Bol. Soc. Zool. Uruguay,
2ª época, 2006. 15: 1-7

ASSESSING THE COMPLETENESS OF BIODIVERSITY INVENTORIES: AN EXAMPLE FROM BAÑADOS DEL ESTE BIOSPHERE RESERVE, URUGUAY

Alvaro Soutullo.

School of Biological Sciences, University of East Anglia. Norwich NR4 7TJ, UK. Email: a.soutullo@adinet.com.uy

ABSTRACT

The performance of five non-parametric estimators of species richness (ICE, Jack1, Jack2, Chao2 and Bootstrap) was compared using available inventories for three vertebrate groups within the Biosphere Reserve Bañados del Este, Uruguay. The data used were compiled from zoological collections, bibliography, field notes and other unpublished sources. All estimators provided good estimations of true richness when approximate 50% of the localities sampled were considered, but only Jack2, ICE and Chao2 provided good estimations with lower samples. For data that was not systematically gathered and shows varying degrees of accuracy, I suggest simultaneously using the indexes ICE, Jack1 and Jack2 to evaluate inventories' completeness. Together these indexes provide good estimations of the true species richness, even when sampling effort is low, and they define a range of values within which true richness is likely to fall.

KEY-WORDS: Biosphere reserve, *EstimateS*, non-parametric estimators, species richness, Uruguay.

RESUMEN

Evaluando cuán exhaustivos son los inventarios de biodiversidad: un ejemplo de la Reserva de Biósfera Bañados del Este, Uruguay.

Se evaluó el rendimiento de cinco índices no paramétricos de riqueza de especies (ICE, Jack1, Jack2, Chao2 y Bootstrap) utilizando inventarios disponibles para tres grupos de vertebrados dentro de la Reserva de Biósfera Bañados del Este, Uruguay. Los datos utilizados fueron compilados de colecciones zoológicas, bibliografía, notas de campo y otras fuentes no publicadas. Todos los índices brindaron buenas estimaciones de la verdadera riqueza una vez que cerca del 50% de las localidades estudiadas fueron consideradas, pero sólo Jack2, ICE y Chao2 brindaron buenas estimaciones con muestras más pequeñas. Para datos que no han sido colectados sistemáticamente, y que poseen diferentes grados de exactitud, sugiero utilizar simultaneamente los índices ICE, Jack1 y Jack2 para evaluar cuán exhaustivos son los inventarios disponibles. Juntos, estos índices brindan buenas estimaciones de la verdadera riqueza, incluso cuando el esfuerzo de muestreo es pequeño, y definen un rango de valores dentro del cual es esperable que se encuentre la verdadera riqueza de especies.

PALABRAS CLAVE: Reserva de Biósfera, *EstimateS*, estimadores no paramétricos, riqueza de especies, Uruguay.

INTRODUCTION

Ensuring that the maximum numbers of species are properly protected is often an important goal in conservation planning (Sutherland, 2000). Nevertheless, species richness is generally an elusive quantity to measure properly (May, 1988). The problem of how many species occur in an area is quite resistant to statistical solution. This is because no matter how many species have been recorded already, a number of very rare species may still remain unnoticed (Bunge & Fitzpatrick, 1993).

A common technique to estimate species richness is to extrapolate true richness using species accumulation curves (Soberón & Llorente, 1993; Sutherland, 2000). Although this is a valuable technique for well-designed ecological studies, its applicability is more limited when the quality of the available data is poor or the sampling has not been standardised (Colwell & Coddington, 1994; Peterson & Slade, 1998; Gotelli & Colwell, 2001; Gómez de Silva & Medellín, 2001; Moreno & Halffter, 2001; Willot, 2001). Unfortunately, the data on which conservation and management decisions have to be made are often far less accurate than that from well designed ecological studies (especially if the area of concern is large), which limits the value of species accumulation curves as a conservation tool.

An alternative to species accumulation curves is the use of non-parametric methods to estimate species richness from samples (Colwell & Coddington, 1994; Colwell, 1999). However, these methods have not been widely evaluated yet (Colwell & Coddington, 1994) and their relative merits have been poorly explored so far (Palmer, 1990, 1991; Colwell & Coddinton, 1994). The aim of this note is to evaluate the efficiency of these methods to assess the completeness of species inventories from data that were not systematically gathered, come from different sources, and have varying degrees of accuracy. To do that, data of species from three vertebrate groups recorded within UNESCO's Bañados del Este Biosphere Reserve, Uruguay was used.

