

Sampling completeness in seed dispersal networks: When enough is enough



José M. Costa^{a,b,*}, Luís P. da Silva^{a,b}, Jaime A. Ramos^b, Ruben H. Heleno^a

^aCentre for Functional Ecology, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

^bMARE—Marine and Environmental Sciences Centre, Department of Life Sciences, University of Coimbra, Coimbra, Portugal

Received 23 December 2014; accepted 16 September 2015
Available online 1 October 2015

Abstract

Ecological networks are an increasingly popular tool to explore community assembly rules and frame practical conservation issues. However, most described networks vary largely in sampling effort, hampering the distinction of true biological patterns from artefacts caused by poor sampling. Identifying entire seeds in the droppings of mist-netted birds is generally considered a preferred sampling method for building unbiased, quantitative seed dispersal networks. We retrieved seeds from the droppings of 936 mist-netted birds captured during five days in seven sites in Portugal and estimated sampling completeness as the diversity of seed species, disperser species, and links detected with respect to those predicted by Chao 2 estimator. In one of those sites, sampling effort was extended to 25 days to evaluate the sensitivity of ten network structure descriptors to increasing sampling effort. After five sampling days we detected 93% of the seed species, 97% of the disperser species, and 79% of the links predicted by Chao 2, however sampling for 25 days resulted in the detection of more seeds, dispersers, and links than those estimated at day 5. Most network descriptors only began to stabilize around day 8, except for Connectance and Weighted Connectance that stabilized earlier. Similarly, only after 8 days most networks descriptors significantly departed from the confidence interval estimated by null models exclusively constrained by species abundances, thus reflecting independent ecological patterns. Nestedness was the only exception, as it never departed from the null models. We suggest that Chao 2 may slightly underestimate the real diversity and that in our case at least eight sampling days were needed to build a sound seed dispersal network as 67% of the seeds, 88% of the dispersers, and 71% of the links were detected. Our results have important implications for the interpretation of seed dispersal networks because under-sampled networks may produce biased descriptors that do not suitably characterize the focal communities.

Zusammenfassung

Ökologische Netzwerke werden immer beliebter als eine Methode, um die ‘assembly rules’ für Gemeinschaften zu untersuchen und praktische Naturschutzregeln zu entwerfen. Indessen variieren die meisten Netzwerke erheblich hinsichtlich des Probenahmeaufwandes, was die Unterscheidung zwischen tatsächlichen biologischen Mustern und Artefakten, die sich aus unzureichender Probenahme ergeben, erschwert. Die Bestimmung von Samen aus dem Kot von Vögeln aus Japannetzfängen gilt als eine bevorzugte Sammelmethode, um unverfälschte, quantitative Samenausbreitungsnetzwerke zu konstruieren. Wir separierten Samen aus dem Kot von 936 gefangenen Vögeln (sieben Fangplätze in Portugal; Fangzeitraum: i.d.R. fünf Tage)

*Corresponding author at: Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, Coimbra, Portugal. Tel.: +351 239 855 210.

E-mail address: jmgncosta@gmail.com (J.M. Costa).

und bestimmten die Vollständigkeit der Erfassung als die Diversität der festgestellten Samenarten, Vogelarten und Verbindungen in Relation zum nach Chao 2 berechneten Schätzwert. An einem der Fangplätze wurde die Probenahme auf 25 Tage ausgedehnt und die Reaktion von zehn Netzwerkdeskriptoren auf die zunehmende Stichprobengröße ermittelt. Nach fünf Probetagen erhielten wir 93% der Samenarten, 97% der Vogelarten und 79% der vorhergesagten (Chao 2) Interaktionen. Der Fangzeitraum von 25 Tagen erbrachte indessen noch mehr Samenarten, Vogelarten und Interaktionen. Die meisten Netzwerkdeskriptoren begannen sich nach etwa acht Tagen zu stabilisieren. Nur Konnektanz und gewichtete Konnektanz stabilisierten sich früher. In gleicher Weise wichen die meisten Netzwerkdeskriptoren erst nach acht Tagen signifikant von den Vertrauensbereichen ab, die für Nullmodelle mit limitierten Artenabundanzen berechnet wurden, was unabhängige ökologische Muster anzeigt. Die ‘nestedness’ war die einzige Ausnahme und wich in keinem Fall von den Nullmodellen ab. Wir meinen, dass Chao 2 die tatsächliche Diversität leicht unterschätzen könnte und dass in unserem Fall mindestens acht Fangtage benötigt wurden, um ein vernünftiges Ausbreitungsnetzwerk zu konstruieren, da 67% der Samenarten, 88% der Vogelarten und 71% der Interaktionen (bezogen auf den 25-tägigen Fangzeitraum) entdeckt wurden. Unsere Ergebnisse haben wichtige Konsequenzen für die Interpretation von Samenausbreitungsnetzwerken, weil unzureichend beprobte Netzwerke verfälschte Deskriptoren hervorbringen können, die zu unbrauchbaren Beschreibungen der untersuchten Gemeinschaften führen.

© 2015 Gesellschaft für Ökologie. Published by Elsevier GmbH. All rights reserved.

Keywords: Asymptotic estimators; Cumulative sampling; Endozoochory; Food-webs; Mist-nets; Sampling effort

Introduction

The study of community-wide species interaction networks experienced a tremendous growth in recent years and proved a powerful tool to explore many processes in ecology, being particularly valuable in disentangling the relationships between the structure and function of nature’s “entangled bank” (Heleno et al. 2014; Olesen, Dupont, Hagen, Rasmussen, & Trøjelsgaard 2012; Darwin 1859). The use of ecological networks deepened our understanding of important conservation issues such as environmental degradation (e.g. Traveset et al. 2013; Tylianakis, Tscharntke, & Lewis 2007) and ecological restoration (e.g. Heleno, Lacerda, Ramos, & Memmott 2010).

