More species, less effort: Designing and comparing sampling strategies to draft optimised floristic inventories

Marco D'Antraccoli[a,*], Giovanni Bacaro[b], Enrico Tordoni[b], Gianni Bedini[a], Lorenzo Peruzzi[a]

[a] Department of Biology, Unit of Botany, University of Pisa, Via Derna 1, 56126, Pisa, Italy
[b] Department of Life Sciences, University of Trieste, Via L. Giorgeri 10, 34127 Trieste, Italy

ARTICLE INFO

Keywords:
Algorithms
Flora
NDVI
Sampling optimisation
Species diversity
Species inventory

ABSTRACT

Floristic inventories are an essential part of basic and applied research in botany. Despite their long history, floristic research is still carried out following non-objective (preferential) sampling approaches. Accordingly, final outputs (i) are extremely variable in the quality and quantity of collected data and hardly repeatable, (ii) rely on the researcher ability, and (iii) miss the basic assumptions to allow inferential statistical analyses. The aim of this work is to explore the drafting of a floristic inventory by means of geostatistical approaches to locate sampling units (plots) in the study area. We planned, carried out and then compared two different sampling strategies: (i) 'basic strategy', a stratified random sampling design based solely on a spatial optimization criterion (no prior information is available), and (ii) 'advanced strategy', a sampling design based on the maximisation of the spectral heterogeneity among sampling units, quantified in terms of Normalized Difference Vegetation Index values (NDVI). The strategy that maximises collected floristic information was assessed based on a combination of descriptive and quantitative statistics, such as (i) the completeness of the floristic inventory, (ii) the steepness of the rarefaction curves, (iii) the sampling time effort, and (iv) the plot contribution to the total β diversity. The 'advanced strategy' detects more species than the 'basic strategy' in all the sampling sites. The 'advanced strategy' accumulates species more quickly than the 'basic strategy'. The 'advanced strategy' selects sampling units more homogeneously contributing to total β diversity; in addition, they are better spatially arranged across the study area to capture environmental peculiarities of sampling sites. The 'advanced strategy' needs a little more effort in the design of the sampling strategy, but it is more effective than the 'basic strategy' in drafting a species inventory. We provide here the R routine to perform the 'advanced strategy', which can be profitably and freely used in any other geographic location and vegetation context.

1. Introduction

Documenting plant diversity distribution through appropriate and efficient sampling strategies is a crucial step to acquire a reliable knowledge of natural resources and to promote its efficient conservation (Gaston, 2000; Newmaster et al., 2005). On the other hand, measuring plant diversity is an incredibly complex task, considering spatial and temporal variability (Rocchini et al., 2018) and the large number of existing measurement metrics (Chiarucci et al., 2011; Scheiner, 2019). Among these metrics, the species richness (α diversity) and the variation in community composition among sites in a region of interest (β diversity; Whittaker, 1972) are commonly used.

The traditional approach to document vascular plant diversity of a given area consists in the preparation of floristic lists, commonly known as 'floras' (Palmer et al., 2002; D'Antraccoli et al., 2019). Floras can be defined as species inventories for a given territory in a given time, aimed to link taxonomic, geographic, and temporal information (Palmer et al., 1995; Chiarucci and Bonini, 2005; Bedini et al., 2016). From a sampling perspective, floristic inventories are usually compiled by subjectively searching for and collecting plants, trying to obtain an exhaustive list of species over a large area, even though this task is often very challenging (Chiarucci and Palmer, 2005; Vondrák et al., 2016; Chiarucci et al., 2018). With this approach, generally referred to as 'preferential' or as 'opportunistic', surveyors select specific sites or habitats based on their subjective expectations (Chiarucci et al., 2018). Palmer et al. (2002) defined the highly subjective combination of ability, experience, expertise and intuition that guide the botanist in the field as the "*botanic internal algorithm*". Wishing to increase the overall

* Corresponding author.
E-mail addresses: marco.dantraccoli@biologia.unipi.it (M. D'Antraccoli), gbacaro@units.it (G. Bacaro), etordoni@units.it (E. Tordoni), gianni.bedini@unipi.it (G. Bedini), lorenzo.peruzzi@unipi.it (L. Peruzzi).

number of species per survey, botanists tend to move to another area when the time required to find a new species becomes too long (ter Steege et al., 2011). Furthermore, the ratio of common to rare species is likely to vary with many factors, such as time availability, familiarity of the botanist with the local flora, botanist weariness, understory cover, proportion of rare or elusive species (Scott and Hallam, 2002; Archaux et al., 2006), as well as habitat and population features (Morrison, 2016).

On the other hand, in a probabilistic sampling approach species are identified and listed only if occurring in *a priori* selected sampling units, generally corresponding to plots (Chiarucci et al., 2018). In this approach, independent sampling units can be selected, thus satisfying the condition that each site has the same probability of being sampled. The most commonly used probabilistic sampling strategies are the 'simple random' design, the 'systematic sampling' and the 'stratified random sampling'. While the first two sampling approaches require minimal *a priori* information and are easy to use in terms of design complexity (Elzinga et al., 1998; Daniel, 2012), the stratified random sampling can be applied only when environmental information about the study area is available. Considering current limited resources allocated to field surveys, the selection of sampling design parameters is crucial to optimise the amount of information collected within the time and budget constraints set for field work. Several parameters may affect this trade-off, such as the number, size, shape, and spatial configuration of sampling units (Chiarucci et al., 2001; Bacaro et al., 2015; Güler et al., 2016; Hoffmann et al., 2019). However, an 'ideal sampling' does not exist. Hoffmann et al. (2019) recommended an operational approach, in which the sampling scale is chosen based on clear and repeatable criteria rather than vegetation characteristics. Accordingly, sampling optimisation issues are of both theoretical and practical concern.

Although a great deal of research is being conducted on conceiving efficient sampling methods in ecology (Stohlgren, 2007; Gonzalez-Oreja et al., 2013; Hoffmann et al., 2019), this issue is almost unexplored in floristic research. Palmer et al. (2002) highlighted that preferential sampling is expected to be more efficient than probabilistic sampling, since botanists "generally have a strong intuition or 'educated guess' about where to direct one's effort". Again, the total area sampled by probabilistic methods is severely lower than that covered by preferential methods, as quantified by Golodets et al. (2013). Accordingly, the same authors inferred that using just over one-tenth of the workforce and resources required by a probabilistic method, a preferential strategy detected twice the number of species. On the other hand, the use of probabilistic approaches to draft a floristic inventory is encouraged (Palmer et al., 2002; Chiarucci and Bonini, 2005; Wiser et al., 2011), in order to allow rigorous statistical analyses (Chiarucci et al., 2018). Indeed, only probabilistic approaches meet the conditions of randomness, known probability and independence of statistical sampling (Chiarucci, 2007), thus (i) allowing comparisons among different regions and times or among unequal sampling efforts (Colwell and Coddington, 1994; Gotelli and Colwell, 2001; Koellner et al., 2004; Kalkhan et al., 2007), (ii) avoiding the onset of artefacts on the analyses (Palmer et al., 2008), and (iii) facilitating the reproducibility of methods. Maximising species obtained per unit of sampling effort reduces the cost of the inventorying activities in terms of personnel time, supplies, and environmental impact (Yoccoz et al., 2001; Baffetta et al., 2007).

One of the major issues in collecting field data is the autocorrelated structure of communities and ecosystem processes, which can lead to spurious results during data analysis (Legendre, 1993). Indeed, according to 'distance decay of similarity' law (Tobler, 1970; Nekola and White, 1999) which states that "*everything is related to everything else, but near things are more related than distant things*", we can assume that species replacement along ecological gradients (Legendre, 2014) tends to increase by increasing the distance among sampling units, because habitats and environmental conditions are expected to be less similar (Kunin, 1997; Stohlgren, 2007; Chiarucci et al., 2009; Dengler, 2009). Spatial autocorrelation may strongly influence analyses (Kühn, 2007) and affect cost efficiency (Bacaro, 2008) also in a sampling perspective. Accordingly, spatially-explicit methods should be preferred in order to decrease the number of samples (i.e., reduce sampling effort), to improve sampling accuracy (Haining, 2003; Wang et al., 2012), and to provide robust estimates of biodiversity at different hierarchical levels, from individuals to communities (Rocchini et al., 2018 and literature therein). In the perspective of the above-cited spatially-explicit methods, remote sensing approaches allow "*the acquisition of information about Earth's surface without being in physical contact with it*" (Wegmann and Leutner, 2016) thus representing a promising tool in biodiversity studies (Rocchini et al., 2018) and providing important ecological data that are explicit in the spatial domain. Remote sensing samples reflect and emit electromagnetic radiation from the Earth's ecosystems. Spectral data can be easily acquired from a plethora of online portals, and then processed by the user in several ways, depending on the study purposes. The most frequently used vegetational index is the Normalised Difference Vegetation Index (NDVI), first proposed by Rouse et al., 1974, and based on the use of near-infrared and red wavebands. NDVI quantifies vegetation greenness, and is useful for understanding vegetation cover types. It has been shown that remote sensing is both efficient and appropriate to plan an objective sampling design (see Rocchini et al., 2010 for a review). Palmer et al. (2002) formulated the so-called 'Spectral Variation Hypothesis' (hereafter SVH), which states that spectral variation in remotely sensed images is expected to be related to environmental heterogeneity and therefore could serve as a proxy of species diversity. Numerous strategies have been tested to quantify the spectral variation of a satellite imagery to provide a link with species diversity (e.g., Podolsky, 1994; Gould, 2000; Foody and Cutler, 2003; Rocchini, 2007; Rocchini et al., 2009; Oldeland et al., 2010). The validity of these approaches is highly dependent on several factors, such as the type of vegetation investigated, and the metrics derived from remotely sensed information to estimate spectral heterogeneity (Rocchini et al., 2010, 2018).

