with random assignments. For each of the 273 surveys, we drew an association at random within the 228 possibilities of the phytosociological repository, assigned the suited alliance and class, and then calculated the agreement ratio compared to the suggestions of the five expert organizations. We repeated this 1,000 times, and calculated the mean value of the 1,000 agreement ratios and corresponding standard deviation.

All associations had the same probability to be picked up in this random assignment. Nevertheless, some associations are more common and widespread than others. To take this parameter into account, we weighted the probability of each association to be picked up by its frequency in the assignments made by the expert organizations. An association mentioned 10 times has a probability of $\frac{10}{n*i}$ to be picked up. A non-mentioned association has a probability of 0. We simulated 1,000 random draws using this weighted phytosociological repository, and then we calculated the agreement ratio of each draw and the mean of the 1,000 agreement ratios and standard deviation for this weighted random assignment.

RESULTS

Comparison among expert organizations

Comparison of the agreement ratios

Each expert organization reported the time it spent to fulfill the mission. The accuracy of the answers was uneven, and dependent on the number of people involved in the assignment process. It took each organization one to three weeks to assign an association to each of the 273 surveys.

The mean value of the five agreement ratios of the expert organizations was 0.27 for association assignment (Fig. 2), i.e., one expert organization agreed with another on association assignment approximately one time in four on average, and the highest possible agreement ratio of the experts was around 0.52 (Appendix S2). For alliance assignment, the average value

of the five agreement ratios of the expert organizations was 0.48 (Fig. 2), i.e., one expert organization agreed with another one on alliance assignment slightly less than one time in two on average, and the highest possible agreement ratio of the experts was around 0.67 (Appendix S2).For class assignment, the average value of the five agreement ratios of the expert organizations was 0.90 (Fig. 2), i.e., one expert organization agreed with another one on alliance assignment nine times in ten on average, and the highest possible agreement ratio of the experts was around 0.94 (Appendix S2).

Similar levels of skills among the expert organizations

The five expert organizations made no suggestion 48 times, and they made a suggestion out of the provided repository 24 times. They fully agreed on 7% (n=20) of the vegetation plots and fully disagreed on 19% (n=51) of the vegetation plots (Fig. 3) for association assignment. The agreement ratios of all expert organizations for association assignment were really close to one another (Fig. 2). The difference between the highest and lowest agreement ratios was 0.05. We observed the same pattern for alliance assignment, where the difference was 0.03 (Fig. 2) and the five expert organizations fully agreed on 19 % of the vegetation plots and fully disagreed on 4 % (Fig. 3). For class assignment, the agreement ratios of all expert organizations were even closer to one another: the difference between the highest and lowest agreement ratios was 0.02 (Fig. 2).

When we simulated the inclusion of an inexperienced phytosociology expert in the panel of experts, the agreement ratios of the five expert organizations were even closer, with differences between the highest and lowest agreement ratios of 0.04 for association assignment and 0.02 for alliance and class assignment (Appendix S3). The agreement ratios were nearly 20% lower than the initial reference agreement ratios for association and alliance assignment. Nevertheless, the mean agreement ratios of the random draws were far below the ones of the expert organizations. In addition, a pairwise comparison of the expert organizations showed similar agreement ratios regardless of the organizations for association and alliance assignment: they all agreed and disagreed with each other in the same proportion (Tab. 1). Therefore, the closeness of the agreement ratio was not an artefact of the index, but showed that the skills of the five reference expert organizations were quite comparable regardless of their practice and background in phytosociology.

Sources of disagreement among expert organizations

Using an analysis of variance, we showed that plot species richness did not have a significant influence on the number of expert organizations that agreed on association assignment ($p = 0.87$), or on alliance assignment ($p = 0.36$). Therefore, a small or a large amount of species did not influence convergence among the expert organizations. The correlation of the typicality index with the number of expert organizations that agreed on association and alliance

assignment was significant ($p < 0.001$). Experts tended to disagree about plots with a low typicality index, which were transitional plots or plots with a low number of diagnostic species.

The expert organizations tended to agree more on precisely defined (e.g. *Rhododendro-Vaccinion*) or frequent alliances, for example *Fagion* or *Quercion* alliances (Fig. 4). Nevertheless, as regards frequent alliances, we failed to highlight whether it was because they were widespread and thus well-known, or because they were well defined.

Agreement ratios of the automatic classification programs *vs.* agreement ratios of the expert organizations

The highest possible agreement ratios of the five reference expert organizations and of any automatic classification program were really close, with a difference between the highest and lowest maxima of 0.03 at most (Appendix S2). Therefore, the increase of the number of comparisons performed to calculate the agreement ratio of any automatic classification program did not disadvantage or advantage them.

