BIAS, PRECISION, AND ACCURACY OF FOUR MEASURES OF SPECIES RICHNESS

JESSICA J. HELLMANN[1] AND GARY W. FOWLER

School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan 48109 USA

Abstract. Species richness is a widely used surrogate for the more complex concept of biological diversity. Because species richness is often central to ecological study and the establishment of conservation priorities, the biases and merits of richness measurements demand evaluation. The jackknife and bootstrap estimators can be used to compensate for the underestimation associated with simple richness estimation (or the sum of species counted in a sample). Using data from five forest communities, we analyzed the simple measure of richness, the first- and second-order jackknife, and the bootstrap estimators with simulation and resampling methods to examine the effects of sample size on estimator performance. Performance parameters examined were systematic under- or overestimation (bias), ability to estimate consistently (precision), and ability to estimate true species richness (accuracy).

For small sample sizes in all studied communities (less than ~25% of the total community), the least biased estimator was the second-order jackknife, followed by the first-order jackknife, the bootstrap, and the simple richness estimator. However, with increases in sample size, the second-order jackknife, followed by the first-order jackknife and the bootstrap, became positively biased. The simple richness estimator was the most precise estimator in all studied communities, but it yielded the largest underestimate of species richness at all sample sizes. The relative precision of the four estimators did not differ across communities, but the magnitude of estimator variance is dependent on the sampled community. Differences in accuracy among the estimators were not independent of community, and accuracy patterns were associated with community species diversity. The results of this study can assist policy makers, researchers, and managers in the selection of appropriate sample sizes and estimators for richness estimation and should facilitate the ongoing assessment of local, and ultimately global, biodiversity.

Key words: biological diversity; bootstrap estimator; jackknife estimator, first- and second-order; Monte Carlo simulation; quadrat sampling; resampling procedures; sample size determination; simple richness estimator; species diversity; species richness.

INTRODUCTION

Assessments of species diversity can facilitate or prioritize management and can serve as starting points for basic research in community ecology. For these reasons, effective methods for estimation of species diversity need to be identified and implemented. Many complicated measures of diversity have been suggested (e.g., Simpson 1949, Shannon and Weaver 1963, Peet 1974), but some ecologists have argued for the use of the simplest measure of biological diversity, the number of species counted in a community, or species richness. This preference stems from the relative ease with which richness can be assessed, the lack of documented success of more complicated indices, and the lack of agreement between researchers and managers as to

which index should be used (see Peet 1974, Magurran 1988).

Yet, problems with simple species richness estimation also exist. In particular, the richness estimator shows sample size dependency (Smith and van Belle 1984). Simple richness estimators, such as the sum total of observed species in a sample (called S_{sim} in this paper), underestimate richness at all sample sizes, and this underestimation decreases with increases in sample size. Fortunately, methods exist that allow the reduction of the underestimation associated with sampling-based richness estimation. These methods include the family of jackknives and the bootstrap estimators. Here, we analyze the performance of S_{sim}, the first- and second-order jackknife, and the bootstrap estimators so that estimator and sample size recommendations can be made for species richness estimation. Further, we analyze the dependence of estimator performance on the ecological properties of a sampled community.

Estimator and sample size selection is largely based upon consideration of the following questions: (1) Does the estimator chosen systematically over or underes-

[1] Present address: Center for Conservation Biology, Department of Biological Sciences, Stanford University, Stanford, California 94305-5020 USA.
E-mail:jessicah@leland.stanford.edu

timate species richness? (bias); (2) If the estimation procedure were repeated a second time, would the second answer be similar to the first? (precision); and (3) How close to the true species richness value is the richness estimate? (accuracy). All estimation procedures differ with respect to their performance in each of these three categories. Further, the performance of an individual estimator can be either dependent or independent of the parameters of a sampled community, and estimator performance may or may not be predictable.

Definitions of bias, accuracy, and precision

Insufficient knowledge of the role that bias, accuracy, and precision play in richness estimation can lead to inappropriate estimator choice, inconsistent estimation results, and poor decision making. For species richness estimation, the definitions of bias, accuracy, and precision are as follows, and each are examined in this study.

Bias is the difference between the expected value of the estimator (the mean of the estimates of all possible samples that can be taken from the population) and the true, unknown, population value of species richness. Accuracy is the difference between a sample estimate and the true population value. For species richness, accuracy is the difference between the estimate of species richness based on sample data and the true species richness of the population or community being sampled. Precision is the difference between a sample estimate and the mean of the estimates of all possible samples that can be taken from the population. For species richness, precision is the difference between an estimate of species richness based on sample data and the mean of all possible estimates of species richness based on all possible samples of the same size from the population or community being sampled. Precision is measured by the variance of the estimator, and accuracy is measured by the mean square error of the estimator. When the bias is zero, precision and accuracy are identical. Because bias is never zero, due to innate estimator bias and data collection error, accuracy may be difficult to achieve. Accuracy, however, is the most desirable of the three parameters because a highly accurate estimate is neither biased nor highly variable.

