

ESTIMATING SPECIES RICHNESS OF TROPICAL BIRD COMMUNITIES FROM RAPID ASSESSMENT DATA

SEBASTIAN K. HERZOG,^{1,4} MICHAEL KESSLER,² AND THOMAS M. CAHILL³

¹*Institut für Vogelforschung "Vogelwarte Helgoland", An der Vogelwarte 21, 26386 Wilhelmshaven, Germany;*

²*Albrecht-von-Haller Institut für Pflanzenwissenschaften, Abteilung Systematische Botanik, Untere Karspüle 2, 37073 Göttingen, Germany; and*

³*Canadian Environmental Modelling Centre, Trent University, 1600 West Bank Drive, Peterborough, Ontario K9J 7B8, Canada*

ABSTRACT.—Rapid assessment surveys of tropical bird communities are increasingly used to estimate species richness and to determine conservation priorities, but results of different studies are often not comparable due to the lack of standardization. On the basis of computer simulations and six years of field testing, we evaluated the recently proposed "20-species-list" survey method and statistical estimators for assessing species richness of tropical bird communities. This method generates a species-accumulation curve by subdividing consecutive observations of birds into lists of 20 species, thus relating cumulative species richness to the number of observations rather than time or space and thereby accounting for moderate differences in observer qualification and field conditions. Species accumulation curves from computer-simulated communities and two empirical data sets from Bolivia were analyzed with nine species richness estimators to evaluate estimator accuracy with respect to variations in species-list size, sample size, species-pool size, and community structure. For empirical and most simulated data sets, the MMMEAN estimator performed best, but it was more sensitive to differences in community structure than most other estimators. The CHAO 2 estimator, which was recommended by previous studies, performed reasonably well but was considerably more sensitive to sample size than MMMEAN. The bootstrap and first- and second-order jackknife estimators performed poorly. We recommend using MMMEAN or, when standard deviations of richness estimates are indispensable, CHAO 2 with 10-species lists for estimating species richness of tropical bird communities and propose a set of standard survey rules. Careful examination of estimator accumulation curves is required, however, and a technique based on the ratio between estimator and species accumulation curve is suggested to control for the confounding effects of sampling effort. Overall, the species-list method combined with statistical richness estimation is doubtlessly much more standardized and valuable than simple comparisons of one-dimensional locality lists and represents a promising tool for conservation assessment and the study of avian diversity patterns in the tropics. *Received 29 June 2001, accepted 15 April 2002.*

RESUMEN.—Cada vez se usan con mayor frecuencia evaluaciones rápidas de comunidades de aves tropicales para estimar la riqueza de especies y para determinar prioridades de conservación, pero los resultados de diferentes estudios a menudo no son comparables debido a la falta de estandarización. Basados en simulaciones realizadas en computadoras y en seis años de evaluaciones de campo, evaluamos el método de muestreo "lista de 20 especies" recientemente propuesto y los estimadores estadísticos para determinar la riqueza de especies de comunidades de aves tropicales. Este método genera curvas de acumulación de especies subdividiendo observaciones consecutivas de aves en listas de 20 especies. Así, relaciona la riqueza acumulada de especies con el número de observaciones y no con tiempo o espacio, incorporando de este modo diferencias moderadas en la habilidad del observador y en las condiciones del tiempo. Analizamos curvas acumuladas de especies originadas a partir de comunidades simuladas y de dos juegos de datos empíricos de Bolivia empleando nueve estimadores de la riqueza de especies para evaluar la exactitud de los estimadores con relación al tamaño de las listas de especies, tamaño de la muestra, tamaño del set de especies y estructura de la comunidad. Para juegos de datos empíricos y para la mayoría de los simulados, el estimador MMMEAN fue el mejor, pero fue más sensible que otros estimadores

⁴ E-mail: skherzog@compuserve.com

a diferencias en la estructura de la comunidad. El estimador CHAO 2, recomendado por estudios previos, funcionó razonablemente bien, pero fue considerablemente más sensible al tamaño de la muestra que MMMEAN. El "bootstrap" y los estimadores de primer y segundo orden de "jackknife" funcionaron mal. Recomendamos usar MMMEAN o, cuando se requieren estimaciones de la desviación estándar de las estimaciones de riqueza, CHAO 2 con listas de 10 especies para estimaciones de riqueza de especies de comunidades de aves tropicales, y proponemos una serie de reglas estándares de muestreo. Sin embargo, es necesario examinar cuidadosamente los estimadores de curvas acumulativas, y sugerimos una técnica basada en el cociente entre el estimador y la curva acumulada de especies para controlar la distorsión creada por efectos de esfuerzo de muestreo. En resumen, el método de la lista de especies combinado con la estimación estadística de la riqueza es sin duda mucho más estándar y válido que comparaciones simples de listas de especies unidimensionales, y representa una herramienta prometedora para evaluaciones de conservación y el estudio de los patrones de diversidad de aves en el trópico.

QUANTIFYING THE SPECIES richness of bird communities has gained increasing importance in environmental impact assessments (e.g. Fjeldsø 1999), conservation planning (Bibby et al. 1992, Stotz et al. 1996), and ecological research (Huston 1994, Rosenzweig 1995). In Holarctic regions, where species richness is low and avian communities are well characterized, standardized count and census methods are available (Holmes et al. 1986, Bibby et al. 2000). In the tropics, however, species-rich communities inhabit highly complex, heterogeneous environments and those methods are often difficult to apply (Terborgh et al. 1990, Remsen 1994, Poulsen et al. 1997a). Detailed quantitative studies of tropical bird communities are highly labor intensive (e.g. 12 person-months by Terborgh et al. 1990) and require a variety of methods such as mist-netting, spot-mapping, point counts, and observation of mixed species flocks to achieve a high degree of completeness (Remsen and Parker 1983, Terborgh et al. 1990, Poulsen 1994, Remsen 1994, Gram and Faaborg 1997). They are therefore limited to few sites in selected habitats (e.g. Blake et al. 1990, Terborgh et al. 1990, Cohn-Haft et al. 1997, Robinson et al. 2000).

In the light of increasing forest destruction (Dale et al. 1994) and wide gaps in the understanding of tropical bird communities, several researchers have recently applied a rapid assessment approach to maximize data collection with limited funds, time, and personnel (Parker and Bailey 1991, Parker et al. 1993, Poulsen et al. 1997a, Poulsen and Krabbe 1998). To find general patterns in responses of avian species richness to disturbance, many spatial and temporal replicates will be necessary, instead of conducting exact counts at a few sites. This is a

principal reason for developing rapid assessment methods. Unfortunately, lack of standardization with respect to survey method, observer qualification, area, time, weather, and season have largely precluded quantitative comparisons of rapid assessment studies.

Until recently, the most frequently used method for inventories of tropical bird communities was mist-netting, which was often combined with unstandardized visual observations and tape recordings (Karr 1981, 1990; Poulsen 1994; Schmitt et al. 1997). Mist-netting undoubtedly has an advantage in reducing biases introduced by varying observer experience and qualification, but it is nonetheless subject to a variety of other biases, such as net avoidance, weather, habitat structure, and behavioral differences between species and individuals of the same species (Karr 1981, 1990; Jenni et al. 1996; Remsen and Good 1996). Other major disadvantages of mist-netting are high labor intensity, low time efficiency, the comparatively small proportion of the total community sampled, and a strong bias towards understory species (Gram and Faaborg 1997, Whitman et al. 1997).

The recent increase in knowledge of vocalizations (Parker 1991, Budney and Grotke 1997), commercially available reference recordings (e.g. Mayer 2000), and high-quality field guides has resulted in an increased use of acoustical and visual observations in rapid assessments. However, even highly experienced observers are subject to a variety of biases, mainly relating to varying detectabilities between species or between seasons for any particular species (e.g. Karr 1981, 1990; Oniki and Willis 1982a, b, 1983; Verner 1985; Verner and Milne 1990), which are compounded by differences in ob-

server skill and observation technique (Sauer et al. 1994). Thus, purely observational data must always be considered with caution (Cohn-Haft et al. 1997). On the other hand, auditory-visual methods are much more time efficient, report a considerably larger portion of local avifaunas than mist netting (Whitman et al. 1997), and the presence of most species at any given locality can be documented by tape recording (Parker 1991).

Some auditory-visual rapid assessments employed standard survey techniques such as point counts (Poulsen and Krabbe 1998) or line transects (Karr 1971). Although those methods generate quantifiable data with well-established sampling protocols, they have several disadvantages for rapid assessments: they (1) are often difficult or impossible to apply under tropical field conditions (e.g. in steep, inaccessible montane forests), (2) require the exclusion of observations between sample intervals from the analysis, which further reduces the already limited amount of obtainable data, (3) require high observer qualification (Poulsen et al. 1997a), and (4) tend to underestimate richness and abundance of some bird groups (e.g. nocturnal species; Bibby et al. 2000). Thus, an ideal survey or analysis method should lack the limitations of rigid standardization while ensuring comparability between studies.

MacKinnon and Phillips (1993) suggested a quantitative approach to analyzing auditory-visual survey data that accounts, at least to some degree, for differences in effort, observer qualification, and weather (Poulsen et al. 1997a). In that method, observations are grouped into consecutive lists of 20 species and a species accumulation curve is generated from adding those species not recorded on any previous list to the total species number, which is then plotted as a function of list number. It is crucial to include even observations that cannot be positively identified at first (Poulsen et al. 1997a). Hence, observers can devote the necessary time to become completely familiar with the avifauna and are forced to tape record and track down unidentified vocalizations. Because the method relates species richness to the number of observations rather than time or area, this method allows comparison of data obtained by different observers or under varying field conditions. Of course, no method will ac-

count for strong deficiencies in observer qualification or extreme weather.

Although attractive due to its simplicity and the possibility for comparing different studies, the "20-species-list method" (Poulsen et al. 1997a) remains largely untested. Poulsen et al. (1997a) concluded that "20-species lists are biased like any other bird count, but not more than other methods" and that the method is "suitable for judging (a) when a site is adequately surveyed, (b) the magnitude of the species richness, (c) the relative abundance of each species and (d) an α -index of diversity." These conclusions, however, are based on a variety of untested assumptions about frequency distributions. In addition, neither MacKinnon and Phillips (1993) nor Poulsen et al. (1997a) suggested standardized survey rules or assessed the possibilities of statistical data analysis. In a later note, Poulsen et al. (1997b) considered potential biases and called for more rigorous testing before the method can be recommended.

Quantitative comparisons of species accumulation curves have been widely explored (e.g. Palmer 1990, Baltanás 1992, Bunge and Fitzpatrick 1993, Colwell and Coddington 1994, Walther and Morand 1998, Gotelli and Colwell 2001) but nonetheless are seldom used in biodiversity studies (Boulinier et al. 1998). To estimate species richness from accumulation curves, three types of analyses have been employed: (1) extrapolation, (2) fitting species-abundance distributions, and (3) nonparametric estimators (Soberón and Llorente 1993, Colwell and Coddington 1994, Walther et al. 1995, Chazdon et al. 1998, Walther and Morand 1998). However, the performance of estimators and their sensitivity to variations in sampling protocol, sample size, species richness, and other variables remains largely unknown.

Based on computer simulations and six years of field testing, this article presents a quantitative assessment of the species-list method and recommendations for a standardization of surveys to allow quantitative comparisons between studies. We further explore possibilities of estimating species richness from data gathered with the species-list method by comparing the performance of nine statistical richness estimators included in the program ESTIMATES 5.0.1 (Colwell 1997). Specifically, we assessed the following parameters for each estimator: (1) accuracy, (2) sensitivity to sample

size (i.e. number of species lists), (3) sensitivity to true species richness, and (4) the influence of underlying species-abundance or species-detectability distributions. An ideal estimator should have constant, high accuracy irrespective of changes in any of the other parameters.