Seed dispersal is one of the research areas where ecological networks attracted greatest interest (e.g. Heleno, Olesen, Nogales, Vargas, & Traveset 2013; Carlo & Yang 2011; Donatti et al. 2011; Mello et al. 2011). The dispersal of seeds away from the mother plant is a key process, allowing plants to avoid competition, find suitable conditions for germination, and expand their distribution range (Traveset, Heleno, & Nogales 2014). Due to their high mobility, frugivorous birds are the main seed dispersers in most ecosystems (Herrera 1984; Traveset et al. 2014). Avian seed dispersal data can be collected by: (1) recording fruit consumption by birds on focal plants (Olesen et al. 2011), (2) identifying entire seeds in the droppings of mist-netted birds (Heleno et al. 2013), and (3) identifying seeds in droppings collected in the field and identifying the disperser species with molecular techniques (González-Varo, Arroyo, & Jordano 2014).

Most network structure descriptors are affected by sampling effort to some degree (Blüthgen 2010; Blüthgen, Fründ, Vázquez, & Menzel 2008; Rivera-Hutinel, Bustamante, Marín, & Medel 2012), with qualitative indices being more sensitive to sample size than quantitative analogues (Banasek-Richter, Cattin, & Bersier 2004). Specifically, poor

sampling underestimates the real diversity of links, truncating estimated trophic breath and leading to a biased network structure (Blüthgen et al. 2008). On the other hand, as implementing such a holistic approach is inherently highly labour-intensive, it is important to know when further effort will not significantly increase the accuracy of the community description, thus avoiding unnecessary work load (Hegland, Dunne, Nielsen, & Memmott 2010). This effort is of utmost importance to allow meaningful cross-study comparisons (Heleno et al. 2014).

Richness estimators based on species and link accumulation curves are a powerful way to evaluate sampling completeness (Chacoff et al. 2012; Olito & Fox 2015; Rivera-Hutinel et al. 2012), where the number of missing species is estimated as those needed to reach a theoretical asymptote (Chacoff et al. 2012). Although some statistical methods have been suggested to ease the problem (e.g. Bartomeus 2013), there are no satisfactory solutions for poor sampling, and the sampling effort needed to build comprehensive seed dispersal networks from which theoretical and applied conclusions can be drawn has not yet been evaluated. In this study we aim to evaluate (1) the effect of sampling effort on the completeness of seed dispersal networks based on the analysis of droppings from mist-netted birds, and (2) the sampling effort needed to build high quality seed dispersal networks so that reliable network descriptors can be calculated.

Materials and methods

Study sites and data collection

On five consecutive days in the first half of September 2012, birds were captured in seven sites throughout Portugal ranging from agro-forestry systems to secondary native forest (Fig. 1) (see Costa et al. 2014). In each site/day,

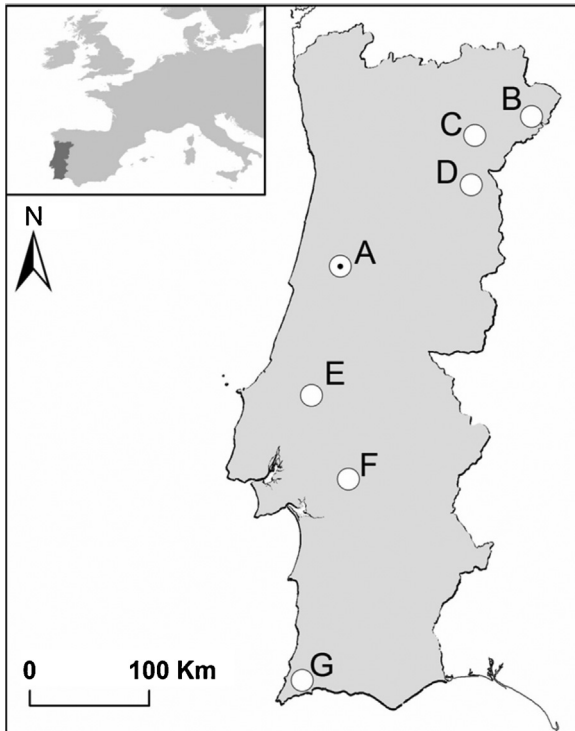


Fig. 1. Location of the study sites. (A) Larçã, Souselas; (B) Atenor, Miranda do Douro; (C) Nozelos, Torre de Moncorvo; (D) Santa Maria de Aguiar dam, Figueira de Castelo Rodrigo; (E) Casais da Estrada, Achete; (F) Herdade de Freixo do Meio, Foros de Vale Figueira; (G) Bensafrim, Lagos. Site A was sampled for 25 consecutive days; all other sites were sampled for five days.

80–100 m of mist-nets (according to the site specific conditions) were operated during 5 h after dawn. Nets were visited every 30 min and captured birds were individually placed in ringing bags to produce droppings (Heleno et al. 2013). Intact seeds were later extracted from the droppings and identified under a dissecting microscope by comparison with a reference collection. Interaction frequency was quantified as the number of droppings from each bird species containing at least one intact seed of any of the plant species. We built quantitative seed dispersal networks for each site by considering the cumulative samples collected up to day i . We used the number of days as a measure of sampling effort because there were no significant differences in the number of droppings collected per day (average \pm SD droppings collected per day = 124.8 ± 22.7 , see Appendix A: Table 1). The availability of fleshy fruits was independently assessed at each site by counting all ripe standing fruits along three 2×25 m transects parallel to the mist-nets. In the site with highest fruit diversity ($n = 10$; Fig. 1, Site A—Larçã), sampling continued for another consecutive 20 days.