History of the jackknife and bootstrap

The species richness estimators investigated here have a long history in statistics and ecology and have broad application in their pure form. The jackknife and bootstrap first originated as generic nonparametric estimators of bias and standard error. The jackknife was introduced by Quenoille in 1949 and was later advanced by Tukey in 1958. Generalized equations for the kth-order jackknife were derived by Gray and Schucany (1972) and were discussed in Miller (1974). Suggestions for the use of the jackknife estimator in species richness estimation originated when Zahl (1977) treat-

ed rectangular vegetation plots, or quadrats, as independent samples that could be jackknifed for diversity estimation. This procedure assumes a random sample of independent quadrats rather than a random sample of individuals (Smith and van Belle 1984, Palmer 1990). The random selection of quadrats is, in fact, a random sample of space (Heltshe and Forrester 1983). Furthermore, nonparametric estimators, such as the jackknife and bootstrap, do not assume relationships among species within a quadrat and make no assumptions about underlying species distributions.

The simplest jackknife species richness estimator, the first-order, is a function of the number of rare species found in a community. Its calculation involves the number of species that occur in one and only one quadrat. The second-order jackknife is a function of both the number of species in only one quadrat and the number of species in only two quadrats. The theoretical premise of the jackknife is that estimates of the parameter of interest are obtained from n samples of size $n - k$, where each sample is formed by deleting k of the original n quadrats. In the first- and second-order jackknife, $k = 1$ and 2, respectively. Jackknife estimators exist for each deletion sample size up to the original sample size, $k = n - 1$. Hence, the first through kth-order jackknives form a "family" of estimators. For species richness, however, a closed, noniterative form of the jackknife exists, allowing an estimate to be calculated from a single sample to reduce richness underestimation (see the *Appendix*).

The bootstrap estimator was proposed by Efron (1979, 1981) and is a resampling procedure where bootstrap samples of size n are randomly selected from n quadrats with replacement. A noniterative equation also exists for the bootstrap, enabling one to calculate a bootstrap richness estimate from a single sample (see the *Appendix*).

Other, more recently developed, nonparametric estimators of species richness include the Chao1 and Chao2 estimators, an abundance-based coverage estimator (ACE), and an incidence-based coverage estimator (ICE) (Chao 1984, 1987, Chao and Lee 1992, Lee and Chao 1994; see Robert K. Colwell's "EstimateS" [1997, available online][2]). Chao2 and ICE use presence–absence data, as are the data reported in this study, but these estimators are not the focus of analysis here. The Chao2 estimator uses the observed number of species in a sample, combined with the number of species appearing in only one and two samples (see Colwell and Coddington 1994). ICE can be challenging to calculate and does not produce estimates in many cases (Walther and Morand 1998). The Chao2 estimator and ICE are mentioned here so that readers are aware of their potential application to presence–absence data sets.

[2] URL: ⟨http://viceroy.eeb.uconn.edu/estimates⟩

TABLE 1. Locations of the five study systems.

Site	Location by Section, Township (T), and Range (R)
S1	University of Michigan Biological Station; Section 33, T37N, R3W, Cheboygan County, Michigan, USA.
S2	Section 26, T44N, R37W, Iron County, Michigan, USA.
S3	University of Michigan's Stinchfield Woods; Section 12, T1S, R4E, Washtenaw County, Michigan, USA.
S4 and S5	University of Michigan's Radrick Forest; Sections 24 and 25, T2S, R6E and Sections 19 and 20, T2S, R7E, Washtenaw County, Michigan, USA.

Performance analysis and recommendations

Nonparametric estimators, such as the jackknife and bootstrap, show promise as effective and reliable measures of species richness (Colwell and Coddington 1994). Therefore, using computer resampling procedures to analyze thousands of hypothetical samples in five distinct ecological communities, this paper further analyzes the first- and second-order jackknife, the bootstrap, and a simple measure of species richness, S_{sim}, so that these procedures can be better applied in biodiversity surveys and analyses. The bias, precision (Monte Carlo variance), and accuracy (mean square error) of these four estimators are compared across community types, and recommendations for estimator and sample size choice are presented. It is the goal of this study to provide improved understanding of the relationship between estimator behavior, sample size, and ecological community properties in order to assist researchers, managers, and policy makers in their species diversity estimation and evaluation efforts.

METHODS

Study sites

Data were collected in 1994 and 1995 from five 0.4 ha (80 × 50 m) forested plots (communities) in Michigan. These five sites are labeled S1 through S5, and their geographic locations are given in Table 1. Each of the five plots contained 160 quadrats measuring 5 × 5 m, and the presence or absence of woody species was recorded in every quadrat. The abundance of stems ≥1.5 cm in diameter (dbh) in each quadrat was also recorded. The total species richness (*S*) of each of the

five plots, or communities, was considered the "true" species richness of the community and was used for comparison with all calculated estimates of richness. The five sites comprising this study were all northern forested ecosystems selected for their differences in total number of species, numbers of rare and common species, and degree of dominance or evenness (see Tables 2–4).

The S1 site is located on a sandy, well-drained high-level outwash lake plain and is a predominately red oak, red pine, and white pine forest; red maple can be found in the understory (see Tables 3 and 4). The S2 site is mesic, on a moraine of fertile silt, and is dominated by sugar maple with a sparse understory. The S3 site has loamy to clay soils and is a fire-dominated oak–hickory forest that is slowly being replaced with maple. The S4 site is on moderately well-drained soils and is an old-growth forest dominated by white oak; black maple, ironwood, and slippery and American elm can be found in the understory. Finally, the S5 site is mesic with fine-textured clay soils. It is dominated by oaks and red maples; cherry, maple, ironwood, and dogwood can be found in the understory.

