Journal of Threatened Taxa

Building evidence for conservation globally

Open Access

10.11609/jott.2022.14.1.20311-20538
www.threatenedtaxa.org

26 January 2022 (Online & Print)
14 (1): 20311-20538
ISSN 0974-7907 (Online)
ISSN 0974-7893 (Print)
Estimating the completeness of orchid checklists and atlases: a case study from southern Italy

Antonio Croce

GIROS (Gruppo Italiano per la ricerca sulle Orchidee Spontanee- Italian Group for the Research on Wild Orchids)
Via Chiesa - Tuoro, 44 - 81057 Teano, Caserta, Italy.
antocrx@gmail.com

Abstract: Checklists and atlases are important tools for knowledge of the biodiversity of a geographic unit. Nevertheless, they often suffer from bias due to preferential sampling. It is important to assess the level of completeness of the data collected during such research to allow comparison of the biodiversity of different areas, or to use them for macroecology, biogeography or conservation purposes. This assessment is not trivial, especially when information from heterogeneous sources is used (e.g., herbaria specimens, field observations, literature data). The author suggests some simple methods to assess the completeness of floristic database and to represent the distribution of the completeness at a scale level appropriate to the size of the studied area or, on another hand, to the precision level of the available data. Such information is useful to direct the surveys identifying less explored areas or habitats and thereby correcting the sampling biases. Adding information about sampling effort or completeness could be very useful to make floristic research more objective.

Keywords: European orchids, floristic studies, sampling effort, species richness estimators, completeness, citizen science.

Riassunto: le checklist e gli atlanti floristici sono strumenti importantissimi per la conoscenza della biodiversità. Tuttavia essi sono realizzati senza un design sperimentale e sono soggetti a bias dovuto soprattutto al campionamento preferenziale. È comunque importante, soprattutto quando questi studi si basano su informazioni derivanti da fonti eterogenee (campioni d'erbario, osservazioni in campo, dati bibliografici, ecc.) valutare il loro grado di completezza per poter confrontare la biodiversità di diverse aree geografiche o per eseguire analisi macroecologiche, biogeografiche e per la valutazione dello stato di conservazione. L'autore propone alcuni semplici metodi per stimare l'esauritezza dei dati floristici, rappresentare la distribuzione della completezza a scale adeguate da una parte alla dimensione dell'area oggetto di studio e dall'altra al livello di precisione dei dati a disposizione. Tali informazioni sono utili anche per orientare le ricerche nel territorio, individuando aree o habitat meno esplorati e correggendo i bias di campionamento. L'aggiunta di informazioni sullo sforzo di campionamento e la completezza delle ricerche può essere utile a conferire agli studi floristici di base una maggiore oggettività.
INTRODUCTION

Floristic inventories or check lists and atlases are important tools for assessing biodiversity and addressing its conservation (Vallet et al. 2012). They are often the result of careful and time-consuming researches conducted in specific geographic units, focused on vascular plants or on smaller taxonomic group such as Orchidaceae, one of the largest and most widespread family of flowering plants (Dressler 1981; WCSFP 2019). The presence and distribution of species of this family have been assessed at different scales as most of them are rare, threatened or endangered (Cribb et al. 2003). A checklist is a “card collection” aiming at listing all the taxa belonging to the studied taxonomic group and reporting whether they are observed, collected or reported in literature for a given area (e.g., Mathew & George 2015; Aung et al. 2020; Popovich et al. 2020). The taxa are typically identified at species or subspecies level, some sites of growth are reported together with other information on the habitats, variety, rarity, ecology, chorology, systematic or taxonomic issues. Atlases are more focused on the geographic distribution of the taxa, instead. To be accomplished they require a field work aiming not only at listing all the different taxonomic entities, but also at detecting as more sites of growth as possible for each taxon. The result of such work is a checklist with cartographic references or distribution maps and, sometimes, their elaborations (e.g., Crain & Fernández 2020; Efimov 2020). Due to the long time needed for exhaustive surveys, at a local scale this kind of research is increasingly carried out by non academics, the so called ‘citizen scientists’. This is particularly true for the inventories and atlases of the European terrestrial orchids, often published in specialized journals (e.g., Galesi & Lorenz 2010; Frangini et al. 2019; Katopodi & Tsiftsis 2019; Marrero et al. 2019).

The huge amount of work, even when results in detailed distribution maps, almost never follows an experimental design, and currently data are affected by bias caused by a preferential sampling approach, e.g., data collector tends to sample protected areas or to collect more data along the roads (Croce & Nazzaro 2017). Furthermore, none of the above mentioned floristic studies is usually provided with a clear reference to the sampling effort or to the level of completeness of the surveys. The absence of a repeatable background and of a standardized approach is not a trivial issue, as such collections of data are of great value for macroecology, ecology, biogeography or conservation research (Soberón et al. 2000, 2007; Rocchini et al. 2011; Weigelt et al. 2020).

In order to make inventories and atlases useful tools for biogeographical or ecological research it is thus necessary to take into account these issues and support floristic works with appropriate measures of the degree of uncertainty (Rocchini et al. 2011). In the same context, maps of floristic richness should be accompanied by maps of knowledge, “maps of ignorance” or maps of completeness. These can be realized considering that the number of species (namely the species richness) recorded in a given period and in a given area is partial and lower than the real number of species present (Gotelli & Colwell 2011). The more the sample effort increases the more the number of observed species approaches the theoretical, real number of species. On the contrary, a sampling activity carried over a too long time could detect the species turnover (e.g., for habitat change due to socio-economic or ecological reasons or for climate changing) resulting in an overestimation of the number of species than the existing habitats could theoretically host in a given time. The real floristic richness and its distribution in an area can be estimated with different methods (Gotelli & Colwell 2001; Vallet et al 2012). The most suitable for the kind of data recorded in the field by orchidologists is the use of ‘sample based species rarefaction-curves’ (Gotelli & Colwell 2001). Given that the sampling order in an area is not important, data are resampled and curves are built. While the shape of accumulation curves depends upon the order in which the samples are considered, the rarefaction curves show smoother lines facilitating the comparison among entire datasets or subsets. A species rarefaction curve is plotted starting from the mean number of species of the smallest sample size. Then the mean number of species is calculated for all combinations of the next sample size (i.e., the mean number of species of two random samples, then three random samples, etc.).