Our study aims at comparing the efficiency of two floristic sampling strategies (hereafter named as 'basic' and 'advanced' strategies), based on probabilistic approaches, which show an increasing design complexity. In the 'basic strategy', the study area is divided into sampling sites coarsely homogeneous on vegetational grounds, then sampled using a purely spatial optimisation algorithm. On the contrary, the 'advanced strategy' is based on the SVH, and explores the integration within the sampling design of additional, more detailed, ecological information from remote sensing data. To offer the widest chances of application of the method, we designed sampling strategies assuming no availability of previous vegetation or habitat cartography.

As an indicator of sampling efficiency, we compared three different criteria between the selected strategies: (i) the trend of plot-based rarefaction curves, (ii) the time-effort, and (iii) the analysis of the spatial structure of β diversity among sampling units. The resulting 'optimal' sampling strategy should allow the sampling of communities as diverse as possible from the compositional point of view. The latter point is crucial in sampling optimisation, since a redundant sampling increases the risk of obtaining similar floristic inventories, determining a waste of time and financial resources, which could be allocated to other sampling localities.

2. Material and methods

2.1. Study area

The study area is the central and northern portions of the Migliarino-San Rossore-Massaciuccoli Regional Park (MSRM Park; Fig. 1), an area well documented in floristic literature. Within this area, we chose three sampling sites coarsely homogeneous on vegetational grounds, as a proxy for the general habitat variation found in the whole area. Specifically, the selected sites were: Site A (sand dune vegetation,

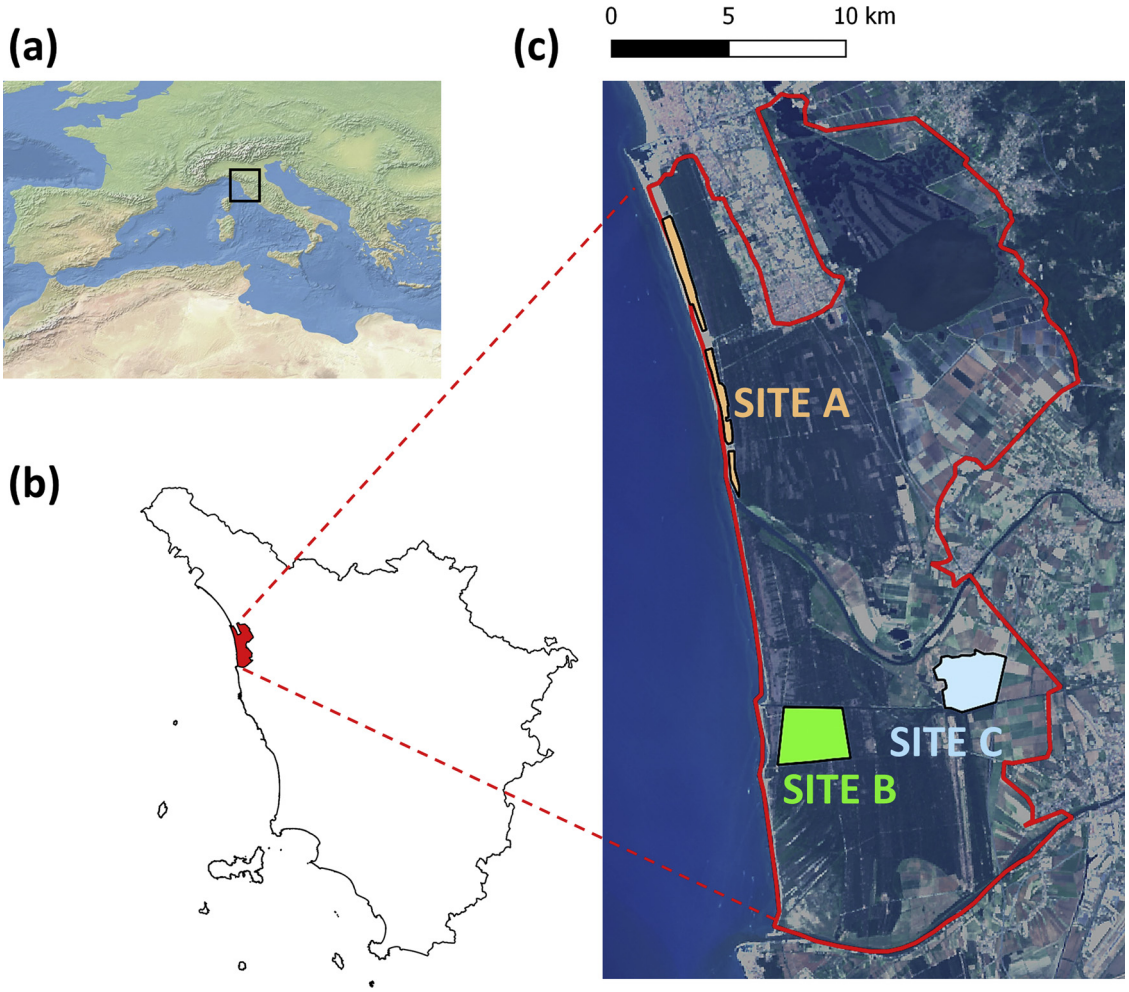

Fig. 1. (a) Location of the study area with respect to Mediterranean basin. (b) Tuscany region, in red the macro-area where sampling sites were selected. (c) Particular of the study area showing the three sampling sites. To draft the satellite map, we used the spectral images (red, green, and blue bands) from Landsat 8 as provided by U.S. Geological Survey (U.S. Geological Survey (USGS), 2018) (for interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

1.28 Km^2), Site B (thermophilous forests and maquis, 2.39 Km^2), and Site C (patches of thermophilous forests, hygrophilous forests, and open spaces, 2.07 Km^2). Previous information on floristic inventories in the sampling sites were extracted from the online geodatabase Wiki-plantbase #Toscana (Peruzzi and Bedini, 2015), in which all the floristic records available for the MSRM have been stored (see Supplemental Material S1 for the species lists).

2.2. Sampling strategies

Two different probabilistic sampling strategies characterised by an increasing design complexity were adopted: in the 'basic strategy', we applied a spatially optimised sampling design by minimising spatial autocorrelation among sampling units (hereafter, plots), by just assuming a gross *a priori* selection of the three environmentally diverse sampling sites. In the 'advanced strategy', we added information about environmental heterogeneity applying the principles of the SVH. Further details on both strategies are described below.

For each strategy, we selected 15 square plots of 100 m^2 per site, for a total of 45 plots. The number of plots and plot size were chosen to guarantee a good trade-off between statistical robustness and sampling effort (Stohlgren et al., 1997; Maccherini et al., 2020). We selected 10 × 10 m plots considering this size as able to capture diversity in different habitat types (from sand dunes to forests) and to ensure plots inventories as complete as possible. The plot orientation was randomly selected from the four cardinal directions (North, South, East, West). All plots were located in the field with a GPS Garmin Oregon 600 (accuracy 4 m); sampling frequency was bimestrial, from March to November (5 temporal replicates in total), in order to fully cover the phenological season and to avoid temporal sampling bias. The 'basic strategy' was carried out in 2017, and the 'advanced strategy' in 2018. Following our sampling design, a total of 450 sampling replicates were recorded.

For each plot, we drafted a species list, along with the time needed to reach the plot, to lay the quadrat and to collect species occurrences. Sampling time was recorded as the time needed to draft the floristic inventory of a plot in the field, whereas the time to reach a plot was measured by recording the travelling time spent from a plot to the next one (for the first sampling unit in a given sampling day, we used the access to the sites as fixed point).