Computer simulation procedures

A FORTRAN computer program was written to randomly sample, and then resample the sample, from each of the five plots/communities for sample sizes ranging from 2.5 to 70% of the population of 160 quadrats. For each random sample, quadrats were selected without replacement using a pseudorandom-number generator. For each of the five communities, the FORTRAN pro-

TABLE 2. Descriptive statistics for the five sites related to the number of species per plot (i.e., community). Each plot contains 160 quadrats. These data include plants in the over- and understory as well as ground cover species.

Community	No. of species/plot	Species abundance categories (No. of species found in *x* quadrats)					Mean no. of species/quadrat	Variance in species/quadrat*	Variance–mean ratio*	Range in no. of species/quadrat*
		"Very Rare" 1–10	"Rare" 11–50	"Average" 51–100	"Common" 101–140	"Very Common" 141–160				
S1	14	2	2	0	3	7	9.2	1.41	0.153	5
S2	17	3	5	1	4	4	8.2	2.69	0.326	9
S4	42	20	11	3	3	5	11.1	4.73	0.424	11
S3	45	25	10	4	2	4	9.4	4.96	0.530	13
S5	56	29	16	8	2	1	9.8	8.60	0.879	15

* Pearson correlation against number of species per plot, $P < 0.05$.

TABLE 3. (A) Species presence and abundance, and (B) basic statistics (see Magurran 1988) of woody plants in the understory (1.5 ≤ dbh < 9.1 cm) and overstory (dbh ≥ 9.1 cm) of each study site.

A) Species abundance

		No. of stems/site				
Scientific name	Common name	S1	S2	S3	S4	S5
Ulmus americana	American elm	...	...	3	70	31
Tilia americana	Basswood	...	10	1	42	13
Populus grandidentata	Bigtooth aspen	4	...	...	...	...
Carya cordiformis	Bitternut hickory	...	...	...	3	5
Prunus serotina	Black cherry	...	...	...	29	83
Acer nigrum	Black maple	...	...	14	137	23
Quercus velutina	Black oak	...	...	33	3	17
Juglans nigra	Black walnut	...	...	...	...	2
Prunus virginiana	Choke cherry	...	...	38	3	18
Malus pumila	Common apple	...	...	...	1	...
Rhamnus cathartica	Common buckthorn	...	...	...	29	1
Celtis tenuifolia	Dwarf hackberry	...	...	1	...	...
Tsuga canadensis	Eastern hemlock	...	2	...	...	...
Cornus florida	Flowering dogwood	...	...	...	1	91
Vitis spp.	Grape	...	...	3	...	3
Ostrya virginiana	Ironwood	...	3	...	308	133
Lonicera maackii	Maack's honeysuckle	...	...	...	12	...
Carpinus caroliniana	Musclewood	...	...	...	1	...
Acer platanoides	Norway maple	...	...	...	1	6
Carya glabra	Pignut hickory	...	...	71	21	11
Toxicodendron radicans	Poison ivy	...	...	...	2	
Acer rubrum	Red maple	58	2	35	1	161
Quercus rubra	Red oak	167	...	...	4	8
Pinus resinosa	Red pine	192	...	...	...	...
Sassafras albidium	Sassafras	...	...	2	26	5
Amelanchier spp.	Serviceberry	...	...	4	...	1
Carya ovata	Shagbark hickory	...	...	...	17	2
Ulmus rubra	Slippery elm	...	...	...	29	60
Acer saccharum	Sugar maple	...	213	191	58	6
Prunus avium	Sweet cherry	...	...	...	...	1
Juglans spp.	Walnut	...	...	...	...	1
Fraxinus americana	White ash	...	...	12	9	3
Quercus alba	White oak	...	...	30	57	21
Pinus strobus	White pine	148	...	...	...	...
Crataegus spp.	Hawthorn	...	...	1	...	...
Hamamelis virginiana	Witch hazel	...	...	2	...	...
Betula alleghaniensis	Yellow birch	...	5	...	...	...
Total number of stems		569	235	441	864	706

B) Basic statistics

	Site				
Statistical quantity	S1	S2	S3	S4	S5
No. of species (S)	5	6	16	24	25
Density (No. of stems/ha)	1422.50	587.50	1102.50	2160.00	1765.00
Shannon Index	1.34	0.44	1.86	2.23	2.38
Simpson Index	3.61	1.21	4.25	5.72	7.71
Log Series α	0.75	1.12	3.25	4.57	5.05
Berger–Parker	2.96	1.10	2.32	2.81	4.39
Shannon Evenness	0.83	0.25	0.67	0.70	0.74

Notes: These abundance data were used to calculate measures of dominance and evenness for each community for over- and understory data only; abundance data for ground-cover plants are not available for several sites. See Table 2 for a nonindex-based comparison of dominance and evenness that includes ground-cover woody plants.

gram calculated richness estimates based upon 5000 samples (iterations) using S_{sim}, the first- and second-order jackknife, and the bootstrap for all tested sample sizes.

Estimates of the expected value of species richness ($E(\hat{S})$) were calculated as a mean:

$$E(\hat{S}) = \left(\sum_{i=1}^{n} \hat{S}_i \right) \Big/ n$$

where $\hat{S}_i$ are individual estimates of species richness based on one of the $n = 5000$ random samples. The bias of each estimator was estimated as follows:

$$\text{bias}(\hat{S}) = \left[\left(\sum_{i=1}^{n} \hat{S}_i \right) \Big/ n \right] - S$$

where S is the true species richness value of a community. The Monte Carlo variance (precision) was also calculated for each estimator as follows:

TABLE 4. Very common species found in the studied communities in the understory, overstory, and ground cover (see Table 2).