METHODS

COMPUTER SIMULATIONS FOR ESTIMATION TESTS

To evaluate performance and biases of statistical richness estimators, we used a computer model that created lists of consecutive observations drawn at random from species pools of different sizes with a species-abundance or species-detectability distribution according to the exponential decline model:

$$N_x = N_0 e^{-rx} \quad (1)$$

where N is the number of individuals in species x , N_0 the number of individuals in the most common species, and r the exponential rate constant determined by the number of individuals in the most common and least common species for a given species-pool size. Several other models have been proposed to describe species-abundance distributions in nature (see Gotelli and Graves 1996 for a recent summary). However, surveys of tropical bird communities are subject to a variety of biases, such as the considerable differences in detectability between species (Karr 1971, 1981, 1990; Boulinier et al. 1998), and it is thus unlikely that the species-abundance distribution of rapid assessment data accurately reflects the sample community's true structure in most cases. We examined species-abundance distributions of 15 empirical data sets collected with the species-list method (see below) using chi-square tests. Three distributions were significantly different from the exponential decline and four were significantly different from the log normal model. Exponential decline was a better fit than log normal in nine cases and an equal fit in three cases. The log normal model was a better fit in only three cases. We thus used the exponential decline model as the underlying distribution in our simulations unless otherwise stated.

Species richness of model communities and the number of individuals in the most common and rarest species were user defined (and thereby the total number of individuals). Abundance of the latter was always set at 1 and that of the former was set to obtain a total of about 7,640 individuals, which is $4 \times$ the number of individuals found by Terborgh et al. (1990) to nest in 100 ha of Amazonian forest. Because Terborgh et al. (1990) determined that 26% of the nesting species had population densities of <1 pair per 100 ha, we assumed 400 ha to completely contain the home range of 0.5 pair of the lowest-density breeding species.

From the list of consecutive observations, the program compiled 20-species lists following MacKinnon and Phillips (1993): the first list consists of the first 20 species observed, the second list includes the next 20 species and may contain species already found on the first list, and so on. Original information on the abundances of species was maintained on the 20-species lists. Poulsen et al. (1997a) found that 20-species lists were too long in species-poor communities and suggested using 10-species lists. Therefore, we used list sizes of 5, 10, and 20 species. Different list sizes obviously signify different sample sizes, so a given number of m_1 -species lists is thus not directly comparable to the same number of m_2 -species lists. For the statistical analysis with ESTIMATES, each species list was treated as a separate sample.

To evaluate estimator performance, we varied the following parameters: (1) species richness (pools of 50, 100, 250, 500 species), (2) list size (5, 10, 20 species per list), and (3) number of lists (10, 20, 50, 100; 100 only for 5-species lists). For each combination of parameters, we computed 40 replications that were analyzed with nine richness estimators included in the program ESTIMATES 5.0.1 (Colwell 1997; see also Colwell and Coddington 1994): seven nonparametric statistics (ACE: Chao et al. 1993, Chazdon et al. 1998; ICE: Lee and Chao 1994, Chazdon et al. 1998; Chao 1: Chao 1984; Chao 2: Chao 1987; jackknife 1: Burnham and Overton 1978, 1979; Heltshel and Forrester 1983; Smith and van Belle 1984; jackknife 2: Burnham and Overton 1978, 1979; Smith and van Belle 1984; Palmer 1991; bootstrap: Smith and van Belle 1984) and two statistics that extrapolate species accumulation curves (MMRuns, MMMean: Raaijmakers 1987). ACE and Chao 1 are abundance-based estimators, whereas all other statistics are based on the incidence of species in samples. For details on all estimators, including equations of the seven nonparametric statistics, see Colwell (1997). MMRuns and MMMean are based on the Michaelis-Menten model (Raaijmakers 1987), which was evaluated recently by Keating and Quinn (1998). For comparison we also included S_{obs} , which is the "raw" species accumulation curve that itself is an estimator with a strong, negative bias (Colwell and Coddington 1994), in the analysis of estimator accuracy.

ESTIMATES parameters were set to the default values (50 randomized runs, random number seed 17, 10 incidence classes for ICE, 10 abundance classes for ACE). When compiling species accumulation curves and computing richness estimates, ESTIMATES randomizes the order of species lists, so the original sample order is irrelevant to all analyses.

To evaluate estimator biases with respect to community structure, we calculated richness estimates as described above with a uniform distribution of species (i.e. all species were equally abundant) for the following parameter combinations: (1) pools of 50,

100, 250, and 500 species; (2) lists of 5 and 20 species; and (3) 10, 20, and 50 lists. These results were then compared to those obtained from the exponential decline distribution. Although a uniform species-abundance distribution has not been documented for natural bird communities, determining estimator performance under such extreme, unrealistic conditions is nonetheless helpful for a general understanding of estimator behavior within the natural range of circumstances.

The accuracy of all richness estimates for both exponential decline and even distributions was calculated as the mean (\pm SD) of 40 replications for each parameter combination and estimator, and expressed as percentage of the predetermined total species richness.

To determine effectiveness of list size, sets of 400 consecutive observations were drawn from 50- and 500-species pools conforming to the exponential decline model and then subdivided into 5-, 10-, and 20-species lists. The resulting number of lists for each list size and species pool was averaged (\pm SD) over 20 replications.

EMPIRICAL DATA FOR ESTIMATION TESTS

Forty-five forest localities were surveyed by S.K.H. and M.K. in the Bolivian Andes and adjacent lowland areas with the species-list method totaling ~400 person days of data collection from 1995 to 2000. From that we derived our recommendations for standardizing the species-list method, which are detailed in the below. Data from two localities, the Botanical Garden of Santa Cruz de la Sierra (departamento Santa Cruz) and Cerro Asunta Pata/Calabatea (departamento La Paz), were analyzed here with the same nine estimators used on the computer-simulated data sets (see above) for a quantitative test of the method. The avifauna of both sites was inventoried intensively by a number of field workers and the resulting species lists were used for comparison with data obtained during rapid-assessment surveys using the species-list method.

Surveys were conducted from dawn to mid-day and often again from late afternoon until after dusk. While walking slowly and quietly along existing roads or trails and "randomly" through the habitat wherever feasible, all visual and acoustical observations of birds within 50 m of the observer (Schieck 1997) were recorded continuously, including the number of individuals per species for each encounter; observations of birds at camp sites also were recorded. Tape recordings of dawn choruses, mixed-species flocks, and individual birds were made to supplement visual observations and for later identification of unknown voices (Parker 1991; see also Haselmayer and Quinn 2000). The observer's rate of movement depended largely on the level of bird activity. When necessary, an hour or more was spent in

one spot to observe mixed-species flocks or high vocal activity at dawn. In such cases and during occasional resampling of the same trail area, repeated counts of obviously territorial individuals were avoided. The resulting master list of temporally consecutive observations from each site was then subdivided into 5-, 10-, and 20-species lists as outlined above. All tape recordings were treated like other observations and were integrated into the master list. Further details on the survey methodology are given in the below.

The Santa Cruz Botanical Garden (17°47'S, 63°04'W, 450 m elevation) was a remnant 187 ha fragment of tall lowland semideciduous forest and chaco thorn scrub surrounded by urban and agricultural areas ~13 km east of Santa Cruz city (Parker et al. 1993). It was inventoried over a 38 month period by constant-effort mist netting (six days per month from September 1995 to December 1997, three days per month from January to October 1998), nest searches, and opportunistic observations by S. E. Davis (pers. comm.) and coworkers of the Museo de Historia Natural Noel Kempff Mercado, Universidad Autónoma Gabriel René Moreno, Santa Cruz. Additionally, tape recordings and observations were made during repeated visits by S. Mayer and S.K.H. (see Mayer 2000, Herzog and Kessler 2002). The resulting cumulative species list for the area was provided by S. E. Davis (unpubl. data). The rapid assessment data analyzed here were obtained from 17 to 20 August 1999 by S.K.H. in an area of 45 ha including the two major habitat types. Because the cumulative locality list contained data from all seasons, we extracted, in collaboration with S. E. Davis, those species known to occur in the forest and thorn scrub area during the mid- to late dry season (July to September). Very rare species of uncertain status that probably represent vagrants were also excluded. The resulting list of 140 species was considered the area's actual species richness during the survey period and was used to calibrate results of the statistical analysis of our rapid assessment data.

Cerro Asunta Pata (15°03'S, 68°29', 850–1,500 m) and Calabatea (14°59'S, 68°28'W, 1,300–1,600 m) comprise an area of evergreen montane forest on the southwest and northeast side, respectively, of the Rio Yuyo (850 m) along a dirt road from Charazani to Apolo. The area's vegetation consisted of up to 30 m tall evergreen forest in a transition from lowland to montane elevational belts (Parker and Bailey 1991). Some small-scale forest clearing had occurred along the road, but in general the area was covered by pristine forest. Calabatea was surveyed (tape recording and observation) by T. A. Parker from 7 to 12 June 1990 (Parker and Bailey 1991, Parker et al. 1991). Cerro Asunta Pata was inventoried (specimen collection, tape recording) by a team of the Museum of Natural Science, Louisiana State University, Baton Rouge, from 11 July to 17 August 1993, totaling ~10,500 net-

meter days and 175 observer days (S. W. Cardiff and J. V. Remsen in litt.). Our rapid assessment data were collected from 31 May to 7 June, 13 to 16 June, and 20 to 23 June 1997 in an area of 130 ha at 1,000–1,500 m (Cerro Asunta Pata) and at 1,000–1,100 m on the northeast side of the Río Yuyo. Because all field work was conducted during the same season, we combined the species lists of the three expeditions. To ensure comparability with our data, we included only species that were found between 1,000 and 1,500 m by the other surveys. Because Parker and Bailey (1991) gave no lower elevational limits of species at Calabatea, these were inferred from published (e.g. Parker et al. 1996) and our own unpublished information on the elevational distribution of forest birds in Bolivia. The resulting list of 302 species was considered the actual species richness of the Asunta Pata–Calabatea area during the given season and was used to calibrate results of the statistical analysis of our rapid assessment data.

RESULTS

COMPUTER SIMULATIONS

Accuracy of estimators.—All estimators underestimated species richness in communities conforming to the exponential decline model for all but two parameter combinations (Appendix 1). MMRuns considerably overestimated species richness with 10 5-species lists in 500-species pools ($144.1 \pm 41.1\%$ accuracy), and jackknife 2 slightly overestimated species richness with 50 20-species lists in 50-species pools ($101.9 \pm 8.4\%$). For all other combinations, accuracy varied from as low as $12.0 \pm 0.2\%$ (bootstrap, 10 5-species lists) and $16.5 \pm 0.2\%$ (jackknife 1, 10 5-species lists) in 500-species pools to $98.7 \pm 6.3\%$ (jackknife 1, 50 20-species lists) and $98.6 \pm 7.8\%$ (Chao 2, 50 20-species lists) in 50-species pools (Appendix 1). Averaging all replications of all parameter combinations for each estimator, overall accuracy values were (1) MMRuns, $77.5 \pm 17.0\%$; (2) MMMean, $75.3 \pm 11.0\%$; (3) jackknife 2, $71.8 \pm 18.9\%$; (4) Chao 2, $71.1 \pm 14.1\%$; (5) Chao 1, $70.4 \pm 14.4\%$; (6) ICE, $68.8 \pm 12.1\%$; (7) ACE, $67.2 \pm 13.0\%$; (8) jackknife 1, $63.9 \pm 19.5\%$; (9) bootstrap $55.5 \pm 19.6\%$; and (10) S_{obs} , $48.6 \pm 19.5\%$.

Sensitivity to sample size.—Accuracy always increased in model communities with increasing sample size and thereby with the number of species observed for all estimators except MMRuns and MMMean (Appendix 1). For all list sizes, highest sensitivity to sample size was

observed in 500-species pools for bootstrap, jackknife 1, and jackknife 2 (e.g. 10-species lists: 2.6-, 2.4-, and 2.1-fold increase in accuracy, respectively, from 10 to 50 samples); those estimators also had the highest sensitivity to sample size in smaller pools. MMMean was the least sensitive estimator with respect to sample size. In 250- and 500-species pools, it had remarkably constant accuracy nearly independent of sample size or slightly decreasing accuracy with sample size (except for 20-species lists in 250-species pools, where accuracy increased slightly with sample size). MMRuns performed similarly well except for 5-species lists in 500-species pools, where it was highly sensitive to sample size (Appendix 1).