Network descriptors

The consistency of network structure descriptors to increasing sampling effort was evaluated with the data

collected during five days in all sites (Table 1) and with the data collected during 25 consecutive days at the site with the highest fruit diversity. We explored the effect of increasing sampling effort in six qualitative and four quantitative network descriptors. Qualitative descriptors included: Connectance, the proportion of realised links from all possible links in the network (Jordano 1987); Nestedness (NODF) (Almeida-Neto, Guimarães, Guimarães, Loyola, & Ulrich 2008), which reflects the degree of organization of interactions around a core of generalist species; Disperser richness; Seed richness; Link richness; Qualitative Modularity, reporting the existence of clusters of tightly interacting species (Olesen, Bascompte, Dupont, & Jordano 2007). Qualitative Modularity was estimated with the algorithm QuaBiMo (Dormann & Strauss 2014) using both binary (qualitative) and weighted (quantitative) matrices. Although this algorithm has been specifically developed for weighted interaction matrices, we applied the same algorithm to binary versions of the original matrices in order to evaluate the effect of the input data (binary/weighted). The quantitative descriptors included: Weighted Connectance, the ratio between linkage density (mean number of links per species) and the number of species in the network (Bersier, Banasek-Richter, & Cattin 2002); Weighted Nestedness (WNODF), as NODF but takes into account interaction frequency (Almeida-Neto & Ulrich 2011); Network specialisation index (H_2'), measuring the degree of the partner's selectivity as the departure from a theoretical non-discrimination of interactions (Blüthgen, Menzel, & Blüthgen 2006); Quantitative Modularity, as Qualitative Modularity but based on the original, weighted interaction matrices. The significance of each network descriptor was assessed by comparison with the 95% confidence interval of a set of 1000 matrix randomizations (100 for Modularity) with marginal totals equal to the observed matrix. This allows us to distinguish if a certain parameter is driven by biological/ecological characteristics of the interacting species or results simply from random interaction patterns driven mostly by species abundances. Piecewise regressions were used to estimate the sampling effort needed to achieve asymptotic values (i.e. slope of partial regression line not significantly different from zero).

Network descriptors were calculated with packages bipartite 2.05 (Dormann, Gruber, & Fründ 2008) and vegan 2.2 (Oksanen et al. 2015), while piecewise regressions were performed with package segmented 0.5 (Muggeo 2008) in R 3.0.2 (R Development Core Team 2014).

Sampling completeness

The minimum expected asymptotic richness of seed species, disperser species and links was estimated with the Chao 2, a non-parametric estimator based on the proportion of uniques (species captured on a single day) relative to duplicates (species captured on two days) (Chao 1984; Colwell & Coddington 1994). The expected richness was

Table 1. Percentage of disperser species, seed species, and links detected during five consecutive sampling days across the seven sites. Estimated values of richness were computed using the Chao 2 estimator. Estimates marked with * were computed based on the classic formula of Chao 2 because the coefficient of variation for incidence distribution >0.5. *O*—observed richness, *E* (95% C.I.)—estimated richness (95% confidence intervals), %—sampling completeness (= $O/E \times 100$).

Site	No. of days	Seeds			Dispersers			Links		
		<i>O</i>	<i>E</i> (95% C.I.)	%	<i>O</i>	<i>E</i> (95% C.I.)	%	<i>O</i>	<i>E</i> (95% C.I.)	%
A—Larçã	25	12	12 (12–16)	100	8	8* (8–21)	100	38	52 (42–93)	73
A—Larçã	5	7	7 (7–12)	100	6	6 (6–7)	100	17	30 (20–68)	57
B—Atenor	5	5	7* (5–24)	71	8	8 (8–15)	100	13	18 (14–40)	72
C—Nozelos	5	7	8 (7–19)	88	11	12 (11–21)	92	18	23 (19–44)	79
D—Santa Maria de Aguiar dam	5	7	7 (7–18)	100	11	12 (11–24)	92	25	29 (26–45)	86
E—Achete	5	10	11 (10–22)	91	11	11 (11–15)	100	33	41 (35–62)	81
F—Freixo do Meio	5	2	2* (2–3)	100	5	5 (5–6)	100	6	6 (6–14)	100
G—Bensafrim	5	4	4 (4–5)	100	6	6 (6–14)	100	10	13 (10–28)	77

computed using the program EstimateS 9.1.0 (Colwell 2013). The Chao 2 estimator was chosen for being more robust to reduced sample size (Colwell & Coddington 1994), performing considerably better than other estimators (Walther & Moore 2005).

Sampling completeness was considered as the percentage of observed richness relative to the estimated asymptotic richness. By extrapolation of the estimated rarefaction curves (Colwell et al., 2012), we estimated the additional number of sampling days needed to detect 80%, 90%, and 100% of the total estimated richness.

To evaluate whether sampling completeness is affected by fruit diversity across sites, a generalized linear model with quasibinomial error distribution (due to data underdispersion) was used.

Results

Network descriptors

During the five sampling days, different network descriptors showed different trends regardless of whether they were based on qualitative or quantitative links (Fig. 2). Both Connectance and Weighted Connectance of all networks stabilized at day 3. NODF, Weighted NODF and H_2' showed a high site-specific variability (Fig. 2).

All but two (Seed Species and Link Richness) descriptors stabilised before the end of the study at the site sampled for 25 days (see Appendix A: Table 2) but different metrics showed distinct patterns in respect to increasing sampling effort. The number of detected seed species and links was still increasing as revealed by the positive slopes of the partial regressions (seeds: $0.19 < \beta_{95\% \text{ CL}} < 0.42$; links: $0.62 < \beta_{95\% \text{ CL}} < 0.80$) (Fig. 2, see Appendix A: Table 2).

However, disperser's diversity stabilised at day 8 with no further detected species ($-0.02 < \beta_{95\% \text{ CL}} < 0.02$, Fig. 2, see Appendix A: Table 2). Both Connectance and Weighted Connectance remained relatively stable throughout the study period (Connectance = 40–50% ($-0.02 < \beta_{95\% \text{ CL}} < 0.0005$, Fig. 3A); Weighted Connectance = 21–27% ($-0.001 < \beta_{95\% \text{ CL}} < 0.0002$, Fig. 3B). NODF stabilised after day 8, (day 8–21: $-0.01 < \beta_{95\% \text{ CL}} < 1.17$; day 21–25: $-5.09 < \beta_{95\% \text{ CL}} < 2.01$; Fig. 3C, see Appendix A: Table 2), while Weighted NODF tended to stabilize at day 8 but formally reached its asymptotic value only on day 16 ($-0.51 < \beta_{95\% \text{ CL}} < 0.58$, Fig. 3D). The Qualitative and Quantitative Modularity as well as Network Specialization H_2' stabilized around day 8 (Qualitative Modularity: $-0.004 < \beta_{95\% \text{ CL}} < 0.00002$; Quantitative Modularity: $-0.002 < \beta_{95\% \text{ CL}} < 0.0008$; H_2' : $-0.004 < \beta_{95\% \text{ CL}} < 0.001$) (Fig. 3E, F and H; see Appendix A: Table 2).