The agreement ratios of the three programs for association assignment were below the mean agreement ratio of the five expert organizations by 0.02 to 0.13 points (Fig. 2). Nevertheless, for alliance assignment, the agreement ratio of the Phi-program was slightly higher than the mean agreement ratio of the five expert organizations (by 0.01 points) (Fig. 2), and the agreement ratios of the IndVal and FPFI programs were slightly below (by 0.04 and 0.06 points, respectively). Thus, the agreement ratios of the automatic classification programs were really close to the agreement ratios of the expert organizations for alliance assignment. For class assignment, the agreement ratios of the three automatic classification programs were higher than the mean agreement ratio of the five expert organizations (by 0.01 to 0.02 points) (Fig. 2).

Random assignment *vs.* expert organizations

Comparing 1,000 random draws with the five expert organizations, we observed on average 5.5 identical assignments to associations, 62.3 to alliances and 813.0 to classes, with mean agreement ratios of 0.004, 0.046 and 0.596, respectively (Fig. 2). Therefore, an automatic classification program could agree with the five expert organizations by chance less than one time in two hundred and fifty for association assignment, and less than one time in twenty for alliance assignment.

Considering that common and widespread associations had more chances to be picked up, we also compared 1,000 weighted random draws with the five expert organizations. We found 54.9 identical assignments to associations, 222.8 to alliances and 1,180.3 to classes on average, with mean agreement ratios of 0.040, 0.163 and 0.865, respectively (Fig. 2). It follows that an automatic classification program using a weighted phytosociological repository could agree

with the five expert organizations by chance less than one time in twenty for association assignment, and less than one time in six for alliance assignment.

DISCUSSION

Lack of consistency among phytosociology experts

This study shows uniform and low agreement ratios among five expert organizations regarding the assignment of vegetation plots to phytosociological units specifically targeted by public policies and currently used in local management, such as associations and alliances. The number and the panel of expert organizations involved in the study on the one hand, and the close agreement ratios on the other hand, suggest that the low agreement ratios did not result from differences in the experts' skills in phytosociology.

As expected, we found high agreement ratios for class assignment, and better agreement ratios among expert organizations regarding the assignment of higher than lower hierarchical levels in the classification. The probability for two identical assignments increased when the number of units decreased. Nevertheless, we cannot conclude that using a high hierarchical level is the best option because using large units is a double-edged sword due to the trade-off between precision and error (Ullerud et al., 2018). Errors about adjacent vegetation types made at higher hierarchical levels involve larger deviations among possibilities than at lower levels. Accordingly, the choice of the hierarchical level should be guided by the needs for consistency and for detailed information on the assigned vegetation units (Eriksen et al., 2019).

Several studies have compared mappings of land-cover types by experts, and they showed that experts agreed among one another on average less than one time in two for the lower hierarchical levels of the classification (Eriksen et al., 2019; Hearn et al., 2011; Stevens et al., 2004; Ullerud et al., 2018). We showed similar agreement ratios among expert judgments for the alliance level, and lower ones for associations. However, the classification levels did not go down to phytosociological associations in these previous studies, and unit assignments were made in the field in a definite ecological context limiting the range of possible choices. In addition, the study areas were limited, whereas a large number of vegetation plots spread across mainland France were used for the present study, with a large range of 228 possible associations, 43 alliances and eight classes covering the diversity of the vegetation types of temperate and mountainous forests of Western Europe. This study shows a lack of consistency among phytosociology experts regarding the assignment of vegetation plots to vegetation units currently used in local management for the first time at a national scale and for the lower hierarchical levels of the phytosociological classification.

Sources of variability among phytosociology experts

Disagreements over habitat classifications are probably difficult to avoid completely because vegetation is inherently variable and complex (Cherrill, 2016; Rodwell et al., 2018). The sources of variability among phytosociology experts are manifold, and range from vegetation plots to determination keys or lack of one, and to the definition of vegetation units. The low agreement ratios for association assignment can be partly explained by the context of this study, as the expert organizations had to assign vegetation units mostly on an ex-post basis using floristic and plot location data instead of doing this on site. This hindered the use of the ecological context for assigning a vegetation unit, especially in forest ecosystems where the environmental context helps for assignment. Nevertheless, giving the same surveys to all the experts helped to attribute potential differences to their judgment and not to the potential variability in the lists of species observed on site. Secondly, the plots were located randomly on the map before the fieldwork phase although the phytosociological method recommends choosing homogeneous floristic and local ecological conditions to carry out a phytosociological survey (Braun-Blanquet et al., 1932). The experts tended to disagree on transitional surveys. However, NFI field agents had been reporting for a few years that approximately 5% of the randomly located surveys fell into two or more associations (NFI 2018, pers. comm.). Therefore, the random location may have destabilized the phytosociology experts but only for a minor part of the vegetation plots. The disagreement could be explained much more by the heterogeneous floristic composition resulting from dynamic stages than by spatial variability, which seemed to occur only for a minor part of the vegetation plots. Classification inconsistency was also pointed out as the main source of inter-observer differences in previous comparative studies of land-cover maps (Cherrill and McClean, 1999; Hearn et al., 2011; Ullerud et al., 2018). Gaps in knowledge about plant communities and overlapping typology categories (Willner, 2011) are some of the most likely sources of variability. We showed that experts tended to agree more on precisely defined or frequent units, the latter being widespread or well defined. The absence of explicit assignment rules is also a major source of variability because it leads to diverging definitions about part of the vegetation units. Last of all, we used identical assignments among expert organizations as an index of consistency. This methodological approach decreased the estimation of convergence because studying inter-observer variation (Cherrill and McClean, 1999, 1995; Hearn et al., 2011) systematically underestimates the congruence of the suggestions as compared to estimating a deviation from an ideal, 'true' value (Eriksen et al., 2019).