Sensitivity to species-pool size.—All estimators were sensitive to species richness in model communities. In 50- and 100-species pools, highest accuracy generally was obtained by jackknife 2, Chao 2, and Chao 1, respectively (Appendix 1); overall accuracy (all list and sample sizes combined) of those estimators was: $82.5 \pm 14.8\%$ (50-species pools) and $75.5 \pm 14.5\%$ (100-species pools) for jackknife 2; $82.2 \pm 14.6\%$ (50-species pools) and $71.6 \pm 13.6\%$ (100-species pools) for Chao 2; and $82.3 \pm 14.6\%$ (50-species pools) and $71.0 \pm 13.3\%$ (100-species pools) for Chao 1. Lowest overall accuracy was obtained by MMMean ($69.0 \pm 10.3\%$), MMRuns ($69.1 \pm 10.2\%$), and bootstrap ($69.9 \pm 15.0\%$) in 50-species pools and by bootstrap ($61.0 \pm 14.6\%$) in 100-species pools. In 250-species pools, highest accuracy was calculated by jackknife 2 for large sample sizes, but due to that estimator's high sensitivity to sample size, its overall accuracy was only $68.6 \pm 16.9\%$, whereas MMRuns ($77.3 \pm 5.9\%$) and MMMean ($75.9 \pm 4.7\%$) had highest and bootstrap ($50.6 \pm 16.8\%$) had lowest overall accuracy. In 500-species pools, MMRuns ($95.6 \pm 22.0\%$) and MMMean ($88.7 \pm 6.0\%$) had highest overall accuracy and bootstrap had lowest overall accuracy ($40.6 \pm 18.2\%$). It should be noted here that, when measuring α -diversity, pools of 500 species are much more species-rich than any natural bird community. Therefore, estimator performance in 500-species pools is considerably less important than the performance in smaller species pools.

Both MM estimators had a tendency of increased overall accuracy with increasing species richness (although that was not the case for

empirical data, see below), whereas that pattern was reversed in the remaining estimators. Comparing mean overall accuracy for each of the four species pools, ICE was the least sensitive to species richness (10.8% difference between highest and lowest mean overall accuracy), followed by ACE (13.2%), Chao 2 (17.0%), Chao 1 (18.9%), MMMean (21.2%), jackknife 2 (21.9%), jackknife 1 (26.1%), MMRuns (27.7%), and bootstrap (29.3%).

Sensitivity to community structure.—All estimators consistently computed higher estimates for pools with equally abundant species (Appendix 2) than for pools conforming to the exponential decline model (Appendix 1). MMRuns and MMMean overestimated species richness considerably for most parameter combinations, especially in 250- and 500-species pools, but overestimation decreased with increasing sample size. Contrarily, with 20-species lists Chao 1 and ACE (and to a lesser degree Chao 2 and ICE) produced highly accurate estimates with little or no sensitivity to sample size, which contrasted with the performance of those estimators in simulated exponential decline communities. Jackknife 1 and 2 and bootstrap performed well in 50-species pools, but were moderately to highly sensitive to sample size in the remaining species pools. Thus, all estimators were sensitive to differences in community structure. That was most pronounced in estimates of MMRuns and MMMean, which, for the same parameter combination, were as much as 3.5 and 3.1 \times higher, respectively, in evenly distributed species pools (10 5-species lists in 250-species pools for MMRuns, 10 5-species lists in 100-species pools for MMMean).

Effectiveness of list size.—For 50-species pools, 68.2 \pm 2.6 5-species lists, 25.9 \pm 1.1 10-species lists, and 6.7 \pm 0.7 20-species lists were compiled from 400 consecutive observations. For 500-species pools, the respective values were 79.4 \pm 0.3, 39.3 \pm 0.2, and 19.2 \pm 0.1. Because 100 5-species lists achieved accuracies very similar to those of 50 10-species lists and 10 or 20 20-species lists for any given species pool and estimator (Appendix 1), 5-species lists made the most effective use of the raw data without being more biased than either 10- or 20-species lists when excluding the extremely small sample size of 10 5-species lists.

Figure 1 illustrates estimator performance for 10-species lists. An ideal estimator would

have a level surface with accuracy close to 100%. Surfaces of bootstrap, jackknife 1, and jackknife 2 closely paralleled that of S_{obs} but at higher accuracy levels and, for jackknife 2, with less sensitivity to species richness, illustrating the high sensitivity to sample size of those estimators. The similar surfaces of ICE, ACE, Chao 1, and Chao 2 had their largest deviation from the ideal estimator along the transition from small to moderately species-rich communities. MMRuns and MMMean had the perhaps most level surfaces, but with a pronounced tendency towards low accuracy in small species pools for small sample size.

EMPIRICAL DATA

At Santa Cruz, 91 species (65.0% of the actual species richness) were observed and 85 5-species lists, 41 10-species lists, and 18 20-species lists were compiled. At Asunta Pata, 224 species (74.2% of the actual species richness) were recorded and 324 5-species lists, 157 10-species lists, and 74 20-species lists were compiled. Jackknife 2 overestimated richness by 6.3% at Asunta Pata, but all other estimators underestimated actual species richness to varying degrees (Table 1). Overall accuracy (mean \pm SD of all estimates from Table 1 for each statistic) was (1) MMRuns, 76.6 \pm 9.1%; (2) jackknife, 2 76.3 \pm 20.8%; (3) MMMean, 74.6 \pm 8.8%; (4) Chao, 2 71.6 \pm 15.5%; (5) ICE, 70.4 \pm 14.0%; (6) jackknife, 1 67.1 \pm 20.3%; (7) ACE, 59.5 \pm 17.0%; (8) Chao, 1 59.4 \pm 17.8%; (9) bootstrap, 57.7 \pm 19.4%; and (10) S_{obs} , 49.5 \pm 18.2%.

Figures 2 and 3 illustrate the performance of seven estimators for 10-species lists. Curves of MMRuns (except for an initial spike) and Chao 1 resembled those of MMMean and ACE, respectively, and are not shown. ACE, jackknife 1 and 2, and bootstrap mostly paralleled the S_{obs} curve (Figs. 2A and 3A), illustrating their high sensitivity to sample size, which was invariably higher at Asunta Pata. Here, both jackknife statistics and bootstrap performed worse than S_{obs} , increasing in accuracy from 10 lists to maximum sample size by >65%, >55%, and >40% with 5-, 10-, and 20-species lists, respectively (Table 1). Those estimators also were the most sensitive at Santa Cruz with only slightly better performance than S_{obs} . MMRuns and MMMean were by far the least sensitive to sample size at both sites; for 5-species lists at Santa Cruz, ac-

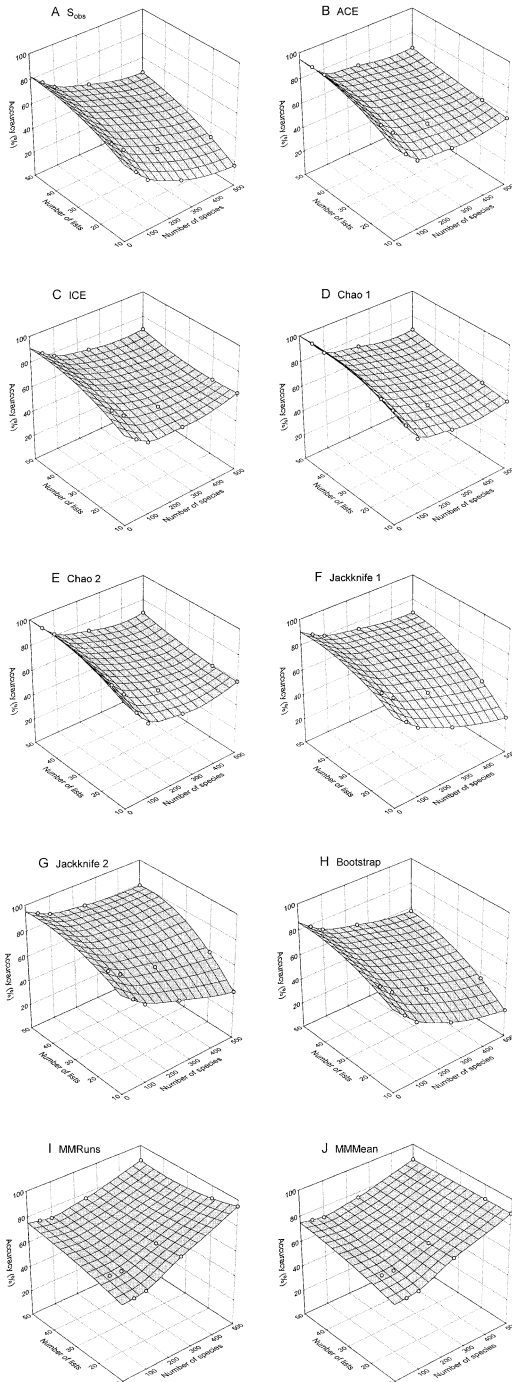


FIG. 1. Performance of (A) S_{obs} and (B–J) nine statistics in estimating species richness using the MacKinnon and Phillips (1993) species-list approach (10 species per list) for four computer-simulated species pools (50, 100, 250, 500 species) with species-abundance distributions conforming to the exponen-

curacy decreased by 13.6 and 5.0%, respectively, from 10 to 85 lists. At Asunta Pata, MMRuns performed similarly with 5-species lists, except that after dropping to a minimum at intermediate sample size accuracy increased again, whereas MMMean accuracy gradually increased with sample size (Table 1). With 10- and 20-species lists, both MM estimators were highly insensitive to sample size at Santa Cruz (see Fig. 2B for MMMean); at Asunta Pata, however, accuracy increased by approximately 20–25% from 10 lists to maximum sample size (see Fig. 3B for MMMean). ICE and Chao 2 performed better than their respective abundance-based counterpart, but both had inconsistent estimates at small sample size, were overall more sensitive to sample size than MMMean, and ICE overestimated species richness after two and three lists at Santa Cruz (Figs. 2B and 3B).

The MM statistics outperformed all other estimators also with respect to sensitivity to species richness, computing nearly identical accuracy at both sites for maximum sample size, whereas the remaining estimators had considerably higher final accuracy at Asunta Pata (Table 1). Contrarily, for small and medium sample size, all estimators computed higher accuracy at Santa Cruz. Averaging the three maximum sample size accuracy values for each site and estimator, ICE (19.9% higher accuracy at Asunta Pata), jackknife 2 (18.9%), and ACE (18.2%) had highest sensitivity to species richness. List size had little or no influence on predicted species richness for maximum sample size except in Chao 2 and jackknife 2 at Santa Cruz. Here, 10-species lists resulted in 5.7 and 4.3% higher accuracy, respectively, than 5-species lists.

Standard deviations of richness estimates were computed by ESTIMATES only for S_{obs} and five estimators (Table 1). Standard deviations mostly decreased with increasing sample size, and S_{obs} , ACE, and ICE returned no standard deviation at maximum sample size. S_{obs} and jackknife 1 computed low values, whereas

tial decline model. Accuracy values are percentage of the predetermined species richness. Open circles represent raw data points (mean of 40 replications with 50 randomizations each). Surfaces were fitted by distance-weighted least-squares smoothing.

those of Chao 2 were highest. ICE and Chao 2 had higher standard deviations than their respective abundance-based counterpart.

We consider MMMean the overall most robust estimator (see below), but two contrasting patterns in its performance for simulated versus empirical data require further analysis. First, accuracy steadily increased with sample size in small species pools but quickly reached an asymptote in large species pools for simulated data sets (Fig. 1J, Appendix 1). For empirical data, the situation was reversed: accuracy quickly reached an asymptote at Santa Cruz (Fig. 2B), but still ascended at maximum sample size at Asunta Pata (Fig. 3B). Second, accuracy increased with species-pool size for simulated data (Fig. 1J, Appendix 1) and hence tended to inflate true differences in species richness. For empirical data, respective estimates were lower for the more species-rich data set for small to intermediate sample size; for maximum sample size, however, MMMean had a slightly higher accuracy at Asunta Pata than at Santa Cruz. Because accuracy still increased at maximum sample size (Fig. 3B), continued surveying at Asunta Pata probably would have resulted in a still higher final accuracy value.