Sampling completeness

After five sampling days, we detected, on average, 93% (range: 71–100%) of the estimated seed species, 97% (92–100%) of the estimated disperser species, and 79% (57–100%) of the estimated links on the seven study sites (Table 1). Fruit diversity at a given site did not significantly affect sampling completeness (GLM: $\beta = -0.06$, $t = -0.464$, $p = 0.662$; see Appendix A: Fig. 1). After the first five sampling days in Larçã, we detected only 58% of the seed species, 75% of the disperser species, and 45% of the links detected within 25 days. The detection rate improved to 67% of the seeds, 88% of the dispersers, and 71% of the links by day 8. After 25 sampling days, we estimated to have recorded all seed species ($n = 12$), all disperser species ($n = 8$), and 73% (38 out of 52) of the estimated number of links at site A

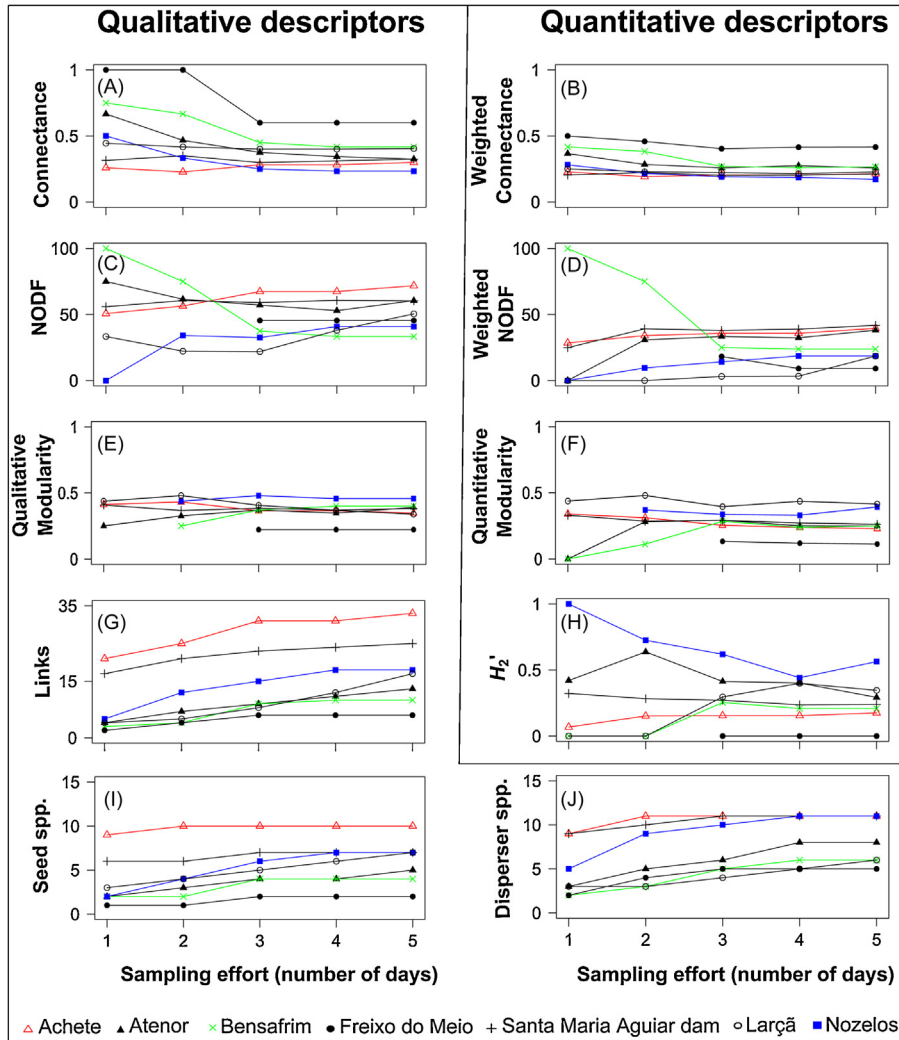


Fig. 2. Effect of sampling effort (measured as the number of days) on the Connectance (A), Weighted Connectance (B), Nestedness (NODF (C) and Weighted NODF (D)), Modularity (Qualitative (E) and Quantitative (F)), Specialization (H_2' (H)), richness of seeds (I), dispersers (J), and links (G) detected in seven seed-dispersal networks sampled simultaneously during five consecutive days throughout Portugal. Missing values correspond to days/sites where the network was too small to allow the calculation of those network descriptors. For the definition of network descriptors see the Materials and methods section.

(Table 1). The observed richness in this site after 25 days was higher than that estimated using Chao 2 at day 5 for all levels considered (Seeds, Dispersers, and Links; Table 1). While the final number of detected seed species and links was

within the confidence interval estimated after five sampling days, the same was not true for disperser species (Table 1). A total of 50 days were considered necessary to detect 90% of the links and 72 days would be needed to detect all links

Table 2. Number of additional sampling days needed to detect 80%, 90%, and 100% of the diversity estimated by the Chao 2 estimator.

