Ann. Limnol. - Int. J. Lim. 46 (2010) 101–110
© EDP Sciences, 2010
DOI: 10.1051/limn/2010008

Available online at:
www.limnology-journal.org

Assessment of richness estimation methods on macroinvertebrate communities of mountain ponds in Castilla y León (Spain)

Carlos Martínez-Sanz*, Francisco García-Criado, Camino Fernández Aláez and Margarita Fernández Aláez

Area of Ecology, Faculty of Biology and Environmental Science, University of León, Campus de Vegazana, s/n. C.P. 24071, León, Spain

Received 27 August 2009; Accepted 25 January 2010

Abstract – Complete inventories of the fauna at a given place, for a specific community or geographical area are often exceedingly hard to get. In recent years a number of estimation techniques have emerged that can be used to extrapolate from these samples to the true number of species in an area. These estimation models are based on different mathematical approaches and can be classified as either species accumulation curves or nonparametric estimators (Brose *et al.*, 2003, *Ecology*, *84*, 2364–2377). In this paper, we have tested the performance of some of the richness estimators on nineteen mountain ponds in Castilla y León (Spain) in order to provide guidance on their potential use in future researches. We collected benthic macroinvertebrate of these ponds from the littoral zone with a pond net by kicking and sweeping. Ten-second samples were collected in each pond up to a total time of 3 to 5 minutes per pond, depending on the pond size. In addition, two of the ponds, were sampled in 2004, 2006 and 2007 providing a three-year time series. The results of this study showed that Jackknife 2 was the best of the evaluated methods based on all chosen criteria and also performed well across all studied ponds. Jackknife 1, Chao 1 and Chao 2 also presented good results and they were inferior to Jackknife 2 mainly because of the requirement for larger sub-sample sizes.

Key words: Richness / estimators / macroinvertebrate / mountain ponds / accumulation curves / non-parametric estimators

Introduction

Wetlands and small water bodies such as many mountain ponds have frequently been degraded or destroyed by human alteration (Chapin *et al.*, 2000; Schindler *et al.*, 2001). Shallow aquatic habitats are particularly vulnerable to impacts from anthropogenic inputs (Karakoç *et al.*, 2003; Schippers *et al.*, 2006; Søndergaard and Jeppesen, 2007). This poses a serious threat to the biodiversity of systems which, in the case of ponds, are considered to support a high richness of organisms, particularly macroinvertebrates (Oertli *et al.*, 2002; Williams *et al.*, 2004), both on a local and regional basis (Toro *et al.*, 2006). Nobody questions the need to preserve such systems and their diversity although only recently research programs have focused on them (Biggs *et al.*, 2005; Oertli *et al.*, 2005; Bilton *et al.*, 2009).

Taxon richness, especially species richness, is the simplest and the most intuitive concept for characterizing biodiversity (Gaston, 1996; Chao *et al.*, 2005). It is currently the most used parameter, not only in biodiversity studies (Magurran, 1988; Colwell and Coddington, 1994; Flather, 1996; Keddy and Drummond, 1996) but also in community and trophic ecology (Martinez *et al.*, 1999; Williams and Martinez, 2000) for conservation (Prendergast *et al.*, 1993; Pressey *et al.*, 1993; Conroy and Noon, 1996; Kerr, 1997) and macroecology (Gaston, 2000; Whittaker *et al.*, 2001). Therefore, richness measures are becoming valuable tools for a number of scientific applications as well as a means to assess environmental degradation. The assessment of richness in a given area requires a count of observed species or taxa (Melo and Froehlich, 2001) but, unfortunately, in biological and ecological sciences, the compilation of complete species census and inventories is costly or even impossible (Foggo *et al.*, 2003; Hortal *et al.*, 2006). In addition, biodiversity data suffer from heterogeneity in sampling strategies and/or

*Corresponding author: cmars@unileon.es

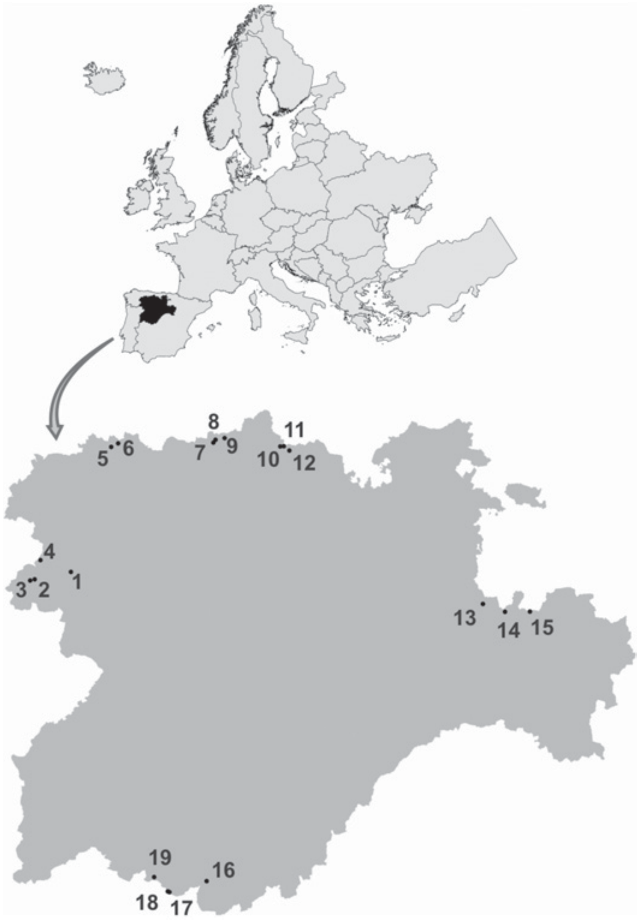

Fig. 1. Study area and mountain ponds: 1. Truchillas (León); 2. Roya (Zamora); 3. Camposagrado (Zamora); 4. Baña (León); 5. Grande de Babia (León); 6. Verdes (León); 7. Robledo (León); 8. Isoba (León); 9. Tronisco (León); 10. Hoyos de Vargas (León); 11. Fuentescarrionas (Palencia); 12. Curavacas (Palencia); 13. Pardillas (Soria); 14. Helada (Soria); 15. Cebollera (Soria); 16. Cervunal (Ávila); 17. Caballeros (Ávila); 18. Cuadrada (Ávila); 19. Trampal (Ávila).

sample size: the larger the sampling effort, the larger will be the number of observed taxon (Walther *et al.*, 1995).