Such contrasts in performance could be related to differences in community structure. The analysis of simulated pools with evenly abundant species shows that all statistics, and especially the MM estimators, are sensitive to the data set's underlying species distribution. Because MMMean estimates are based on the incidence of species in samples, we plotted rank-frequency curves for 10-species lists for three simulated and both empirical data sets (Fig. 4A, B). Community structure at Santa Cruz was intermediate between the structure of model communities of 100 and 250 species, but it approximated that of 250-species pools more than that of 100-species pools (Fig. 4A). Because estimates quickly reached an asymptote for both Santa Cruz and 250-species pools, it could be concluded that MMMean is sensitive to sample size when data sets have many high-frequency and few low-frequency species, such as in 100-species pools (Fig. 4A). However, the Asunta Pata curve (Fig. 4B) closely matched that of 250-species pools, except for a longer "tail" of low-frequency species, but despite that similarity, Asunta Pata estimates did not reach an asymptote quickly. A long tail of low-

frequency species alone does not appear to cause the poorer estimator performance because such a tail also characterized the 500-species-pool curve (Fig. 4B).

DISCUSSION

STANDARDIZATION OF THE SPECIES-LIST METHOD

The species-list method is a useful technique for rapid assessments of species richness in tropical bird communities. Fjeldså (1999) used a similar random walk approach in humid montane forests of Tanzania and found highly significant correlations between random walk and point-count data sets and concluded that random walking does not appear to be more biased than other observational methods. Its main advantages are time efficiency (the entire daylight period may be used to generate data) and relative observer independence compared to any timed species-count method (e.g. point counts; Fjeldså 1999).

We agree with Poulsen et al. (1997a) that the method is suitable for appraising (1) the magnitude of the species richness, (2) when a locality has been adequately sampled, and (3) for determining the relative abundance of each species. However, comparisons of relative abundances should be made only within species (e.g. across sites or seasons) because across-species comparisons are hampered by considerable differences in detectability between species (Karr 1971, 1981, 1990; Boulinier et al. 1998). For the same reason, opposed to Poulsen et al. (1997a), we consider data gathered with the species-list method as unsuitable for calculating indices of α -diversity. As a cautionary note, we recommend further testing of the accuracy of relative abundance estimates obtained with the species-list method in comparison with point-count and spot-mapping data especially in Neotropical forests, and we consider that only spot-mapping (Kendeigh 1944, Bibby et al. 2000) during longer field sessions will yield reliable measurements of absolute abundance for most tropical forest birds (see Terborgh et al. 1990, Robinson et al. 2000).

A number of recommendations for the standardization of the species-list method are suggested here (see also above). The method is by no means "fool proof" and a certain level of experience with visual and vocal identification of

TABLE 1. Observed (S_{obs}) and estimated species richness calculated by nine statistics for two Bolivian bird communities surveyed with the MacKinnon and Phillips (1993) species-list approach. Values are percentage (\pm SD) of the actual species richness at each site.

		Estimator				
Number of lists	Santa Cruz (140 species)			Asunta Pata (302 species)		
	5-species lists	10-species lists	20-species lists	5-species lists	10-species lists	20-species lists
S_{obs}						
10	25.7 \pm 1.4	39.3 \pm 2.1	55.0 \pm 2.1	13.2 \pm 0.7	21.9 \pm 1.3	33.8 \pm 2.3
20	39.3 \pm 2.1	52.9 \pm 2.1	—	21.9 \pm 1.3	32.8 \pm 1.7	46.7 \pm 2.3
50	56.4 \pm 1.4	—	—	36.8 \pm 2.0	50.0 \pm 2.0	65.6 \pm 1.7
100	—	—	—	50.0 \pm 2.0	64.6 \pm 1.7	—
Max ^a	65.0 \pm 0.0	65.0 \pm 0.0	65.0 \pm 0.0	74.2 \pm 0.0	74.2 \pm 0.0	74.2 \pm 0.0
ACE						
10	37.9 \pm 7.1	47.9 \pm 4.3	60.0 \pm 3.6	26.2 \pm 6.3	36.1 \pm 6.6	46.4 \pm 5.3
20	49.3 \pm 6.4	57.9 \pm 3.6	—	35.4 \pm 5.6	46.4 \pm 5.3	58.3 \pm 4.6
50	61.4 \pm 2.1	—	—	49.7 \pm 4.6	61.9 \pm 4.0	77.8 \pm 3.0
100	—	—	—	61.6 \pm 3.6	76.8 \pm 3.0	—
Max ^a	67.9 \pm 0.0	67.9 \pm 0.0	67.9 \pm 0.0	86.1 \pm 0.0	86.1 \pm 0.0	86.1 \pm 0.0
ICE						
10	62.1 \pm 15.7	62.1 \pm 8.6	69.3 \pm 5.7	47.4 \pm 17.9	51.0 \pm 11.3	59.6 \pm 9.3
20	62.9 \pm 10.7	66.4 \pm 5.7	—	47.4 \pm 9.9	58.3 \pm 8.9	70.2 \pm 6.6
50	68.6 \pm 3.6	—	—	59.9 \pm 8.3	71.2 \pm 6.3	86.8 \pm 4.3
100	—	—	—	70.5 \pm 5.0	84.4 \pm 4.0	—
Max ^a	73.6 \pm 0.0	75.0 \pm 0.0	75.7 \pm 0.0	94.7 \pm 0.0	94.4 \pm 0.0	95.0 \pm 0.0
Chao 1						
10	35.7 \pm 6.4	47.1 \pm 5.0	61.4 \pm 4.3	25.8 \pm 7.3	33.8 \pm 5.3	44.4 \pm 4.3
20	47.9 \pm 5.0	57.9 \pm 3.6	—	33.1 \pm 5.0	45.0 \pm 5.0	59.6 \pm 5.0
50	61.4 \pm 3.6	—	—	49.0 \pm 5.0	62.6 \pm 5.0	78.5 \pm 4.6
100	—	—	—	62.9 \pm 5.0	77.2 \pm 4.6	—
Max ^a	69.3 \pm 2.9	69.3 \pm 2.9	69.3 \pm 2.9	85.8 \pm 4.3	85.8 \pm 4.3	85.8 \pm 4.3
Chao 2						
10	60.0 \pm 19.3	60.7 \pm 10.0	70.7 \pm 7.9	45.7 \pm 18.5	50.3 \pm 12.3	57.9 \pm 8.6
20	64.3 \pm 11.4	67.0 \pm 7.1	—	47.0 \pm 10.6	56.6 \pm 8.6	71.5 \pm 8.3
50	70.7 \pm 7.1	—	—	58.6 \pm 7.6	72.5 \pm 7.6	89.1 \pm 7.6
100	—	—	—	71.9 \pm 7.3	86.8 \pm 7.3	—
Max ^a	76.4 \pm 6.4	82.1 \pm 9.3	80.7 \pm 8.6	96.7 \pm 7.3	96.7 \pm 7.3	96.7 \pm 7.3
Jackknife 1						
10	42.1 \pm 2.1	57.9 \pm 2.9	72.1 \pm 3.6	22.8 \pm 0.7	35.1 \pm 1.7	50.3 \pm 2.0
20	59.3 \pm 2.9	69.3 \pm 3.6	—	35.8 \pm 1.3	50.0 \pm 2.3	65.6 \pm 2.6
50	73.6 \pm 3.6	—	—	54.6 \pm 2.0	69.2 \pm 2.6	86.8 \pm 3.0
100	—	—	—	69.5 \pm 2.3	85.1 \pm 2.6	—
Max ^a	79.3 \pm 2.9	80.7 \pm 3.6	80.0 \pm 3.6	95.0 \pm 2.6	95.0 \pm 2.6	95.4 \pm 3.0
Jackknife 2						
10	52.9	66.4	77.9	29.8	44.0	59.6
20	70.0	75.0	—	44.7	59.9	76.5
50	79.3	—	—	64.6	79.5	98.0
100	—	—	—	79.8	96.0	—
Max ^a	85.0	89.3	87.9	106.3	106.3	106.3
Bootstrap						
10	32.9	47.9	62.9	17.2	27.5	41.1
20	48.6	67.9	—	27.5	40.4	55.0
50	64.3	—	—	44.7	58.6	75.2
100	—	—	—	58.9	73.8	—
Max ^a	71.4	72.1	72.1	83.8	83.4	83.8
MMRuns						
10	94.3	80.7	80.7	89.4	59.3	62.3
20	87.9	80.7	—	66.9	62.9	67.5

TABLE 1. Continued.

Number of lists	Estimator					
	Santa Cruz (140 species)			Asunta Pata (302 species)		
	5-species lists	10-species lists	20-species lists	5-species lists	10-species lists	20-species lists
50	81.4	—	—	65.2	69.9	77.8
100	—	—	—	69.9	76.8	—
Max ^a	80.7	80.7	80.7	82.1	82.5	82.8
	MMMean					
10	85.7	79.3	80.0	62.3	57.3	61.3
20	85.0	80.7	—	64.2	61.9	67.5
50	81.4	—	—	64.9	69.5	77.5
100	—	—	—	69.9	76.8	—
Max ^a	80.7	80.7	80.7	82.1	82.5	82.8

^a Maximum number of lists compiled: Santa Cruz 85 5-species, 41 10-species, 18 20-species lists; Asunta Pata 324 5-species, 157 10-species, 74 20-species lists.

most species present is necessary. Data collected by an observer largely unfamiliar with a given avifauna are not comparable with those obtained by experienced observers. It is crucial that names are assigned to species not confidently identified (by sight or voice) at first (Poulsen et al. 1997a). The extensive use of a tape recorder is indispensable (Parker 1991); recordings should be made of dawn choruses, mixed-species flocks, individual birds, and all unfamiliar vocalizations and integrated into the consecutive master list by an expert after a thorough revision using reference recordings. Five-, 10-, or 20-species lists should be compiled only during data analysis to ensure flexibility (see Poulsen et al. 1997a). Because not all species are active at the same time of day, surveys should cover most of the daylight period as well as dawn, dusk, and early evening. Mixed-species flock encounters should be treated like other observations and detected individuals should be recorded consecutively, although especially in larger flocks the number of conspecific flock members may not be assessable for some species at first. In such cases, we suggest assigning additional individuals to the same species list as the first individual of that species.

Because the detectability of most forest bird vocalizations drops considerably at distances of >50 m (Schieck 1997), observations beyond that distance should be excluded from the analysis. In our experience, this results only very occasionally in a rare but loud species having to be dropped entirely from the analysis because it never was recorded within 50 m of the observer, but it avoids filling species lists with

common, noisy birds (e.g. Screaming Piha [*Liopaugus vociferans*]) and overestimating their relative abundance. If longer time periods are spent in one spot or when resampling a given section of the study area, repeated counts of obviously territorial individuals should be avoided, because this also tends to overestimate relative abundances of those species. In sexually dimorphic species such as many antbirds, the male and the female in a given territory may each be counted once. Obviously, it will occasionally be difficult to determine whether a territorial bird has already been counted, and we have no ready solution to that problem except not to survey sections of a study area more than once, but in most cases that will be impractical or even impossible. Because activity and detectability levels of most species show diurnal variations (Blake 2000), resampling should ideally be carried out at a different time of day than previous surveys in the same section to minimize probability of encountering the same individuals more than once. When in doubt, it might be best to adopt a conservative approach and omit a given observation from the analysis.

The size and, in mountains, elevational range of survey areas should be held constant or, if difficult logistically, should at least be quantified; the same applies to habitat diversity (Remsen 1994, Cohn-Haft et al. 1997). If, for example, species richness in Bolivian dry forests is to be assessed, data from nonzonal vegetation types (e.g. gallery forest) must be excluded because nonzonal habitats tend to be represented unevenly at different sites and thus influence total species richness to varying degrees (Herzog and Kessler 2002). The dis-

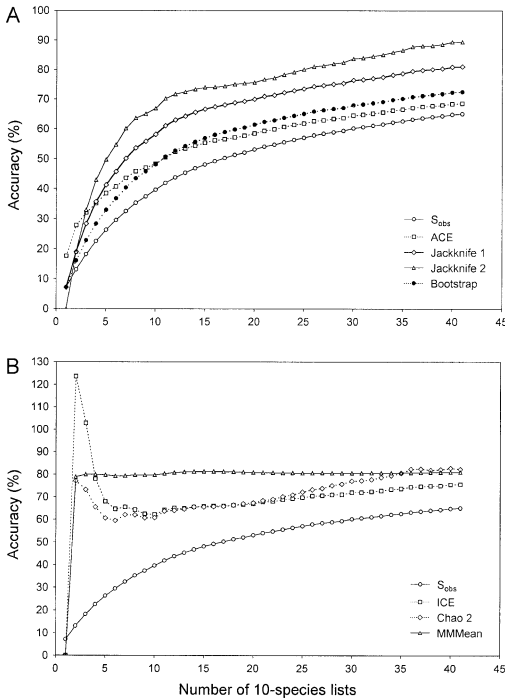


FIG. 2. Performance of seven estimators of species richness for an empirical bird data set collected with the species-list method at the Botanical Garden of Santa Cruz de la Sierra, Bolivia. (A) ACE, jackknife 1, jackknife 2, bootstrap. (B) ICE, Chao 2, MMMean. Curves of MMRuns (except for an initial spike) and Chao 1 resembled those of MMMean and ACE, respectively, and are not shown. Values are percentage of the total species richness (140) at the site. The species accumulation curve (S_{obs}) and estimator curves indicate accuracy as a function of the number of 10-species lists. Sample accumulation order of all curves was randomized 50 times, and each point represents the mean of the resulting 50 estimates.