The Biosphere Reserve Bañados del Este was created in 1976 to protect a diverse mosaic of coastal and terrestrial habitats including grasslands, woodlands, wetlands, a system of coastal lagoons, rocky shores, and sandy beaches. The reserve also includes one of the most important productive regions of the country (PROBIDES, 1999). Despite records of the reserve's fauna have been collected for more than a century, there are still several areas that have been very poorly studied, and no systematic assessment of the reserve's species richness has been conducted so far (Soutullo *et al*., 2003).

MATERIALS AND METHODS

The data used for this analysis was extracted from a database compiled by the Uruguayan museum of natural history, PROBIDES and the NGO VIDA SILVESTRE (Uruguay) in 2001. This database was compiled from zoological collections, bibliography, field notes and other unpublished sources, and is the most complete account of mammals, amphibians and reptiles recorded within the reserve between 1812 and 2001. It includes more than 3,000 records of 31 species of amphibians from 162 different localities, 69 species of mammals from 276 localities, and 43 species of reptiles from 159 localities. Accuracy and precision of these data, as well as the methods of data gathering, vary notably. While for a given locality the list of known species may be the outcome of several year of intense sampling, for others it is the result of only one or a few visits, or haphazard observations. All these factors undermine the ability of different estimators of species richness to provide accurate estimations (Palmer, 1990; Colwell & Coddington, 1994), but at the same time provide an ideal scenario for comparing their performance when data quality is poor.

The software *EstimateS* (Colwell, 1999) was used to analyse the completeness of the inventories available for these groups using five non-parametric incidence-based estimators of true species richness: ICE, Jack1, Jack2, Chao2 and Bootstrap (Colwell & Coddington, 1994; Colwell, 1999). As suggested by Colwell (1999), rarefaction curves were plotted to compare the performances of these indexes for different numbers of sampled localities. The advantage of using rarefaction curves instead of an arbitrary ordering of samples is that they represent the means of repeated re-sampling of all pooled samples (Gotelli & Colwell, 2001). They show how the expected (or observed) number of species changes with an increase in the number of samples drawn at random from the pool of all the samples taken. Thus, as they represent the statistical expectation for the corresponding accumulation curves, they allow meaningful standardisation and comparisons (Gotelli & Colwell, 2001). For this analysis *EstimateS* was used to calculate the mean values for each estimator and sample size after re-sampling the pool of localities 50 times.

RESULTS

Fig. 1 shows the performance of the five estimators when different numbers of localities have been sampled. Table I shows the final species richness estimated for each group using these indexes. Important differences in the performance of the estimators were observed. Bootstrap consistently estimated the lowest true number of species, and was a poor estimator until approximate 50% of the samples were considered. Jack2 showed the opposite pattern. Once roughly 25% of the samples were examined, it consistently estimated the largest number of species. Jack1 showed a performance that was intermediate between those of these two indexes. However, final values tended to be closer to those estimated by Jack2. Indexes Chao2 and ICE behaved in a similar fashion to one another. Even when less than 15% of the samples were considered, they both estimated values of true richness that were close to the values estimated when all of the samples were considered. Their estimations systematically lied between those of Bootstrap and Jack2. However, Chao2 fluctuates more than ICE. Fig. 2 shows a comparison of the standard deviation of their values computed for different sample sizes.

Table I - Number of species of amphibians, mammals and reptiles (top to bottom) recorded within Bañados del Este Biosphere Reserve (Uruguay) since 1812 (Sobs), and estimations of the true richness in the area, according to five different non-parametric estimators of species richness.

	Sobs	ICE	Chao2	Jack1	Jack2	Bootstrap
Reptiles	43	45.84	46	48.96	49	46.13
Mammals	69	75.53	76.56	79.96	82.97	74.44
Amphibians	31	31.28	31.5	31.99	32	31.61