S1	S2	S4	S3	S5
Red oak, *Quercus rubra*	Sugar maple, *Acer saccharum*	Sugar maple, *Acer saccharum*	Sugar maple, *Acer saccharum*	Red maple, *Acer rubrum*
White pine, *Pinus strobus*	Black cherry, *Prunus serotina*	Black maple, *Acer nigrum*	Black cherry, *Prunus serotina*	Black cherry, *Prunus serotina*
Low sweet blueberry, *Vaccinium angustifolium*	Red-berried elder, *Sambucus pubens*	Choke cherry, *Prunus virginiana*	White ash, *Fraxinus americana*	Choke cherry, *Prunus virginiana*
Wintergreen, *Gaultheria procumbens*	Yellow birch, *Betula alleghaniensis*	White ash, *Fraxinus americana*	Gooseberry, *Ribes* spp.	
Huckleberry, *Gaylussacia baccata*		Downy arrowwood, *Viburnum rafinesquianum*	Hickory, *Carya* spp.	
Trailing arbutus, *Epigea repens*				
Bush-honeysuckle, *Diervilla lonicera*				

Notes: Commonness is based on the number of quadrats a species occupies per plot. Species listed here that are not listed in Table 3 are species found only in the ground cover. (Common species for S5 are also listed; see Table 2.) The communities are presented in the order of increasing number of species per plot (*S*).

$$\text{var}(\hat{S}) = \left[\sum_{i=1}^{n} \hat{S}_i^2 - \left(\sum_{i=1}^{n} \hat{S}_i \right)^2 \Big/ n \right] \Big/ (n - 1)$$

and mean square error (accuracy) was calculated as

$$\text{MSE}(\hat{S}) = \text{var}(\hat{S}) + \left[\left(\sum_{i=1}^{n} \hat{S}_i \Big/ n \right) - S \right]^2.$$

(Note that the mean square error is equivalent to $\text{MSE}(\hat{S}) = \text{var}(\hat{S}) + (\text{bias}(\hat{S}))^2$.)

Analysis

The true number of species in an community (*S*) was correlated (Pearson's *r*) with the mean number of species per plot, the variance in the number of species per plot, the variance–mean ratio in number of species per plot, and the range of the number of species per plot for the 160 quadrats in each plot (i.e., community). The sample size at which each estimator begins to overestimate species richness was compared across communities. The magnitude of the Monte Carlo variance of the estimators was compared with the true species richness of each community. Finally, the sample size where the mean square error for a given estimator becomes larger or smaller than another estimator (called the "crossover point") was compared with true species richness and the variance–mean ratio in each of the five systems.

RESULTS

Ecosystem differences

The true species richness values (*S*) of each of the five communities are 14, 17, 42, 45, and 56 species for S1, S2, S4, S3, and S5, respectively (Table 2). A diversity of species can be found on the five sites, and no two sites would be generally classified as the same forest community type (Tables 3 and 4). The five communities also differ with respect to their relative numbers of rare and common species (Tables 2 and 3). S1 and S2 have the fewest rare species and a high percentage of common species. The S5 site and the S3 stand have a high number of rare species and a small percentage of common species. Table 3 shows the dominance and evenness calculations for overstory (dbh $\geq$ 9.1 cm) and understory ($1.5 \leq$ dbh $\leq$ 9.1 cm) trees in each of the five sites. Abundance data were not collected for ground cover (dbh $\leq$ 1.5 cm) species in several sites, so abundance-based indicators could not be calculated for the sum total (understory, overstory, and ground cover) of woody plants in each community.

The mean number of species per quadrat for the five communities ranges from 8.2 to 11.1 (Table 2); the difference between the minimum and maximum number of species found in a quadrat ranges from 5 to 15 for the five communities; the variance in the number of species per quadrat ranges from 1.41 to 8.6; and the variance–mean ratio in the number of species per quadrat ranges from 0.153 to 0.879. Qualitatively, the communities seem similar in their mean number of species per quadrat but differ in their ranges of species per quadrat, variance in species per quadrat, and variance–mean ratios.

Based upon the variance–mean ratios, S1 is the most uniform of the five communities, and S5 is the most random (Table 2). A strong positive linear relationship exists between true species richness and the variance–mean ratio ($n = 5$, $r = 0.889$, $P = 0.044$), the variance in the number of species per plot ($n = 5$, $r = 0.943$, $P = 0.016$), and the range in the number of species per plot ($n = 5$, $r = 0.938$, $P = 0.018$). There is not a significant relationship between true species richness and the mean number of species per plot ($n = 5$, $r = 0.597$, $P = 0.288$).

FIG. 1. Species richness as a function of sample size for the four estimators of S3: S(sim) = S_{sim}; Jack1 = first-order jackknife; Jack2 = second-order jackknife; Boot = bootstrap. The dashed line represents the true species richness of S3; estimates below the dashed line are negatively biased, and estimates above the dashed line are positively biased.
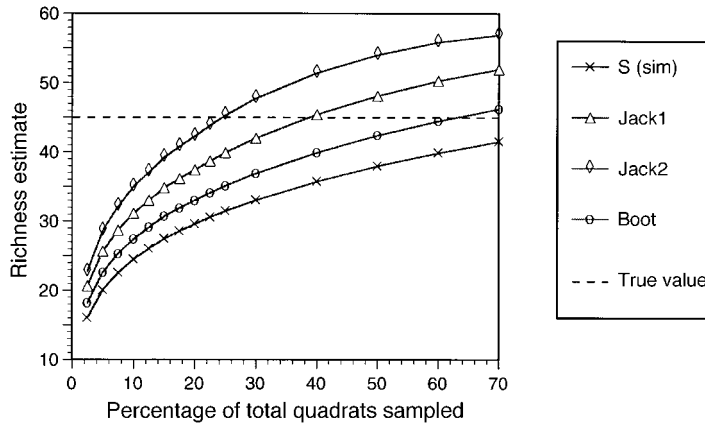
Bias