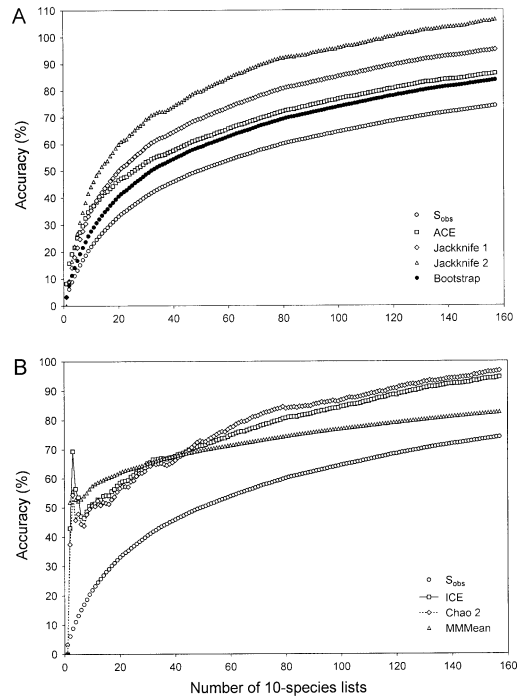


FIG. 3. Performance of seven estimators of species richness for an empirical bird data set collected with the species-list method at Cerro Asunta Pata, Bolivia. (A) ACE, jackknife 1, jackknife 2, bootstrap. (B) ICE, Chao 2, MMMean. Curves of MMRuns (except for an initial spike) and Chao 1 resembled those of MMMean and ACE, respectively, and are not shown. Values are percentage of the total species richness (302) at the site. The species accumulation curve (S_{obs}) and estimator curves indicate accuracy as a function of the number of 10-species lists. Sample accumulation order of all curves was randomized 50 times, and each point represents the mean of the resulting 50 estimates.

inction of the avifauna into core and noncore species (Remsen 1994) is ideal but often not achievable to a satisfactory degree unless an observer is very familiar with a given habitat or area. However, if survey durations are similar, proportions of noncore species should also be comparable. Obvious noncore species (e.g. a heron in flight over dry forest) should always be excluded.

To standardize survey effort and to determine equal stopping points for data collection, we suggest using the Chao 1 estimator in the field by comparing estimated with observed species richness. That should be done every evening until the observed species richness is

>90% of the respective Chao 1 estimate. Although not the most robust estimator (see below), Chao 1 has a practical advantage over all other estimators: it can be applied directly to the raw data without the time consuming subdivision of observations into species lists, and it is so simple that it can be calculated by hand. For practical reasons, we include the Chao 1 formula:

$$S_{Chao1} = S_{obs} + F_1^2 / 2F_2 \quad (2)$$

where S_{obs} is the number of species observed, F_1 the number of singletons (species with only one individual), and F_2 the number of doubletons (species with exactly two individuals) in the

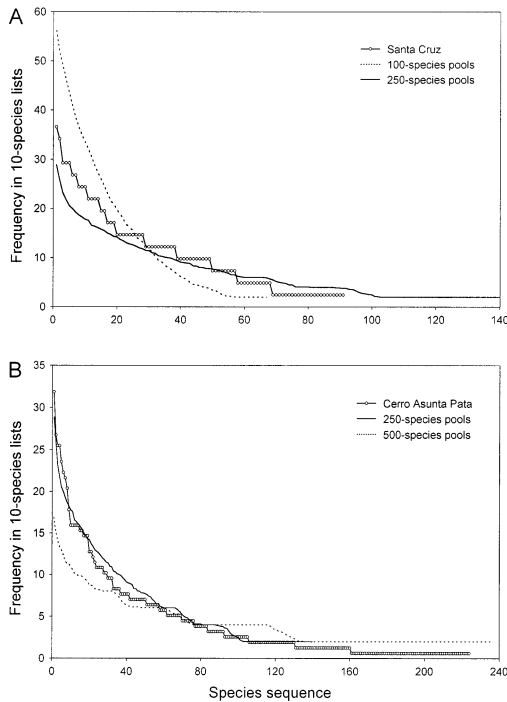


FIG. 4. Frequency-distribution graphs for 10-species lists for simulated (100-, 250-, 500-species pools, $n = 50$ lists per pool) and empirical (Santa Cruz, $n = 41$; Asunta Pata, $n = 157$) data sets. Simulated data sets were drawn from a species-abundance distribution conforming to the exponential decline model, and each point represents the mean of 40 replications. (A) Community structure and species richness at Santa Cruz is intermediate between that of simulated 100- and 250-species pools. (B) Community structure at Asunta Pata is similar to that of simulated 250-species pools except for a longer "tail" of low-frequency species.

data set (Chao 1984, Chazdon et al. 1998). This simple measure can also be applied to other count and census methods (e.g. point counts) to monitor the completeness of a survey.

Finally, the species-list method is entirely compatible with point counts (or line transects) if point-count observations are recorded in the same consecutive order as in the species-list approach. Thus, where logistical conditions permit, experienced observers familiar with a given bird community could use both methods in conjunction, that is, record point-count data consecutively to allow subdivision into species lists and use the species-list approach between point-count intervals. That method combination would gain the advantages of rigid stan-

dardization of point counts while still maintaining the flexibility of species lists.

USE OF ESTIMATION METHODS

Performance varied considerably between the nine estimators and between simulated and empirical data sets for most estimators. Bootstrap, jackknife 1, and jackknife 2 performed so poorly with simulated (Fig. 1, Appendices 1 and 2) and empirical (Table 1, Figs. 2 and 3) data sets that they will not be considered any further. The remaining estimators basically fall into two groups: the first includes the nonparametric statistics ACE, ICE, Chao 1, and Chao 2 (Fig. 1B–E), and the second includes the two estimators based on the Michaelis-Menten model (Fig. 1I–J). With simulated exponential decline and empirical data sets, all estimators of the first group had lower overall accuracy than the MM estimators. However, accuracy as such is not the key parameter because an estimator with low but constant accuracy regardless of variation in species richness or sample size would be easy to calibrate. Thus, sensitivity to both sample size and species richness are more crucial.

With respect to sample size, MMEan and—when excluding the smallest sample size in 500-species pools—MMRuns again performed better than either ACE, ICE, Chao 1, or Chao 2 with both simulated exponential decline (Fig. 1, Appendix 1) and empirical (Table 1, Figs. 2 and 3) data sets. In contrast, regarding sensitivity to species richness, ACE, ICE, Chao 1, and Chao 2 obtained better results for simulated data than both MM estimators, and MMEan performed better than MMRuns. However, the MM estimators clearly outperformed all other statistics with empirical data (Table 1, Figs. 2 and 3).

Thus, with simulated exponential decline and empirical data sets MMEan and MMRuns were the overall most robust of the nine estimators tested here. Although both generally obtained very similar results, MMRuns was more sensitive to species richness with simulated data than MMEan and also tended to be more sensitive to sample size for 5-species lists with species-rich data sets (Appendix 1). We therefore consider MMEan as the overall least biased, preferable estimator, despite its poor performance in model communities with equally

abundant species. Although that indicates that the Michaelis-Menten model is more sensitive to inherent differences in abundance or frequency distributions than the other estimators tested here, an absolutely even community is an extreme case that has not been documented for birds. Our simulated exponential decline communities and empirical data sets covered a relatively wide range of degrees of evenness (Fig. 4) likely to be found in rapid assessment data sets, under which MMMean performed best. Nonetheless, species-frequency distributions should be examined before statistical analysis, and if nearly uniform structures are found, MMMean should not be used. ICE and Chao 2 both had very similar performances (Figs. 1C and E, 2B, and 3B) and were somewhat more accurate and less biased than their abundance-based counterparts (ACE and Chao 1, respectively). Because Chao 2 also was recommended by other studies (see below), we consider it the best alternative to MMMean. However, the greater sensitivity to sample size of Chao 2 compared to MMMean (Figs. 2B and 3B, Appendix 1) clearly is a disadvantage, and it could be objected that the sampling protocol of the species-list method violates that estimator's assumption of random incidence samples (Chao 1987, R. K. Colwell in litt.).

The good performance of MMMean is in disagreement with the few other comparative studies that have tested all or most of the estimators considered here. For simulated and five empirical data sets of parasite communities, Walther and Morand (1998) found that Chao 2 and jackknife 1 were the overall most accurate and least biased estimators. In that same study, MMMean obtained good results for two empirical data sets, but its overall performance was intermediate, and MMRuns performed badly. On the basis of analyses of empirical seed bank and tree seedling data sets, Colwell and Coddington (1994) and Chazdon et al. (1998) also recommended Chao 2. Walther and Martin (2001) tested 17 estimation methods on an empirical bird data set and determined Chao 2 and Chao 1 as the overall least biased and most precise estimators with MMMean returning only intermediate bias and precision values. However, all of those studies used species-poor data sets (from three to 40 species) based in at least some cases on different community structures than employed here, which

may largely account for differences in estimator performance. That indicates that no single estimator may work well for all taxonomic groups due to inherent differences in community structure.

At present, we are unable to fully explain variations in the performance of MMMean with respect to species-frequency distributions. Keating and Quinn (1998) also found that the Michaelis-Menten model was not robust to differences in community structure: it performed well for at least moderately large (≥ 100 species) communities conforming to the broken-stick model, but it yielded poor estimates for less even community structures resembling the random-fraction model. Keating and Quinn (1998) further concluded "that the Michaelis-Menten model implicitly assumes a highly even community structure." That is supported by our results for exponential decline communities: species-frequency distributions became increasingly even (Fig. 4A, B; data for 50-species pools not shown) and estimator performance improved (Fig. 1J, Appendix 1) with species-pool size. The substantial overestimation of MMMean for species-rich simulated communities with a uniform species-abundance distribution (Appendix 2) suggests that an evenness threshold exists where the estimator computes highly accurate values. However, the relatively poor performance for the Asunta Pata data remains unexplained.

Colwell (1997) briefly discussed the occasionally erratic performance of MMRuns for small sample size that may occur when some samples are much more species-rich than others. However, for any given analysis in this article, all samples contained the same number of species (either 5, 10, or 20). Thus, an initial spike in the MMRuns curve probably occurred when the first lists drawn during a given randomization had little or no overlap in species composition, a likely scenario when using small lists in species-rich communities. MMMean circumvents that problem by computing estimates only once for the mean accumulation curve, whereas MMRuns averages estimates over all randomizations (Colwell 1997).

Despite the lack of a consistently good performance of MMMean with species-poor data sets (Colwell and Coddington 1994, Chazdon et al. 1998, Walther and Morand 1998, Walther and Martin 2001) and unevenly structured

communities (Keating and Quinn 1998), on the basis of our results we nonetheless consider it a useful and promising technique for analyzing data gathered in diverse tropical bird communities with the species-list method. This approach to assessing species richness is doubtlessly much more standardized and valuable than simple, one-dimensional locality lists accompanied by some measure of survey effort (e.g. see Remsen 1994, Cohn-Haft et al. 1997). The lack of a standard deviation, however, may be problematic when absolute values of estimated species richness have to be compared directly between sites or between surveys of the same site (e.g. before and after selective logging). In this case, Chao 2 may be used provided that sample size is similar.