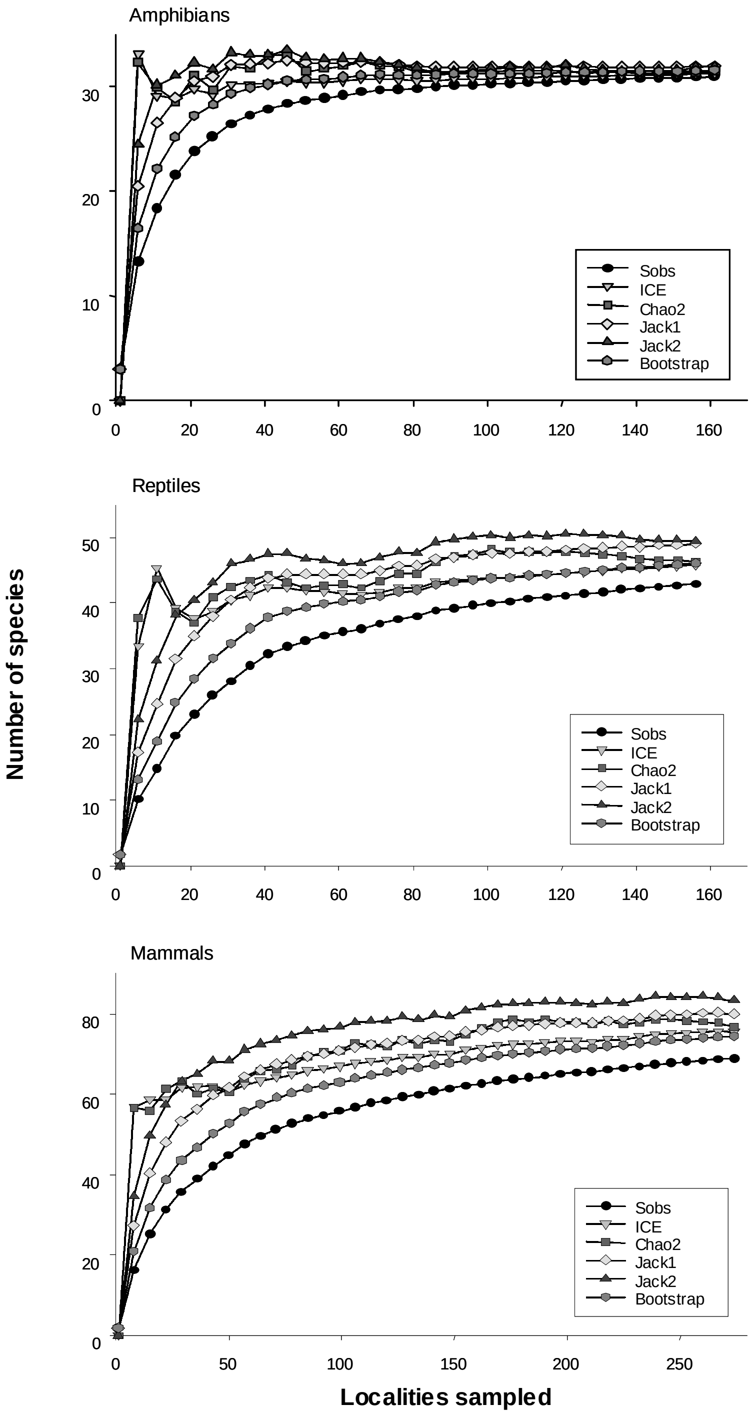

Fig. 1 - Performance of five non-parametric estimators of species richness for three vertebrate groups in Bañados del Este Biosphere Reserve (Uruguay). Sobs indicates the observed number of species as a function of the number of pooled samples from all the localities sampled. The other curves display the estimated total number of species based on successively larger number of samples (from top to bottom the graphs show amphibians, reptiles and mammals richness)