Whenever possible, species were identified directly in the field; when direct identification was uncertain or not possible, we collected specimens outside the sampling area. Identification in the laboratory was performed through analytical keys available in floras (e.g., Pignatti, 1982; Tison and de Foucault, 2014) and in monographs of given taxonomic groups. Collected specimens were deposited at PI herbarium (acronym follows Theirs, 2019 onwards). Nomenclature and taxonomic circumscription follow the Italian Checklists of vascular flora (Bartolucci et al., 2018; Galasso et al., 2018).

2.3. Basic sampling strategy

The 'basic strategy' was applied using the algorithm provided by the R package 'spcosa' (Walvoort et al., 2010), which computes stratified random samplings. Specifically, the algorithm partitions a spatial object into compact strata of equal area using k-means clustering, with the objective function to minimise the mean squared shortest distance (MSSD) among strata centroid (for further analytical details, see Walvoort et al., 2010). This approach ensures a spatially optimised sampling design, since it drafts sampling strata showing optimal coverage and disposition across the sampling site, hence reducing spatial autocorrelation among sampling units. Furthermore, using this algorithm, strata are determined on an equally-sized partition of the whole sampling area, thus avoiding any undesirable edge effect. The number of strata was set up as n = 15 for each site, and one plot per stratum was randomly selected to match the total sampling effort planned per site.

2.4. Advanced sampling strategy

To perform the 'advanced strategy', we firstly acquired satellite images of the study area in the form of raster layers, with a spatial resolution of 30 m from Landsat 8, as provided by U.S. Geological Survey (U.S. Geological Survey (USGS), 2018). Red and infrared bands (see details at Supplemental S2.1) were used to obtain the NDVI maps of the three sampling sites, in the form of raster layers through the following formula.

$$NDVI_{[x,y]} = \frac{NIR_{[x,y]} - RED_{[x,y]}}{NIR_{[x,y]} + RED_{[x,y]}}$$

RED and NIR are the values of red and near infrared bands at spatial location [*x, y*], respectively. NDVI can vary between -1.0 and 1.0: negative values correspond to water, values close to zero (~ -0.1–0.1) correspond to barren rocky, sandy, or snowy areas; sparse vegetation such as shrubs and grasslands may result in moderate NDVI values (~ 0.2–0.5), whereas high NDVI values (~ 0.6–0.9) correspond to dense vegetation such as forests (Neigh et al., 2008; Wegmann and Leutner, 2016).

Then, we wrote an R routine (R Core Team, 2019; code available in Supplemental S2.2) to iteratively displace random points on sampling sites and then to select the best spatial configuration following two given conditions: (i) maximising NDVI variance among plots and ii) maximising their dispersion within the sampling area. These two conditions should ensure that the higher NDVI variance cope with spatial dispersion of sampling units, in order to reduce as much as possible habitat similarities among plots, resulting thus in a more complete final floristic inventory for the study sites selected. This function works in 4 fundamental steps, as follows (names follow the R code): once defined the number of iterations '*perm*', the number of plots to sample '*samp*', and an arbitrarily chosen quantile '*quant*', the algorithm (i) explores *perm* random configurations of *samp* plots within the NDVI map (here, *perm* = 10^5 and *samp* = 15); (ii) for each configuration, it measures both the NDVI variance and the distance of each plot from the plots' centroid; (iii) it filters all the solutions with NDVI variance above the *quant* value (here, *quant* = 0.99); (iv) it selects the solution showing the maximum spatial median value of plots distances, according to the algorithm published by Vardi and Zhang (1999). The spatial configuration of sampling points for both strategies is shown in Fig. 2.

2.5. Data analysis

Floristic inventory completeness was estimated for each strategy and site by comparing the number of species detected in our work against that reported in Wikiplantbase #Toscana.

In order to assess differences in floristic compositions returned by 'basic strategy' and 'advanced strategy' at each spatial scale (whole study area, sampling sites), we used multivariate permutational analysis of variance (PERMANOVA, Anderson, 2001) with Jaccard dissimilarity and 4999 permutations of the residuals under the reduced model and Type III sums of squares. We tested the following factors: 'Site' (fixed, three levels), 'Strategy' (fixed, two levels), along with the interaction term 'Strategy × Site'. In case of significant terms, these were investigated using *a posteriori* pairwise comparisons with *t* statistic and 4999 permutations. The relationship among plots in terms of floristic composition and adopted strategy was visualised through a Non-metric MultiDimensional Scaling (NMDS; Kruskal and Wish, 1978). Moreover, the correlation between mean plot sampling time and plot species richness was assessed through the Spearman's correlation test, whereas differences of sampling times were tested through the analysis of variance (ANOVA).

The rate of accumulation of species against sampling effort was assessed through Spatially Explicit Rarefactions (SERs, Bacaro et al., 2016), which unlike standard plot-based rarefaction curves take into account the spatial structure of the data (Bacaro et al., 2012). However, to improve the accounting for the spatial configuration of sampling units, we chose to apply a new algorithm, namely the here-defined function *centroid_pattern()*, written in R programming language (see Supplemental S2.3 for the code and further details).

To estimate the quota of additional plots to sample for inventory completeness (A), we used this formula:

$$A = \frac{N_{tot} - n_{samp}}{\frac{\sum_{i=1}^{n-1} nspecies_{i+1} - nspecies_i}{(n-1)}}$$

where n = number of sampled plots, N_{tot} = number of species occurring in the site, n_{samp} = number of detected species, $nspecies_i$ = mean number of detected species in i plots according to the rarefaction curve. Accordingly, the estimated number of additional plots to reach the inventory completeness (A) is determined by the number of species still undetected over the mean number of additional species found when a new plot is added. The assumption that the accumulation rate of species remains constant during the addition of new plots is a severe overestimation, thus determining an extremely conservative estimate. Accordingly, in our intentions the estimate does not play a predictive role, but instead it is just useful to compare the two sampling strategies.

In order to allow a visual comparison of α and β diversity spatial patterns between strategies, we performed a stochastic kriging procedure based on a Gaussian Spatial Simulation algorithm (Goovaerts, 1999), using the 'gstat' R package (Gräler et al., 2016). This method was selected because of its ability to grasp heterogeneity area while preserving its variance with respect to classical kriging methods (Zhao et al., 2018 and references therein), and allowing also the realisation of Uncertainty Maps. Semivariogram parameters selected to draft the maps are available in Supplemental S2.4. Concerning α diversity, species richness was selected as response variable, whereas the contribution of each plot to the total β diversity detected by each strategy was quantified using Local Contributions to Beta Diversity (hereafter LCBD, Legendre and De Cáceres, 2013). Let Y be a presence-absence dataset structured as plots × species, having p species and n sampling units, the LCBD of plots is then defined as follows:

$$\text{LCBDs} = \frac{\sum_{j=1}^{p} (y_{ij} - \bar{y}_i)^2}{\sum_{i=1}^{n} \sum_{j=1}^{p} (y_{ij} - \bar{y}_i)^2}$$

In this formula, i represents the ith sampling unit, j the jth species, y_{ij} individual values in the dataframe Y. LCBD represents the degree of uniqueness of a plot in terms of floristic composition, and larger values of LCBDs indicate sites whose species composition strongly differs from that of a mean site. This statistic was computed through the R package 'adespatial' (Dray et al., 2019). In our expectations, the best strategy should be able (i) to increase LCDB absolute values (increase plot

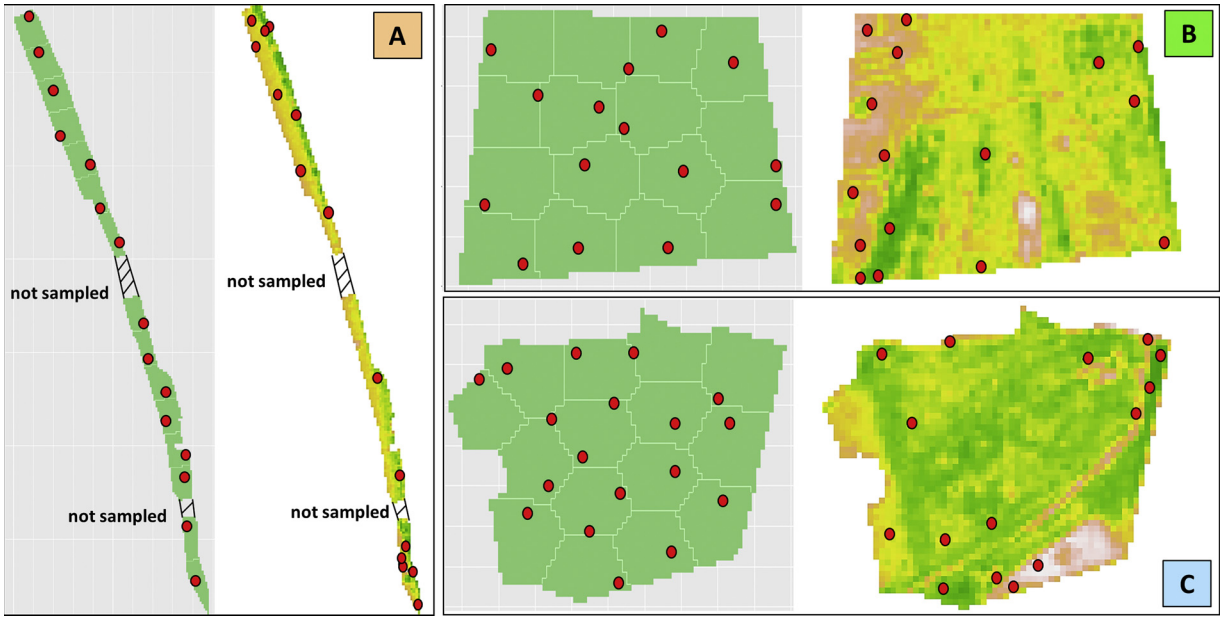

Fig. 2. Spatial pattern of sampling units in the three sites (A), (B), and (C). In each box: on the left the 'basic strategy', on the right the 'advanced strategy'.