However, two basic aspects need to be considered to ensure comparability of results. First, when comparing MMMean (or Chao 2) estimates between sites, a close inspection of curve shapes is required. If some or all curves do not reach an asymptote quickly (i.e. after 10 to 15 lists), a standardized cut-off point must be determined to control for the confounding effects of sampling effort (see Gotelli and Colwell 2001). Rather than taking the estimated richness after a given number of individuals sampled as proposed by Gotelli and Colwell (2001; sample-based rarefaction with a rescaled x -axis from samples to individuals), we suggest determining the cut-off point from the relation between the S_{obs} and the MMMean curve by expressing each S_{obs} value as the proportion of the respective MMMean value. Our Santa Cruz data set contained 41 10-species lists (Table 1), and after 41 lists the S_{obs} value constituted 80.2% of the MMMean value. At Asunta Pata, the equivalent cut-off point is found after 78 10-species lists, where the S_{obs} value constituted 80.1% of the MMMean value. The respective richness estimates were 113 species at Santa Cruz and 224 at Asunta Pata, or 80.7 and 74.2% of each site's actual species richness, respectively. Although a smaller difference between the two values is desirable (ideally they should be equal), it is lower than most Chao 2 standard deviations for the empirical data (Table 1).

In this particular case, determining the cut-off point based on a given number of individuals after rescaling the x -axis from samples to individuals (Gotelli and Colwell 2001) obtains

a nearly identical result (estimates of 113 species at Santa Cruz and 223 species at Asunta Pata). However, tropical habitats often contain one to several species occurring in large numbers in intraspecific aggregations, such as flocks of parakeets or swifts. The presence of such species can lead to a rapid and dramatic increase in the number of individuals observed, which in turn will bias the rarefaction based on a rescaled x -axis causing an artificial inflation of survey effort. Preliminary analyses indicate that the more abundant the most common species, the lower will be the estimate obtained by individual-based rarefaction when compared to rarefaction based on the relation between the S_{obs} and the MMMean curve, and that relationship appears to be significant.

Second, we strongly advocate the use of a standard list size unless it can be shown that that leads to more biased results than other list sizes for a given data set. In species-poor habitats, 20-species lists often are too long (Poulsen et al. 1997a), and sample size can be very small with 20-species lists when the size of the survey area is limited (S. K. Herzog pers. obs.). Estimator curves for 5-species lists had several basic shapes (asymptotic, gradual increase, minimum at intermediate sample size, initial spike), whereas curves based on 10-species lists were less variable. Ten-species lists therefore appear to be a good intermediate solution.

Conclusion.—We believe that with a minimum degree of standardization and careful data analysis, the species-list method and the quantitative comparison of the resulting species accumulation curves are useful tools for both conservation assessment and the study of avian species richness patterns in the tropics. It can be argued that it simply is the nature of estimation and extrapolation that obtained values are highly speculative and that comparisons of estimates are not reliable. From a conservation viewpoint it must be emphasized, though, that complete avifaunal inventories in the tropics simply cannot be conducted as often as they are required and that standardized survey methods employed in temperate regions are often difficult or impossible to apply. Rapid assessments are a necessity dictated by the urge for conservation action and by limitations of time, personnel, and funding. Here, the use of statistical estimators will certainly aid in comparing

different studies and in making more informed conservation decisions.

ACKNOWLEDGMENTS

This paper benefited from discussions with R. K. Colwell, B. Fisher, J. Fjeldsá, R. MacLeod, C. Rahbek, J. Umbanhowar, and B. A. Walther. We are grateful to S. W. Cardiff, S. E. Davis, and J. V. Remsen for sharing their unpublished data, to S. Mayer and B. M. Whitney for help with identification of "mystery" tape recordings, and to R. K. Colwell for making ESTIMATES available to a wide audience. S.K.H. would like to thank F. Bairlein for guidance and support. F. Bairlein, J. Fjeldsá, R. MacLeod, P. C. Stouffer, and an anonymous reviewer made valuable comments on earlier drafts of this paper. Financial support was provided to S.K.H. by the German Academic Exchange Service (DAAD, 1998), Fauna and Flora International (1998), the DIVA project under the Danish Environmental Programme (1997), and the Gesellschaft für Tropenornithologie (1996), and to M.K. by the German Research Council (Deutsche Forschungsgemeinschaft) and the A.F.W. Schimper-Stiftung (1995).

LITERATURE CITED

- BALTANÁS, A. 1992. On the use of some methods for the estimation of species richness. *Oikos* 65:484-492.
- BIBBY, C. J., N. J. COLLAR, M. J. CROSBY, M. F. HEATH, C. IMBODEN, T. H. JOHNSON, A. J. LONG, A. J. STATTERSFIELD, AND S. J. THERGOOD. 1992. Putting biodiversity on the map: Priority areas for global conservation. ICBP, Cambridge, United Kingdom.
- BIBBY, C. J., D. A. HILL, N. D. BURGESS, AND S. MUSTOE. 2000. *Bird Census Techniques*. Academic Press, London.
- BLAKE, J. G. 2000. Temporal variation in point counts of birds in a lowland wet forest in Costa Rica. *Condor* 94:265-275.
- BLAKE, J. G., F. G. STILES, AND B. A. LOISELLE. 1990. Birds of La Selva Biological Station: Habitat use, trophic composition, and migrants. Pages 161-182 in *Four Neotropical Forests* (A. H. Gentry, Ed.). Yale University Press, New Haven, Connecticut.
- BOULINIER, T., J. D. NICHOLS, J. R. SAUER, J. E. HINES, AND K. H. POLLOCK. 1998. Estimating species richness: The importance of heterogeneity in species detectability. *Ecology* 79:1018-1028.
- BUNGE, J., AND M. FITZPATRICK. 1993. Estimating the number of species: A review. *Journal of the American Statistical Association* 88:364-373.
- BUDNEY, G. F., AND R. W. GROTKE. 1997. Techniques for audio recording vocalizations of tropical birds. *Ornithological Monographs* 48:147-163.
- BURNHAM, K. P., AND W. S. OVERTON. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* 65:623-633.
- BURNHAM, K. P., AND W. S. OVERTON. 1979. Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60:927-936.
- CHAO, A. 1984. Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11:265-270.
- CHAO, A. 1987. Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43:783-791.
- CHAO, A., M.-C. MA, AND M. C. K. YANG. 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80:193-201.
- CHAZDON, R. L., R. K. COLWELL, J. S. DENSLow, AND M. R. GUARIGUATA. 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. Pages 285-309 in *Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies* (F. Dallmeier and J. A. Comiskey, Eds.). Parthenon Publishing, Paris, France.
- COHN-HAFT, M., A. WHITTAKER, AND P. C. STOUFFER. 1997. A new look at the "species-poor" central Amazon: The avifauna north of Manaus, Brazil. *Ornithological Monographs* 48:205-235.
- COLWELL, R. K. 1997. ESTIMATES: Statistical estimation of species richness and shared species from samples, version 5. [Online.] Available at <http://viceroy.eeb.uconn.edu/EstimateS>.
- COLWELL, R. K., AND J. A. CODDINGTON. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of London* 345:101-118.
- DALE, V. H., S. M. PEARSON, H. L. OFFERMAN, AND R. V. O'NEILL. 1994. Relating patterns of land-use change to faunal biodiversity in the central Amazon. *Conservation Biology* 8:1027-1036.
- FJELDSÁ, J. 1999. The impact of human forest disturbance on the endemic avifauna of the Udzungwa Mountains, Tanzania. *Bird Conservation International* 9:47-62.
- GOTELLI, N. J., AND R. K. COLWELL. 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4:379-391.
- GOTELLI, N. J., AND G. R. GRAVES. 1996. *Null Models in Ecology*. Smithsonian Institution Press, Washington, D.C.
- GRAM, W. R., AND J. FAABORG. 1997. The distribution of Neotropical migrant birds wintering in El Cielo Biosphere Reserve, Tamaulipas, Mexico. *Condor* 99:658-670.

- HASELMAYER, J., AND J. S. QUINN. 2000. A comparison of point counts and sound recording as bird survey methods in Amazonian southeast Peru. *Condor* 102:887–893.
- HELTSHE, J., AND N. E. FORRESTER. 1983. Estimating species richness using the jackknife procedure. *Biometrics* 39:1–11.
- HERZOG, S. K., AND M. KESSLER. 2002. Biogeography and composition of dry forest bird communities in Bolivia. *Journal für Ornithologie* 143:171–204.
- HOLMES, R. T., T. W. SHERRY, AND F. W. STURGES. 1986. Bird community dynamics in a temperate deciduous forest: Long-term trends at Hubbard Brook. *Ecological Monographs* 56:201–220.
- HUSTON, M. A. 1994. *Biological Diversity*. Cambridge University Press, Cambridge, United Kingdom.
- JENNI, L., M. LEUENBERGER, AND F. RAMPAZZI. 1996. Capture efficiency of mist nets with comments on their role in the assessment of passerine habitat use. *Journal of Field Ornithology* 67:263–274.
- KARR, J. R. 1971. Structure of avian communities in selected Panama and Illinois habitats. *Ecological Monographs* 41:207–231.
- KARR, J. R. 1981. Surveying birds with mist nets. *Studies in Avian Biology* 6:62–67.
- KARR, J. R. 1990. Birds of tropical rainforest: comparative biogeography and ecology. Pages 215–228 in *Biogeography and Ecology of Forest Bird Communities* (A. Keast, Ed.). SPB Academic Publishing, The Hague, The Netherlands.
- KEATING, K. A., AND J. F. QUINN. 1998. Estimating species richness: The Michaelis-Menten model revisited. *Oikos* 81:411–416.
- KENDEIGH, S. C. 1944. Measurement of bird populations. *Ecological Monographs* 14:67–106.
- LEE, S.-M., AND A. CHAO. 1994. Estimating population size via sample coverage for closed capture–recapture models. *Biometrics* 50:88–97.
- MACKINNON, S., AND K. PHILLIPPS. 1993. *A Field Guide to the Birds of Borneo, Sumatra, Java and Bali*. Oxford University Press, Oxford.
- MAYER, S. 2000. *Birds of Bolivia 2.0*. [CD-ROM.] Bird Songs International B.V., Westerland, The Netherlands.
- ONIKI, Y., AND E. O. WILLIS. 1982a. Breeding records of birds from Manaus, Brazil: Formicariidae to Pipridae. *Revista Brasileira de biologia* 42:563–569.
- ONIKI, Y., AND E. O. WILLIS. 1982b. Breeding records of birds from Manaus, Brazil, II. Apodidae to Furnariidae. *Revista Brasileira de biologia* 42: 745–752.
- ONIKI, Y., AND E. O. WILLIS. 1983. Breeding records of birds from Manaus, Brazil, IV. Tyrannidae to Vireonidae. *Revista Brasileira de biologia* 43:45–54.
- 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.
- PARKER, T. A., III. 1991. On the use of tape recorders in avifaunal surveys. *Auk* 108:443–444.
- PARKER, T. A., III, AND B. BAILEY, EDs. 1991. A biological assessment of the Alto Madidi region and adjacent areas of northwest Bolivia, May 18–June 15, 1990. Conservation International, RAP Working Papers, no. 1.
- PARKER, T. A., III, A. CASTILLO U., M. GELL-MANN, AND O. ROCHA. 1991. Records of new and unusual birds from northern Bolivia. *Bulletin of the British Ornithologists' Club* 111:120–131.
- PARKER, T. A., III, A. H. GENTRY, R. B. FOSTER, L. H. EMMONS, AND J. V. REMSEN, JR. 1993. The lowland dry forests of Santa Cruz, Bolivia: A global conservation priority. Conservation International, RAP Working Papers, no. 4.
- PARKER, T. A., III, D. F. STOTZ, AND J. W. FITZPATRICK. 1996. Ecological and distributional databases. Pages 131–436 in *Neotropical Birds: Ecology and Conservation* (D. F. Stotz, J. W. Fitzpatrick, T. A. Parker III, and D. K. Moskowitz, Eds.). University of Chicago Press, Chicago.
- POULSEN, B. O. 1994. Mist-netting as a census method for determining richness and abundances in an Andean cloud forest bird community. *Gerfaut* 84:39–49.
- POULSEN, B. O., AND N. KRABBE. 1998. Avifaunal diversity of five high-altitude cloud forests on the Andean western slope of Ecuador: Testing a rapid assessment method. *Journal of Biogeography* 25:83–93.
- POULSEN, B. O., N. KRABBE, A. FRØLANDER, M. HINOJOSA B., AND C. QUIROGA O. 1997a. A rapid assessment of Bolivian and Ecuadorian montane avifaunas using 20-species lists: Efficiency, biases and data gathered. *Bird Conservation International* 7:53–67.
- POULSEN, B. O., N. KRABBE, A. FRØLANDER, M. HINOJOSA B., AND C. QUIROGA O. 1997b. A note on 20-species lists. *Bird Conservation International* 7:293.
- RAAIJMAKERS, J. G. W. 1987. Statistical analysis of the Michaelis-Menten equation. *Biometrics* 43:793–803.
- REMSSEN, J. V., JR. 1994. Use and misuse of bird lists in community ecology and conservation. *Auk* 111:225–227.
- REMSSEN, J. V., JR., AND D. A. GOOD. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. *Auk* 113:381–398.
- REMSSEN, J. V., JR., AND T. A. PARKER, III. 1983. Contribution of river-created habitats to bird species richness in Amazonia. *Biotropica* 15:223–231.
- ROBINSON, W. D., J. D. BROWN, AND S. K. ROBINSON. 2000. Forest bird community structure in central