This paper analyzes some typical aspects of local scale inventories and atlases hitherto neglected. Here, we propose simple approaches, accessible even for the non-academic, citizen scientists to answer the following specific questions:

i. How can the richness of a floristic database be assessed and how can different database be compared?

ii. Which richness estimator is more suitable for terrestrial European orchids, given its intrinsic difficulties of observation in field?

iii. When is the sampling of an area sufficiently complete?

iv. How can completeness maps be realised and how they can be useful to identify where to address further explorations?
MATERIALS AND METHODS

Source of the data and study areas

We used three datasets reporting the presence of orchids in three areas in southern Italy, in northern Campania region, (Figure 1; Table 1) about 50 km north of Naples and 150 south of Rome. The first dataset includes 3,046 records collected from 1996 to 2019 on the Roccamonfina volcano (Croce & Nazzaro 2012 and following observations). It covers an area of about 210 km² and lists 46 taxa (species and subspecies). The second dataset consists of 278 records collected from 2002 to 2005 on the little limestone mountain range of Vairano Patenora and Pietravairano municipalities (Croce 2012 and following observations), hereafter called “Vairanese”. It covers an area of 17 km² and lists 32 taxa. The third dataset consists of 305 records collected mainly in 2005 and then from 2013 to 2019 on the limestone mountain ranges of the western Matese area, hereafter called “W–Matese”. It covers an area of 20 km² and lists 33 taxa.

Data collection

Only the observations geolocated with a precision level lower than 100 m (punctual data according to Croce & Nazzaro 2017) were included in the analysis. Nomenclature was revised and, when needed, standardised and hybrids were excluded from the analysis. To avoid the oversampling bias (i.e., a single population of plants sampled in different sampling units) the records have been clumped to represent the presence of the taxa in 100 x 100 m squares, connected to the geographic grid of the used coordinates system (WGS 84 / UTM zone 33N, EPSG 32633). Each sampling unit (plot) is univocally identified, therefore, by the geographic position of the square and by the sampling date so that two sampling activities that took place in two different date but inside the same square have been considered as two different plots. In this way, I take into account the sampling effort in terms of time, very important for species requiring observations at different times to be correctly observed and identified. In the end I get, for each dataset, a matrix taxon × plot that I used for the elaborations and further analysis.
Data analysis

To compare the three datasets in terms of sampling effort and observed specific richness ($S_{obs}$), I have mapped the specific richness for each area using a grid with 1 km$^2$ resolution (i.e., 1 x 1 km UTM cells) intersecting the study areas (i.e., the three geographic units as defined above) and calculating both the number of plots and the number of observed species in each cell. A regression analysis between the number of plots and the number of species per each cell has been performed to correlate the sampling effort to the observed species richness and therefore to validate the density of plots as an indicator of the sampling effort. Then for each area I built a sample-based rarefaction curve using the plots as samples. The curves have been limited to the lower number of plots in the three datasets for a better comparison of the observed species richness and its pattern among the three studied areas. Being drawn with resampling statistical methods, the curves allow the calculation of the 95% confidence limits or the standard deviations.

Among the methods used to estimate the species richness of an area starting from presence-absence data, the most appropriate for floristic inventories and atlases is the relation between number of species and sampling effort (Vallet et al. 2012). This relation is investigated mainly using non parametric estimators, less sensitive to the sampling effort (Palmer 1990; Brose et al. 2003). Such indexes give an estimate of the species richness for a given geographic unit, based upon the considered sample and, therefore, upon its species assemblage. Once an estimate value is obtained, the completeness for each of the three datasets can be calculated by means of the completeness index proposed by Soberón et al. (2000). Such index (C) is expressed as a percentage value of the ratio between the number of observed species ($S_{obs}$) and the number of estimated species ($S_{est}$):

$$C = \frac{S_{obs}}{S_{est}}$$

The most used non parametric estimators for presence/absence data or incidence data are Jackknife, Chao, Bootstrap, and ICE (Gotelli & Colwell 2011; Vallet et al. 2012). While the first of these indexes could represent a good compromise (Brose et al. 2003), several other authors prefer to compare more than one index (Martinez-Sanz et al. 2010; Bruno et al. 2012; Garcia-Marquez et al. 2012; Vallet et al. 2012; Archer 2019). It is therefore noted that the Jackknife estimator gives higher values of estimated richness and, accordingly, lower completeness values than the Bootstrap estimator (Garcia-Marquez et al. 2012). Nevertheless, it is particularly effective in estimating the richness of small sample size (Hortal et al. 2006). Another very used estimator is Chao2 (Ugland et al. 2003; Chao & Chiu 2016; Ildohou et al. 2015; Asase & Peterson 2016) that gives more emphasis to the presence of singletons species (i.e., present in only one plot of the set or subset) or doubletons (i.e., present in only two plots). Considered that many orchid species are locally rare and the number of rare species increases with decreasing the size of the sampled area, I calculated the completeness index (C) choosing as value of estimated richness ($S_{est}$) the maximum value between Chao2 ($S_{chao}$) and Jackknife1 ($S_{jack1}$) estimates. For each of the three study areas I calculated the total value of completeness (C) and the completeness of each cell of the 1 km$^2$ UTM grid, using the plots as sampling units. Only for Roccamonfina area the completeness has been calculated also for each cell of a 4 km$^2$, 9 km$^2$, 16 km$^2$, 25 km$^2$, and 36 km$^2$ UTM grid intersectoring the study area. Then I aggregated the data into 1 x 1 km cells and the obtained taxon x cells matrix has been used to recalculate the estimated species richness and the completeness of each study area. This was intended to test the reliability of such atlases built mapping the presence of the species in grids with cells of 1 km$^2$ or more, to estimate the species richness of the study areas. In order to test the estimators robustness when even larger sample units are used, the above mentioned aggregation method has been repeated using grids of 4 km$^2$, 9 km$^2$, 16 km$^2$, 25 km$^2$, and 36 km$^2$ cells, only for the larger area of Roccamonfina volcano. In other terms, I used increasing size cells as sampling units. Such cells size can be useful to analyse atlases produced with bibliographic data whose precise geolocation is not possible. The completeness of each cell, for all the grids of different cells size, has been classified into four levels: 0–25 %, 25–50 %, 50–75 %, and 75–100 %. The cell with less than six plots have not been analysed and have been classified as “not evaluable” (n.e.). These limits have been set considering for all the datasets used an average number of five plot sampled in a day. According to the method used in Bruno et al. (2012), the cells with completeness >65% have been considered sufficiently studied squares (SSS).