compositional distances from the plots' centroid) and (ii) to minimise differences among LCDB values (compositional distances from the centroid are similar): such a pattern would denote a tendency to select sampling units contributing more homogeneously to total β diversity.

All the above-mentioned analyses were performed in R 3.6.0 (R Core Team, 2019), except for PERMANOVA and NMDS analyses, computed with the software PRIMER 6 (Clarke and Warwick, 2005) with the add-on package PERMANOVA + (Anderson et al., 2008).

3. Results

3.1. Floristic composition of the inventories

Considering the whole study area, plot species richness ranged from 3 to 35 for the 'basic strategy', with a mean value of 18.5. Concerning the 'advanced strategy', species richness ranged from 6 to 46, with a mean value of 26.8. In Site A, mean species richness was 20.6 (range: 9–29) for the 'basic strategy', and 27.1 (range: 14–46) for the 'advanced strategy'. In Site B, mean species richness was 20.9 (range: 11–35) for the 'basic strategy', and 30.1 (range: 22–43) for the 'advanced strategy'. In Site C, mean species richness in plots were 13.9 (range: 3–26) for the 'basic strategy', and 23.1 (range: 6–38) for the 'advanced strategy'. As showed in Table 1, the 'advanced strategy' consistently allowed the detection of a higher number of species.

PERMANOVA outputs showed a significant interaction between the terms 'Site' and 'Strategy' (pseudo-F = 1.48, $p < 0.01$, see Table 2); floristic composition of plots among sites and strategies is graphically shown by NMDS (Fig. 3). *Post-hoc* tests highlighted that sites differ in both strategies ($p < 0.001$), whereas floristic composition yielded by the two strategies did not differ in Site A ($t = 1.05$, $p > 0.05$). On the contrary, in Site B ($t = 1.28$, $p < 0.05$) and Site C ($t = 1.47$, $p < 0.01$) significant differences were highlighted (see Supplemental S2.5 for *a posteriori* pairwise comparisons).

Table 2
PERMANOVA outcome based on Jaccard dissimilarities between plots along with the proportion of variance explained by each factor (***$p < 0.001$; **$p < 0.01$; *$p < 0.05$).

Source	df	SS	MS	Pseudo-F	Var (%)
Site	2	81344.00	40672.00	12.55***	26.77
Strategy	1	6265.20	6265.2.00	1.93***	1.44
Site × Strategy	2	4811.10	4811.10	1.48**	2.24
Residual	84	272340.00	3242.10	–	69.55
Total	89	369570.00	–	–	–

3.2. Rarefaction curves

SERs at largest spatial scale show an increasing pattern in all sites for both strategies (Fig. 4). At every sampling effort level, the rarefaction curve obtained with the 'advanced strategy' lies above that obtained with the 'basic strategy'. All curves showed no asymptotic pattern.

Concerning additional plots needed to reach a putative 'inventory completeness' at study area scale, the estimation was 222 for the 'basic strategy' and 138 for the 'advanced strategy'. In site A 63 ('basic') and 29 ('advanced') plots, in site B 23 ('basic') and 15 ('advanced'), whereas in Site C 85 ('basic') and 41 ('advanced').

Table 1
Summary statistics of inventories returned by the two sampling strategies ('basic', 'advanced') at the scales of the whole study-area and single sampling sites. '% complet.' indicates the % proportion of inventory completeness with respect to available floristic knowledge for the sites.

	Study-area		Site A		Site B		Site C	
	Basic	*Advanced*	*Basic*	*Advanced*	*Basic*	*Advanced*	*Basic*	*Advanced*
n° plot	45	45	15	15	15	15	15	15
Taxa detected	195	285	86	142	100	128	70	121
% complet.	17.00	24.85	22.69	37.47	43.29	55.41	17.11	29.58

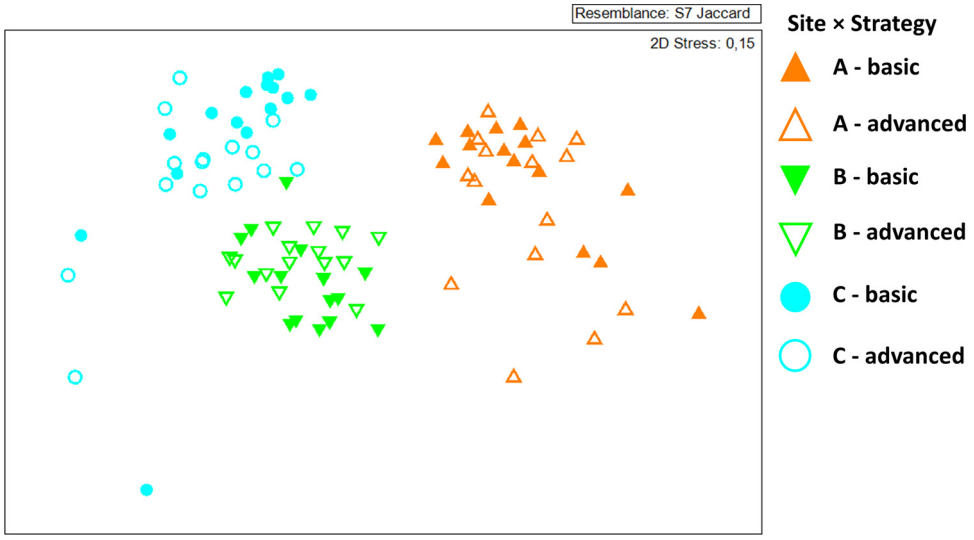

Fig. 3. Non-metric MultiDimensional Scaling ordination of floristic composition of plots based on Jaccard dissimilarity (stress = 0.15). Colours corresponds to factor 'site' (three levels: 'A', 'B', and 'C), whereas symbols represent different sampling strategies (two levels: 'basic' and 'advanced').

3.3. Time effort

There was a positive correlation (Spearman's $\rho = 0.70$, $p < 0.001$) between the mean time to sample a plot and its richness in species. Nevertheless, no significant difference was observed across sampling sites between the two strategies ($F = 2.00$, $p > 0.05$; Supplemental S2.6). This allowed to calculate a mean total time to sample a plot (mean time to sample a plot × n° of sampling sessions) irrespective of strategy and site, which is 85 ± 34 min. The mean total travelling time to reach a plot was 76.9 ± 24.7 min and 50.7 ± 30.2 min for the 'basic' and 'advanced strategy', respectively, showing significant differences within all the sites (Site A: $p < 0.001$; Site B: $p < 0.05$; Site C:

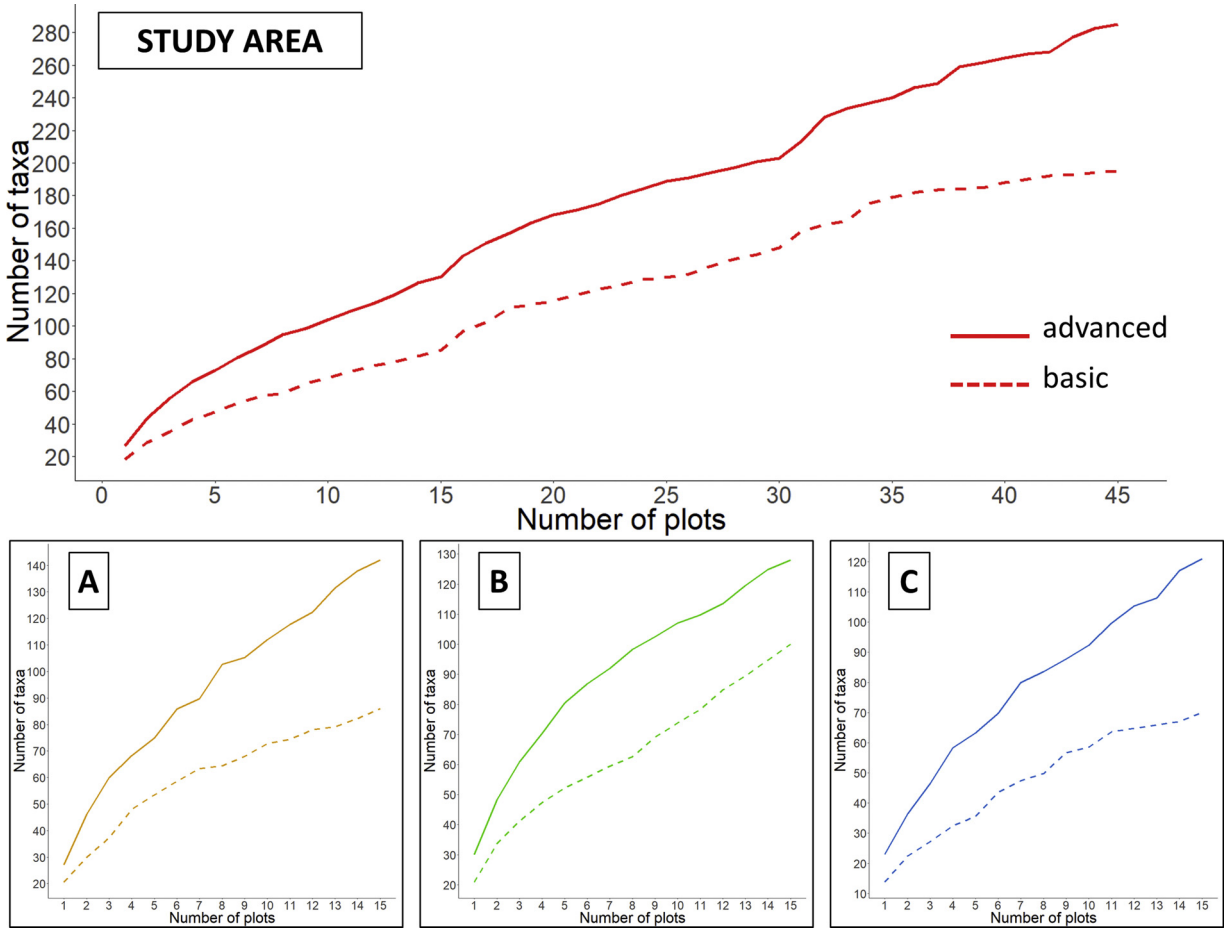

Fig. 4. Spatially Explicit Rarefaction curves for the study area (graph above) and samplings sites (graphs below, from left to right: Site A, Site B, and Site C). Dashed and continuous lines represent the 'basic' and 'advanced' strategy, respectively.

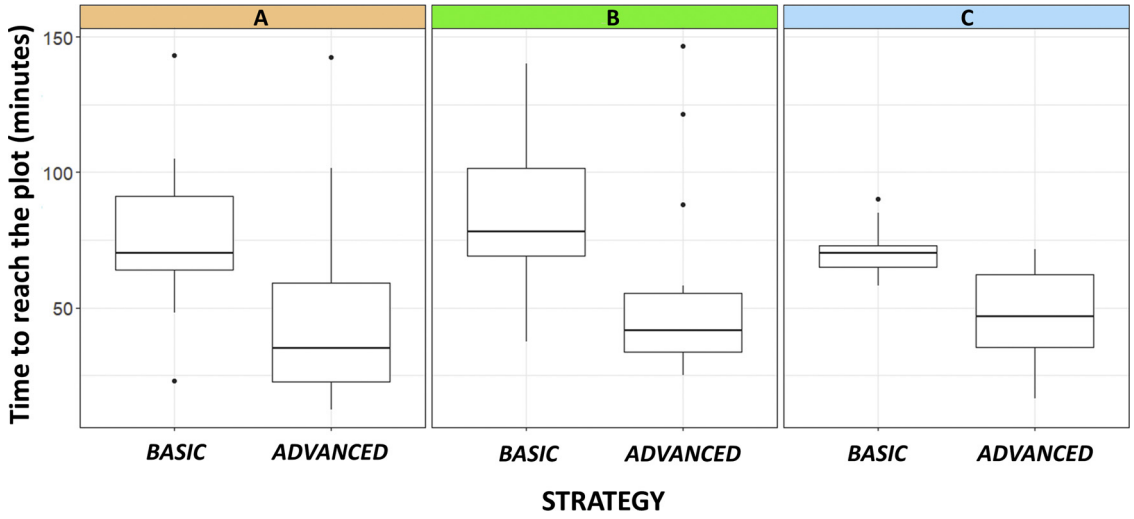

Fig. 5. Boxplots showing the total mean time to reach a plot (mean time to reach a plot × n° of sampling sessions) in all the sampling sites.

$p < 0.001$) (Fig. 5). On the contrary, for a given sampling strategy, no difference in travelling time to reach plots within each site was detected.

3.4. Patterns of spatial configuration of α and β diversity

Species richness (α diversity) values across the study sites showed a wider variation range in the 'advanced strategy' and no consistent pattern emerged between the two strategies (maps are available in Supplemental Material S2.7). In all the sites, the 'advanced strategy' scores higher LCBD median values and smaller variation range with respect to the 'basic strategy' (see Supplemental S2.8 for the complete table). Fig. 6 showed the spatial pattern of LCBD values across the study sites (Uncertainty Maps are available in Supplemental Material S2.9).

4. Discussion

4.1. Balancing sampling effort and inventory completeness: which strategy to follow?

The 'advanced strategy' is more effective in detecting additional species than the 'basic' one, irrespective of the spatial scale (study area, sites) or the vegetational context considered. Accordingly, our results concerning sampling efficiency are not scale-dependent. In addition, considering (i) the good floristic knowledge of the sampling sites before the start of our field surveys, and (ii) the multiple sampling sessions per plot along the entire phenological season, we can exclude that having performed the two sampling strategies in different years could have introduced some bias, i.e. 'favouring' in some way the 'advanced strategy' (performed in the second year).

The sampling design in the 'advanced strategy' is just a little more complex, since it requires spectral data gathering (two spectral bands in

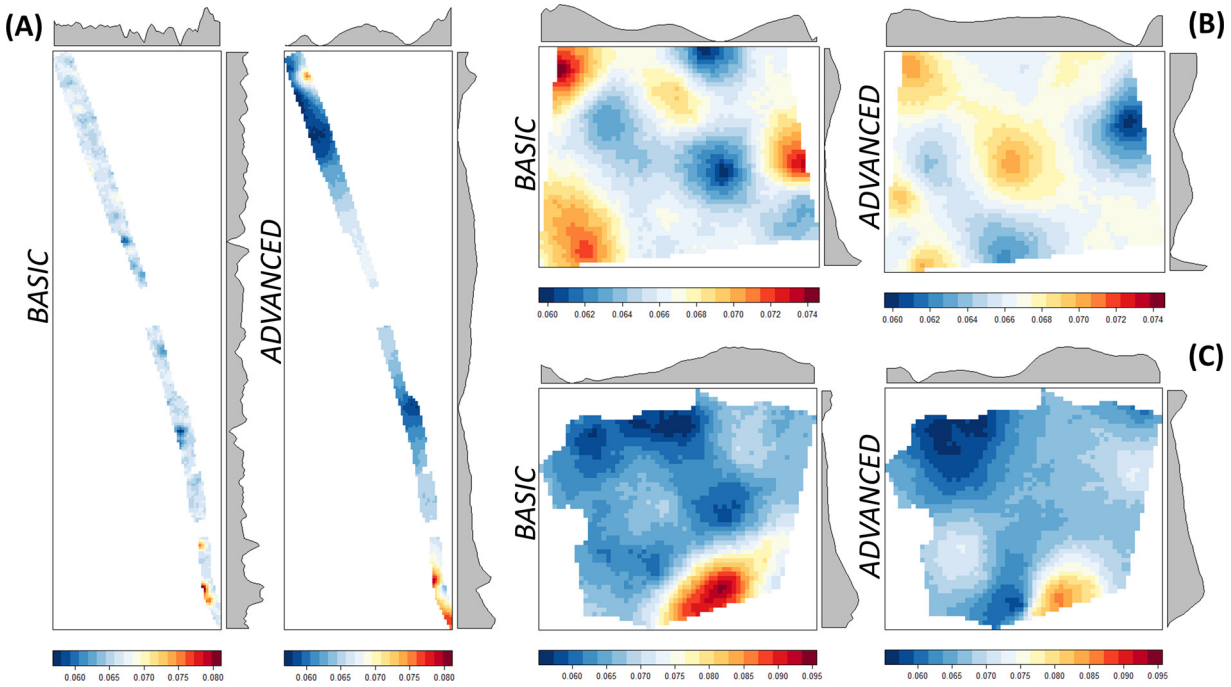

Fig. 6. Kriging maps of LCBD values for the three sampling sites: (A), (B), and (C), respectively. For each site, left panel refers to the 'basic strategy' whereas the right panel to the 'advanced strategy'. Graphs on each plot margin show the mean value of LCBD at a given value of longitude (x-axis) or latitude (y-axis).

the form of raster layers, easily available for any other part of the world) and geoprocessing (performed by the R code we are providing). Anyway, the net gain in the number of sampled species in function of sampling effort completely justifies the addition of this level of complexity. As resulted from several perspectives, the 'advanced strategy' performs better than its 'basic' counterpart, being able to inventory a greater number of species for every sampling effort threshold. However, pairwise comparisons (see PERMANOVA outputs in Supplemental Material) of floristic lists returned by both strategies of the same sampling site did not detect differences between inventories. Both samplings strategies missed rare species known to occur within the study sites, such as for instance *Hypericum elodes* L., whose unique Italian population falls in Site C, or *Utricularia australis* R.Br., whose unique population in MSRM Park falls in SITE A.