To circumvent this problem, it is possible to use estimation methods (Burnham and Overton, 1979; Smith and van Belle, 1984; Gotelli and Colwell, 2001). Taxon richness estimation techniques have been developed, mainly in the past 20 years, and are emerging as a powerful tool for providing a cost-effective method of assessing richness in an area without the need for full inventories. These estimation methods have been the subject of a growing body of literature (Colwell and Coddington, 1994; Hammond, 1994; Gotelli and Colwell, 2001), with validation of the methods for a variety of taxa and habitats remaining a priority (Colwell and Coddington, 1994; Melo and Froehlich, 2001; Brose *et al.*, 2003). Richness estimation models are based on different mathematical approaches and can be classified as either species accumulation curves or nonparametric estimators (Gotelli and Colwell, 2001; Brose *et al.*, 2003). Taxon accumulation curves extrapolate species richness *vs.* sample size data to an asymptote of total richness (Soberón and Llorente, 1993; Colwell and Coddington, 1994). The most often used accumulation curves are the exponential equation (Holdridge *et al.*, 1971) and the Michaelis-Menten model (Michaelis and Menten, 1913). Nonparametric estimators are sampling theoretic extrapolation methods that only require the number of samples in which each taxon is found rather than any parametric information about their abundance.

The performance of several of these estimators have been carried out and compared in a number of studies (Chazdon *et al.*, 1998; Keating *et al.*, 1998; Peterson and Slade, 1998; Walther and Morand, 1998; Chiarucci *et al.*, 2001; Walther and Martin, 2001; Brose, 2002; Longino *et al.*, 2002; Borges and Brown, 2003; Chiarucci *et al.*, 2003; Melo *et al.*, 2003; Brose and Martinez, 2004; see review in Walther and Moore, 2005). However, in freshwaters their use has mostly been restricted to stream invertebrates (Melo and Froehlich, 2001) and lake zooplankton (Dumont and Segers, 1996; Arnott *et al.*, 1998). Few researches of this kind have been undertaken in ponds (Foggo *et al.*, 2003), but none in mountain ponds. Current and future research conducted in these ecosystems would be aided by knowing which of the estimators can preferably be chosen for a reliable assessment of true richness. It would be particularly useful to know whether estimators based on single samples can provide such information. In this study we have addressed this issue for macroinvertebrates in mountain ponds from a large Spanish region: Castilla y León. We have tested the performance of some of the richness estimators on nineteen mountain ponds in Castilla y León (Spain) in order to provide guidance on their potential use in future researches.

Materials and methods

Study area

Castilla y León (north Spain) is a vast region (94 223 km^2) consisting of a large central plateau (700–1000 m a.s.l.) surrounded by a set of mountain ranges (altitudes up to 2600 m a.s.l.). Nineteen mountain ponds were selected for this study (Fig. 1). The pond selection was intended to record a wide gradient of environmental conditions of altitude (from 1400 to 2200 m a.s.l.), area (between 0.5 and 8 ha), depth (maximum depths between 0.3 and 14 m), and water permanece (temporary and permanent systems). Furthermore, most of the shallow, usually temporary, systems supported dense stands of vegetation (either submerged, emergent or both) whereas littoral zone in deep ponds was mostly sandy to stony and was poorly vegetated except for *Isoetes*, a species present in most of the ponds.

Sampling and sorting of macroinvertebrates

All ponds were sampled once in June/July 2006 or 2007. Benthic macroinvertebrates were collected from

Table 1. Summary of the eleven evaluated estimators. NP: non-parametric estimators, ESAC: estimators based on the extrapolation of species accumulation curves.

Abbreviation	Estimator	Type	References
ACE	Abundance-based coverage estimator of species richness	NP	Chao and Lee (1992); Chao *et al.* (2000); Chazdon *et al.* (1998) in Colwell (2004)
ICE	Incidence-based coverage estimator of species richness	NP	Lee and Chao (1994); Chao *et al.* (2000); Chazdon *et al.* (1998) in Colwell (2004)
Chao 1	Abundance-based estimator of species richness	NP	Chao (1984) in Colwell (2004)
Chao 2	Incidence-based estimator of species richness	NP	Chao (1984, 1987); Colwell (2004)
Jackknife 1	First-order Jackknife richness estimator	NP	Burnham and Overton (1979); Heltshe and Forrester (1983) in Colwell (2004)
Jackknife 2	Second-order Jackknife richness estimator	NP	Smith and van Belle (1984) in Colwell (2004)
Bootstrap	Bootstrap richness estimator	NP	Smith and van Belle (1984) in Colwell (2004)
MMRuns	Transformation of Michaelis-Menten hyperbole by Raaijmakers. Estimate curves averaged over randomizations (runs)	ESAC	Raaijmakers (1987); Colwell (1997)
MMMean	Transformation of Michaelis-Menten hyperbole by Raaijmakers. Estimate curve computed once for mean species	ESAC	Raaijmakers (1987); Colwell (1997)
Clench	Estimation of Michaelis-Menten function asymptote	ESAC	Clench (1979) in Soberón and Llorente (1993)
Exp Neg	Estimation of negative exponential function asymptote	ESAC	Miller and Wiegert (1989) in Soberón and Llorente (1993)

the littoral zone with a pond net (FBA standard, mesh size 500 μm) by kicking and sweeping. Ten-second samples were collected in each pond up to a total time of 3 (18 samples), 4 or 5 (30 samples) minutes per pond depending on the size (< 1 ha, 1–5 ha or > 5 ha, respectively). Total sampling time was proportionally distributed among the main habitats according to their surface in the pond. Macroinvertebrates were separated from the plant material and counted under a binocular microscope (10×). The specimens were identified to genus (often the lowest attainable taxonomic level) excepting Diptera (to subfamily) and Oligochaeta (class).