Once I knew the less explored cells, to which priority in the future research should be given, I could assess the level of completeness of our datasets among different habitats. So, I assigned a kind of vegetation to each plot on the basis of the collected field information and therefore I estimated the completeness of each vegetation type for each study area as explained above.

The cartographic elaborations have been performed by the software Qgis3 (QGIS Development Team 2019),
RESULTS

In Table 1 the data about the three study areas are reported, including the list of the taxa considered. The Roccamonfina area has the highest species richness, average number of records/plot and plot/km². Vairanese and W-Matese show comparable values of the number of records/plot (higher values for W-Matese) and number of plots/km² (higher values for Vairanese). Nevertheless, the distribution of the number of plots (Figure 2a) and observed species richness (Figure 2b) in the 1 km² cells is extremely heterogeneous with a very high standard deviation of the plots/cells ratio (6.9 for Roccamonfina, 6.5 for Vairanese and 5.9 for Matese areas). Such values underline a sampling effort not uniformly distributed in the studied areas.

The regression analysis (Figure 3) shows, for all the three areas, a statistically significant (p <0.001) positive correlation between the number of plots and the number of species inside the 1 km² cells. The two variables are statistically correlated according to the Kendall’s tau test.

The rarefaction curves (Figure 4) indicate a similar pattern for all the three areas: limited to 121 plots, they show slight differences with a higher species richness for the W-Matese area (32.84 average observed species) followed by the Vairanese area (32 average observed species) and the Roccamonfina volcano (31.32 average observed species).

The total estimated floristic richness, computed using the plots as sampling units (Table 2) for each of the three areas, gives completeness values between 78.2% (Vairanese) and 88.5% (Roccamonfina). Using the 1 km² cells as sampling units (Table 3), we get identical values for Roccamonfina area, a slightly higher value for Vairanese area and slightly lower for W-Matese area.

The completeness of the 1 km² cells in the three areas (Figure 2c) is distributed in a similar way in the Roccamonfina and Vairanese areas (Table 4): the 35.6% and 33.3% of the 1 km² cells, respectively, have a completeness higher than 65% and therefore are considered as Sufficiently Studied Squares (SSS). For the W-Matese area only the 25% of the 1 km² cells are SSS. It is relevant, for each area, the great number of cells with data not allowing further elaborations (‘n.e.’ cells).

The estimated richness for the Roccamonfina area, calculated using sampling units of increasing size (Figure 5) shows a general stability of the two estimators chosen, always with higher values for Jackknife1 estimator (51.82–53.47) compared to Chao2 estimator (48.11–49.34). Both the estimators feature variations included within 1.65 unity, a value lower than the standard deviations calculated by the software. The completeness of the cells of increasing size, calculated for Roccamonfina area (Table 5) using the plots as sampling units, gives a gradual increase of the number of SSS, up to over 50% of the 9 km² cells and 80% of the 36 km² cells.

In Table 6 the observed and estimated species richness and the completeness of the different habitats using the plots as sampling units are reported. For the Roccamonfina area the completeness of the habitats is high except for agricultural environments. The chestnut orchards host the higher species richness (38 species, 82% of the whole area), followed by the open habitats such as meadows and shrublands (33 species). In the other study areas the completeness is relatively low for the broadleaved woodlands of Vairanese and open habitats of the W-Matese, indicating a still not adequate sampling for such habitats. For a better comparison of the species richness among the different habitats, considering that more than 70% of the plots are located inside chestnut orchards, the rarefaction curves were plotted for Roccamonfina habitats (Figure 6), limited to 100 plots. The richness curve rises in a steeper way in the chestnut orchards but it is overtaken by artificial habitats around 30 plots and by open habitats around 50 plots. The richness of broadleaved woodlands and chestnut coppices is always lower, as expected.