Network descriptor	Sampling period	Target sampling completeness		
		80%	90%	100%
Seed species	5 Days (mean of 7 sites)	0	0	2
	25 Days (1 site)	0	0	0
Disperser species	5 Days (mean of 7 sites)	0	0	1
	25 Days (1 site)	0	0	0
Link diversity	5 Days (mean of 7 sites)	1	2	5
	25 Days (1 site)	8	25	47

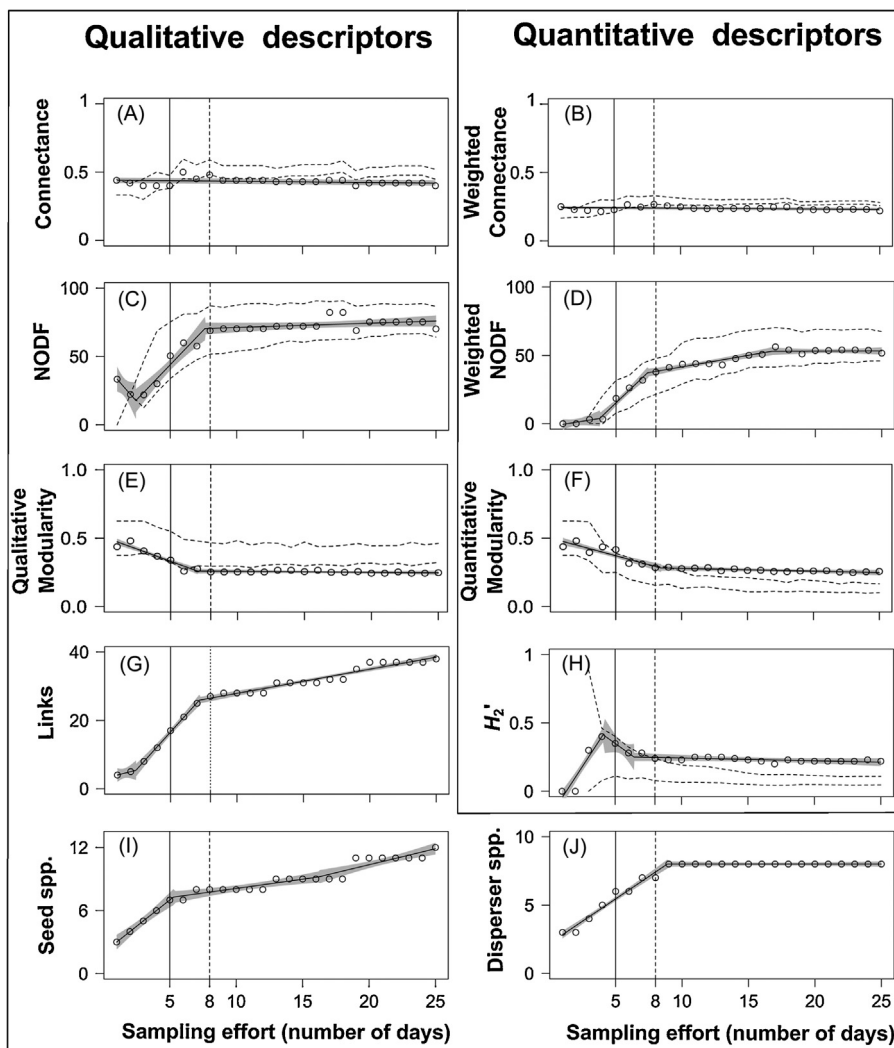


Fig. 3. Effect of the sampling effort (measured as the number of days) on the Connectance (A), Weighted Connectance (B), Nestedness (NODF (C) and Weighted NODF (D)), Modularity (Qualitative and Quantitative), Specialization (H_2' (H)), richness of seeds (I), dispersers (J), and links (G) detected on the seed-dispersal network at Larçã (Fig. 1). The dashed lines indicate the 95% confidence interval of 1000 runs (100 for Modularity) of a null model based exclusively on species abundances. The fitted lines represent the theoretical values estimated by piecewise linear regressions, except for Connectance, which was fitted with linear regression, with the grey area indicating the 95% confidence interval. The dashed vertical line at day 8 indicates the minimum adequate sampling effort suggested for this site, while the full line at day 5 indicates the duration of the sampling period on the other sites. For the definition of network descriptors see the Materials and methods section.

(Table 2). An animation of accumulation of species and links at the most intensively sampled site (Larçã) is available as supporting information (see Appendix B).

Discussion

In this study we show that while some network descriptors can be accurately estimated with lower sampling effort (number of Disperser Species, Connectance, and Weighted Connectance), most network metrics only started to stabilize after eight mist-netting days (number of Links and Seed Species, NODF, Weighted NODF, Modularity (both Quantitative and Qualitative), and H_2').

Network properties

While five sampling days suitably captured the biological diversity across the seven study sites, it was not sufficient to deliver a completely sound representation of the community structure measured by all but two network descriptors (Connectance and Weighted Connectance). Contrary to our expectation we could not detect a trend for quantitative indices performing better to lower sampling efforts. In fact, Weighted NODF took slightly longer to stabilize than its qualitative counterpart.

Both Connectance and Weighted Connectance remained relatively stable throughout the study period and therefore were accurately predicted even with low/moderate sampling

effort. However, only after day 7/8 the network is significantly less connected than we would expect according to a random association among species (Fig. 3). Still, given the small difference between observed and predicted values, such significance may be explained by the intrinsic mathematical behaviour of the used null model, which tends to generate matrices more connected than that observed (Dormann, Fründ, Blüthgen, & Gruber 2009). Additionally, these descriptors showed similar trends in all of our study sites, suggesting that such robustness to sample size is not network-specific but also that these descriptors are of limited informative value to compare seed dispersal networks. Previous studies on pollination networks found a tendency for an initial increase in Connectance towards an asymptote that might be reached at relatively low (Nielsen and Bascompte 2007) or higher (Rivera-Hutinel et al. 2012) levels of sampling completeness. Our results suggest that, despite both Connectance and Weighted Connectance of seed dispersal networks may be stable even with a low sampling effort, they are of limited use if we want to compare at least similar seed dispersal networks. Moreover, despite its wide use in ecological network studies and apparent robustness to sample size, Connectance seems to have no relation between its value and the network conservation status (Heleno, Devoto, & Pocock 2012). Further work is required to confirm the extent of this generalization.