Selection and calculation of richness estimations methods

The performances of 11 different richness estimators (Table 1) were compared. Nine of them, ACE, ICE, Chao 1, Chao 2, Jackknife 1, Jackknife 2, Bootstrap, MMRuns and MMMeans were calculated with the software EstimateS version 7.0 (Colwell, 1997). They have all been widely used and studied (Chazdon *et al.*, 1998; Brose *et al.*, 2003; Chiarucci *et al.*, 2003). Chao 1 and Jackknife 1 are designed to estimate richness from single samples while the rest require several samples. In addition, we evaluated the performance of two asymptotic accumulation functions, Clench and Negative Exponential functions. In these functions, richness is calculated as the asymptote value of the function fitted to the smoothed taxon accumulation curve provided by EstimateS version 7.0 (100 randomizations; Colwell, 2004). This ideal curve represents an unbiased description of the sampling process, where possible effects due to the order by which the samples have been taken or listed are removed by randomizing their order of entrance in the curve. We used the software Statistica 6.0 to fit each function to the data and calculate the asymptote value from the obtained parameters (Soberón and Llorente, 1993; Hortal *et al.*, 2004). These accumulation functions are able to predict estimate richness when they are close to the asymptote. They are all standard, widely used methods that have been previously tested and discussed by a number of authors under different circumstances (*e.g.* Palmer, 1990; Colwell and Coddington, 1994; Coddington *et al.*, 1996; Condit *et al.*, 1996; Carlton and Robison, 1998; Chazdon *et al.*, 1998; Walther and Morand, 1998; Gotelli and Colwell, 2001; Walther and Martin, 2001; Petersen and Meier, 2003).

Evaluation of estimator performance

The concepts of bias, precision and accuracy (see a review of Bruno and Joslin, 2005) are frequently used to assess the performance of richness estimation methods (Brose *et al.*, 2003; Foggo *et al.*, 2003; Walther and Moore, 2005). These approaches include measures of bias and accuracy of the estimated richness in relation to the true richness using an *a priori* chosen sub-sample size. However, the estimated richness is strongly dependent on sample size (Colwell and Coddington, 1994; Melo and Froehlich, 2001; Petersen and Meier, 2003) and different sub-samples sizes will produce different bias and accuracy values (Hellmann and Fowler, 1999). In addition, these approaches require data for maximum species numbers for their calculation, so they cannot be used here. Instead, we used some criteria we argue are more practical and functional (see also in Melo and Froehlich, 2001). Mere comparisons of the differences between estimated and observed richness are quite useful: estimators whose final estimation values do not even reach the observed richness (richness measured by sampling) cannot be considered good estimators because they obviously underestimate true richness. The behaviour of the curve shape in each pond (erratic or non-erratic) and the similarity in the curve shapes across ponds (measured as a simple scale: 1, low; 2, intermediate; 3, high) give information about the reliability and constancy of the estimators. The minimum number of samples required to attain the observed richness and the constancy of this number, measured as standard deviation (SD), are an indication of how great the sampling effort (number of samples) must be to obtain a reliable estimation. Other features being equal, an estimator requiring low sampling effort should be preferred.

No doubt, the best way to assess the performance of the estimators would be comparing the estimations with true richness values. Unfortunately, such values are difficult to obtain and are not available for any of the ponds in the study area. We have tried to partially solve this shortcoming by using information from additional samplings when these were available. Two of the ponds, Grande de Babia and Helada, were sampled in summer 2004, 2006 and 2007 (whether as twenty-four 10-second samples or as a single 4-minute one, depending on the year). Pooling data from three visits to each pond provided a more comprehensive (although probably not complete yet) taxon list. We have taken the global richness over these three years as the best available approximation to true richness of these ponds. We refer to this value as "*true richness*" in the text. This made it possible to compare estimated richness with this assumed "*true richness*". This comparison only has been used here as an additional criterion.

Table 2. Summary of observed richness (*S*), number of ten-second samples (*n*) and total number of individuals of the ponds included in the study.

Ponds	*S*	*n*	Total individuals
Baña	27	24	1418
Caballeros	26	24	3354
Camposagrado	38	18	2083
Cebollera	25	19	3110
Cuadrada	23	18	2318
Curavacas	21	24	1549
Cervunal	22	18	3480
Fuentes Carrionas	18	24	611
Grande de Babia	12	24	1589
Helada	26	24	3823
Hoyos de Vargas	22	18	2888
Isoba	24	24	694
Verdes	28	18	31 719
Pardillas	30	18	2565
Robledo	18	24	664
Roya	42	24	1101
Trampal	30	24	4546
Tronisco	14	18	1171
Truchillas	26	30	3130

Results

The nine ponds differed in macroinvertebrate abundance, observed richness and number of samples collected (Table 2). Despite the differences in the sampled assemblages, the performance of the estimators to all criteria was similar through all ponds (*e.g.* Figs. 2A–2F).

MMMeans, MMRuns and bootstrap were the estimators provided by the EstimateS software requiring highest number of samples to attain observed richness, with average values of 70%, 55%, and 66%, respectively, of the total number of samples collected. Moreover, these average values were highly variables in MMMeans and MMRuns, showing standard deviations of 6.84 and 7.43, respectively (Table 3). Bootstrap, however, showed consistent patterns, absence of erratic behaviours and high similarity curves shapes across ponds (see Figs. 2A–2K). In contrast, the dissimilarity of curve shapes and the erratic behaviours of MMRuns were evident (see Figs. 2G–2K).

Jackknife 2 performed satisfactorily in all the criteria. It only required 30% of total number of samples to attain the observed richness. Similar values were obtained for Chao 2 and ICE (31% and 37% respectively), but Jackknife 2 displayed higher similarity in curve shapes (see Figs. 2A–2K). In addition, ICE was quite erratic and the curve shapes were less similar across ponds. The response of Jackknife 1 was almost as good as that of Jackknife 2 in minimum sample number and constancy (SD) and it did not display erratic behaviours either. Finally, the performance of ACE estimator was intermediate between the previous estimators for all the criteria.

Asymptotic accumulation functions required a high number of samples to attain observed richness and these values had low constancy (Table 3). Expo, in particular, would never reach such value while Clench always needs more that 100% of samples collected to assess observed richness. However, these functions provided curves similar across ponds and with little erratic behaviours (see *e.g.* Fig. 3).