DISCUSSION

The higher species richness is correlated to the sampling effort, expressed as number of plots, as well as the ecological features of the areas and their extension. This parameter is known, in ecology as the species/area relationship (SAR - Preston 1962) and it could be used to compare and estimate species richness of floristic atlases only under certain conditions that, if disregarded impede its extrapolation (Vallet et al. 2012). The correlation analysis here performed confirms that the higher is the number of sampling units (plot) in an area, the higher will be the observed species richness. Comparing the richness of the three studied areas plotted by rarefaction curves, highlights that with the same sampling effort...
Figure 2. Distribution maps: a—number of plots | b—number of observed species | c—Completeness level, for the 1 km² cells covering the three study areas.
| Table 1. Data of the three study areas and list of the taxa considered for the analysis. |
|-------------------------------------------------|-----------------|-----------------|-----------------|
| | Roccamonfina | Vairanese | W-Matese |
| Sobs | 46 | 32 | 33 |
| Area (km²) | 210 | 18 | 20 |
| 1 km² cells | 163 | 18 | 20 |
| Altitude (min-max) | 150–1005 | 125–588 | 150–811 |
| Number of plots | 1184 | 121 | 124 |
| Database-records | 3046 | 263 | 296 |
| Records/plot | 2.57 | 2.17 | 2.39 |
| Plot/km² | 7.26 | 6.72 | 6.2 |

- **Anacamptis coriophora** (L.) R.M.Bateman, Pridgeon & M.W.Chase
- **Anacamptis maria** (L.) R.M.Bateman, Pridgeon & M.W.Chase
- **Anacamptis papilionacea** (L.) R.M.Bateman, Pridgeon & M.W.Chase
- **Anacamptis pyramidalis** (L.) Rich.
- **Cephalanthera damasonium** (Mill.) Druce
- **Cephalanthera longifolia** (L.) Frisch
- **Cephalanthera rubra** (L.) Rich.
- **Dactylorhiza maculata** (L.) Soó subsp. *saccifera* (Brongn. ) Diklić
- **Dactylorhiza romana** (Sebast.) Soó subsp. *romana*
- **Dactylorhiza sambucina** (L.) Soó
- **Epipactis exilis** P.Delforge
- **Epipactis helleborine** (L.) Crantz subsp. *helleborine*
- **Epipactis microphylla** (Ehrh.) Sw.
- **Epipactis muelleri** Godfery
- **Epipactis mariae** (Croce, Bongiorni, De Vivo & Fori) Presser & S.Hertel
- **Epipactis placenta** Bongiorni & Grünanger
- **Gymnadenia conopsea** (L.) R.Br.
- **Himantoglossum adriaticum** H.Baumann
- **Limodorum abortivum** (L.) Sw.
- **Neottia nidus-avis** (L.) Rich.
- **Neottia ovata** (L.) Bluff & Fingerh.
- **Ophrys apifera** Huds.
- **Ophrys argolica** H.Fleischm. ex Vierh. subsp. *crabronifera* (Mauri) Faurh.
- **Ophrys bertoloni** Moretti subsp. *bertoloni*
- **Ophrys bombyliiflora** Link
- **Ophrys exaltata** Ten. subsp. *monts-leonis* (O.Danesch & E.Danesch) Soca
- **Ophrys holosericea** (Burm.f.) Greuter subsp. *gracilis* (Büel, O.Danesch & E.Danesch) O.Danesch & E.Danesch
- **Ophrys holosericea** (Burm.f.) Greuter subsp. *holosericea*
- **Ophrys incubacea** Bianca
- **Ophrys insectifera** L.
- **Ophrys lutea** Cav.
- **Ophrys promontorii** O.Danesch & E.Danesch
- **Ophrys sphegodes** Mill. subsp. *sphegodes*
Orchid checklists and atlases: a case study from southern Italy

(i.e., the same number of plots), the richest area can host a relatively lower number of species than the less rich area. Nevertheless, such kind of analysis requires the same exhaustivity of the studies for each area. The overall completeness of the study areas gives values close to 90% and consistently above 70%. Also, very interesting is the data emerging from the estimates of the richness and the completeness calculated using the 1 km$^2$ cells of the UTM grid as sampling units. Such size could be very useful to study larger areas or to include lower precision data in the analysis and the completeness values did not differ significantly from the resulting estimates obtained using 100 x 100 m sampling units (plots). For the Roccamonfina area, in addition, even using increasing size cells as sampling units, the estimates do not vary significantly. This result can be taken into account whenever we have to choose the better grid resolution to draw atlases from non punctual data (e.g., literature data or observations with low location accuracy). The elaborations should follow, in this case, a reverse path: starting from a large sampling unit (e.g., a 10 x 10 km cells UTM grid), decreasing the size of the sampling units and calculating the completeness for the study area. Since small size cells will have more probability to hold ‘singletons’ (unique presence data) for a bigger number of species, the used estimators will

---

**Table 2. Total completeness values for the three study areas using 100 x 100 m plots as sampling units.**

|                | Roccamonfina | Vairanese | W-Matese |
|----------------|--------------|-----------|----------|
| n. Plots (100 x 100 m) | 1184 | 121 | 124 |
| Sobs            | 46           | 32        | 33       |
| $S_{n=0}$       | 49           | 35.72     | 37.9     |
| $S_{n=1}$       | 51.99        | 40.93     | 39.94    |
| Completeness %  | 88.5         | 78.2      | 82.6     |

**Table 3. Total completeness values for the three study areas using 1 km$^2$ cells as sampling units.**

|                | Roccamonfina | Vairanese | W-Matese |
|----------------|--------------|-----------|----------|
| No. of 1 km$^2$ cells | 163 | 18 | 20 |
| Sobs            | 46           | 32        | 33       |
| $S_{n=0}$       | 48.98        | 40.5      | 36.33    |
| $S_{n=1}$       | 51.96        | 40.5      | 40.3     |
| Completeness %  | 88.5         | 79.0      | 81.9     |

**Table 4. Levels of completeness values of the 1 km$^2$ cells, for the three study areas.**

|                | Roccamonfina | Vairanese | W-Matese |
|----------------|--------------|-----------|----------|
| Completeness level % | n. cells | % | n. cells | % | n. cells | % |
| n.e.           | 84 | 51.5 | 7 | 38.9 | 9 | 45.0 |
| 0–25          | 3 | 1.8 | 2 | 11.1 | 2 | 10.0 |
| 25–50        | 5 | 3.1 | 1 | 5.6 | 1 | 5.0 |
| 50–75       | 35 | 21.5 | 4 | 22.2 | 5 | 25.0 |
| 75–100     | 36 | 22.1 | 4 | 22.2 | 3 | 15.0 |
| Total       | 163 | 18 | 20 | |
| 555        | 58 | 35.6 | 6 | 33.3 | 5 | 25.0 |
give higher estimates of richness and, therefore, lower values of completeness.