The bias of an estimator is the difference between the expected value of the estimator and the true species richness value. Positive bias is identical to the overestimation of true species richness, while negative bias is the underestimation of that value. The percent of sampled quadrats at which the first- and second-order jackknife and the bootstrap begin to overestimate true species richness (e.g., Fig. 1) is qualitatively similar for each of the five communities (Table 5). The simple estimator, S_{sim}, yields the largest underestimate of species richness for all sample sizes. It always underestimates species richness with the underestimate decreasing with increasing sample size. The bootstrap estimator yields estimates of species richness uniformly larger than S_{sim} and underestimates species richness until sample sizes are >63–69% of the population, where it begins overestimating. The first-order jackknife estimator yields estimates of species richness uniformly larger than the bootstrap and underestimates species richness until sample sizes are >37–44%. The second-order jackknife estimator yields estimates of species richness uniformly larger than the first-order jackknife estimator and underestimates species richness until samples sizes are >22.5–30%, where this estimator begins to overestimate. Each estimator yields increasing estimates of species richness as sample size increases, and it should be noted that for a given sample size, the magnitude of each estimator is positively correlated with the true species richness values across the five communities.

Precision

S_{sim} is the most precise (least variable) of the four estimators at all sample sizes tested (Fig. 2). The bootstrap estimator is the second most precise estimator followed by the first- and the second-order jackknife, respectively. The general trend for the first-order jackknife, the bootstrap, and S_{sim} in all five communities is a steady but slow decline in variability with an increase in sample size. However, the variance of the second-order jackknife estimator is larger and the trend is more hump-shaped than that of the other three estimators. While the order of precision of the four estimators is the same in all communities, the magnitude of the variance differs across communities (Fig. 2). S1 is the least variable community followed, in increasing order, by S2, S4, S3, and S5.

Accuracy

S_{sim} is the least accurate of all the estimators for all sample sizes <7.5–40%, depending on the community (Fig. 3). The second-order jackknife is the least accurate of the estimators for sample sizes >40%. The second-order jackknife is the most accurate estimator for sample sizes <22.5–25% for S4, S3, and S5 (Fig. 3c–e). The first-order jackknife estimator, however, is the most accurate estimator for sample sizes <12.5% for S1 (Fig. 3a) and for sample sizes <30% for S2 (Fig. 3b). The bootstrap estimator is the most accurate of the estimators for sample sizes >40–50% for S2, S3, S4, and S5. S_{sim}, however, is the most accurate of the estimators for sample sizes >50% for S1. There are mixed results for sample sizes within ~20–60% in all study sites.

The crossover points of the mean square error curves for each of the four estimators describe changes in accuracy for the estimators within a sampled community. The mean square error curve of the second-

TABLE 5. Linear interpolations of the crossover point (i.e., sample size as the percentage of 160 quadrats in plot) where the first- and second-order jackknife and bootstrap estimates begin to overestimate the true species richness.

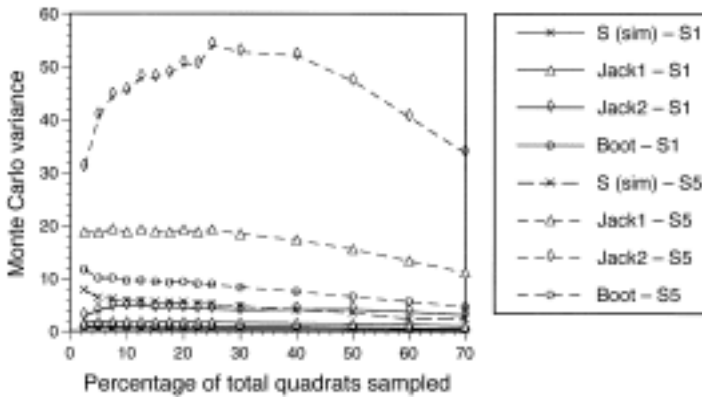| Community	Sample size (%)		
	Second-order jackknife	First-order jackknife	Bootstrap
S1	24.897	41.030	63.901
S2	29.122	43.902	69.025
S4	23.830	39.386	65.728
S3	24.532	38.835	63.095
S5	22.593	36.805	63.464

FIG. 2. Monte Carlo variance (i.e., precision) as a function of sample size for the four estimators of S1 (least variable of the five study sites) and S5 (most variable of the five study sites). Low Monte Carlo variance indicates high precision. See Fig. 1 legend for meanings of abbreviations in key.

order jackknife crosses that of the bootstrap at smaller sample sizes for those communities with smaller values of species richness and smaller variance–mean ratios (Table 2, Fig. 3). For those communities with larger values of species richness and variance–mean ratios, the crossover of the second-order jackknife and the bootstrap occurs at larger sample sizes. A crossover of the second-order jackknife mean square error curve and the first-order jackknife curve occurs only in S4, S3, and S5, the three richest of the five sites (Fig. 3c–e). The sample sizes at which the first-order jackknife crosses and exceeds the bootstrap mean square error curve differs for those communities with small and large species richness and variance–mean ratio values. Again, crossover points occurred at smaller sample sizes for those communities with low richness, while crossover points occurred at larger sample sizes for those communities with higher richness.