- Panama: Influence of spatial scale and biogeography. *Ecological Monographs* 70:209–235.
- ROSENZWEIG, M. L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, United Kingdom.
- SAUER, J. R., B. G. PETERJOHN, AND W. A. LINK. 1994. Observer differences in the North American Breeding Bird Survey. *Auk* 111:50–62.
- SCHIECK, J. 1997. Biased detection of bird vocalizations affects comparisons of bird abundance among forested habitats. *Condor* 99:179–190.
- SCHMITT, C. G., D. C. SCHMITT, AND J. V. REMSEN, JR. 1997. Birds of the Tambo area, an arid valley in the Bolivian Andes. *Ornithological Monographs* 48:701–716.
- SMITH, E. P., AND G. VAN BELLE. 1984. Nonparametric estimation of species richness. *Biometrics* 40: 119–129.
- SOBERÓN, M., AND J. LLORENTE B. 1993. The use of species accumulation functions for the prediction of species richness. *Conservation Biology* 7: 480–488.
- STOTZ, D. F., J. W. FITZPATRICK, T. A. PARKER III, AND D. K. MOSKOVITS, EDs. 1996. *Neotropical Birds: Ecology and Conservation*. University of Chicago Press, Chicago.
- TERBORGH, J., S. K. ROBINSON, T. A. PARKER III, C. A. MUNN, AND N. PIERPONT. 1990. Structure and organization of an Amazonian bird community. *Ecological Monographs* 60:213–238.
- VERNER, J. 1985. Assessment of counting techniques. *Current Ornithology* 2:247–302.
- VERNER, J., AND K. A. MILNE. 1990. Analyst and observer variability in density estimates from spot mapping. *Condor* 92:313–325.
- WALTHER, B. A., P. COTGREAVE, R. D. PRICE, R. D. GREGORY, AND D. H. CLAYTON. 1995. Sampling effort and parasite species richness. *Parasitology Today* 11:306–310.
- WALTHER, B. A., AND J.-L. MARTIN. 2001. Species richness estimation of bird communities: How to control for sampling effort? *Ibis* 143:413–419.
- WALTHER, B. A., AND S. MORAND. 1998. Comparative performance of species richness estimation methods. *Parasitology* 116:395–405.
- WHITMAN, A. A., J. M. HAGAN III, AND N. V. L. BROKAW. 1997. A comparison of two bird survey techniques used in a subtropical forest. *Condor* 99:955–965.

Associate Editor: S. J. Hackett

APPENDIX 1. Accuracy of S_{obs} and nine richness estimators using the MacKinnon and Phillips (1993) species-list approach (5, 10, 20 species per list) in four computer-simulated species pools (50, 100, 250, 500 species) with species-abundance distributions conforming to the exponential decline model. Values (mean \pm SD of 40 replications with 50 randomizations each) are percentage of each pool's predetermined species richness.

Estimator						
Number of lists	50 species			100 species		
	5-species lists	10-species lists	20-species lists	5-species lists	10-species lists	20-species lists
S_{obs}						
10	37.6 \pm 1.5	51.6 \pm 3.4	70.9 \pm 1.7	28.3 \pm 0.7	40.7 \pm 1.1	55.2 \pm 1.3
20	47.2 \pm 2.1	60.6 \pm 1.8	79.3 \pm 2.3	39.2 \pm 1.3	51.5 \pm 1.8	66.0 \pm 1.8
50	60.1 \pm 3.0	73.6 \pm 3.5	90.9 \pm 3.9	53.6 \pm 2.2	66.0 \pm 3.0	79.4 \pm 3.0
100	70.9 \pm 4.6	—	—	64.4 \pm 3.4	—	—
ACE						
10	50.2 \pm 2.7	64.7 \pm 3.1	82.3 \pm 3.5	46.8 \pm 2.7	55.4 \pm 2.6	69.0 \pm 2.3
20	59.5 \pm 3.2	74.7 \pm 4.7	89.9 \pm 4.9	54.7 \pm 3.3	64.7 \pm 3.6	78.0 \pm 3.5
50	72.5 \pm 6.1	85.0 \pm 10.8	96.6 \pm 6.3	65.9 \pm 3.8	75.2 \pm 6.4	88.1 \pm 7.0
100	82.3 \pm 10.5	—	—	74.0 \pm 5.6	—	—
ICE						
10	52.4 \pm 3.6	64.0 \pm 2.8	79.8 \pm 2.9	49.5 \pm 2.9	57.3 \pm 3.3	69.3 \pm 2.7
20	60.6 \pm 4.1	73.5 \pm 4.3	90.3 \pm 4.8	55.3 \pm 3.4	65.4 \pm 3.5	78.9 \pm 3.9
50	73.0 \pm 6.0	83.1 \pm 9.2	97.0 \pm 6.8	66.7 \pm 4.7	77.3 \pm 6.9	89.0 \pm 6.3
100	82.7 \pm 9.6	—	—	75.4 \pm 7.0	—	—
Chao 1						
10	57.1 \pm 3.8	74.1 \pm 4.9	90.0 \pm 6.5	49.6 \pm 3.1	59.6 \pm 2.9	73.4 \pm 2.9
20	66.6 \pm 4.6	82.0 \pm 7.4	96.0 \pm 6.1	59.0 \pm 4.1	69.0 \pm 4.6	81.9 \pm 6.0
50	80.2 \pm 9.8	90.2 \pm 12.4	97.8 \pm 6.5	70.0 \pm 4.4	79.2 \pm 11.6	91.3 \pm 9.5
100	88.5 \pm 11.8	—	—	77.6 \pm 7.4	—	—
Chao 2						
10	58.7 \pm 5.4	72.1 \pm 4.9	89.8 \pm 5.2	50.6 \pm 3.3	59.4 \pm 2.8	73.4 \pm 3.2
20	68.0 \pm 6.0	80.4 \pm 6.0	96.0 \pm 6.9	58.2 \pm 3.7	69.3 \pm 4.5	83.0 \pm 5.6
50	80.8 \pm 8.5	89.7 \pm 14.1	98.6 \pm 7.8	71.2 \pm 5.3	81.1 \pm 11.4	91.4 \pm 8.7
100	88.0 \pm 11.5	—	—	78.8 \pm 8.7	—	—
Jackknife 1						
10	50.7 \pm 2.7	64.5 \pm 2.4	83.4 \pm 2.8	42.8 \pm 1.4	55.6 \pm 1.9	70.0 \pm 2.2
20	61.1 \pm 3.5	74.0 \pm 3.0	91.3 \pm 3.8	54.5 \pm 2.3	66.3 \pm 3.6	80.4 \pm 3.7
50	74.2 \pm 4.9	84.4 \pm 6.8	98.7 \pm 6.3	68.6 \pm 4.0	79.6 \pm 6.0	91.7 \pm 5.7
100	83.2 \pm 7.8	—	—	77.9 \pm 5.9	—	—
Jackknife 2						
10	57.1 \pm 3.3	70.8 \pm 3.4	88.8 \pm 3.8	50.4 \pm 2.0	62.6 \pm 2.4	76.8 \pm 2.7
20	67.9 \pm 4.3	80.1 \pm 4.6	96.3 \pm 5.3	62.0 \pm 3.4	73.5 \pm 4.4	87.0 \pm 4.7
50	81.0 \pm 7.0	90.2 \pm 10.8	101.9 \pm 8.4	75.6 \pm 5.2	85.9 \pm 9.9	96.7 \pm 9.1
100	90.8 \pm 10.7	—	—	84.4 \pm 8.5	—	—
Bootstrap						
10	43.8 \pm 1.8	57.2 \pm 1.9	76.8 \pm 2.3	34.8 \pm 1.0	47.6 \pm 1.4	62.0 \pm 1.6
20	53.8 \pm 2.6	66.9 \pm 2.5	85.0 \pm 3.0	46.3 \pm 1.7	58.3 \pm 3.0	72.6 \pm 3.0
50	66.4 \pm 3.7	78.5 \pm 4.6	94.5 \pm 4.7	60.6 \pm 2.9	72.3 \pm 4.1	85.2 \pm 3.9
100	76.4 \pm 5.8	—	—	70.7 \pm 4.5	—	—
MMRuns						
10	55.3 \pm 3.0	61.0 \pm 2.3	75.1 \pm 2.1	62.5 \pm 3.5	62.3 \pm 2.9	67.4 \pm 2.2
20	58.3 \pm 3.3	65.7 \pm 2.5	80.6 \pm 2.4	62.0 \pm 3.1	65.0 \pm 3.0	72.3 \pm 2.4
50	64.3 \pm 3.7	73.1 \pm 3.2	87.8 \pm 3.4	65.5 \pm 3.4	71.3 \pm 3.6	80.2 \pm 2.9
100	70.3 \pm 4.4	—	—	70.1 \pm 4.1	—	—
MMMean						
10	54.7 \pm 3.2	60.8 \pm 2.3	75.1 \pm 2.1	60.5 \pm 3.2	61.8 \pm 2.8	67.3 \pm 2.1
20	58.1 \pm 3.3	65.6 \pm 2.5	80.5 \pm 2.5	61.5 \pm 3.0	65.0 \pm 3.1	72.3 \pm 2.5
50	64.4 \pm 3.7	73.1 \pm 3.1	87.8 \pm 3.4	65.5 \pm 3.4	71.3 \pm 3.6	80.2 \pm 2.9
100	70.3 \pm 4.4	—	—	70.1 \pm 4.2	—	—

APPENDIX 1. Extended.