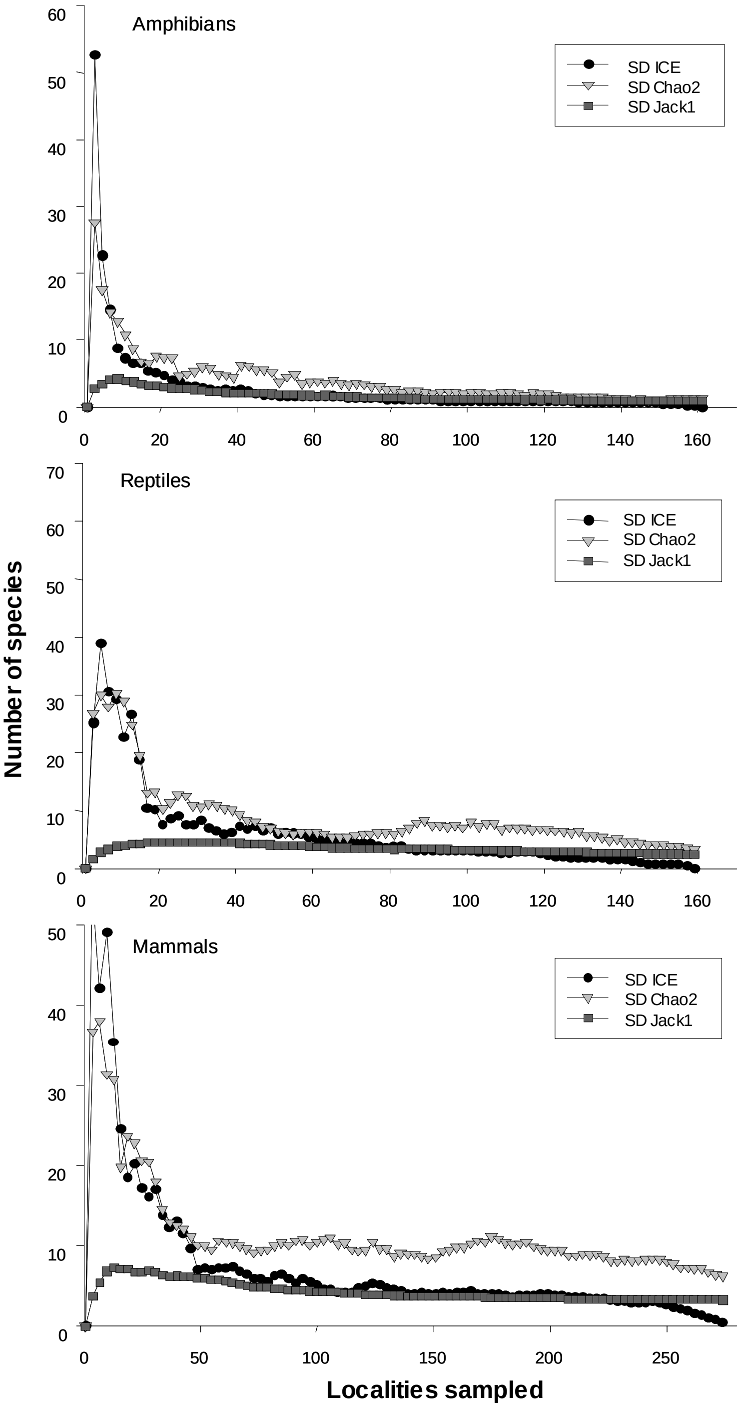

Fig. 2 – Changes in the standard deviation (SD) of the estimated number of amphibians, reptiles and mammals (top to bottom) in Bañados del Este Biosphere Reserve (Uruguay) for different sample sizes according to three non-parametric estimators of species richness (ICE, Chao2 and Jack1).

DISCUSSION

The ability of these estimators to accurately estimate the true number of species in an area depends of the accuracy with which rare species are identified, the exhaustiveness of the information available from the localities sampled, and the representativeness of these localities of the range of habitats included within the study area. If species have been misclassified as rare due to poor sampling, or if localities sampled do not represent the heterogeneity of the study area, then the estimators' ability to properly estimate the actual number of species in the area decreases (Palmer, 1990; Colwell & Coddington, 1994). Furthermore, species richness and composition are assumed to be constant in time, something that is not necessarily true for the data set used here. The poor quality of the data used for this analysis is far from the ideal expected for a good performance of these indexes. However, the data set does represent the kind of data available for most biodiversity assessments of large areas (e.g., Ponder *et al.*, 2001; Johnson & Ward, 2002). Hence, comparing the performance of these index with these data, may provide some insight about how robust they are to deviations from proper sampling design, and thus, how useful they are for conservation planning and assessment in many real–word situations.

Bunge and Fitzpatrick (1993) suggest that in the absence of precise information about sampling and population structure, the more appropriate estimator to use is Chao2. This is because it tolerates deviations from the sampling plan, but adjusts upwards to account for non-homogeneity in class sizes. Peterson and Slade (1998), and Colwell and Coddington (1994) also suggest the use of Chao2, as it is a good estimator of true richness even when sampling effort has been low. Colwell and Coddington (1994) also highlight the appropriateness of using Jack2, as its performance is similar to that of Chao2. Contrasting, Palmer (1990, 1991) suggests that Jack1 gets closer to the true number of species than the other estimators he used (Bootstrap and Jack2 among them), but also suggests that Jack2 is the least biased (Palmer, 1991).