Automatic classification programs perform similarly to phytosociology experts in vegetation unit assignment

The present study assesses variability among several phytosociology experts and automatic classification programs for the first time. To estimate the efficiency and relevance of the

automatic classification programs, we did not wonder if they would suggest the same assignments as one phytosociology expert providing the "right" assignment, but if they could agree with several expert organizations more, less, or as often as one phytosociological expert would. The approach gave us the possibility to go beyond the levels of the phytosociology experts by being more consensual. The agreement ratios of the classification programs were lower than the agreement ratios of the expert organizations for association and alliance assignment. The Phi-program stood out by being really close to the experts for association assignment, and even above four of the five expert organizations for alliance assignment. Furthermore, for class assignment, the three automatic classification programs were above the average value of the five agreement ratios of the expert organizations, and two programs (the FPFI-program and the Phi-program) even outperformed all five expert organizations. Therefore, automatic classification programs can be more consensual than expert organizations at the highest hierarchical level of the classification. The ability of the automatic classification programs to reproduce expert judgment and suggest consensual assignments was all the better as the classification level was higher and the number of units decreased. Such increasing performance with the increase of the hierarchical level of the classification means that when the automatic classification program assigns a different association than the phytosociological experts, the assigned associations are 'nearby' and frequently belong to the same alliance and class. For the vegetation units specifically targeted by public policies and currently used in local management such as associations and alliances, although automatic classification programs can still be improved, this study confirms that they are efficient and that using them to assign floristic surveys to vegetation units is relevant. A further study with an *ad hoc* design will be needed to i) establish the exact vegetation units about which the experts tend to agree or disagree, and ii) compare all the existing automatic classification methods and indexes to establish their scope of validity and relevance according to the objectives.

<u>Complementarity of manual and automatic assignments</u>

Floristic surveys still have to be made manually in the field and are time-consuming. In contrast, it usually takes only a few minutes for a surveyor to assign a vegetation plot to a vegetation unit *in situ* once a floristic survey has been made. Therefore, in local studies when new observations have to be made, an automatic classification program is a complementary approach to field assignment. It can also be used to coach and check new phytosociology experts, or to make a choice in doubtful conditions. Classification programs could also be used for the unprecedented assignment of tens of thousands of pre-existing floristic surveys–the time needed for the expert organizations to fulfil the mission is incompatible with this survey size –, and for attuning databases merged from different organizations (Schaminee et al., 2009), creating standardized habitat assignment data. Creating habitat time series is crucial for the monitoring and the

conservation of endangered habitats. In particular, this could help to implement the mandatory monitoring of Natura 2000 habitats across the European territory, as requested by the Habitats Directive (EC Council 2006).

Towards unified and clarified vegetation classifications

The large number of vegetation units, of vegetation plots distributed across mainland French forests and the large panel of phytosociology experts involved in the study likely indicate that the results can be extended to other countries and to other types of ecosystems in Western Europe, in particular those with definitions focused on species. We suggest perspectives and corrective measures to address the possible lack of consistency of the sources used for assigning vegetation plots to vegetation units, and to avoid consequences on environmental assessments and professional decisions as described by Cherrill (2016). The results support the current trend toward unifying and filling the gaps of the existing classification (De Cáceres et al., 2018; Guarino et al., 2018; Mucina et al., 2016; Peterka et al., 2017), which will ultimately lead to fewer overlapping categories in typologies. This task should be an opportunity to specify and clarify the assignment rules of existing typologies, and to implement guiding tools such as vegetation keys. Lastly, the results support the implementation of intercalibration sessions among typology users at the regional and national scales to improve consistency among phytosociology experts and promote the use of assignment rules and vegetation keys when assigning plots to vegetation units.