Comparison with "true richness"

Total richness found through three years of sampling at Helada and Grande was considerably higher than the observed richness measured in a single year: 35 against 26 in Helada, 19 against 16 in Grande (Table 4). If we assume this accumulated richness to be a better approach to true richness, as exposed above, we may test which estimators assess richness more precisely. Two of them, Jackknife 2 and Chao 2, provided richness estimations very close to this assumed true richness (Table 4). Estimations by Jackknife 1 are quite close to "true richness" (Table 4) and are only slightly lower than those obtained by Jackknife 2. Chao 1, in contrast, shows a heterogeneous response, with values well below "true richness" (Helada) or above it (Grande de Babia). The relationship between Chao 1 and Chao 2 is not consistent either (Table 4). MMRuns, MMMeans, Bootstrap and accumulation curves provided estimations well below these values. All the richness estimators, excepting species accumulation curves, provided values closer to three-year accumulated richness than to richness measured in a single year (Table 4).

Discussion

There is a general agreement on the advisability of using richness estimators as a better approach to richness than mere lists of observed taxa (Palmer, 1990; Colwell and Coddington, 1994; Bruno and Joslin, 2005; Walther and Moore, 2005). The real debate is about choosing the best estimator for a particular study, taxonomical group or data set. This is a controversial issue, still under study, because different authors reach different conclusions on which is the best estimator, as shown by contrasting results obtained in their studies by, for example, Colwell and Coddington (1994); Walther and Morand (1998); Chiarucci *et al.* (2003); Foggo *et al.* (2003) and Hortal *et al.* (2006). Therefore, until more conclusive information is available, checking the suitability of several estimators seems a convenient starting point when studying little known communities.

Overall, non-parametric methods performed better than accumulation curves in our study. There is a general consensus on the advantages of using nonparametric estimators. They are usually less biased and more precise than accumulation curves, as reported by Brose *et al.* (2003) from a study of landscape simulation under three sampling intensities. Curve-fitting models have been extensively tested and usually perform worse than non-parametric estimators (Melo and Froehlich, 2001; Walther and Moore, 2005). The reason for the usually superior performance of non-parametric estimators might be due to the fact that they, unlike curve models, have been developed from several and underlying models of detection probability (Cam *et al.*, 2002).

In spite of the use of different approaches and different data sets by different authors to evaluate estimation methods, there is some congruence in these results. Expo was the worst estimator found by Peterson and Slade (1998) on data sets derived from state automobile license plates observed in Mexico City and Lawrence, Kansas. These data sets had the advantage of providing known “communities” to be sampled. MMClench was also considered inadequate by Keating *et al.* (1998) on simulated and real data sets. In Melo and Froehlich (2001) the worst of the evaluated methods were MMRuns, MMMeans and Expo and, just as in our study, MMRuns produced erratic behaviour at small sub-samples sizes. In accordance with these studies, MMRuns, MMMeans and the two asymptotic accumulation functions (negative exponential function and Clench function), cannot be recommended for richness estimations on mountain ponds like those in our study area.

The ACE, ICE and Bootstrap methods performed at an intermediate level in our study. In relation to these estimators, the literature is extremely varied and it could be explained from the differences in their sample coverage: different sampling intensity, community evenness, and consequently different sample coverage yield different results concerning estimator ranking (Brose *et al.*, 2003). Bootstrap was a poor estimator in the studies of Colwell and Coddington (1994) and Chazdon *et al.* (1998). This estimator also needed an unacceptably large subsample (65.3%) to estimate sample richness in the study by Melo and Froehlich (2001). Bootstrap and ACE perform poorly at low sampling efforts, which is the portion of the effort curve of greatest interest in practical terms (Foggo *et al.*, 2003). However, Bootstrap has been considered a good estimator in studies with few rare species, such as parasite species richness (Walther and Morand, 1998), or for intensely sampled assemblages (Smith and van Belle, 1984). Hortal *et al.* (2006) showed that although Bootstrap estimates were highly precise, they clearly underestimated species richness (also reported by Chiarucci *et al.*, 2003). ICE and ACE methods have also been found inadequate for species-poor assemblages (Walther and Morand, 1998) but very useful when richness is high (Chazdon *et al.*, 1998). We have not found any relation between the behaviour of the different richness estimators and any of the characteristics of the studied ecosystems (altitude, latitude, temporality, nature of the substrate, depth and vegetal cover). Neither we have found any evidence of a variable behaviours of those richness estimators based on the nature of data set (total abundance; observe richness and percentage of rare taxa). Therefore, with the obtained results in our study, we can not make a solid discussion about the why of the obtained disparity of the conclusions by different authors in relation to these estimators. In the future, we recommend carry out a study more methodical, concrete and specialized that clarifies the true utility of these estimators.

On the other hand, Jackknife 2 was the best of the evaluated methods based on all criteria and also performed well across all studied ponds. Jackknife 1, Chao 1 and Chao 2 also presented good results and they were inferior to Jackknife 2 mainly because of the requirement for larger sub-sample sizes. Similar results have been obtained in previous studies. In Melo and Froehlich (2001) the range of subsample sizes of Jackknife 2, Jackknife 1 and Boot estimators needed to estimate richness in total samples were, respectively, 22.4–26.7, 35.6–41.3 and 63.6–66.7% of total samples. In a similar study, Hellmann and Fowler (1999) found that for Jackknife 1, Jackknife 2 and Boot, the sub-sample size needed to estimate richness in the total sample were, respectively, 22.6–29.1, 36.8–43.9 and 63.1–69.0% of total samples. In both cases, these values are very close to those found in our study. Melo and Froehlich (2001) also concluded that Jackknife 2 was the best estimator methods based on four criteria: (1) the smallest sub-sample size required to estimate the observed richness in the total sample; (2) constancy of the sub-sample size needed to estimate the observed richness in the total sample, measured as 1 standard deviation (SD) of the previous criterion; (3) lack of erratic behaviour in curve