250 species			500 species		
5-species lists	10-species lists	20-species lists	5-species lists	10-species lists	20-species lists
16.0 ± 0.3	26.7 ± 0.5	40.6 ± 0.8	9.0 ± 0.1	16.5 ± 0.2	28.4 ± 0.4
26.3 ± 0.7	39.3 ± 1.0	53.7 ± 1.1	16.4 ± 0.2	27.9 ± 0.5	42.9 ± 0.9
43.3 ± 1.6	56.4 ± 2.0	70.4 ± 2.0	32.0 ± 0.7	47.3 ± 1.4	62.9 ± 1.7
56.6 ± 2.3	—	—	46.7 ± 1.4	—	—
50.1 ± 4.2	52.2 ± 3.0	60.0 ± 1.9	62.7 ± 7.1	55.3 ± 2.9	58.3 ± 2.7
51.2 ± 3.5	59.4 ± 3.2	69.7 ± 2.5	55.5 ± 4.3	58.0 ± 3.6	65.6 ± 2.3
62.7 ± 3.6	71.4 ± 4.6	82.8 ± 5.0	58.4 ± 3.7	67.5 ± 4.7	78.6 ± 3.5
72.9 ± 4.0	—	—	66.6 ± 4.4	—	—
56.2 ± 4.6	56.3 ± 3.6	62.6 ± 2.2	72.9 ± 9.1	62.6 ± 3.8	64.9 ± 3.7
53.2 ± 3.1	60.0 ± 3.5	71.0 ± 2.3	58.4 ± 4.1	61.4 ± 4.0	67.9 ± 2.7
62.6 ± 3.6	71.0 ± 4.3	83.4 ± 4.3	59.8 ± 3.8	69.4 ± 5.9	80.6 ± 4.4
72.1 ± 4.6	—	—	67.7 ± 4.6	—	—
50.5 ± 4.4	53.7 ± 3.1	61.3 ± 2.2	62.9 ± 6.5	55.2 ± 3.0	58.7 ± 2.6
52.1 ± 3.5	60.6 ± 3.5	72.3 ± 3.2	55.6 ± 4.4	58.4 ± 3.7	66.3 ± 2.1
64.4 ± 3.4	73.0 ± 4.9	85.0 ± 5.7	58.9 ± 3.5	68.5 ± 6.0	81.2 ± 5.4
76.7 ± 7.8	—	—	68.6 ± 6.7	—	—
53.8 ± 4.9	54.1 ± 3.4	61.6 ± 2.3	68.6 ± 8.3	58.6 ± 3.6	60.7 ± 2.8
52.8 ± 3.1	60.0 ± 3.4	72.5 ± 2.6	56.8 ± 4.0	59.5 ± 3.8	67.2 ± 2.5
63.8 ± 3.5	72.8 ± 6.2	85.2 ± 6.2	59.8 ± 3.7	69.2 ± 5.7	83.0 ± 6.1
75.2 ± 8.5	—	—	68.9 ± 5.6	—	—
27.7 ± 0.7	42.7 ± 1.2	58.6 ± 1.4	16.5 ± 0.2	28.9 ± 0.4	46.2 ± 0.9
42.7 ± 1.5	57.6 ± 2.1	71.9 ± 1.9	29.3 ± 0.5	46.3 ± 1.2	64.3 ± 1.7
62.3 ± 2.8	73.9 ± 3.5	87.0 ± 3.6	51.8 ± 1.6	68.9 ± 3.2	83.7 ± 3.4
74.8 ± 3.9	—	—	68.0 ± 2.8	—	—
36.1 ± 1.1	52.5 ± 1.8	67.4 ± 2.0	22.5 ± 0.4	38.1 ± 0.7	57.6 ± 1.4
53.0 ± 2.2	66.7 ± 3.0	80.4 ± 2.5	39.3 ± 0.9	58.4 ± 2.0	75.4 ± 2.4
71.7 ± 3.7	81.3 ± 5.4	93.6 ± 5.9	64.0 ± 2.6	79.4 ± 5.3	93.0 ± 5.8
83.3 ± 6.5	—	—	78.7 ± 4.5	—	—
20.8 ± 0.5	33.7 ± 0.8	48.8 ± 1.0	12.0 ± 0.2	21.6 ± 0.3	36.0 ± 0.6
33.4 ± 1.0	47.6 ± 1.4	62.2 ± 1.4	21.6 ± 0.3	35.7 ± 0.7	52.5 ± 1.2
52.0 ± 2.1	64.5 ± 2.6	78.2 ± 2.6	40.4 ± 1.1	57.1 ± 2.1	72.6 ± 2.3
65.0 ± 2.9	—	—	56.3 ± 2.0	—	—
85.9 ± 9.2	76.1 ± 5.0	74.9 ± 2.8	144.1 ± 43.1	94.9 ± 6.9	89.8 ± 4.6
76.9 ± 6.1	74.4 ± 3.9	76.2 ± 2.5	97.8 ± 7.9	90.1 ± 5.6	88.2 ± 3.9
75.1 ± 4.8	75.8 ± 3.8	81.2 ± 2.7	88.3 ± 5.5	88.1 ± 5.2	89.0 ± 3.5
76.9 ± 4.6	—	—	85.9 ± 6.7	—	—
76.5 ± 6.6	74.4 ± 4.6	74.5 ± 2.8	92.4 ± 10.7	88.7 ± 5.8	88.6 ± 4.4
75.0 ± 5.6	74.0 ± 3.9	76.1 ± 2.4	89.8 ± 6.4	88.8 ± 5.4	88.0 ± 3.8
74.8 ± 4.7	75.8 ± 3.8	81.2 ± 2.7	87.4 ± 5.4	87.9 ± 5.2	89.0 ± 3.6
76.9 ± 4.6	—	—	86.3 ± 5.0	—	—

APPENDIX 2. Accuracy of S_{obs} and nine richness estimators using the MacKinnon and Phillips (1993) species-list approach (5, 20 species per list) in four computer-simulated species pools (50, 100, 250, 500 species) with uniform species-abundance distributions. Values (mean \pm SD of 40 replications with 50 randomizations each) are percentage of each pool's predetermined species richness.

Estimator	Num-ber of lists	50 species				100 species				250 species				500 species			
		5-species lists	20-species lists	50-species lists	20-species lists	5-species lists	20-species lists	50-species lists	20-species lists	5-species lists	20-species lists	50-species lists	20-species lists	5-species lists	20-species lists		
S_{obs}	10	63.9 \pm 1.2	98.0 \pm 0.0	39.6 \pm 0.7	88.8 \pm 0.7	18.1 \pm 0.3	56.3 \pm 0.6	9.5 \pm 0.1	33.4 \pm 0.2								
	20	86.6 \pm 1.9	99.8 \pm 0.7	63.9 \pm 1.3	98.4 \pm 0.5	33.0 \pm 0.5	80.8 \pm 0.9	18.1 \pm 0.2	55.6 \pm 0.4								
ACE	50	99.2 \pm 1.6	100.0 \pm 0.0	92.7 \pm 2.5	100.0 \pm 0.0	63.0 \pm 1.2	98.4 \pm 0.7	39.7 \pm 1.0	86.8 \pm 1.1								
	10	108.1 \pm 4.4	99.1 \pm 1.0	113.4 \pm 10.8	100.3 \pm 0.7	129.8 \pm 14.5	102.4 \pm 3.4	124.2 \pm 14.7	104.4 \pm 3.8								
ICE	20	100.9 \pm 2.2	99.7 \pm 0.7	104.1 \pm 5.0	99.8 \pm 0.4	105.9 \pm 4.0	100.6 \pm 2.2	123.6 \pm 21.1	101.1 \pm 2.6								
	50	99.3 \pm 1.3	100.0 \pm 0.0	100.3 \pm 4.0	100.0 \pm 0.0	100.4 \pm 7.1	99.8 \pm 0.6	109.4 \pm 12.6	100.1 \pm 2.5								
Chao 1	10	114.1 \pm 5.7	100.0 \pm 0.0	126.8 \pm 11.8	104.0 \pm 1.0	150.2 \pm 15.3	112.8 \pm 3.0	133.8 \pm 14.4	119.4 \pm 4.1								
	20	102.0 \pm 2.4	99.8 \pm 0.7	108.2 \pm 5.3	100.0 \pm 0.0	111.6 \pm 5.7	103.2 \pm 2.1	126.4 \pm 18.8	106.3 \pm 2.5								
Chao 2	50	99.2 \pm 1.7	100.0 \pm 0.0	100.9 \pm 4.1	100.0 \pm 0.0	101.6 \pm 7.2	100.1 \pm 0.7	108.9 \pm 15.0	100.7 \pm 2.4								
	10	107.5 \pm 5.0	99.9 \pm 0.4	112.1 \pm 12.9	100.2 \pm 0.7	127.5 \pm 17.5	100.9 \pm 3.2	132.2 \pm 20.6	102.4 \pm 3.8								
Jackknife 1	20	100.5 \pm 2.1	99.7 \pm 0.7	102.9 \pm 4.8	100.0 \pm 0.0	103.5 \pm 3.9	100.1 \pm 2.0	120.2 \pm 19.6	100.0 \pm 2.6								
	50	99.9 \pm 2.2	100.0 \pm 0.0	99.8 \pm 5.0	100.0 \pm 0.0	98.0 \pm 8.9	99.9 \pm 0.8	107.9 \pm 12.7	99.9 \pm 3.3								
Jackknife 2	10	108.6 \pm 5.8	100.0 \pm 0.0	118.9 \pm 13.7	101.2 \pm 0.7	139.4 \pm 15.9	105.8 \pm 2.7	134.7 \pm 16.4	109.9 \pm 3.8								
	20	100.9 \pm 2.4	99.8 \pm 0.7	104.7 \pm 4.9	100.0 \pm 0.0	106.5 \pm 5.3	101.3 \pm 1.9	121.6 \pm 18.8	102.7 \pm 2.5								
Bootstrap	50	100.5 \pm 6.1	100.0 \pm 0.0	100.3 \pm 5.0	100.0 \pm 0.0	98.8 \pm 8.4	100.1 \pm 0.8	107.0 \pm 15.8	99.9 \pm 3.1								
	10	99.2 \pm 3.8	102.1 \pm 0.3	68.0 \pm 1.6	112.7 \pm 0.9	33.2 \pm 0.4	90.1 \pm 1.1	17.7 \pm 0.2	58.3 \pm 0.5								
MMRuns	20	112.2 \pm 2.9	99.9 \pm 0.4	99.9 \pm 2.9	104.1 \pm 0.6	58.5 \pm 1.1	112.1 \pm 1.9	33.9 \pm 0.6	90.5 \pm 1.1								
	50	100.6 \pm 2.9	100.0 \pm 0.0	111.8 \pm 5.1	100.0 \pm 0.0	98.5 \pm 3.2	105.1 \pm 1.4	69.9 \pm 2.7	113.1 \pm 2.5								
MMMean	10	116.1 \pm 3.9	96.5 \pm 1.0	88.1 \pm 2.7	112.4 \pm 1.2	45.4 \pm 0.7	109.6 \pm 1.8	24.7 \pm 0.4	76.7 \pm 0.8								
	20	111.6 \pm 3.1	98.2 \pm 0.5	118.0 \pm 4.4	97.1 \pm 0.9	78.3 \pm 1.7	118.4 \pm 2.5	47.5 \pm 1.1	111.2 \pm 1.7								
MMMean	50	97.0 \pm 6.8	100.0 \pm 0.0	105.5 \pm 9.9	99.9 \pm 0.3	115.5 \pm 6.8	97.7 \pm 4.1	92.8 \pm 4.8	112.7 \pm 5.2								
	10	79.8 \pm 2.0	102.0 \pm 0.0	51.7 \pm 1.0	101.9 \pm 0.8	24.1 \pm 0.3	71.1 \pm 0.8	12.7 \pm 0.1	43.6 \pm 0.3								
MMMean	20	100.7 \pm 2.1	100.0 \pm 0.0	80.0 \pm 1.9	103.2 \pm 0.4	43.2 \pm 0.7	96.4 \pm 1.3	24.3 \pm 0.3	70.7 \pm 0.7								
	50	101.2 \pm 2.0	100.0 \pm 0.0	104.2 \pm 3.3	100.0 \pm 0.0	78.8 \pm 1.8	104.0 \pm 0.6	51.9 \pm 1.6	101.1 \pm 1.6								
MMMean	10	173.5 \pm 9.4	123.4 \pm 0.9	212.3 \pm 25.3	148.3 \pm 1.9	300.0 \pm 88.8	176.5 \pm 6.7	312.3 \pm 161.3	191.4 \pm 7.3								
	20	151.8 \pm 5.6	112.0 \pm 0.0	178.3 \pm 11.4	129.9 \pm 0.8	199.6 \pm 16.7	159.2 \pm 4.3	283.9 \pm 152.7	176.8 \pm 5.1								
MMMean	50	126.2 \pm 2.7	104.1 \pm 0.3	148.8 \pm 6.2	113.0 \pm 0.0	168.8 \pm 8.0	133.5 \pm 1.6	193.4 \pm 26.0	154.0 \pm 3.5								
	10	168.2 \pm 8.6	123.4 \pm 0.9	187.5 \pm 16.9	148.0 \pm 1.9	195.3 \pm 16.4	174.5 \pm 6.1	205.8 \pm 37.7	185.5 \pm 6.7								
MMMean	20	151.2 \pm 5.4	112.0 \pm 0.0	174.2 \pm 10.9	129.9 \pm 0.8	185.5 \pm 13.4	158.9 \pm 4.3	199.7 \pm 30.7	175.6 \pm 5.1								
	50	126.2 \pm 2.7	104.1 \pm 0.3	148.7 \pm 6.3	113.0 \pm 0.0	168.0 \pm 8.0	133.5 \pm 1.7	190.9 \pm 25.3	153.9 \pm 3.5								