Acknowledgements

We thank the providers of the NFI vegetation plots, as well as the people who collected floristic data and were involved in the assignment of the vegetation plots to the vegetation units at the NFI (Andre G., Bernard S., Bircker L., Blond W., Boithias B., Bourrinet F., Cano R., Daubigney L., Daviaud J.-F., De Taxis Du Poet T., Delayat J.-M., Delhaye S., Delquaire P., Hirsch B., Hugerot Y., Leclaire G., Ledeme G., Magnette F., Malemanche L., Michel C., Paque G., Payen D., Pedrot L., Pihou O., Rives J.-F., Vaillot J.-B.) and at the national forestry office (Blin M., Blondel F., Darnis T., Fallour-Rubio D., Gattus J.-C., Holveck P., Hum P., Keller J., Lathuillière L., Loustalot-Forest F., Ritz F., Rollier C., Témoin J.-L.). We also thank Anne Gégout-Petit for her help with statistics.

Data accessibility statement:

The data and the code used in this study are available from the corresponding author upon request.

REFERENCES

Archaux, F., Gosselin, F., Bergès, L., Chevalier, R. (2006) Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science*, 17, 299–306. https://doi.org/10.1111/j.1654-1103.2006.tb02449.x

Bardat, J., Bioret, F., Botineau, M., Boullet, V., Delpech, R., Géhu, J.-M., *et al*. (2004) *Prodrome des végétations de France.* Paris, Publication scientifique du Muséum national d'Histoire naturelle.

Bioret, F., Bensettiti, F., Royer, J.-M. (2014) Presentation of the Prodrome of French vegetation, from the first version (PVF1) to the second, declension at plant community level (PVF2). *Acta Botanica Gallica* 161, 395–402. https://doi.org/10.1080/12538078.2014.972978

Braun-Blanquet, J., Fuller, G.D., Conard, H.S. (1932) *Plant sociology: the study of plant communities*. New York and London, McGraw-Hill book company, inc.

Černá, L., Chytrý, M. (2005) Supervised classification of plant communities with artificial neural networks. J*ournal of Vegetation Science* 16, 407–414. https://doi.org/10.1111/j.1654-1103.2005.tb02380.x

Cherrill, A. (2016) Inter-observer variation in habitat survey data: investigating the consequences for professional practice. *Journal of Environmental Planning and Management* 59, 1813–1832. https://doi.org/10.1080/09640568.2015.1090961

Cherrill, A., McClean, C. (1999) Between-observer variation in the application of a standard method of habitat mapping by environmental consultants in the UK. *Journal of Applied Ecology* 36, 989–1008. https://doi.org/10.1046/j.1365-2664.1999.00458.x

Cherrill, A., McClean, C. (1995) An Investigation of Uncertainty in-Field Habitat Mapping and the Implications for Detecting Land-Cover Change. *Landscape Ecology* 10, 5–21. https://doi.org/10.1007/BF00158550

Chytrý, M., Tichý, L., Holt, J., Botta-Dukát, Z. (2002) Determination of diagnostic species with statistical fidelity measures. *Journal of Vegetation Science* 13, 79–90. https://doi.org/10.1111/j.1654-1103.2002.tb02025.x

Couvreur, J.-M., Fiévet, V., Smits, Q., Dufrêne, M. (2015) Evaluation of the "observer effect" in botanical surveys of grasslands. *Biotechnologie, Agronomie, Société et Environnement* 19, 132–142.

De Cáceres, M., Font, X., Vicente, P., Oliva, F. (2009) Numerical reproduction of traditional classifications and automatic vegetation identification. *Journal of Vegetation Science* 20, 620–628. https://doi.org/10.1111/j.1654-1103.2009.01081.x

De Cáceres, M., Franklin, S.B., Hunter, J.T., Landucci, F., Dengler, J., Roberts, D.W. (2018) Global overview of plot-based vegetation classification approaches. *Phytocoenologia* 48, 101–112. https://doi.org/10.1127/phyto/2018/0256

Dufrêne, M., Legendre, P. (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* 67, 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

EC Council (2006) *Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora*.

Eriksen, E.L., Ullerud, H.A., Halvorsen, R., Aune, S., Bratli, H., Horvath, P. *et al*. (2019) Point of view: error estimation in field assignment of land-cover types. *Phytocoenologia* 49, 135–148. https://doi.org/10.1127/phyto/2018/0293

European Commission DG Environment (2013) *Interpretation manual of European Union habitats. EUR 28*.

Gégout, J.-C., Coudun, C. (2012) The right relevé in the right vegetation unit: a new typicality index to reproduce expert judgement with an automatic classification programme. *Journal of Vegetation Science* 23, 24–32. https://doi.org/10.1111/j.1654-1103.2011.01337.x

Gégout, J.-C., Rameau, J.-C., Renaux, B., Jabiol, B., Bar, M. (2009) *Les habitats forestiers de la France tempérée ; typologie et caractérisation phytoécologique*. Nancy, AgroParisTech-ENGREF. ed.