For the same reason linked to the presence of singletons, in our study the number of sufficient studied squares (SSS) increases as their size become bigger. In the case of Roccamonfina area, using a grid of 9 km\(^2\) cells, a half of them are classified as SSS. The distribution of the completeness for a grid of 1 km\(^2\) cells (Table 4), on the other hand, is comparable for Roccamonfina and Vairanese, with more than 33% of the squares classified as SSS while for W-Matese area this value reaches only 25%. To assess whether these rates represent a good result (i.e., the area is exhaustively well studied), we can refer to the choice of the limit of 65% to consider a cell as sufficiently studied. In Bruno et al. (2012) this completeness limit has been chosen to select a useful number of squares to perform further analysis. These authors, for all the four considered taxonomic groups, get lower portion of squares SSS compared to the portion we get for our studied areas. Nevertheless, the absolute number of SSS for both Vairanese and W-Matese areas (respectively six and five squares) is too low and recall the need to continue the study in these two areas. The stratified analysis by habitat types underlines firstly what habitats need more studies or are less suitable for orchids. For example, agricultural habitats for Roccamonfina would need further sampling since their completeness is only 55% (Table 6). It could be expected that, adding further sampling, the completeness would increase even without an increasing of the species...
These habitats are in fact less suitable to host orchids as they are affected by frequent and strong ecological changes (e.g., soil tillage, switching to other crops, supply of nutrients). Such considerations could be made for the broadleaved woodlands of the Vairanese area, mostly represented by Holm oaks woodlands with very low light in the understory since orchids abundance is highly correlated to light regime (Djordjević & Tsiftsis 2020; Hrivnák et al. 2020). On the contrary we expect that the low completeness value for the open habitats of the W-Matese area is due to a high theoretical richness of such habitats, not fully detected by the sampling activity. In other words, the sampling effort for the open habitats of the W-Matese area is still insufficient.

![Rarefaction curves for the different habitats of the Roccamonfina area.](image)

**Figure 6.** Rarefaction curves for the different habitats of the Roccamonfina area.

---

**Table 5.** Levels of completeness values of the cells of different size, for the Roccamonfina area (n.e. = not evaluated).

|    | 1 km² | 4 km² | 9 km² | 16 km² | 25 km² | 36 km² |
|----|-------|-------|-------|--------|--------|--------|
|    | n. cells | % | n. cells | % | n. cells | % | n. cells | % | n. cells | % | n. cells | % |
| n.e. | 84 | 51.5 | 23 | 37.7 | 10 | 16.4 | 5 | 22.7 | 3 | 20 | 1 | 9.1 |
| 0–25 | 3 | 1.8 | 1 | 1.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25–50 | 5 | 3.1 | 4 | 6.6 | 5 | 8.2 | 1 | 4.5 | 1 | 6.7 | 0 | 0 |
| 50–75 | 35 | 21.5 | 14 | 23 | 6 | 9.8 | 7 | 31.8 | 4 | 26.7 | 2 | 18.2 |
| 75–100 | 36 | 22.1 | 19 | 31.1 | 12 | 19.7 | 9 | 40.9 | 7 | 46.6 | 8 | 72.7 |
| tot | 163 | 61 | 33 | 22 | 15 | 11 | 17 | 51.5 | 11 | 50 | 9 | 60 |
| SSS | 58 | 35.6 | 28 | 45.9 | 17 | 51.5 | 11 | 50 | 9 | 60 | 9 | 81.8 |

**Table 6.** Completeness values of the main habitats in the three areas.

| Habitats | Sobs | Plots | SChao2 | SJack1 | C % |
|----------|------|-------|--------|--------|-----|
| Artificial (incl. Road verges) | 25 | 50 | 28.9 | 33.8 | 73.9 |
| Agriculture | 11 | 18 | 20.0 | 16.7 | 55.0 |
| Open habitats | 33 | 158 | 36.5 | 40.9 | 80.6 |
| Broadleaved woodlands (excl. Chestnut woods) | 25 | 82 | 27.1 | 30.9 | 80.9 |
| Chestnut coppices | 18 | 62 | 19.0 | 21.0 | 85.9 |
| Chestnut orchards | 38 | 839 | 44.0 | 44.0 | 86.4 |
| Open habitats | 26 | 89 | 26.5 | 29.0 | 89.8 |
| Broadleaved woodlands | 24 | 33 | 42.0 | 35.6 | 57.1 |
| Evergreen woodlands | 6 | 8 | 6.7 | 8.6 | 69.6 |
| Open habitats | 28 | 60 | 100.0 | 39.8 | 28.0 |
| Broadleaved woodlands | 27 | 57 | 28.6 | 32.0 | 84.4 |
Also, the rarefaction curves allow ecological considerations (Figure 6). The chestnut orchards represent an ecosystem made of a mosaic between woodlands and meadows, so they are a suitable habitat for the most heliophilous species as well as for the nemoral ones. This explains why their average species richness increases steeply even with a few plots (it is possible to observe more than 20 species in one plot). Nevertheless, on a larger scale, the richness of chestnut orchards is higher than the richness in open habitats only because of the higher area occupied by the former. When the curves are limited to 50 plots, surprisingly the richest habitats are the artificial areas. This result can be explained with the apophyte behavior of many orchids species (Adamowski 2006) and with the fact that we considered the roadsides as artificial habitats. Such environments can host many species characteristics of open habitats such as meadows and grasslands, and constitute important refuge areas for native species (Auestad et al. 2011).