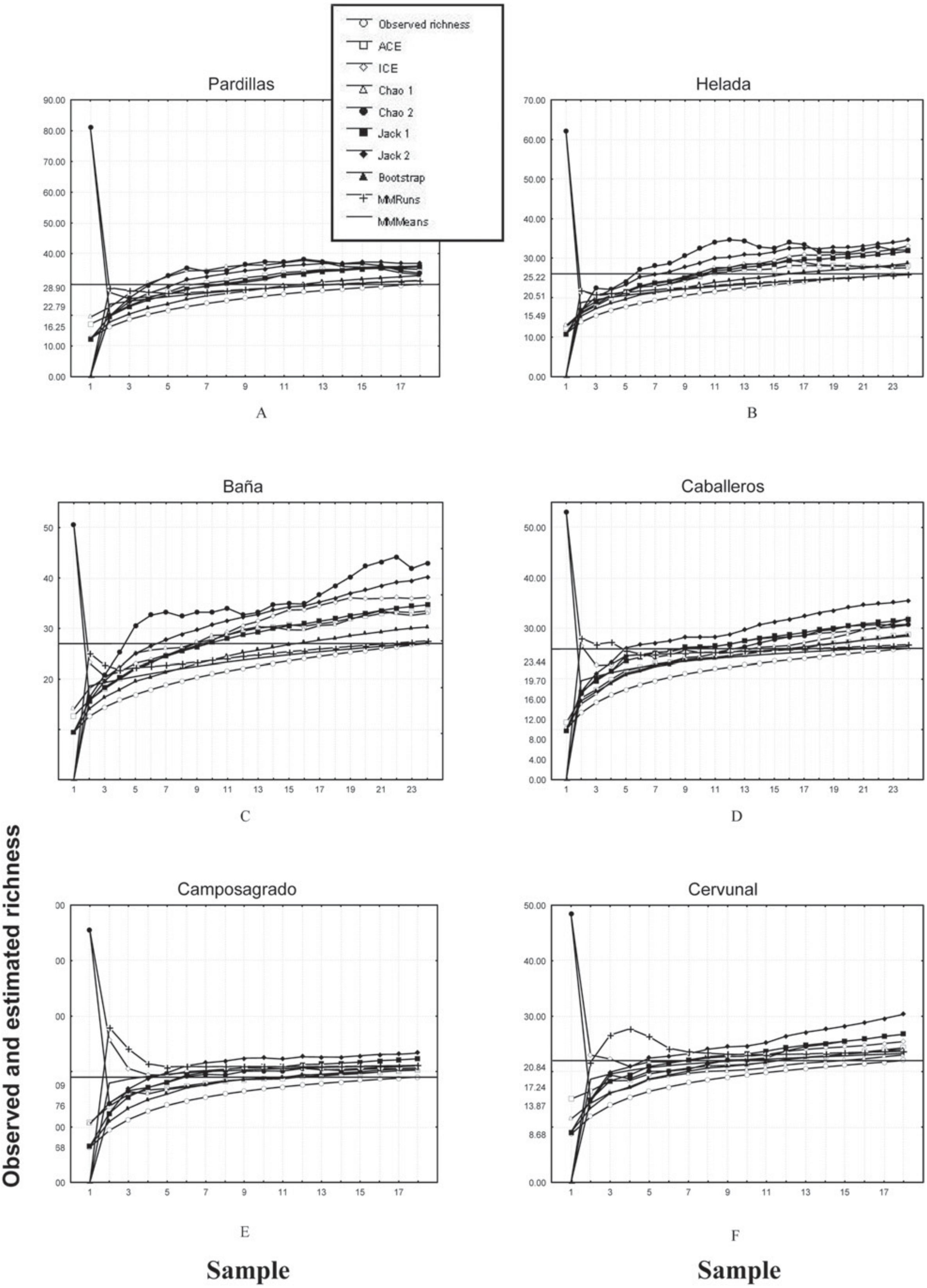

Fig. 2. Curves of accumulation of observed richness and estimates for ponds. The horizontal line indicates the richness observed in the pond. See Table 1 for definitions of the estimators.

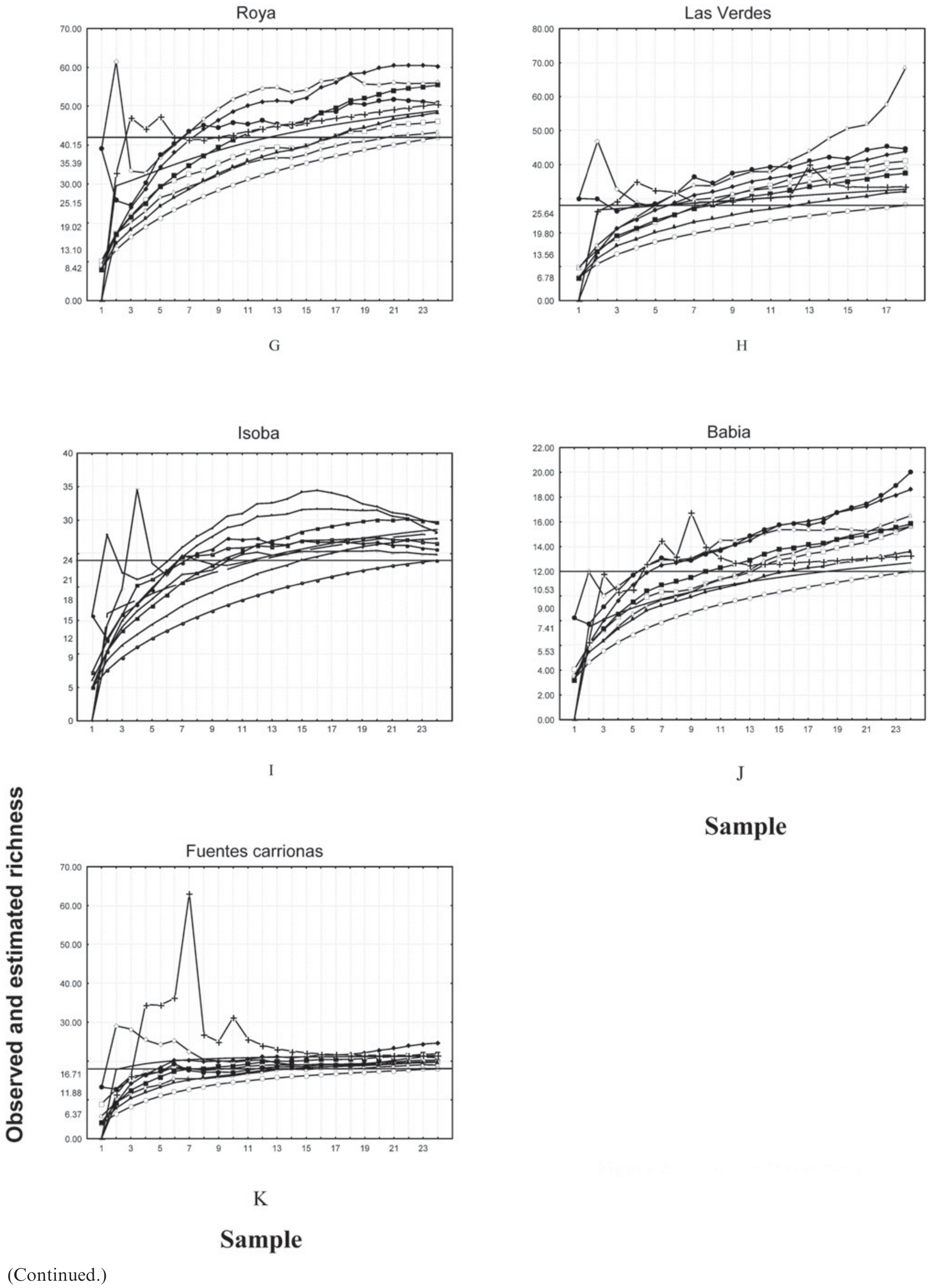

Fig. 2. (Continued.)