Guarino, R., Willner, W., Pignatti, S., Attorre, F., Loidi, J.J. (2018) Spatio-temporal variations in the application of the Braun-Blanquet approach in Europe. *Phytocoenologia* 48, 239–250. https://doi.org/10.1127/phyto/2017/0181

Hearn, S.M., Healey, J.R., McDonald, M.A., Turner, A.J., Wong, J.L.G., Stewart, G.B. (2011) The repeatability of vegetation classification and mapping. *Journal of Environmental Management* 92, 1174–1184. https://doi.org/10.1016/j.jenvman.2010.11.021

Hervé, J.-C. (2016) France. In: *National Forest Inventories - Assessment of Wood Availability*. Springer International Publishing, pp. 385–404.

Kočí, M., Chytrý, M., Tichý, L. (2003) Formalized reproduction of an expert-based phytosociological classification: A case study of subalpine tall-forb vegetation. *Journal of Vegetation Science* 14, 601–610. https://doi.org/10.1111/j.1654-1103.2003.tb02187.x

Milberg, P., Bergstedt, J., Fridman, J., Odell, G., Westerberg, L. (2008) Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science* 19, 633–644. https://doi.org/10.3170/2008-8-18423

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

Mucina, L., Bueltmann, H., Dierssen, K., Theurillat, J.-P., Raus, T., Carni, A., *et al*. (2016) Vegetation of Europe: hierarchical floristic classification system of vascular plant,

bryophyte, lichen, and algal communities. *Applied Vegetation Science* 19, 3–264. https://doi.org/10.1111/avsc.12257

Olden, J.D., Joy, M.K., Death, R.G. (2006) Rediscovering the species in community-wide predictive modeling. *Ecological Applications* 16, 1449–1460. https://doi.org/10.1890/1051-0761(2006)016[1449:RTSICP]2.0.CO;2

Oliver, I., Broese, E.A., Dillon, M.L., Sivertsen, D., McNellie, M.J. (2013) Semi-automated assignment of vegetation survey plots within an a priori classification of vegetation types. *Methods in Ecology and Evolution* 4, 73–81. https://doi.org/10.1111/j.2041-210x.2012.00258.x

Peterka, T., Hajek, M., Jirousek, M., Jimenez-Alfaro, B., Aunina, L., Bergamini, A., *et al*. (2017) Formalized classification of European fen vegetation at the alliance level. *Applied Vegetation Science* 20, 124–142. https://doi.org/10.1111/avsc.12271

Pinto, P.E., Dupouey, J.-L., Hervé, J.-C., Legay, M., Wurpillot, S., Montpied, P., Gégout, J.-C. (2016) Optimizing the bioindication of forest soil acidity, nitrogen and mineral nutrition using plant species. *Ecological Indicators* 71, 359–367. https://doi.org/10.1016/j.ecolind.2016.05.047

Prosser, M.V., Wallace, H.L. (1999) *Countryside survey 2000 quality assurance exercise.* Ecological Surveys (Bangor). Canon Pyon, UK.

Rodwell, J.S., Evans, D., Schaminee, J.H.J. (2018) Phytosociological relationships in European Union policy-related habitat classifications. *Rendiconti Lincei. Scienze Fisiche e Naturali* 29, 237–249. https://doi.org/10.1007/s12210-018-0690-y

Schaminee, J.H.J., Hennekens, S.M., Chytry, M., Rodwell, J.S. (2009) Vegetation-plot data and databases in Europe: an overview. *Preslia* 81, 173–185.

Stevens, J.P., Blackstock, T.H., Howe, E.A., Stevens, D.P. (2004) Repeatability of Phase I habitat survey. *Journal of Environmental Management* 73, 53–59. https://doi.org/10.1016/j.jenvman.2004.05.009

Stiles, J., Lemme, C., Sondur, S., Morshidi, M., Manshardt, R. (1993) Using Randomly Amplified Polymorphic Dna for Evaluating Genetic-Relationships Among Papaya Cultivars. *Theoretical and Applied Genetics* 85, 697–701. https://doi.org/10.1007/BF00225007

Tichý, L. (2005) New Similarity Indices for the Assignment of Relevés to the Vegetation Units of an Existing Phytosociological Classification. *Plant Ecology* 179, 67–72. https://doi.org/10.1007/s11258-004-5798-8

Tichý, L., Chytrý, M., Botta-Dukát, Z. (2014) Semi-supervised classification of vegetation: preserving the good old units and searching for new ones. *Journal of Vegetation Science* 25, 1504–1512. https://doi.org/10.1111/jvs.12193