The two measures of Nestedness (NODF and Weighted NODF) were more sensitive to sampling effort than Connectance. Asymptotic values of NODF were achieved earlier (day 8) than for Weighted NODF (day 16), but overall our results coincide with those of Rivera-Hutinel et al. (2012) who found NODF to be relatively stable if at least 30% of the network had been sampled. However, the observed network was not significantly nested (Fig. 3) even after 25 sampling days. This pattern was not the result of sampling limitations, as both NODF and Weighted NODF values were relatively stable, with as low as half of the full sampling effort and thus not to be expected to change with a higher sample size, but likely driven by species abundances. Taken altogether our results question the ecological relevance of Nestedness indices (see also James, Pitchford, & Plank 2012). Modularity has been increasingly used in network studies as it is strongly related to Nestedness (Fortuna et al. 2010). We found both Qualitative and Quantitative Modularity to stabilize after 7 and 8 sampling days, respectively, and to encompass real ecological information as it cannot be explained by null models based solely on species abundances. However, it is important to note that while Quantitative Modularity revealed a network more modular than expected by chance, the analyses based on binary matrices revealed the opposite pattern, reinforcing the value of weighted interaction networks in the capacity of this algorithm to detect modules (Dormann & Strauss 2014). Network Specialization H_2' stabilized around day 6 to 8 and was higher (i.e. more specialized) than expected under random species associations. These results corroborate the stability of this metric to moderate sampling effort (Blüthgen et al. 2006;

Schleuning, Böhning-Gaese, Dehling, & Burns 2014). These three descriptors (NODF, Weighted NODF, and H_2') showed high site-specific trends, and while they seem to converge to a narrower range during the five sampling days (Fig. 2), this sampling effort is insufficient to provide a clear pattern and highlight the meaningfulness of these descriptors to compare similar seed dispersal networks. The breakpoints identified by the piecewise regression (i.e. those that maximise the variability explained by the regression lines) should also be regarded as conservative given that quasi-asymptotic values are often reached a few days before the estimated breakpoint (Fig. 3).

From Fig. 3, and considering the sampling effort and data quality trade-off, a minimum of eight sampling days are required to adequately sample our seed dispersal network (detection of 67% for seed species, 88% for disperser species, and 71% for links). After this day, all descriptors either reach an asymptote or have a highly reduced slope and show a consistent pattern of significance with respect to null models. Our results are restricted to seven sites in Portugal and to a particular sampling method (identification of seeds on the droppings of mist-netted birds), so further studies testing different sampling strategies and different interaction types, particularly in hyper-diverse ecosystems, are important to test the generality of these results.

An assessment of each descriptor's quality is critical to derive meaningful network descriptors that reveal true ecological attributes of biological communities and not mathematical artefacts resulting from poor sampling or exclusively from species abundances. Since rare species and links require a higher sampling effort to be recorded, the typically low evenness of species abundances alone can drive some observed properties of the structure of some networks (Vázquez & Aizen 2004). Network descriptors can be more or less sensitive to these sampling artefacts (Rivera-Hutinel et al. 2012; Blüthgen 2010; Blüthgen et al. 2008), thus its values may either reflect sample effects or real nature processes. The use of appropriate null models is a way to disentangle if the processes structuring the observed networks are mostly mathematical or biological (e.g. Araujo, de Almeida, Cardoso, & Corso 2010; Vázquez, Chacoff, & Cagnolo 2009; Vázquez et al. 2007; Vázquez & Aizen 2004). These analyses surpass the scope of this study but are crucial if one wants descriptors that mirror true biological patterns.

The efficiency of mist nets to capture birds depends on several factors, such as habitat structure, weather or bird behaviour and size (Estades, Escobar, Tomasevic, Vukasovic, & Páez 2006; Pagen, Thompson, Frank, & Burhans 2002). Nevertheless, mist-netting is likely the best method to construct quantitative seed dispersal networks as it is largely free of observer bias and allows great taxonomic resolution (species level with few exceptions), allows the detection of inconspicuous dispersers and interactions, makes the study of individual food choice possible

through individual marking (e.g. rings), and allows the evaluation of the effect of seed ingestion on its viability (Costa et al. 2014; Traveset, Robertson, & Rodríguez-Pérez 2007; Heleno, Ross, Everard, Memmott, & Ramos 2011).

Sampling completeness

Five sampling-days detected consistently >71% of the dispersed seed species and >92% of the disperser species across the seven sites. We estimate that during the same period the majority of the links (79%) were also detected, but the sampling completeness was lower (down to 57%) on the most diverse site (Larçã).

As expected, the thorough detection of links required higher sampling intensity than the detection of the species. This is partly a mathematical consequence, once that new links require new species but species do not necessarily report new links, and partly because some links are likely very rare even if they involve common species. A similar effect was already described for pollination networks (Chacoff et al. 2012). However, due to the generally lower size of seed dispersal networks, with lower diversity of dispersers, i.e. animals/plants ratio lower in seed dispersal than in pollination networks, one may expect that seed dispersal studies require less effort to attain a similar sampling completeness (Blüthgen, Menzel, Hovestadt, Fiala, & Blüthgen 2007; Guimarães et al. 2007).

In Larçã, 25 sampling days resulted in the detection of a higher richness of seeds, dispersers, and links than that estimated by Chao 2 with only five sampling days. However, since this estimator computes the minimum expected richness (Chao 1984; Colwell & Coddington 1994) and the seeds and link richness detected after 25 days lay within the confidence interval of that estimated after five days, we consider that this estimator performed relatively well for seeds and links. For dispersers, however, the richness detected after 25 days lay above the confidence interval estimated after five sampling days. This, at least apparent, underestimation of Chao 2 might reflect a poor performance of the estimator for species richness due to low sample size (five capture occasions) (Walther & Moore 2005). This apparent underestimation can be largely explained by an increase in the availability of ripe fruits (advancing fruit phenology) and new potential dispersers (bird migration) entering an open community.

We estimated that all species and 73% of the links between plants and avian dispersers were detected in Larçã with 25 sampling days. Interestingly, some species pairs that are present in this site, but apparently not interacting, are known to interact elsewhere: e.g. *Sylvia atricapilla* with *Daphne gnidium*, *Sylvia melanocephala* with *Smilax aspera* (Olesen et al. 2011). We believe that the independent study of species and link distributions (e.g. interaction distribution modelling)

holds a large potential to understand community assembly rules in ecology.