Table 3. Performance of the estimators scored by principal criteria. Mean of minimum sample number (x) required to estimate observed richness in all ponds and percentage in relation to the total number of samples collected (%); constancy of minimum sample number across the ponds (*SD*); erratic behaviour (*EB*) and similarity in the curve shape (1, low; 2, intermediate; 3, high).

Estimator	x	%	*SD*	*EB*	Similarity
ACE	10	47	3.7	without	2
ICE	9	37	3.2	with	2
Chao 1	10	45	4.5	without	2
Chao 2	7	31	2.6	without	2
Jack 1	9	42	2.1	without	3
Jack 2	6	30	1.3	without	3
Bootstrap	14	66	2.6	without	3
MMRuns	14	55	7.4	with	1
MMMeans	15	70	6.8	without	2
Clench	32	+100	11.2	without	3
Expo	-	-	-	without	3

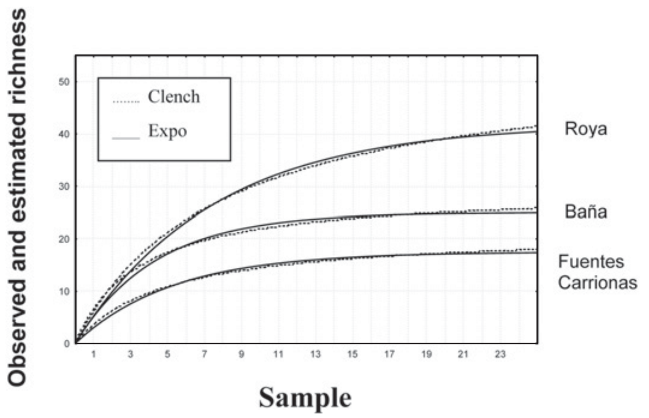

Fig. 3. Asymptotic accumulation functions, Clench and Negative Exponential, of three ponds: Roya, Baña and Fuentes Carrionas.

shape, specially large variations of estimates for closely similar sub-sample sizes; and (4) similarity in curve shape through the six sample data sets. In Petersen and Meier (2003), first and second order Jackknife methods also gave the most accurate estimate of the species richness of the collectable Danish Asilidae using museum label data. They found that first and second order Jackknife and Chao 2 perform better than the other estimation methods because they require less than 50% of the full sample to give stable estimated species richness values. Colwell and Coddington (1994) argued in favour of Chao 2 and Jackknife 2 in a study in which they evaluated the performance of eight methods on a seed-bank data set. Palmer (1990) concluded that Jackknife 1 is the least biased estimator. Chazdon *et al.* (1998) in a study of seedling and sapling diversity found that ICE and Chao 2 were robust to sample size and patchiness. Walther and Morand (1998) evaluation of parasites per host data sets recommended the use of Chao 2 and Jackknife 1. Peterson and Slade (1998) found the Chao 2 method was one of the best (they did not

Table 4. Observed richness on Helada and Grande de Babia through three years, in 2006 and values of methods estimation in 2006.

Estimator	Helada	Babia
True temporal richness (2004, 2006, 2007)	35	19
Observed richness (2006)	26	16
ACE (2006)	28	16
ICE (2006)	33	16
Chao 1 (2006)	27	20
Chao 2 (2006)	32	19
Jack 1 (2006)	32	16
Jack 2 (2006)	35	19
Bootstrap (2006)	29	14
MMRuns (2006)	26	13
MMMeans (2006)	26	13
Clench (2006)	27	14
Expo (2006)	24	11

evaluate Jackknife 1, Jackknife 2 and Chao 1). Using bias and precision as criteria, the two Chao estimators had the best overall performance, followed by the Jackknife 1 and Jackknife 2 estimators, in the study by Walther and Martin (2001). Of course, even Chao and Jackknife estimators may have sometimes a bad perform, the variation of those results depend in many factors that change the structure of the data used on the calculus of the estimators. Total species richness, sample size, and variables that change the aggregation of individuals within samples (Walther and Moore, 2005). In our study these variables are similar and stables in all ponds, this could explain the homogeneity and the constancy of the good behaviour of those two groups of richness estimators in all the studied ecosystems.

An ideal situation for evaluating richness estimators is to compare the estimated value to total species richness in an area. We do not know the total number of taxa of the ponds included in this study, neither have we an expert assessment of species richness, although such expert assessments often are likely to be higher than an estimate based on a collection gathered by non-specialists (Coddington *et al.*, 1996; Longino *et al.*, 2002; Petersen and Meier, 2003). Therefore, we estimated "*true richness*" by pooling the results obtained in three sampling years. Jackknife 2 and Chao 2 were also the best estimators methods. Their values were very close to this assumed true richness. On the other hand, the worst evaluated methods were MMRuns, MMMeans bootstrap and asymptotic accumulation functions. The result based on this variable supports our previous results. The information here provided suggests that estimation of richness in Spanish mountain ponds should probably be based on Jackknife 2 (and Chao 2?) whenever a high number of samples is available. Otherwise, Jackknife 1 (and Chao 1?) may be an acceptable option.

Acknowledgements. We would like to thank all members of the limnology group (University of León). This study was cofunded by the Ministry of Science and Education (CGL2006-03927).

References

Arnott S.E., Magnuson J.J. and Yan N.D., 1998. Crustacean zooplankton species richness: single- and multipleyear estimates. *Can. J. Fish. Aquat. Sci.*, *55*, 1573–1582.