Tichý, L., Chytrý, M., Landucci, F. (2019) GRIMP: A machine-learning method for improving groups of discriminating species in expert systems for vegetation classification. *Journal of Vegetation Science* 30, 5–17. https://doi.org/10.1111/jvs.12696

Ullerud, H.A., Bryn, A., Halvorsen, R., Hemsing, L.Ø. (2018) Consistency in land-cover mapping: Influence of field workers, spatial scale and classification system. *Applied Vegetation Science* 21, 278–288. https://doi.org/10.1111/avsc.12368

van Tongeren, O., Gremmen, N., Hennekens, S. (2008) Assignment of releves to pre-defined classes by supervised clustering of plant communities using a new composite index. *Journal of Vegetation Science* 19, 525–536. https://doi.org/10.3170/2008-8-18402

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/1http://10.0.4.87/j.1654-1103.2007.tb02553.x

Willner, W. (2011) Unambiguous assignment of releves to vegetation units: the example of the Festuco-Brometea and Trifolio-Geranietea sanguinei. *Tuexenia* 271–282.

SUPPORTING INFORMATION

Appendix S1. Method used to estimate the impact on the experts' agreement ratios of the inclusion of an inexperienced phytosociology expert in the panel of reference expert organizations.

Appendix S2. Highest possible agreement ratios that might be reached by each expert organization when comparing the expert organizations among them, given their divergence among them; and highest possible agreement ratios that might be reached by any automatic classification program as compared to the five expert organizations given the divergence among the five expert organizations, regarding association, alliance and class assignment.

Appendix S3. Comparison among the five expert organizations, and among them and 1,000 random draws regarding association, alliance and class assignment: agreement ratios, and mean agreement ratios and corresponding standard deviations, respectively.

TABLE WITH LEGEND

Table 1: Pairwise agreement ratios between expert organizations for the assignment of the 273 vegetation plots to associations (left-down corner, in italic) and alliances (right-up corner, in bold).

	Bot. Cons.[a]	For. Invent.[b]	For. Manag.[c]	Eng. Off.[d]	Museum[e]
Bot. Cons.[a]	*1*	**0.47**	**0.51**	**0.48**	**0.51**
For. Invent.[b]	*0.34*	*1*	**0.50**	**0.44**	**0.45**
For. Manag.[c]	*0.29*	*0.24*	*1*	**0.46**	**0.44**
Eng. Off.[d]	*0.27*	*0.25*	*0.26*	*1*	**0.48**
Museum[e]	*0.32*	*0.22*	*0.25*	*0.29*	*1*

[a] Bot. Cons.: the National Botanical Conservatory of the Massif Central

[b] For. Invent.: the National Forest Inventory of France

[c] For. Manag.: the botanists' network of the National Forestry Office of France

[d] Eng. Off.: Biotope, an engineering office

[e] Museum: the National Museum of Natural History of France

FIGURES WITH LEGENDS

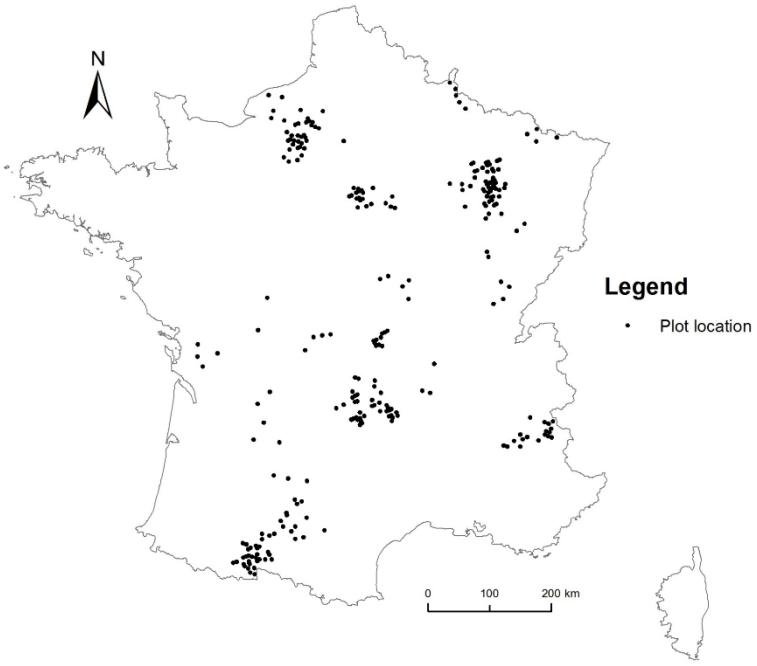

Figure 1: Map of the 273 vegetation plots surveyed between May and September 2013 by six teams of the National Forest Inventory of France (NFI) across the forests of mainland France.