Overall, the analysis of the three datasets allowed the sampling effort to be evaluated and gave useful indications to where and how to conduct the future researches. Moreover, some suggestions on the use of statistical tools to compare different study areas were given. For two areas (Roccamonfina and Vairanese), there is a sufficient level of knowledge of how the orchids richness is distributed, if we assume that a low completeness value in two squares out of three could be due to the lack of suitable habitats (i.e., urban areas or intensive agriculture areas) and to the difficult to locate a sufficient number of sampling units or plots. The squares with no data or with a lower completeness should be regarded as the highest priority areas for the future floristic research. Sampling these areas could increase the level of knowledge (i.e., the completeness value) and could lead to detect new species for the squares or for the studied area. The analysis of the floristic richness and the completeness of every habitat in a less known area would be very useful to prioritize, in each cell of a chosen grid, where to focus the research.

CONCLUSIONS

In conclusion, this study highlights that the quality of a floristic research can benefit from the evaluation of the completeness. Its calculation allows the creation of knowledge/ignorance maps for orchids at different scale using grids at different resolutions (e.g., from cells of 1 km² for small islands and reserves to cells of 100 km² for regions). A randomized and stratified sampling design would reduce the sampling bias, enable the use of abundance indices rather than presence/absence data and allow the investigation on the relation between species richness and environmental variables. It is often necessary, however, to take into account a large amount of data lacking accuracy or uniformity as is the case of data from literature or collected by different and sometimes occasional contributors (e.g., in citizen science projects).

In any case it is desirable in each modern floristic study and particularly orchids distribution study, a quantitative analysis of the work expressing the results not only as the total number of species observed and their distribution but focusing more on the sampling methods and on the distribution of the knowledge. Even if a sampling design avoiding preferential sampling would be desirable but not always possible (e.g., when using data from online platforms or literature), the proposed methods would help the authors to evaluate the sampling effort, identify the less studied areas or postpone the publication of their checklists and atlases until an acceptable level of exhaustivity, or completeness, would be reached.

REFERENCES

Adamowski, W. (2006). Expansion of native orchids in anthropogenous habitats. Polish Botanical Studies 22: 35–44.

Auestad, I., K. Rydgren & I. Austad (2011). Road verges: potential refuges for declining grassland species despite remnant vegetation dynamics. Annales Botanici Fennici 48: 289–303. https://doi.org/10.5735/085.048.0401

Asase, A. & A.T. Peterson (2016). Completeness of digital accessible knowledge of the plants of Ghana. Biodiversity Informatics 11: 1–11. https://doi.org/10.17161/bi.v11i11.5860

Bruno, D., D. Sánchez-Fernández, A. Millán, R.M. Ros, P. Sánchez-Gómez & J. Velasco (2012). Assessing the quality and usefulness of different taxonomic groups inventories in a semiarid Mediterranean region. Biodiversity and Conservation 21(6): 1561–1575. https://doi.org/10.1007/s10531-012-0263-9

Chao, A. & C.H. Chiu (2016). Species richness: estimation and comparison, pp 1–26. In: Balakrishnan, N., T. Colton, B. Everitt, W. Piegorsch, F. Ruggieri & J. Teugels (eds). Wiley StatsRef: statistics reference on-line. John Wiley & Sons, Hoboken, New Jersey. https://doi.org/10.1002/9781118445112.stat03432.pub2

Colwell, R.K. (2013). EstimateS: Statistical estimation of species richness and shared species from samples. Version 9 and earlier.
User’s Guide and application. http://purl.oclc.org/estimates/ Accessed on 1 March 2020.

Crain, B.J. & M. Fernández (2020). Biogeographical analyses to facilitate targeted conservation of orchid diversity hotspots in Costa Rica. Diversity and Distributions 26: 856–866. https://doi.org/10.1111/ddi.13062

Cribb, P.J., S.P. Kell, K.W. Dixon & R.L. Barrett (2003). Orchid Conservation: A global perspective, pp 1–25. In: Dixon, K.W., S.P. Kell, R.L. Barrett & P.J. Cribb (eds). Orchid Conservation. Natural History Publications, Kota Kinabalu, Sabah.

Croce, A. (2012). Le Orchidee dei monti S. Angelo, Caiavola, S. Nicola (Vairano Patenora e Pietravairano, Caserta). Giros Nottizie 49: 60–67.

Croce, A. & R. Nazzaro (2012). The orchid flora of Roccamonfina-foce Garigliano Regional Park. Journal Europäischer Orchideen 44(3): 509–583.

Croce, A. & R. Nazzaro (2017). An atlas of orchids distribution in the Campania region (Italy), a citizen science project for the most charming plant family. Italian Botanist 4: 15–32. https://doi.org/10.3897/italianbotanist.4.14916

Djordjević, V. & S. Tsiftsis (2020). The role of ecological factors in distribution and abundance of terrestrial orchids, pp 1–71. In: Merillon, J.B. & K. Kodja (eds.). Orchids Phytochemistry, Biology and Horticulture. Reference Series in Phytochemistry, Springer International Publishing: Basel, Switzerland. https://doi.org/10.1007/978-3-030-11257-8

Dressler, R.L. (1981). The orchids: natural history and classification. Harvard University Press, Cambridge, MA, USA, 332pp. https://doi.org/10.2307/1219717

Efimov, P.G. (2020). Orchids of Russia: annotated checklist and geographic distribution. Nature Conservation Research 5(Suppl. 1): 1–18. https://doi.org/10.24189/nocr.2020.018

Frangini, G., R. Romolini, F. Sodi, M. Bisti, L. Forbicioni, M.A. Rusci & G. Cortesi (2019). Orchidee dell’isola di Gorgona (Arcipelago Toscano). Giros Orchidee Spontanea d’Europa 62(1): 17–31.