The rarefaction curves obtained with 'basic' and 'advanced' strategies, respectively, differed greatly and consistently across study areas. As hypothesised, the 'basic strategy', based solely on a coarse environmental classification followed by a spatial optimisation, collects less species than the 'advanced strategy'. Moreover, at least in Site A and Site C, the rarefaction curves showed a severe reduction in the slopes of their final part, with respect to the initial part. This denotes that even if the sampling is clearly incomplete, this sampling strategy struggles to select sampling units hosting a relevant *quota* of additional species.

Sampling time is a crucial parameter when dealing with sampling optimisation (Stohlgren et al., 1997). In our work, we quantified and analysed sampling time by considering two different variables: a) the 'time spent for collecting species' in the plot (sampling time) and b) the 'time to reach the sampling unit' (travelling time). As already described in literature (e.g., Gray and Azuma, 2005; Zhang et al., 2014), a positive relationship between the number of species occurring in a plot and the sampling time was observed in our study, with no substantial variation between sampling strategies or sites. On the other hand, the travelling time, a parameter that is rarely taken into account in sampling protocols, exhibited differences between the sampling strategies we tested. Travelling time is indeed recognised as the main factor causing a severe increase of sampling costs (Stohlgren, 2007). The higher travelling time we observed in the 'basic strategy' is probably related to the higher distance among selected plots, related to the adopted algorithm promoting a uniform distribution of sampling units on the whole surface. Conversely, in the 'advanced strategy', we observed a tendency to form clusters, a factor that, on average, reduces travelling time. Another non negligible reason influencing travelling time is the year of data collection: the 'advanced strategy' was carried out in the second year with respect to the 'basic strategy', so that also an increased geographical knowledge of the sites possibly contributed to lower travelling times. Consequently, more efforts should be spent to code formal sampling design aimed to optimise this trade-off, which is crucial in determining the sampling efficiency.

Finally, an interesting pattern of beta diversity characterised the two sampling approaches: in the 'advanced strategy', specifically, the range of variation of LCBD values is smaller. This means that the average contribution of complementary species at the plot level to the total β diversity of the whole sampled area is more uniform. In other words, the spatial pattern of beta diversity reflects a better distribution of plots across different habitats. Conversely. the 'basic strategy' tends to collect more sampling units in common habitats (with a higher proportion of species in common, and, consequently, a lower contribution of LCBD to the total beta diversity), while only a small fraction of plots captures rare habitats (causing high LCBD differences) (see Supplemental S2.10).

4.2. Probabilistic vs. Preferential sampling in floristics: the way out of the sampling dilemma

To draft a species inventory, probabilistic and preferential samplings represent two opposite approaches to reach the same final goal, i.e. an inventory as complete as possible. Probabilistic methods have the great potential – yet mostly unexplored in floristics – to perform objective and flexible sampling methods, then making possible quantitative and inferential analyses. On the other hand, probabilistic sampling does not constitute, alone, the final answer to the sampling dilemma for floristic research. The above-mentioned '*botanic internal algorithm*' represents a valuable source which should not be overlooked. As already documented in literature (Palmer et al., 2002; Hédl, 2007; Golodets et al., 2013), and confirmed by our study, rare species are likely missed by probabilistic strategies.

More theoretical and practical efforts should be spent in '*joining the incompatible*' (quoting Chiarucci et al., 2018), i.e. merging into a standardised and efficient way probabilistic and preferential approaches, preserving their peculiarities and exploiting complementarity to overcome limitations.

In zoological sampling literature, the integration of different sampling methods is a recognised tool to increase the pool of sampled species. This practice is generally known as 'structured inventory' (Gotelli and Ellison, 2013). By transposing and remodulating this concept to floristic inventories, we propose to plan floristic inventories with an integral probabilistic design and then to complement with species (i) detected outside from the sampling units, for instance along trajectories joining sampling units, and (ii) detected from *ad hoc* preferential surveys. The first step of this sampling protocol will allow to obtain a reproducible sample of the investigated area (Golodets et al., 2013), providing primary data to perform statistical inferences. Then, the addition of new species can be boosted through preferential surveys based on subjective choices. In terms of cost-efficiency trade-off, the additional cost required to obtain new information is negligible, being based on a combination of the '*botanic internal algorithm*' and the knowledge of the territory and its flora, acquired/improved during the probabilistic sampling in the first step of the workflow.

5. Conclusions and future prospects

Although floras are generally recognised to have limitations and biases (Palmer et al., 1995; Bennett, 1997; Diggs and Lipscomb, 2002; Palmer, 2005), they represent a fundamental raw material in the conservation practice and to study biodiversity patterns (D'Antraccoli et al., 2019). As argued by Palmer (2005), floras are still underutilised for quantitative analyses. However, an increasing awareness of the key role represented by primary floristic data is spreading at global level in recent years (Kier et al., 2005; Whittaker et al., 2005; D'Antraccoli et al., 2019; Weigelt et al., 2020). Our study aims to explore the application of probabilistic approaches to integrate traditional preferential floristic workflow. We designed, and tested in the field, a sampling method based on the maximisation of the spectral heterogeneity of sampling units, which could be routinely integrated in floristic studies.

Our study performed for the first time a spatial analysis of sampling strategies combining the use of beta diversity captured by each sampling unit with kriging interpolation approaches. The idea behind this approach is to map beta diversity, which is intrinsically multivariate, transforming it in univariate synthetic response data, such as those expressed by LCBD values.

Our approach can be profitably applied to any other geographic location. Compared to existing probabilistic sampling methods, it has the advantage to ensure representativeness of sampling units from different habitats at a fine spatial-scale. This avoids the loss of information intrinsic in traditional probabilistic sampling methods, including those based on the stratification of environmental features. Anyway, standardised probabilistic samplings show some limitations, which are very difficult to overcome when the main goal is to obtain an inventory as complete as possible. Accordingly, codifying a standardised and reproducible protocol able to integrate the former approach

with preferential surveys constitutes a very challenging task for the future of floristic research.

Authors contribution

MD, GB, GB, and LP conceived the idea of this research. ET and MD wrote the algorithms to select the sampling units; MD performed the field surveys. MD and ET analysed the data. MD wrote the manuscript with the contribution of ET. LP supervised the whole project. All authors contributed critically to the manuscript drafts and gave final approval for publication.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgments

The authors thank the Migliarino – San Rossore – Massaciuccoli Regional Park who provided formal and logistic support to access the sampling sites, and Paola De Giorgi for helping in field surveys of the 'advanced' strategy.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ppees.2020.125547.