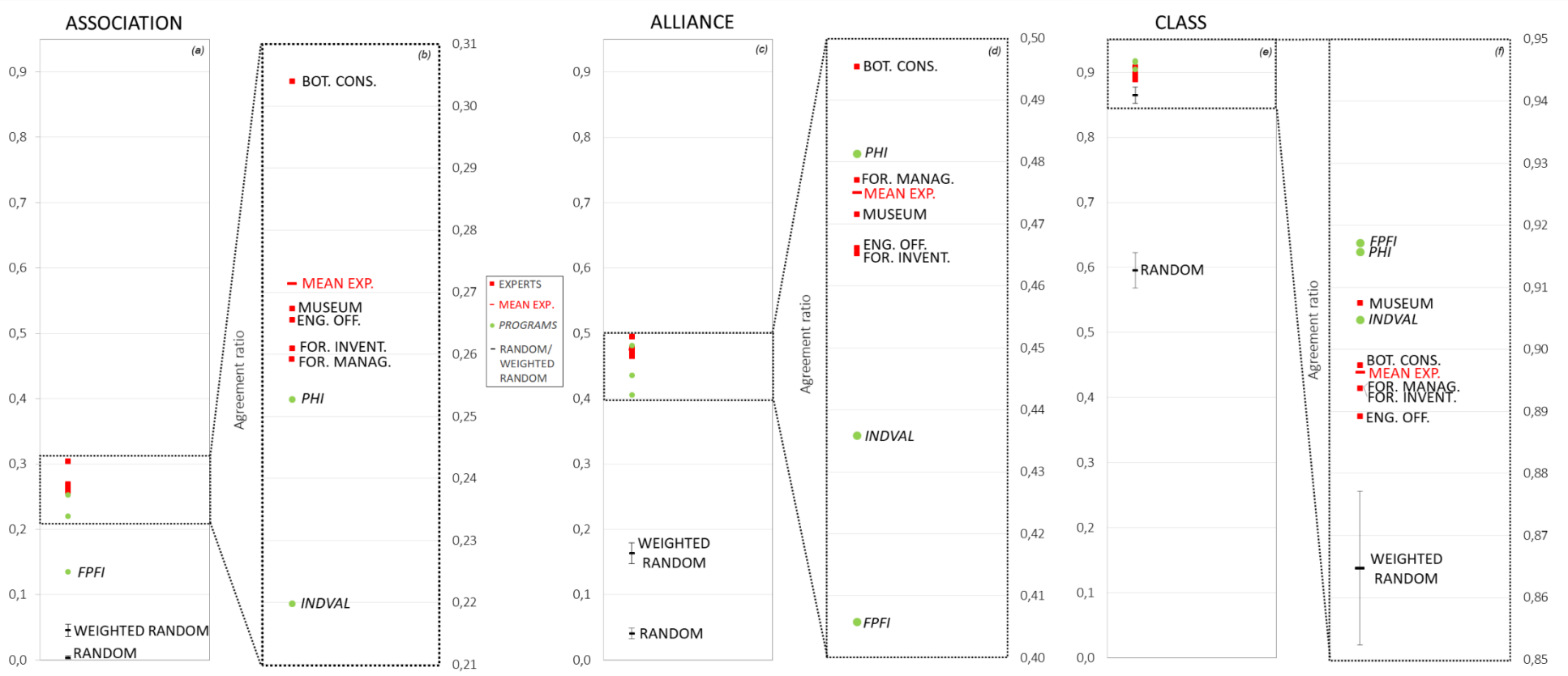
ASSOCIATION
ALLIANCE
CLASS
(a)
(b)
(c)
(d)
(e)
(f)
Agreement ratio
BOT. CONS.
MEAN EXP.
MUSEUM
ENG. OFF.
FOR. INVENT.
FOR. MANAG.
PHI
INDVAL
FPFI
WEIGHTED RANDOM
RANDOM
EXPERTS
MEAN EXP.
PROGRAMS
RANDOM/ WEIGHTED RANDOM

Figure 2: Agreement ratios among expert organizations, between the automatic classification programs and the expert organizations, and between random assignment and the expert organizations regarding association assignment (a, b), alliance assignment (c, d) and class assignment (e, f).

Experts: the National Botanical Conservatory of the Massif Central (Bot. Cons.), the National Forest Inventory of France (For. Invent.), the botanists' network of the National Forestry Office of France (For. Manag.), an engineering office called Biotope (Eng. Off.), and the National Museum of Natural History of France (Museum)

Mean Exp.: average value of the five agreement ratios of the expert organizations.

Programs: automatic classification programs using the Phi-coefficient (*Phi*), IndVal (*IndVal*), or FPFI (*FPFI*).

Random/Random weighted: 1,000 random or weighted random draws.