Biggs J., Williams P., Whitfield M., Nicolet P. and Weatherby A., 2005. 15 years of pond assessment in Britain: results and lessons learned from the work of Pond Conservation. *Aquat. Conserv. Mar. Freshw. Ecosyst.*, *15*, 693–714.

Bilton D.T., McAbendroth L.C., Nicolet P., Bedford A., Rundle S.D., Foggo A. and Ramsay P.M., 2009. Ecology and conservation status of temporary and fluctuating ponds in two areas of southern England. *Aquat. Conserv. Mar. Freshw. Ecosyst.*, *19*, 134–146.

Borges P.A.V. and Brown V.K., 2003. Estimating species richness of arthropods in Azorean pastures: the adequacy of suction sampling and pitfall trapping. *Graellsia*, *59*, 7–24.

Brose U., 2002. Estimating species richness of pitfall catches by non-parametric estimators. *Pedobiologia*, *46*, 101–107.

Brose U. and Martinez N.D., 2004. Estimating the richness of species with variable mobility. *Oikos*, *105*, 292–300.

Brose U., Martinez N.D. and Williams R.J., 2003. Estimating species richness: sensitivity to sample coverage and insensitivity to spatial patterns. *Ecology*, *84*, 2364–2377.

Bruno A.W. and Joslin L.M., 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.

Burnham K.P. and Overton W.S., 1979. Robust estimation of population size when capture probabilities vary among animals. *Ecology*, *60*, 927–936.

Cam E., Nichols J.D., Sauer J.R. and Hines J.E., 2002. On the estimation of species richness based on the accumulation of previously unrecorded species. *Ecography*, *25*, 102–108.

Carlton C.E. and Robison H.W., 1998. Diversity of litter-dwelling beetles in the Ouachita Highlands of Arkansas, USA (Insecta: Coleoptera). *Biodivers. Conserv.*, *7*, 1589–1605.

Chao A., 1984. Non-parametric estimation of the number of classes in a population. *Scand. J. Stat.*, *11*, 265–270.

Chao A., 1987. Estimating the population size for capture-recapture data with unequal catchability. *Biometrics*, *43*, 783–791.

Chao A. and Lee S.-M., 1992. Estimating the number of classes via sample coverage. *J. Amer. Stat. Assoc.*, *87*, 210–217.

Chao A., Hwang W.-H., Chen Y.-C. and Kuo C.-Y., 2000. Estimating the number of shared species in two communities. *Statist. Sinica*, *10*, 227–246.

Chao A., Chazdon R., Colwell R.K. and Shen T.-J., 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.*, *8*, 148–159.

Chapin F.S., Zavaleta E.S., Eviner V.T., Naylor R.L., Vitousek P.M., Reynolds H.L., Hooper D.U., Lavorel S., Sala E.O., Hobbie S.E., Mack M.C. and Díaz S., 2000. Consequences of changing biodiversity. *Nature*, *405*, 234–242.

Chazdon R.L., Colwell R.K., Denslow J.S. and Guariguata M.R., 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. *In*: Dallmeier F. and Comiskey J.A. (eds.), Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies, Parthenon Publishing, Paris, 285–309.

Chiarucci A., Maccherini S. and De Dominicis V., 2001. Evaluation and monitoring of the flora in a nature reserve by estimation methods. *Biol. Conserv.*, *101*, 305–314.

Chiarucci A., Enright N.J., Perry G.L.W., Miller B.P. and Lamont B.B., 2003. Performance of nonparametric species richness estimators in a high diversity plant community. *Divers. Distrib.*, *9*, 283–295.

Clench H., 1979. How to make regional fists of butterflies: some thoughts. *J. Lepid. Soc.*, *33*, 216–231.

Coddington J.A., Young L.H. and Coyle F.A., 1996. Estimating spider species richness in a southern Appalachian cove hardwood forest. *J. Arachnol.*, *24*, 111–128.

Colwell R.K., 1997. Estimates: Statistical Estimation of Species Richness and Shared Species from Samples, Version 5. User's guide and application published at: http://viceroy.eeb.uconn.edu/estimates.

Colwell R.K., 2004. Estimates: Statistical Estimation of Species Richness and Shared Species from Samples, Version 7. User's Guide and application published at: http://purl.oclc.org/estimates, University of Connecticut, Storrs.

Colwell R.K. and Coddington J.A., 1994. Estimating terrestrial biodiversity through extrapolation. *Philos. Trans. R. Soc. Lond. Series B*, *345*, 101–118.

Condit R., Hubbel S.P., Lafrankie J.V., Sukumar R., Manokaran N., Foster R.B. and Ashton P.S., 1996. Species–area and species–individual relationships for tropical trees: a comparison of three 50-ha plots. *J. Ecol.*, *84*, 549–562.

Conroy M. and Noon B.R., 1996. Mapping of species richness for conservation of biological diversity: conceptual and methodological issues. *Ecol. Appl.*, *6*, 763–773.

Dumont H.J. and Segers H., 1996. Estimating lacustrine zooplankton species richness and complementarity. *Hydrobiologia*, *341*, 125–132.

Flather C.H., 1996. Fitting species-accumulation functions and assessing regional land use impacts on avian diversity. *J. Biogeogr.*, *23*, 155–168.

Foggo A., Rundle S.D. and Bilton D.T., 2003. The net result: evaluating species richness extrapolation techniques for littoral pond invertebrates. *Freshwater Biol.*, *48*, 1756–1764.

Gaston K.J., 1996. Species richness: measure and measurement. *In*: Gaston K.J. (ed.), Biodiversity: a biology by numbers and difference, Blackwell Science, Oxford, 77–113.

Gaston K.J., 2000. Global patterns in biodiversity. *Nature*, *405*, 220–227.

Gotelli N.J. and Colwell R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.*, *4*, 379–391.

Hammond P.M., 1994. Practical approaches to the estimation of the extent of biodiversity in speciose groups. *Philos. Trans. R. Soc. Lond. Series B*, *345*, 119–136.

Hellmann J.J. and Fowler G.W., 1999. Bias, precision, and accuracy of four measures of species richness. *Ecol. Appl.*, *9*, 824–834.

Heltshe J. and Forrester N.E., 1983. Estimating species richness using the jackknife procedure. *Biometrics*, *39*, 1–11.