Based on this analysis I would suggest that for data that was not systematically gathered and that shows varying degrees of accuracy, ICE should be used instead of Chao2 to estimate true richness. Improvements in the way it is calculated were introduced after the works of Bunge and Fitzpatrick (1993), Palmer (1990, 1991), and Colwell and Coddington (1994), and have largely increased its ability to accurately estimate true richness (Colwell, 1999). Like Chao2 it is very efficient at estimating richness even when few localities are sampled. However, while it provides similar results when sampling effort increases, it is more stable and shows less variance than Chao2.

Indeed, I suggest using simultaneously the indexes ICE, Jack1 and Jack2 when conducting biodiversity assessments based on poor quality data. Using ICE ensures good estimations of true richness will rapidly arise after a relatively low sampling effort, whereas Jack2 soon provides values that paired with ICE values would define a range in which the true richness is likely to fall. As suggested by Palmer (1990), once enough samples are collected Jack1 may then be used as an indicator of the most likely value within that range for the true number of species.

One of the advantages of simultaneously using several indexes is that their convergence (or lack of convergence) towards a richness value can be used as an indication of how complete a given inventory is. Convergence towards a value close to the number of species actually observed, as in the case of amphibians in this data set, may be interpreted as an indication that the available inventory is rather complete. Convergence towards another richness value, as can be observed for reptiles in this data set, can be seen as an indication of how many new species are still to be recorded. Conversely, lack of convergence of the estimations, as observed for mammals, may suggest not only that the inventory is still incomplete, but that because of the poor quality of the data available, it is unclear how incomplete it is.

ACKNOWLEDGEMENTS

Thanks are due to Johanna Maughan for improving this note English, and to two anonymous referees that help to improve its clarity.

REFERENCES

Bunge, J. & Fitzpatrick, M. 1993. Estimating the number of species: a review. Journal of the American Statistical Association, 88: 364-373.

Colwell, R.K. & Coddington, J.A. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London, Series B, 345: 101-118.

Colwell, R.K. 1999. EstimateS 5: Statistical Estimation of Species Richness and Shared Species from Samples, Version 5.0.1. User's Guide. http://viceroy.eeb.uconn.edu/estimates [accessed on 15/02/03]. University of Connecticut, Storrs, CT.

Gotelli, N.J. & Colwell, R.K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4: 379-391.

Gómez de Silva, H. & Medellin, R.A. 2001. Evaluating completeness of species lists for conservation and macroecology: a case study of Mexican land birds. Conservation Biology, 15: 1384-1395.

Johnson, R.A. & Ward, P.S. 2002. Biogeography and endemism of ants (Hymeoptera: Formicidae) in Baja California, Mexico: a first overview. Journal of Biogeography, 29: 1009-1026.

May, R.N. 1988. How many species on Earth? Science, 241: 1441-1449.

Moreno, C.E. & Halffter, G. 2001. On the measure of sampling effort used in species accumulation curves. Journal of Applied Ecology, 37: 149-158.

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

Palmer, M.W. 1991. Estimating species richness: the second-order jackknife reconsidered. Ecology, 72: 1512-1513.

Ponder, W.F., Carter, G.A., Flemons, P. & Chapman, R.R. 2001. Evaluation of museum collection data for use in biodiversity assessment. Conservation Biology, 15: 649-657.

Peterson, A.T. & Slade, N.A. 1998. Extrapolating inventory results into biodiversity estimates and the importance of stopping rules. Diversity and Distributions, 4: 95-105.

PROBIDES. 1999. Plan Director. Reserva de Biósfera Bañados del Este/Uruguay. PROBIDES, Rocha. 159 pp.

Soberón, J. & Llorente, J. 1993. The use of species accumulation functions for the prediction of species richness. Conservation Biology, 7: 480-488.

Soutullo, A., Chan, E.K.W., Maughan, J., Preuss, S. & Ruiz-Guajardo, J.C. 2003. Distribution of endangered vertebrates in Bañados del Este Biosphere Reserve, Uruguay. Comunicaciones Zoológicas del Museo de Historia Natural de Montevideo, 201: 1-12.

Sutherland, W.J. 2000. The conservation handbook: research, management and policy. Blackwell Science, Oxford. 278 pp.

Willot, S.J. 2001. Species accumulation curves and the measure of sampling effort. Journal of Applied Ecology, 38: 485-487.