Concluding remarks

Our results suggest that while some descriptors can be accurately estimated with lower sampling effort, at least eight sampling days were needed to accurately describe the structure of our seed dispersal network based on the droppings of mist-netted birds. More studies on the effect of sampling intensity on network descriptors are needed to allow a generalization of the conclusions on the effort required to get a realistic overview of the seed dispersal networks' structure and to critically assess sampling limitations in previous studies. The reproducibility of the results is a central tenet of experimental science, and there is no reason why this should not be applied to ecological networks studies, for which the identification of what is a sound sampling effort is paramount.

Acknowledgements

We are grateful to the ringers that contributed to the dataset (Sérgio Timóteo, Pedro M. Araújo, António Rosa, Paulo Tenreiro, and Marcial Felgueiras) and to the Coimbra Botanical Garden, particularly to Arménio Matos for his help with seed identification. We also thank to two anonymous reviewers and the editor Klaus Hövemeyer whose comments greatly improved the manuscript. JMC, LPS, and RH were supported by Fundação para Ciência e Tecnologia (SFRH/BD/96292/2013, SFRH/BD/77746/2011, and IF/00441/2013, respectively) and RH by the Marie Curie action (FP7-PEOPLE-2012-CIG-321794).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.baae.2015.09.008>.

References

- Almeida-Neto, M., Guimarães, P., Guimarães, P. R., Jr., Loyola, R. D., & Ulrich, W. (2008). A consistent metric for nestedness analysis in ecological systems: Reconciling concept and measurement. *Oikos*, *117*, 1227–1239.
- Almeida-Neto, M., & Ulrich, W. (2011). A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software*, *26*, 173–178.
- Araújo, A. I. L., de Almeida, A. M., Cardoso, M. Z., & Corso, G. (2010). Abundance and nestedness in interaction networks. *Ecological Complexity*, *7*, 494–499.

- Banasek-Richter, C., Cattin, M.-F., & Bersier, L.-F. (2004). Sampling effects and the robustness of quantitative and qualitative food-web descriptors. *Journal of Theoretical Biology*, *226*, 23–32.
- Bartomeus, I. (2013). Understanding linkage rules in plant–pollinator networks by using hierarchical models that incorporate pollinator detectability and plant traits. *PLoS ONE*, *8*, e69200.
- Bersier, L.-F., Banasek-Richter, C., & Cattin, M.-F. (2002). Quantitative descriptors of food-web matrices. *Ecology*, *83*, 2394–2407.
- Blüthgen, N. (2010). Why network analysis is often disconnected from community ecology: A critique and an ecologist's guide. *Basic and Applied Ecology*, *11*, 185–195.
- Blüthgen, N., Fründ, J., Vázquez, D. P., & Menzel, F. (2008). What do interaction network metrics tell us about specialization and biological traits? *Ecology*, *89*, 3387–3399.
- Blüthgen, N., Menzel, F., & Blüthgen, N. (2006). Measuring specialization in species interaction networks. *BCM Ecology*, *6*, 9.
- Blüthgen, N., Menzel, F., Hovestadt, T., Fiala, B., & Blüthgen, N. (2007). Specialization, constraints, and conflicting interests in mutualistic networks. *Current Biology*, *17*, 341–346.
- Carlo, T. A., & Yang, S. (2011). Network models of frugivory and seed dispersal: Challenges and opportunities. *Acta Oecologica*, *37*, 619–624.
- Chacoff, N. P., Vázquez, D. P., Lomáscolo, S. B., Stevani, E. L., Dorado, J., & Padrón, B. (2012). Evaluating sampling completeness in a desert plant–pollinator network. *Journal of Animal Ecology*, *81*, 190–200.
- Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, *11*, 265–270.
- Colwell, R. K. (2013). *EstimateS: Statistical estimation of species richness and shared species from samples. Version 9*. User's Guide and application published at (<http://purl.oclc.org/estimates>)
- Colwell, R. K., Chao, A., Gotelli, N. J., Lin, S.-Y., Mao, C. X., Chazdon, R. L., et al. (2012). Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology*, *5*, 3–21.
- Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London, Series B: Biological Science*, *345*, 101–118.
- Costa, J. M., Ramos, J. A., da Silva, L. P., Timoteo, S., Araújo, P. M., Felgueiras, M. S., et al. (2014). Endozoochory largely outweighs epizoochory in migrating passerines. *Journal of Avian Biology*, *45*, 59–64.
- Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray.
- Donatti, C. I., Guimarães, P. R., Galetti, M., Pizo, M. A., Marquitti, F. M. D., & Dirzo, R. (2011). Analysis of a hyper-diverse seed dispersal network: Modularity and underlying mechanisms. *Ecology Letters*, *14*, 773–781.
- Dormann, C. F., Fründ, J., Blüthgen, N., & Gruber, B. (2009). Indices graphs and null models: Analyzing bipartite ecological networks. *The Open Ecology Journal*, *2*, 7–24.
- Dormann, C. F., Gruber, B., & Fründ, J. (2008). Introducing the bipartite package: Analysing ecological networks. *R News*.
- Dormann, C. F., & Strauss, R. (2014). A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, *5*, 90–98.
- Estades, C. F., Escobar, M. A. H., Tomasevic, J. A., Vukasovic, M. A., & Páez, M. (2006). Mist-nets versus point counts in the estimation of forest bird abundances in south-central Chile. *Ornitologia Neotropical*, *17*, 203–212.
- Fortuna, M. A., Stouffer, D. B., Olesen, J. M., Jordano, P., Mouillot, D., Krasnov, B. R., et al. (2010). Nestedness versus modularity in ecological networks: Two sides of the same coin? *Journal of Animal Ecology*, *79*, 811–817.
- González-Varo, J. P., Arroyo, J. M., & Jordano, P. (2014). Who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies. *Methods in Ecology & Evolution*, *5*, 806–814.
- Guimarães, P. R., Machado, G., de Aguiar, M. A. M., Jordano, P., Bascompte, J., Pinheiro, A., et al. (2007). Build-up mechanisms determining the topology of mutualistic networks. *Journal of Theoretical Biology*, *249*, 181–189.
- Hegland, S. J., Dunne, J., Nielsen, A., & Memmott, J. (2010). How to monitor ecological communities cost-efficiently: The example of plant–pollinator networks. *Biological Conservation*, *143*, 2092–2101.
- Heleno, R., Devoto, M., & Pocock, M. (2012). Connectance of species interaction networks and conservation value: Is it any good to be well connected? *Ecological Indicators*, *14*, 7–10.
- Heleno, R., Garcia, C., Jordano, P., Traveset, A., Gómez, J. M., Blüthgen, N., et al. (2014). Ecological networks: Delving into the architecture of biodiversity. *Biology Letters*, *10* (2013100).
- Heleno, R., Olesen, J. M., Nogales, M., Vargas, P., & Traveset, A. (2013). Seed-dispersal networks in the Galápagos and the consequences of plant invasions. *Proceedings of the Royal Society B—Biological Sciences*, *280* (20122112).
- Heleno, R. H., Lacerda, I., Ramos, J. A., & Memmott, J. (2010). Evaluation of restoration effectiveness: Community response to the removal of alien plants. *Ecological Applications*, *20*, 1191–1203.
- Heleno, R. H., Ross, G., Everard, A., Memmott, J., & Ramos, J. A. (2011). The role of avian 'seed predators' as seed dispersers. *Ibis*, *153*, 199–203.
- Herrera, C. M. (1984). A study of avian frugivores, bird-dispersed plants, and their interaction in Mediterranean scrublands. *Ecological Monographs*, *54*, 1–23.
- James, A., Pitchford, J. W., & Plank, M. J. (2012). Disentangling nestedness from models of ecological complexity. *Nature*, *487*, 227–230.
- Jordano, P. (1987). Patterns of mutualistic interactions in pollination and seed dispersal: Connectance, dependence asymmetries, and coevolution. *The American Naturalist*, *129*, 657–677.
- Mello, M., Marquitti, F., Guimarães, P. R., Kalko, E., Jordano, P., & Aguiar, M. D. (2011). The missing part of seed dispersal networks: Structure and robustness of bat–fruit interactions. *PLoS ONE*, *6*, e17395.
- Mugge, V. M. R. (2008). Segmented: An R package to fit regression models with broken-line relationships. *R News*, *8*, 20–25.