Holdridge L.R., Grenke W.C., Hatheway W.H., Liang T. and Tosi J.A., 1971. Forest environments in tropical life zones, Pergamon Press, Oxford.

Hortal J., Garcia-Pereira P. and García-Barros E., 2004. Butterfly species richness in mainland Portugal: predictive models of geographic distribution patterns. *Ecography*, *27*, 68–82.

Hortal J., Borges P.A. and Gaspar C., 2006. Evaluating the performance of species richness estimators: sensitivity to sample grain size. *J. Anim. Ecol.*, *75*, 274–287.

Karakoç G., Erkoç F.U. and Katırcıoglu H., 2003. Water quality and impacts of pollution sources for Eymir and Mogan Lakes (Turkey). *Environ. Int.*, *29*, 21–27.

Keating K.A., Quinn J.F., Ivie M.A. and Ivie L., 1998. Estimating the effectiveness of further sampling in species inventories. *Ecol. Appl.*, *8*, 1239–1249.

Keddy P.A. and Drummond C.G., 1996. Ecological properties for the evaluation, management, and restoration of temperate deciduous forest ecosystems. *Ecol. Appl.*, *6*, 748–762.

Kerr J.T., 1997. Species richness, endemism, and the choice of areas for conservation. *Conserv. Biol.*, *11*, 1094–1100.

Lee S.-M. and Chao A., 1994. Estimating population size via sample coverage for closed capture-recapture models. *Biometrics*, *50*, 88–97.

Longino J.T., Coddington J.A. and Colwell R.R., 2002. The ant fauna of a tropical rain forest: estimating species richness in three different ways. *Ecology*, *83*, 689–702.

Magurran A.E., 1988. Ecological Diversity and its Measurement, Croom-Helm, London.

Martinez N.D., Hawkins B.A., Dawah H.A. and Feifarek B.P., 1999. Effects of sampling effort on characterization of food-web structure. *Ecology*, *80*, 1044–1055.

Melo A.S. and Froehlich C.G., 2001. Evaluation of methods for estimating macroinvertebrate species richness using individual stones in tropical streams. *Freshwat. Biol.*, *46*, 711–721.

Melo A.S., Pereira R.A.S., Santos A.J., Shepherd G.J., Machado G., Medeiros H.F. and Sawaya R.J., 2003. Comparing species richness among assemblages using sample units: why not use extrapolation methods to standardize different sample sizes? *Oikos*, *101*, 398–410.

Michaelis M. and Menten M.L., 1913. Der kinetik der invertinwirkung. *Biochem. Z.*, *49*, 333–369.

Miller R.I. and Wiegert R.G., 1989. Documenting completeness, species-area relations, and the species-abundance distribution of a regional flora. *Ecology*, *70*, 16–22.

Oertli B., Auderset J.D., Castella E., Juge R., Cambin D. and Lachavanne J.B., 2002. Does size matter? The relationship between pond area and biodiversity. *Biol. Conserv.*, *104*, 59–70.

Oertli B., Biggs J., Céréghino R., Grillas P., Joly P. and Lachavanne J.-L., 2005. Conservation and monitoring of pond biodiversity: introduction. *Aquat. Conserv. Mar. Freshw. Ecosyst.*, *15*, 535–540.

Palmer M.W., 1990. The estimation of species richness by extrapolation. *Ecology*, *71*, 1195–1198.

Petersen F.T. and Meier R., 2003. Testing species-richness estimation methods on single-sample collection data using the Danish Diptera. *Biodiv. Conserv.*, *12*, 667–686.

Peterson A.T. and Slade N.A., 1998. Extrapolating inventory results into biodiversity estimates and the importance of stopping rules. *Divers. Distrib.*, *4*, 95–105.

Prendergast J.R., Quinn R.M., Lawton J.H., Eversham B.C. and Gibbons D.W., 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature*, *365*, 335–337.

Pressey R.L., Humphries C.J., Margules C.R., Vane-Wright R.I. and Williams P.H., 1993. Beyond opportunism: key principles for systematic reserve selection. *Tr. Ecol. Evol.*, *8*, 124–128.

Raaijmakers J.G.W., 1987. Statistical analysis of the Michaelis-Menten equation. *Biometrics*, *43*, 793–803.

Schindler D.E., Knapp R.A. and Leavitt P.R., 2001. Alteration of nutrient cycles and algal production resulting from fish introductions into Mountain lakes. *Ecosystems*, *4*, 308–321.

Schippers P., van de Weerd H., Klein J., Jong B. and Scheffer M., 2006. Impacts of agricultural phosphorus use in catchments on shallow lake water quality: About buffers, time delays and equilibria. *Sci. Total Environ.*, *369*, 280–294.

Smith E.P. and van Belle G., 1984. Nonparametric estimation of species richness. *Biometrics*, *40*, 119–129.

Soberón J. and Llorente J., 1993. The use of species accumulation functions for the prediction of species richness. *Conserv. Biol.*, *7*, 480–488.

Søndergaard M. and Jeppesen E., 2007. Anthropogenic impacts on lake and stream ecosystems, and approaches to restoration. *J. Appl. Ecol.*, *44*, 1089–1094.

Toro M., Granados I., Robles S. and Montes C., 2006. High mountain lakes of the Central Range (Iberian Peninsula): Regional limnology & environmental changes. *Limnetica*, *25*, 1-2, 217–252.

Walther B.A. and Martin J.L., 2001. Species richness estimation of bird communities: how to control for sampling effort? *Ibis*, *143*, 413–419.

Walther B.A. and Morand S., 1998. Comparative performance of species richness estimation methods. *Parasitology*, *116*, 395–405.

Walther B.A. and Moore J.L., 2005. The definitions 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.

Walther B.A., Cotgreave P., Gregory R.D., Price R.D. and Clayton D.H., 1995. Sampling effort and parasite species richness. *Parasitol. Today*, *11*, 306–310.

Whittaker R.J., Willis K.J. and Field R., 2001. Scale and species richness: towards a general, hierarchical theory of species richness. *J. Biogeogr.*, *28*, 453–470.

Williams R.J. and Martinez N.D., 2000. Simple rules yield complex food web dynamics. *Nature*, *404*, 180–183.

Williams P., Whitfield M., Biggs J., Bray S., Fox G., Nicolet P. and Sear D., 2004. Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in Southern England. *Biol. Conserv.*, *115*, 329–341.