Galesi, R. & R. Lorenz (2010). Le Orchideae della Riserva Naturale Orientata “Oasi faunistica di Vendicari (Sicilia sud-orientale). Journal Europäischer Orchideen 42(1): 149–166.

García Márquez, J.R., C.F. Dormann, J.H. Sommer, M. Schmidt, A. Thiembiano, S.S. Da, C. Chatelain, S. Dressler & W. Barthlott (2012). A methodological framework to quantify the spatial quality of biological databases. Biodiversity & Ecology 4: 25–39. https://doi.org/10.7809/8-1.00057

Gotelli, N.J. & R.K. Colwell (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecological Letters 4: 379–391. https://doi.org/10.1046/j.1461-0248.2001.00230.x

Gotelli, N.J. & R.K. Colwell (2011). Estimating species richness, pp. 39–54. In: Magurran, A.E. & B.J. McGill (eds.). Biological Diversity: Frontiers in Measurement and Assessment. Oxford University Press, Oxford, United Kingdom, 345 pp.

Hammer, Ø., D.A. Harper & P.D. Ryan (2001). PAST: paleontological statistics software package for education and data analysis. Palaontologica electronica 4(1): 9. https://paleo-electronica.org/2001_1/past/issue1_01.htm Accessed on 01 March 2020.

Hortal, J., P.A.V. Borges & C. Gaspar (2006). Evaluating the performance of species richness estimators: sensitivity to sample grain size. Journal of Animal Ecology 75: 274–287. https://doi.org/10.1111/j.1365-2656.2006.01048.x

Hrivnák, M., M. Slezáč, D. Galváněk, J. Vlček, E. Belanová, V. Rizová, D. Senko & R. Hrivnák (2020). Species Richness, Ecology, and Prediction of Orchids in Central Europe: Local-Scale Study. Diversity 12(4): 154. https://doi.org/10.3390/d12040154

Idohou, R., A. Arino, A. Assogbadjo, R.G. Kakai & B. Sinsin (2015). Diversity of wild palms (Arecaceae) in the Republic of Benin: Finding the gaps in the national inventory combining field and digital accessible knowledge. Biodiversity Informatics 10: 45–55. https://doi.org/10.17161/biivoi.1012.4914

Kapotodi, E. & S. Tsifitsis (2019). Contribution to the knowledge of the orchid flora of Lefkas island (Ionian Islands, Greece). Botanika Chronika 22: 127–143.

Marrero, Á., M. Claessens, D. González, C. Santiago, & J. Claessens (2019). Chorological additions and distribution of the native orchids of Gran Canaria. Botánica Macaronésica 30: 65–88.

Martínez-Sanz, C., F. García-Criado, C.F. Alaez & M.F. Alaez (2010). Assessment of richness estimation methods on macroinvertebrate communities of mountain ponds in Castilla y León (Spain). Anales de linnologio - International Journal of Linneology 46 (2): 101–110. https://doi.org/10.1051/linn/2010008

Mathew, J., & K.V. George (2015). Checklist of Orchids of Kottavasal Hills in Achankovil Forests, southern Western Ghats, (Kollam, Kerala), India. Journal of Threatened Taxa 7(10): 7691–7696. https://doi.org/10.1109/JOTT.2015.7691-6

Palmer, M.W. (1990). The estimation of species richness by extrapolation. Ecology 71(3): 1195–1198. https://doi.org/10.2307/1937387

Popovich, A. V., E.A. Averyanova & L.M. Shagarov (2020). Orchids of the Black Sea coast of Krasnodarsky Krai (Russia): current state, new records, conservation. Nature Conservation Research 5(Suppl. 1): 46–68. https://doi.org/10.24189/nocr.2020.047

Preston, F.W. (1962). The canonical distribution of commonness and rarity: Part I. Ecology 43(2): 185–215. https://doi.org/10.2307/1931976

QGIS Development Team (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.org/. Downloaded on 01 march 2020

Rocchini, D., I. Hortal, S. Lengyel, I.M. Lobas, A. Jimenez-Valverde, C. Ricotta, G. Bacaro & A. Chiarucci (2011). Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. Progress in Physical Geography 35(2): 211–226. https://doi.org/10.1177/0309133311499491

Soberón, J.M., J.B. Llorente & L. Ollate (2000). The use of specimen-label databases in conservation purposes: an example using Mexican Papilionid and Pierid butterflies. Biodiversity and Conservation 9(10): 1441–1466. https://doi.org/10.1023.A:1008987010383

Soberón, J.M., R. Jiménez, J. Golubov & P. Koleff (2007). Assessing completeness of biodiversity databases at different spatial scales. Ecology 88(1): 152–160. https://doi.org/10.1111/j.0012-7692.2006.00467.x

Ugland, K.J., J.S. Gray & K.E. Ellingsen (2003). The species-accumulation curve and estimation of species richness. The Journal of Animal Ecology 72(5): 888–897. https://doi.org/10.1046/j.1365-2656.2003.00748.x

Vallet, J., M. Rambaud, L. Coquel, L. Poncet & F. Hendoux (2012). Effort d’échantillonnage et atlas floristiques-exhaustivité des mailles et caractérisation des lacunes dans la connaissance. Comptes Rendus Biologies 335(12): 753–763. https://doi.org/10.1016/j.crvi.2012.11.005

WCSP (2019). World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. http://wcsp.science.kew.org/ home.do Downloaded on 20 November 2020.