References

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46.
Anderson, M.J., Gorley, R.N., Clarke, K.R., 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E, Plymouth, UK.
Archaux, F., Gosselin, F., Bergès, L., Chevalier, R., 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. J. Veg. Sci. 17, 299–306.
Bacaro, G., 2008. Monitoring Biodiversity in a Regional Natura 2000 Network: From Sampling Design Assessment to the Evaluation of Biodiversity Indicators. PhD Thesis. University of Siena.
Bacaro, G., Rocchini, D., Ghisla, A., Marcantonio, M., Neteler, M., Chiarucci, A., 2012. The spatial domain matters: spatially constrained species rarefaction in a free and open source environment. Ecol. Complex. 12, 63–69.
Bacaro, G., Rocchini, D., Diekmann, M., Gasparini, P., Gioria, M., Maccherini, S., Marcantonio, M., Tordoni, E., Amici, V., Landi, S., Torri, D., Castello, M., Altobelli, A., Chiarucci, A., 2015. Shape matters in sampling plant diversity: evidence from the field. Ecol. Complex. 24, 37–45. https://doi.org/10.1016/j.ecocom.2015.09.003.
Bacaro, G., Altobelli, A., Cameletti, M., Ciccarelli, D., Martellos, S., Palmer, M.W., Ricotta, C., Rocchini, D., Scheiner, S.M., Tordoni, E., Chiarucci, A., 2016. Incorporating spatial autocorrelation in rarefaction methods: implications for ecologists and conservation biologists. Ecol. Indic. 69, 233–238.
Baffetta, F., Bacaro, G., Fattorini, L., Rocchini, D., Chiarucci, A., 2007. Multi-stage cluster sampling for estimating average species richness at different spatial grains. Community Ecol. 8, 119–127.
Bartolucci, F., Peruzzi, L., Galasso, G., Albano, A., Alessandrini, A., Ardenghi, N.M.G., Astuti, G., Bacchetta, G., Ballelli, S., Banfi, E., Barberis, G., Bernardo, L., Bouvet, D., Bovio, M., Cecchi, L., Di Pietro, R., Domina, G., Fascetti, S., Fenu, G., Festi, F., Foggi, B., Gallo, L., Gubellini, L., Gottschlich, G., Iamonico, D., Iberite, M., Jinénez-Mejías, P., Lattanzi, E., Martinetto, E., Masin, R.R., Medagli, P., Passalacqua, N.G., Peccenini, S., Pennesi, R., Pierini, B., Poldini, L., Prosser, F., Raimondo, F.M., Marchetti, D., Roma-Marzio, F., Rosati, L., Santangelo, A., Scoppola, A., Scortegagna, S., Selvaggi, A., Selvi, F., Soldano, A., Stinca, A., Wagensommer, R.P., Wilhalm, T., Conti, F., 2018. An updated checklist of the vascular flora native to Italy. Plant Biosyst. 152, 179–303. https://doi.org/10.1080/11263504.2017.1419996.
Bedini, G., Pierini, B., Roma-Marzio, F., Caparelli, K.F., Bonari, G., Dolci, D., Gestri, G., D'Antraccoli, M., Peruzzi, L., 2016. Wikiplantbase# Toscana, breaking the dormancy of floristic data. Plant. Biosyst. 150, 601–610.
Bennett, J.P., 1997. Nested taxa-area curves for eastern United States floras. Rhodora 99, 241–251.
Chiarucci, A., 2007. To sample or not to sample? That is the question… for the vegetation scientist. Folia Geobot. 42, 209.
Chiarucci, A., Bonini, I., 2005. Quantitative floristics as a tool for the assessment of plant diversity in Tuscan forests. Forest Ecol. Manag. 212, 160–170.
Chiarucci, A., De Dominicis, V., Wilson, J.B., 2001. Structure and floristic diversity in permanent monitoring plots in forest ecosystems of Tuscany. Forest Ecol. Manag. 141, 201–210.
Chiarucci, A., Bacaro, G., Rocchini, D., Ricotta, C., Palmer, M.W., Scheiner, S.M., 2009. Spatially constrained rarefaction: incorporating the autocorrelated structure of biological communities into sample-based rarefaction. Community Ecol. 10, 209–214.
Chiarucci, A., Bacaro, G., Scheiner, S.M., 2011. Old and new challenges in using species diversity for assessing biodiversity. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 366, 2426–2437.
Chiarucci, A., Di Biase, R.M., Fattorini, L., Marcheselli, M., Pisani, C., 2018. Joining the incompatible: exploiting purposive lists for the sample-based estimation of species richness. Ann. Appl. Stat. 12, 1679–1699.
Chiarucci, A., Palmer, M.W., 2005. The inventory and estimation of plant species richness. Encyclopedia of Life Support Systems (EOLSS). EOLSS Publishers, Oxford.
Clarke, K.R., Warwick, R.M., 2005. Primer-6 Computer Program. Natural Environment Research Council, Plymouth.
Colwell, R.K., Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 345, 101–118.
D'Antraccoli, M., Roma-Marzio, F., Carta, A., Landi, S., Bedini, G., Chiarucci, A., Peruzzi, L., 2019. Drivers of floristic richness in the Mediterranean: a case study from Tuscany. Biodivers. Conserv. 28, 1411–1429.
Daniel, J., 2012. Sampling Essentials: Practical Guidelines for Making Sampling Choices; SAGE: Los Angeles, CA, USA, 2012. pp. 125–140.
Dengler, J., 2009. A flexible multi-scale approach for standardised recording of plant species richness patterns. Ecol. Indic. 9, 1169–1178.
Diggs, G.M., Lipscomb, B.L., 2002. What is the writer of a flora to do? Evolutionary taxonomy or phylogenetic systematics? SIDA, Contributions to Botany 20, 647–674.
Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N., Wagner, H.H., 2019. Adespatial: Multivariate Multiscale Spatial Analysis. R Package Version 0.3-7. https://CRAN.R-project.org/package=adespatial.
Elzinga, C.L., Salzer, D.W., Willoughby, J.W., 1998. Measuring and monitoring plant populations. US Department of the Interior, Bureau of Land Management.
Foody, G.M., Cutler, M.E.J., 2003. Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. Ecol. Model. 195, 37–42.
Galasso, G., Conti, F., Peruzzi, L., Ardenghi, N.M.G., Banfi, E., Celesti-Grapow, L., Albano, A., Alessandrini, A., Bacchetta, G., Ballelli, S., Bandini Mazzanti, M., Barberis, G., Bernardo, L., Blasi, C., Bouvet, D., Bovio, M., Cecchi, L., Del Guacchio, E., Di Pietro, R., Domina, G., Fascetti, S., Gallo, L., Gubellini, L., Guiggi, A., Iamonico, D., Iberite, M., Jiménez-Mejías, P., Lattanzi, E., Marchetti, D., Martinetto, E., Masin, R.R., Medagli, P., Passalacqua, N.G., Peccenini, S., Pennesi, R., Pierini, B., Podda, L., Poldini, L., Prosser, F., Raimondo, F.M., Roma-Marzio, F., Rosati, L., Santangelo, A., Scoppola, A., Scortegagna, S., Selvaggi, A., Selvi, F., Soldano, A., Stinca, A., Wagensommer, R.P., Wilhalm, T., Bartolucci, F., 2018. An updated checklist of the vascular flora alien to Italy. Plant Biosyst. 152, 556–592.
Gaston, K.J., 2000. Global patterns in biodiversity. Nature 405, 220–227.
Golodets, C., Kigel, J., Sapir, Y., Sternberg, M., 2013. Quantitative vs qualitative vegetation sampling methods: a lesson from a grazing experiment in a Mediterranean grassland. Appl. Veg. Sci. 16, 502–508.
Gonzalez-Oreja, J.A., Garbisu, C., Mijangos, I., Mendarte, S., Albizu, I., 2013. Reducing costs in biodiversity monitoring: shortcuts for plant diversity in meadows as a case study. Ecol. Indic. 24, 96–104.
Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89, 1–45.
Gotelli, N.J., Colwell, R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4, 379–391.
Gotelli, N.J., Ellison, A.M., 2013. A Primer of Ecological Statistics, 2nd Edition. Sinauer, Sunderland.
Gould, W., 2000. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. Ecol. Appl. 10, 1861–1870.
Gräler, B., Edzer Pebesma, E., Heuvelink, G., 2016. Spatio-Temporal interpolation using gstat. R J. 8, 204–218.
Gray, A.N., Azuma, D.L., 2005. Repeatability and implementation of a forest vegetation indicator. Ecol. Indic. 5, 57–71.
Güler, B., Jentsch, A., Apostolova, I., Bartha, S., Bloor, J.M., Campetella, G., Canullo, R., Házi, J., Kreyling, J., Pottier, J., Szabó, G., Terziyska, T., Uğurlu, E., Wellstein, C., Zimmermann, Z., Dengler, J., 2016. How plot shape and spatial arrangement affect plant species richness counts: implications for sampling design and rarefaction analyses. J. Veg. Sci. 27, 692–703.
Haining, R.P., 2003. Spatial Data Analysis: Theory and Practice. Cambridge University Press, Cambridge, UK.
Hédl, R., 2007. Is sampling subjectivity a distorting factor in surveys for vegetation diversity? Folia Geobot 42, 191–198. https://doi.org/10.1007/BF02893885.
Hoffmann, S., Steiner, L., Schweiger, A.H., Chiarucci, A., Beierkuhnlein, C., 2019. Optimizing sampling effort and information content of biodiversity surveys: a case study of alpine grassland. Ecol. Inform. 51, 112–120.
Kalkhan, M.A., Stafford, E.J., Stohlgren, T.J., 2007. Rapid plant diversity assessment using a pixel nested plot design: A case study in Beaver Meadows, Rocky Mountain National Park, Colorado, USA. Divers. Distrib 13, 379–388. https://doi.org/10.1111/j.1472-4642.2007.00333.x.
Kier, G., Mutke, J., Dinerstein, E., Ricketts, T.H., Küper, W., Kreft, H., Barthlott, W., 2005. Global patterns of plant diversity and floristic knowledge. J. Biogeogr. 32, 1107–1116.
Koellner, T., Hersperger, A.M., Wohlgemuth, T., 2004. Rarefaction method for assessing plant species diversity on a regional scale. Ecography 27, 532–544.
Kruskal, J.B., Wish, M., 1978. Multidimensional Scaling, Sage University Paper Series on Quantitative Application in the Social Sciences. Sage Publications, Beverly Hills and London, pp. 07–011.