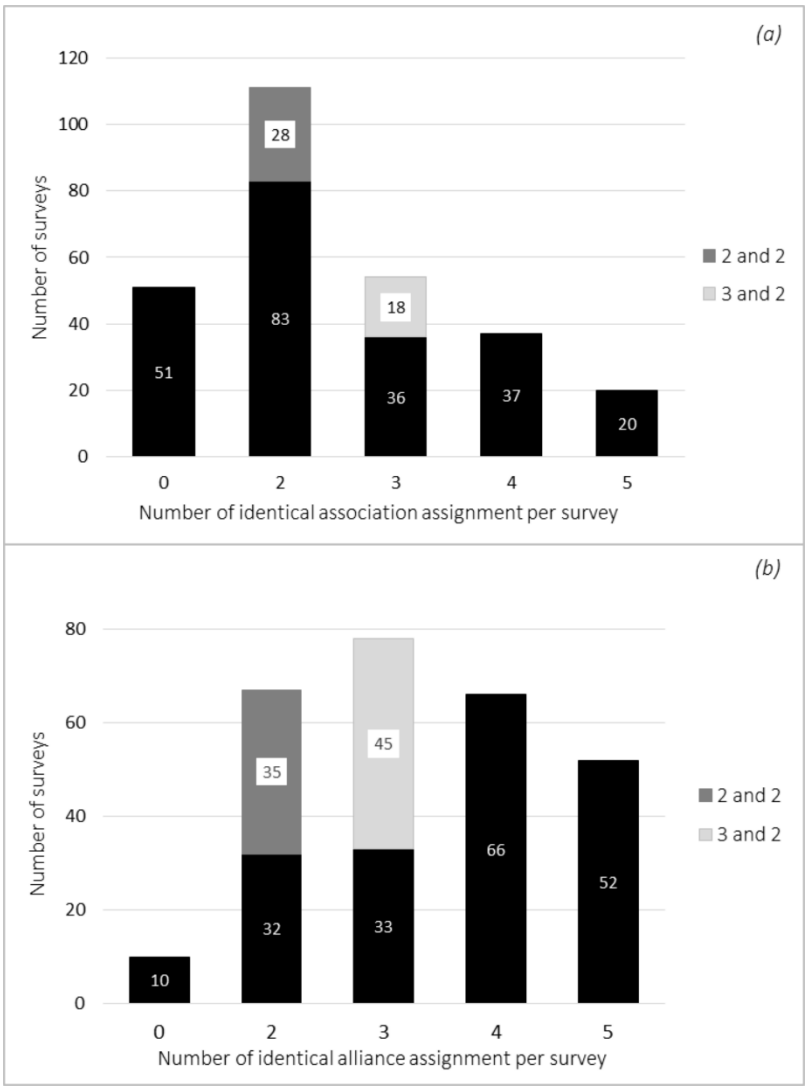

Figure 3: Number of vegetation plots per number of identical association assignments (a) and identical alliance assignments (b). “2 and 2” means that two expert organizations assigned the vegetation plot to the same vegetation unit, and two other expert organizations agreed on another assignment. “3 and 2” means that three expert organizations assigned the vegetation plot to the same vegetation unit, and two other expert organizations agreed on another assignment.

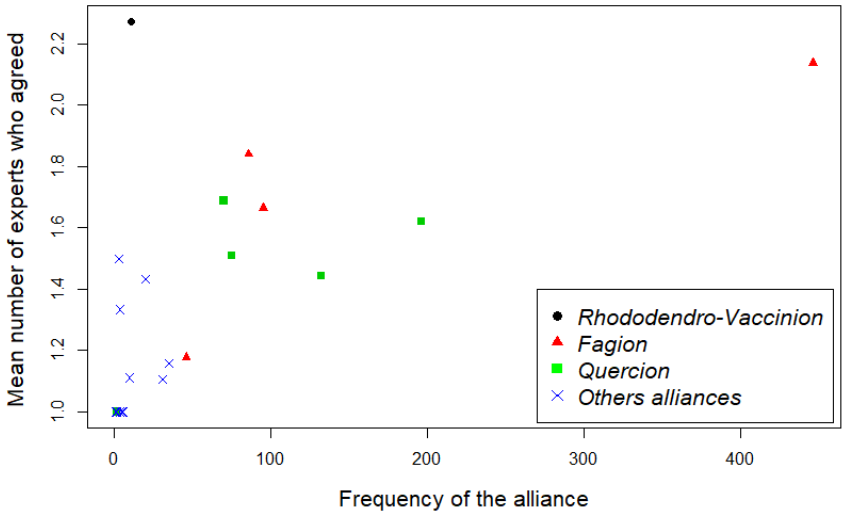

Figure 4: Mean number of experts who agreed *per* alliance as a function of the frequency of the alliance.