DISCUSSION

The desire to assess and describe global biological diversity is a continuing theme in ecology (Magurran 1988). Controversy abounds, however, as to how to go about quantifying biodiversity. Species richness, or the total number of species in a community, is an important measure of biodiversity because it is relatively easy to measure, is comparable across communities, and is well understood by researchers, managers, and the public. Further, species richness is an essential point of consideration in the development and implementation of effective conservation planning. For example, many agree that maximization of total species richness should form one component of our global conservation goals (Hobbs and Lleras 1995).

Given that species richness is both widely used and is an important parameter to measure, researchers and managers need to estimate species richness robustly. Unfortunately, simple species richness estimation depends heavily on sample size. Simple richness estimation, or the enumeration of species in a sample, will result in richness underestimation unless one samples the entire community of interest. Because exhaustive sampling is rarely financially or logistically viable, the

nonparametric jackknife and bootstrap can reduce the biases associated with species richness (Heltshe and Forrester 1983, Smith and van Belle 1984). Our Monte Carlo analysis of the biases, accuracy, and precision of simple species richness, the first- and second-order jackknife, and the bootstrap illustrates that these estimators may have properties that are both dependent and independent of community parameters. An understanding of this dependence and independence will allow selection of the most appropriate estimators and sample sizes for species richness estimation.

Bias

Our study indicates that for any given sample size, there exists a single best estimator with respect to bias. The second-order jackknife estimator is the least biased estimator for sample sizes less than ∼25% in all communities (Table 5, e.g., Fig. 1). The first-order jackknife is the least biased estimator for sample sizes within ∼25–40%. The bootstrap estimator is the least biased estimator for sample sizes within ∼40–65%. The simple estimator, S_{sim}, is the least biased estimator for sample sizes greater than ∼65%. At larger sample sizes, the second-order jackknife, the first-order jackknife, and the bootstrap estimators become positively biased for each of the five communities.

The results presented here are supported by previous research. Smith and van Belle (1984) found, with the use of random spatial distribution models, that when a small number of quadrats are sampled, the expected value of the jackknife is less biased than the bootstrap and the jackknife tends to overestimate for large sample sizes. Colwell and Coddington (1994) also report, for data on germinating rainforest seeds, that the second-order jackknife "clearly provides the least biased estimates" for small sample sizes, followed by the first-order jackknife. Palmer (1991) analyzed richness estimation by random samples of subplots from 30 forested plots in North Carolina and found the second-order jackknife to be the best estimator when compared to the bootstrap and the first-order jackknife with respect to bias. Palmer (1990, 1991) also found
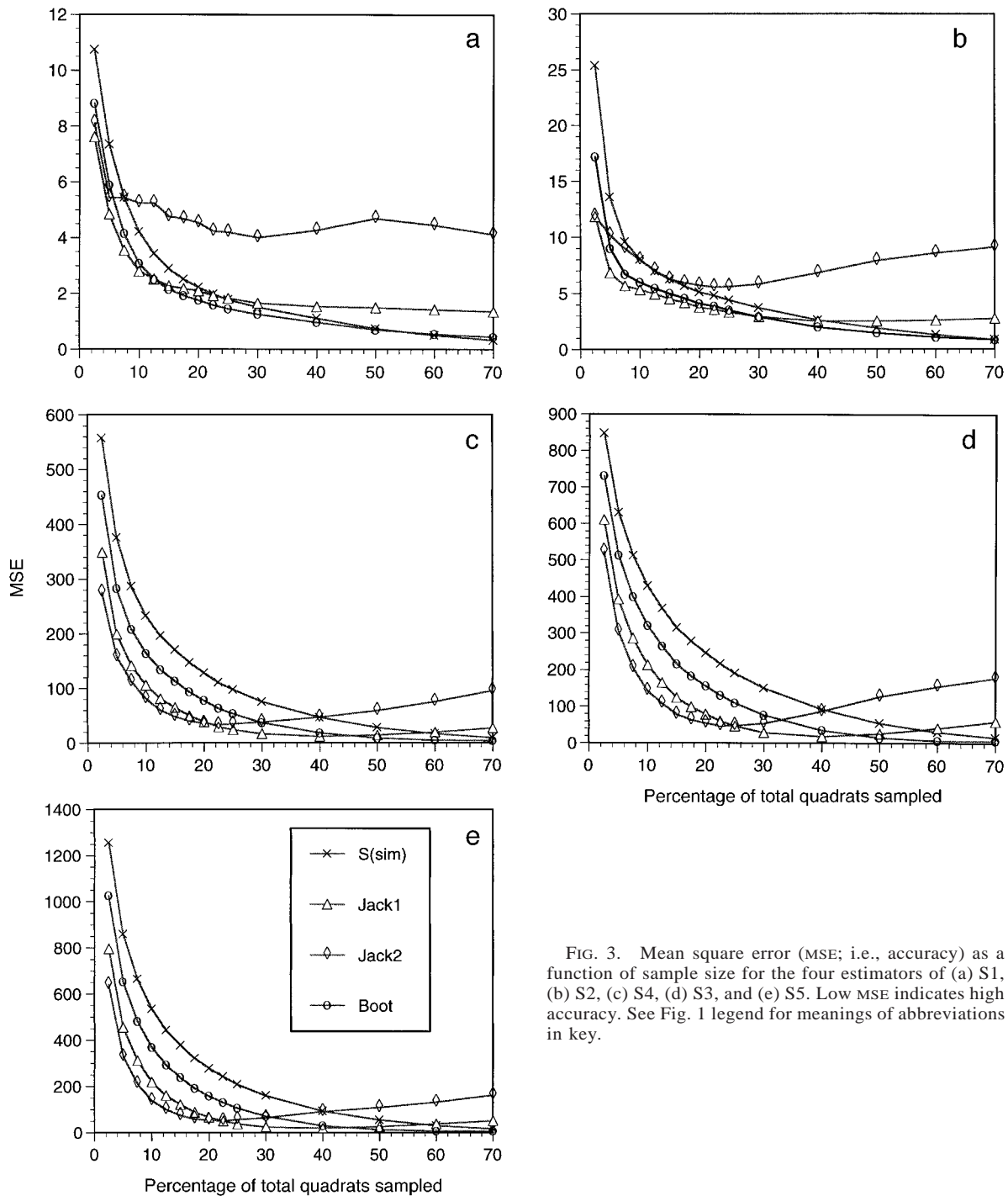
FIG. 3. Mean square error (MSE; i.e., accuracy) as a function of sample size for the four estimators of (a) S1, (b) S2, (c) S4, (d) S3, and (e) S5. Low MSE indicates high accuracy. See Fig. 1 legend for meanings of abbreviations in key.