Kühn, I., 2007. Incorporating spatial autocorrelation may invert observed patterns. Divers. Distrib. 13, 66–69. https://doi.org/10.1111/j.1472-4642.2006.00293.
Kunin, W.E., 1997. Sample shape, spatial scale and species counts: implications for reserve design. Biol. Conserv. 82, 369–377.
Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.
Legendre, P., 2014. Interpreting the replacement and richness difference components of beta diversity. Glob. Ecol. Biogeogr. 23, 1324–1334.
Legendre, P., De Cáceres, M., 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecol. Lett. 16, 951–963.
Maccherini, S., Bacaro, G., Tordoni, E., Bertacchi, A., Castagnini, P., Foggi, B., Gennai, M., Mugnai, M., Sarmati, S., Angiolini, C., 2020. Enough is enough? Searching for the optimal sample size to monitor European habitats: a case study from coastal sand dunes. Diversity 12, 138.
Morrison, L.W., 2016. Observer error in vegetation surveys: a review. J. Plant Ecol. 9, 367–379.
Neigh, C.S.R., Tucker, C.J., Townshend, J.R.G., 2008. North American vegetation dynamics observed with multiresolution satellite data. Remote Sens. Environ. 112, 1749–1772.
Nekola, J.C., White, P.S., 1999. The distance decay of similarity in biogeography and ecology. J. Biogeogr. 26, 867–878.
Newmaster, S.G., Belland, R.J., Arsenault, A., Vitt, D.H., Stephens, T.R., 2005. The ones we left behind: comparing plot sampling and floristic habitat sampling for estimating bryophyte diversity. Divers. Distrib 11, 57–72.
Oldeland, J., Wesuls, D., Rocchini, D., Schmidt, M., Jürgens, N., 2010. Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity? Ecol. Indic. 10, 390–396.
Palmer, M.W., 2005. Temporal trends of exotic species richness in North American floras: an overview. Ecoscience 12, 386–390. https://doi.org/10.2980/i1195-6860-12-3-386.1.
Palmer, M.W., Earls, P., Hoagland, B.W., White, P.S., Wohlgemuth, T.M., 2002. Quantitative tools for perfecting species lists. Environmetrics 13, 121–137.
Palmer, M.W., McGlinn, D.J., Fridley, J.D., 2008. Artifacts and artifictions in biodiversity research. Folia Geobot 43, 245–257. https://doi.org/10.1007/s12224-008-9012-y.
Palmer, M.W., Wade, G.L., Neal, P., 1995. Standards for the writing of floras. BioScience 45, 339–345. https://doi.org/10.2307/1312495.
Peruzzi, L., Bedini, G. (Eds.), 2015 onwards. Wikiplantbase. #Toscana 2 (1). http://bot.biologia. unipi.it/wpb/toscana/index.html.
Pignatti, S., 1982. Flora d'Italia. Edagricole, Bologna.
Podolsky, R., 1994. A method for estimating biodiversity directly from digital earth imagery. Earth Obs. Mag. 3 (June), 30–36.
R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Rocchini, D., 2007. Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. Remote Sens. Environ. 111, 423–434.
Rocchini, D., Ricotta, C., Chiarucci, A., De Dominicis, V., Cirillo, I., Maccherini, S., 2009. Relating spectral and species diversity through rarefaction curves. Int. J. Remote Sens. 30, 2705–2711.
Rocchini, D., Balkenhol, N., Carter, G.A., Foody, G.M., Gillespie, T.W., He, K., Kark, S., Levin, N., Lucas, K., Luoto, M., Nagendra, H., Oldeland, J., Ricotta, C., Southworth, J., Neteler, M., 2010. Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. Ecol. Inform. 5, 318–329.
Rocchini, D., Luque, S., Pettorelli, N., Bastin, L., Doktor, D., Faedi, N., Feilhauer, H., Féret, J.B., Foody, G.M., Gavish, Y., Godinho, S., Kunin, W.E., Lausch, A., Leitão, P.J., Marcantonio, M., Neteler, M., Ricotta, C., Schmidtlein, S., Vihervaara, P., Wegmann, M., Nagendra, H., 2018. Measuring β-diversity by remote sensing: A challenge for biodiversity monitoring. Methods Ecol. Evol. 9, 1787–1798.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. In: In: Freden, S.C., Mercanti, E.P., Becker, M. (Eds.), Third Earth Resources Technology Satellite–1 Syposium Volume I. Technical Presentations, NASA SP-351, NASA, Washington, D.C, pp. 309–317.
Scheiner, S.M., 2019. A compilation of and typology for abundance-, phylogenetic-and functional-based diversity metrics. BioRxiv, 530782.
Scott, W.A., Hallam, C.J., 2002. Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecol. 165, 101–115.
Stohlgren, T.J., 2007. Measuring Plant Diversity. Oxford University Press, Inc., New York, New York, USA.
Stohlgren, T.J., Chong, G.W., Kalkhan, M.A., Schell, L.D., 1997. Multiscale sampling of plant diversity: effects of minimum mapping unit size. Ecol. Appl. 7, 1064–1074.
ter Steege, H., Haripersaud, P.P., Bánki, O.S., Schieving, F., 2011. A model of botanical collectors' behavior in the field: never the same species twice. Am. J. Bot. 98, 31–37.
Theirs, B., 2019. Onwards. Index Herbariorum: a Global Directory of Public Herbaria and Associated Staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/science/ih/.
Tison, J.M., de Foucault, B., 2014. Flora Gallica: Flore De France. Editions Biotope.
Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46, 234–240.
U.S. Geological Survey (USGS), 2018. Earth Explorer. Available at. https://earthexplorer.usgs.gov/.
Vardi, Y., Zhang, C.-H., 1999. The multivariate L1-median and associated data depth. PNAS 97, 1423–1426.
Vondrák, J., Malíček, J., Palice, Z., Coppins, B., Kukwa, M., Czarnota, P., Sanderson, N., Acton, A., 2016. Methods for obtaining more complete species lists in surveys of lichen biodiversity. Nord. J. Bot. 34, 619–626. https://doi.org/10.1111/njb.01053.
Walvoort, D.J.J., Brus, D.J., De Gruijter, J.J., 2010. An r package for spatial coverage sampling and random sampling from compact geographical strata by k-means. Comput. Geosci. 36, 1261–1267.
Wang, J.-F., Stein, A., Gao, B.-B., Ge, Y., 2012. A review of spatial sampling. Spat. Stat. 2, 1–14.
Wegmann, M., Leutner, B., 2016. Introduction to remote sensing and GIS. Wegmann, M., Leutner B., Dech, S., 2016. Remote Sensing and GIS for Ecologists: Using Open Source Software. Exeter: Pelagic Publishing, UK.
Weigelt, P., König, C., Kreft, H., 2020. GIFT–A global inventory of floras and traits for macroecology and biogeography. J. Biogeogr. 47, 16–43.
Whittaker, R.H., 1972. Evolution and measurement of species diversity. Taxon 21, 213–251.
Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E., Willis, K.J., 2005. Conservation biogeography: assessment and prospect. Divers. Distrib. 11, 3–23.
Wiser, S.K., Spencer, N., De Cáceres, M., Kleikamp, M., Boyle, B., Peet, R.K., 2011. Veg-X - an exchange standard for plot-based vegetation data. J. Veg. Sci. 22, 598–609.
Yoccoz, N.G., Nichols, J.D., Boulinier, T., 2001. Monitoring of biological diversity in space and time. Trends Ecol. Evol. 16, 446–453.
Zhang, J., Nielsen, S.E., Grainger, T.N., Kohler, M., Chipchar, T., Farr, D.R., 2014. Sampling plant diversity and rarity at landscape scales: importance of sampling time in species detectability. PLoS One 9, e95334.
Zhao, Y., Lei, J., Tumarbay, H., Xue, J., 2018. Using sequential Gaussian simulation to assess the uncertainty of the spatial distribution of soil salinity in arid regions of Northwest China. Arid Land. Res. Manag. 32, 20–37.