- Nielsen, A., & Bascompte, J. (2007). Ecological networks, nestedness and sampling effort. *Journal of Ecology*, *95*, 1134–1141.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2015). *Vegan: Community ecology package*. In *R package version 2*.
- Olesen, J. M., Bascompte, J., Dupont, Y. L., Elberling, H., Rasmussen, C., & Jordano, P. (2011). Missing and forbidden links in mutualistic networks. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *278*, 725–732.
- Olesen, J. M., Bascompte, J., Dupont, Y. L., & Jordano, P. (2007). The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 19891–19896.
- Olesen, J. M., Dupont, Y. L., Hagen, M., Rasmussen, C., & Trøjelsgaard, K. (2012). Structure and dynamics of pollination networks: The past, present, and future. In S. Patiny (Ed.), *Evolution of plant–pollinator relationships* (pp. 374–391). Belgium: Cambridge University Press.
- Olito, C., & Fox, J. W. (2015). Species traits and abundances predict metrics of plant–pollinator network structure, but not pairwise interactions. *Oikos*, *124*, 428–436.
- Pagen, R. W., Thompson, I., Frank, R., & Burhans, D. E. (2002). A comparison of point-count and mist-net detections of songbirds by habitat and time-of-season. *Journal of Field Ornithology*, *73*, 53–59.
- R Development Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rivera-Hutinel, A., Bustamante, R. O., Marín, V. H., & Medel, R. (2012). Effects of sampling completeness on the structure of plant–pollinator networks. *Ecology*, *93*, 1593–1603.
- Schleuning, M., Böhning-Gaese, K., Dehling, D. M., & Burns, K. C. (2014). At a loss for birds: Insularity increases asymmetry in seed-dispersal networks. *Global Ecology and Biogeography*, *23*, 385–394.
- Traveset, A., Heleno, R., Chamorro, S., Vargas, P., McMullen, C. K., Castro-Urgal, R., et al. (2013). Invaders of pollination networks in the Galápagos Islands: Emergence of novel communities. *Proceedings of the Royal Society B: Biological Sciences*, *280*.
- Traveset, A., Heleno, R. H., & Nogales, M. (2014). The ecology of seed dispersal. In R. S. Gallagher (Ed.), *Seeds: The ecology of regeneration in plant communities* (pp. 62–93). Oxfordshire, UK: CABI.
- Traveset, A., Robertson, A. W., & Rodríguez-Pérez, J. (2007). A review on the role of endozoochory on seed germination. In A. J. Dennis, E. W. Schupp, R. A. Green, & D. A. Westcott (Eds.), *Seed dispersal: Theory and its application in a changing world* (pp. 78–103). Wallingford, UK: CABI Publishing.
- Tylianakis, J. M., Tscharntke, T., & Lewis, O. T. (2007). Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature*, *445*, 202–205.
- Vázquez, D. P., & Aizen, M. A. (2004). Asymmetric specialization: A pervasive feature of plant–pollinator interactions. *Ecology*, *85*, 1251–1257.
- Vázquez, D. P., Chacoff, N. P., & Cagnolo, L. (2009). Evaluating multiple determinants of the structure of plant–animal mutualistic networks. *Ecology*, *90*, 2039–2046.
- Vázquez, D. P., Melián, C. J., Williams, N. M., Blüthgen, N., Krasnov, B. R., & Poulin, R. (2007). Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, *116*, 1120–1127.
- Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, *28*, 815–829.

Available online at www.sciencedirect.com

ScienceDirect