that simple species richness, the first- and second-order jackknife, and the bootstrap were all highly correlated with the true species richness value of the studied plots and were useful for comparing species richness across communities. In a very recent parasitology study, however, the first-order jackknife performed consistently better than both the second-order jackknife and the

bootstrap, especially at low sample sizes (Walther and Morand 1998). The sum total of our study and previous studies suggest that the jackknife is a less biased estimator than the bootstrap for a variety of study systems. These studies do not seem to agree, however, on the relative performance of the first-order vs. the second-order jackknife.

Precision

For all studied communities and sample sizes, the second-order jackknife estimate is the least precise of the four estimators (e.g., Fig. 2). Further, the imprecision of the second-order jackknife increases for more diverse communities. Though less variable than the second-order, the first-order jackknife is relatively imprecise when compared to the simple estimate, S_{sim}, and the bootstrap. For any single community, the above results are supported by Efron and Gong (1983), who report the high variability of the jackknife relative to the bootstrap. Palmer (1990, 1991) also found the first-order jackknife to be superior to the second-order when precision was required. Palmer, however, found the first-order jackknife to be more precise than the bootstrap. Similarly, Walther and Morand (1998) found the first-order jackknife performed better overall than the bootstrap with respect to precision. In sum, therefore, it seems that little consensus on estimator variability has arisen from a broad collection of estimator studies.

Our study further suggests that for sampling in communities with few species, the relative precision of the four estimators may not be important given the low variability of all estimators in systems with low species richness values (Table 2, Fig. 2). For such communities (perhaps those with <20 species), the difference in variability between the four estimators may be relatively insignificant. On the contrary, the richness estimation precision of ecologically diverse communities may require special attention. In rich systems, the magnitude of the variability of each estimator is large, and the difference between estimators with respect to precision grows with increasing richness.

Accuracy

The complicated pattern of accuracy for the four estimators at variable sample sizes is difficult to interpret, and therefore recommendations of estimator and sample size with respect to accuracy can not be easily made. As sample size increases, a change in relative accuracy among the estimators appears (Fig. 3). The relative accuracy of the measures appears to be dependent upon the total number of species in a community (Table 2). For communities with few total species, the relative accuracy of the estimators becomes ''muddled'' at relatively small sample sizes (as small as 7.5%). For diverse communities, the changes in accuracy among the four estimators occur at larger sample sizes ($\sim$22–25%). At very large sample sizes (>50%), either the bootstrap or the simple estimator, S_{sim}, can be the most accurate. To the best of our knowledge, no other study has explicitly examined trends in estimator accuracy with changes in sample size.

Apparent generalities

In summary, the bias of the four studied estimators appears to be relatively constant across the communities studied here, while precision and accuracy differ across systems as related to the properties of the community. This study confirms the performance of the jackknives and the bootstrap at small sample sizes with respect to bias and indicates that their biases may be largely independent of the properties of a site. The defining characteristics of a community (including the total number of species, the number of rare species, and the pattern of species distribution) do not appear to affect the sample size–estimation functional relationship. Precision among the four estimators within a community also appears independent of community properties, but the magnitude of variability is not independent of the community. The accuracy of the four estimators also appears dependent upon the properties of a community. This accuracy and precision dependency on community illustrates that estimator choices could be made, in part, based upon the expected total number of species.

The generality of the results reported here, both to a range of systems and sampling techniques, is largely unresolved. As studies on estimator performance accumulate in the literature, further conclusions about the robustness of estimator performance across systems can be drawn. Certainly, the procedures discussed here can be applied in a large number of cases (e.g., temporal, spatial, and effort-based sampling; Walther and Morand 1998). These cases might include sampling of animals and plants in both terrestrial and aquatic environments. In addition, there is no reason to believe that the outcomes of this study are unique to our sampling strategy or to northern forests. We emphasize caution, however, as estimator performance is a function of the distribution of species in a system of study (Colwell and Coddington 1994). It remains to be seen whether or not patterns in species presence or abundance across systems might lead to consistent estimator performance.