Weigelt, P., C. König & H. Kreft (2020). GIFT–A global inventory of floras and traits for macroecology and biogeography. Journal of Biogeography 47(1): 16–43. https://doi.org/10.1111/jbi.13623
Estimating the completeness of orchid checklists and atlases: a case study from southern India
– Antonio Croce, Pp. 20311–20322

A floristic survey across three coniferous forests of Kashmir Himalaya, India – a checklist
– Ashaq Ahmad Dar, Akhtar Hussain Malik & Narayanawsamy Parthasarathy, Pp. 20323–20345

Associations of butterflies across different forest types in Uttarakhand, western Himalaya, India: implications for conservation planning
– Arun Pratap Singh, Pp. 20346–20370

Comparison of bird diversity in protected and non-protected wetlands of western lowland of Nepal
– Jagan Nath Adhikari, Janak Raj Khatiwada, Dipendra Adhikari, Suman Sapkota, Bishnu Prasad Bhattarai, Deepak Rijal & Lila Nath Sharma, Pp. 20371–20386

Local hunting practices and perceptions regarding the distribution and ecological role of the Large Flying Fox (Chiroptera: Pteropodidae: Pteropus vampyrus) in western Sarawak, Malaysian Borneo
– Jayaishan Mohd-Azlan, Joon Yee Yong, Nabilah Norshuhadah Mohd Hazzrol, Philoveny Pengiran, Ariantul Atong & Sheema Abdul Aziz, Pp. 20387–20399

New distribution record of globally threatened Ocean Turf Grass Halophila beccarii
– Aswathi Anilkumar, Stephen Sequeira, Arun Christy & S.M. Arsha, Pp. 20400–20405

Macrolichens of Mathikettan Shola National Park, Western Ghats: a preliminary investigation with some new records
– Aswathi Anilkumar, Stephen Sequeira, Arun Christy & S.M. Arsha, Pp. 20404–20405

An inventory of new orchid (Orchidaceae) records from Kozhikode, Kerala, India
– M. Sulaiman, C. Murugan & M.U. Sharief, Pp. 20413–20425

Abundance and spatial distribution analyses of Stemonoporus moonii (Dipterocarpaceae) - a critically endangered species endemic to Sri Lanka
– K.A.M.R.P. Atapattu, H.D.D.C.K. Perera, H.S. Kathriarachchi & A.R. Gunawardena, Pp. 20426–20432

Plant diversity of Point Calimere Wildlife Sanctuary and fodder species grazed by the Blackbuck Antilope cervicapra L.
– Ashutosh Kumar Upadhyay, A. Andrew Emmanuel, Ansa Sarah Varghese & D. Narasimhan, Pp. 20433–20443

Raptors observed (1983–2016) in National Chambal Gharial Sanctuary: semi-arid biogeographic region suggestions for parametric studies on ecological continuity in Khathiar-Gir Ecoregion, India
– L.A.K. Singh, R.K. Sharma & Udayan Rao Pawar, Pp. 20444–20460

Nesting success of Sharpe’s Longclaw (Macronyx sharpei Jackson, 1904) around the grasslands of lake O’bolosat Nyandarua, Kenya
– Hamisi Ann Risper, Charles M. Warui & Peter Njorge, Pp. 20461–20468

Population, distribution and diet composition of Smooth-coated Otter Lutrogale perspicillata Geoffroy, 1826 in Hosur and Dharmapur Forest Divisions, India
– Nagarajan Baskaran, Raman Sivaraj Sundarraj & Raveendranathanangiillai Sanil, Pp. 20469–20477

Utilization of home garden crops by primates and current status of human-primate interface at Galigamuwa Divisional Secretariat Division in Kegalle District, Sri Lanka
– Charmalie Anuradhie Dona Nahallage, Dahanakge Ayeshya Madushani Dasanayake, Dilan Thisaru Hewamanna & Dissanayakalage Tharaka Harshani Ananda, Pp. 20478–20487

Revival of Eastern Swamp Deer Rucervus duvaucelii ranjitinsinh (Groves, 1982) in Manas National Park of Assam, India
– Nazrul Islam, Aftab Ahmed, Rathin Barman, Sanatan Deka, Bhaskar Choudhury, Prasanta Kumar Saikia & Jyotishman Deka, Pp. 20488–20493

Tryptosomos evansi infection in a captive Indian Wolf Canis lupus pallipes – molecular diagnosis and therapy
– Manojita Dash, Sarat Kumar Sahu, Santosh Kumar Gupta, Niranjan Sahoo & Debarat Mohapatra, Pp. 20494–20499

COVID-19 and civil unrest undoing steady gains in karst conservation and herpetological research in Myanmar, and an impediment to progress
– Evan S.H. Quah, Lee L. Grismer, Perry L. Wood, Jr., Aung Lin & Myint Kyaw Thura, Pp. 20500–20502

Distribution of Smooth-coated Otters Lutrogale perspicillata (Mammalia: Carnivora: Mustelidae): in Ratnagiri, Maharashtra, India
– Swanand Patil & Kranti Yardi, Pp. 20511–20516

Wildlife at the crossroads: wild animal road kills due to vehicular collision on a mountainous highway in northwestern Himalayan region
– Muzaffer A. Kichloo, Asha Sohil & Neeraj Sharma, Pp. 20517–20522

Counting the cost: high demand puts Bunium persicum (Boiss.) B.Fedtsch. in jeopardy
– Monika Sharma, Manisha Mathela, Rupali Sharma, Himanshu Bargali, Gurinderjit S. Goraya & Amit Kumar, Pp. 20530–20533

First record of Parasitic Jaeger Stercorarius parasiticus (Aves: Charadriiformes: Stercorariidae) from inland freshwater Inle Lake, Myanmar
– Sai Sein Lin Do, Myint Kyaw, L.C.K. Yun, Min Zaw Tun, Yar Zar Lay Naung, Soe Naing Aye & Swan C. Renner, Pp. 20534–20536

Book Review
Capparis of India
– V. Sampath Kumar, Pp. 20537–20538