Other considerations and the politics of biodiversity

Further exploration of the constancy or dependency of bias, precision, and accuracy on ecological parameters could be performed with more extensive computer resampling of data or with simulation of theoretical communities (e.g., Heltshe and Forrester 1983, Smith and van Belle 1984). In particular, we recommend a Monte Carlo analysis of the impact on estimator performance with changes in quadrat shape and size. Changes in the scale of quadrat sampling might complicate some of the generalities uncovered in this study. We also recommend the generation of theoretical communities spanning a greater range of dominance and evenness than the range spanned by the communities studied here. We hypothesize that differences in variability and accuracy may become even more pronounced with strong community contrasts in dominance and evenness.

Finally, this study and others indicate that estimator

and sample size selection is of great importance when estimating species richness. Each field estimation exercise, however, will have specific objectives. Given individual estimation goals, one should consider a priori the desired estimator precision, accuracy, and bias. In addition, there may be political ramifications of richness estimation results. When managers and researchers interpret the biological and political ''meaning'' of an individual richness estimate, a consideration of the underlying biases and properties of the estimator at the selected sample size may be helpful. The two types of richness estimation error, Type A (underestimation) and Type B (overestimation), can be evaluated in the context of the sampling procedure utilized, and the results presented here should be of assistance in that endeavor. The issue, however, as to which of the two error types is more critical is left as a matter for social and political debate.

LITERATURE CITED

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. Scandinavian Journal of Statistics **11**:265–270.

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

Chao, A., and S.-M. Lee. 1992. Estimating the number of classes via sample coverage. Journal of the American Statistical Association **87**:210–217.

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

Efron, B. 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics **7**:1–26.

———. 1981. Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. Biometrika **68**:589–599.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician **37**:36–48.

Gray, H. L., and W. R. Schucany. 1972. The generalized jackknife statistics. Marcel Dekker, New York, New York, USA.

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

Hobbs, R., and E. Lleras. 1995. Protecting and restoring ecosystems, species, populations and genetic diversity. Pages 981–1016 *in* Miller, K., M. H. Allegretti, N. Johnson, and B. Jonsson, editors. Global biodiversity assessment. United Nations Environmental Programme. Cambridge University Press. Cambridge, UK.

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

Magurran, A. E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey, USA.

Miller, R. G. 1974. The jackknife: a review. Biometrika **61**: 1–15.

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

———. 1991. Estimating species richness: the second-order jackknife reconsidered. Ecology **72**:1512–1513.

Peet, R. K. 1974. The measurement of species diversity. Annual Review of Ecology and Systematics **5**:285–307.

Quenoille, M. H. 1949. Approximate testes of correlation in time series. Journal of the Royal Statistical Society, Series B **11**:68–84.

Shannon, C. E., and W. Weaver. 1963. The mathematical theory of communication. University of Illinois Press, Urbana, Illinois, USA.

Simpson, E. H. 1949. Measurement of diversity. Nature **163**: 688.

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

Tukey, J. 1958. Bias and confidence in not quite so large samples. Annals of Mathematical Statistics **29**:614.

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

Zahl, S. 1977. Jackknifing an index of diversity. Ecology **58**:907–913.

APPENDIX

Quadrat-based sampling

In order to use the species richness estimation formulas detailed here, data should be collected in a quadrat-based fashion. The sampled community (i.e., plot) is divided into equally sized quadrats, and a random sample of these quadrats is selected. The number of species is enumerated in each selected quadrat, and the total number of species occurring in all sampled quadrats is determined. (This total richness value is called S_0 in this Appendix and S_{sim} in the body of the paper.) Again, with increases in the percent of total quadrats sampled, simple determinations of species richness (S_0) will approach the true community richness value (i.e., plot value).

Species richness estimators

Species richness estimation formulas given here are in closed form (i.e., single expressions requiring no iteration).

Therefore, computer simulation is not necessary, and the estimates can be calculated using a single simple richness estimation procedure for quadrat-based data.

Jackknife

The closed-form, first-order jackknife estimator of species richness is

$$J_n^1(S) = S_0 + \{r_{1(1)}(n - 1)\}/n$$

where S_0 is the total of the number of species found in a sample, $r_{1(1)}$ is the number of species found in exactly one quadrat, and n is the sample size (Heltshe and Forrester 1983, Smith and van Belle 1984).

The second-order jackknife ($k = 2$) estimator is

$$J_n^2(S) = S_0 + [\{r_{1(1)}(2n - 3)/n\}$$
$$- \{r_{1(2)}(n - 2)^2\}/\{n(n - 1)\}]$$

where $r_{1(1)} = \Sigma_{i=1}^{n} r_{1i}$ is the number of species in only one quadrat, and $r_{1(2)} = \Sigma_{i<1}^{n} r_{1ij}$ is the number of species in only two quadrats (Smith and van Belle 1984, Palmer 1991).

The expression of the kth-order jackknife is

$$J_n^k(S) = S_0 + \frac{1}{k!}\left[\sum_{j=1}^{k} r_{1(j)} \sum_{i=1}^{k} (-1)^{i+1}\binom{k}{i}(n-i)^k \frac{\binom{n-j}{i-j}}{\binom{n}{i}}\right].$$

Bootstrap

The closed-form bootstrap estimate of species richness is

$$B_n(S) = S_0 + \sum_{j=1}^{S_0} (1 - Y_{.j}/n)^n$$

where $Y_{.j}$ is the number of quadrats in which species j is present (Smith and van Belle 1984).