

# Testing species abundance models: a new bootstrap approach applied to Indo-Pacific coral reefs

SEAN R. CONNOLLY,<sup>1,2,3</sup> MARIA DORNELAS,<sup>2</sup> DAVID R. BELLWOOD,<sup>1,2</sup> AND TERENCE P. HUGHES<sup>2</sup>

<sup>1</sup>*School of Marine and Tropical Biology, James Cook University, Townsville, Australia*

<sup>2</sup>*Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Australia*

**Abstract.** Patterns in the commonness and rarity of species are a fundamental characteristic of ecological assemblages; however, testing between alternative models for such patterns remains an important challenge. Conventional approaches to fitting or testing species abundance models often assume that species, not individuals, are the units that are sampled and that species' abundances are independent of one another. Here we test three different models (the Poisson lognormal, the negative binomial, and the neutral, "zero-sum multinomial" [ZSM]) against species abundance distributions of Indo-Pacific corals and reef fishes. We derive and apply several alternative bootstrap analyses of model fit, each of which makes different assumptions about how species abundance data are sampled, and we assess the extent to which tests of model fit are sensitive to such assumptions. For all models, goodness of fit is remarkably consistent, regardless of whether one assumes that species or individuals are the units that are sampled or whether or not one assumes that species' abundances are statistically independent of one another. However, goodness-of-fit estimates are approximately twice as precise and detect lack of model fit more frequently, when based on sampling of individuals, rather than species. Bootstrap analyses indicate that the Poisson lognormal distribution exhibits substantially better fit to species abundance patterns, consistent with model selection analyses. In particular, heterogeneity in species abundances (many rare and few highly abundant species) is too great to be captured by the ZSM model or the negative binomial model and is best explained by models that predict species abundance patterns that are much closer, but not identical, to the lognormal distribution. More broadly, our bootstrap analyses suggest that estimates of model fit are likely to be robust to assumptions about the statistical interdependence of species abundances, but that tests of model fit are more powerful when they assume sampling of individuals, rather than species. Such individual-based tests therefore may be able to identify lack of model fit where previous tests have been inconclusive.

*Key words:* biodiversity; coral reefs; macroecology; neutral theory; parametric bootstrap; species abundance distributions.

## INTRODUCTION

The pattern of commonness and rarity of species is a fundamental attribute of all ecological communities (Magurran 2004). Consequently, such patterns are particularly informative for comparing different communities and for testing the predictions of general models of community structure. At the same time, species abundance distributions contain more information than many other measures of biodiversity (e.g., species counts or evenness metrics), because they incorporate information about species richness, the distribution of rarity, and in some cases the number of unobserved species in a community (McGill et al. 2007).

Although species abundance patterns have engaged ecologists since early in the 20th century (Motomura

1932, Fisher et al. 1943), recently there has been a resurgence of interest in such patterns, as well as in their potential ecological causes. Several alternative models of species abundance have been proposed (see McGill et al. [2007] for an extensive review). Each of these proposed distributions has, in turn, been linked to, or explicitly derived from, one or more ecological mechanism, including niche partitioning (May 1975, Tokeshi 1999), demographic stochasticity (Hubbell 2001, Volkov et al. 2007), environmental stochasticity (May 1975), or some combination thereof (e.g., Engen and Lande 1996a, b, Tilman 2004). For this reason, testing the extent to which alternative species abundance models characterize patterns in the commonness and rarity of species has become increasingly viewed as an important part of assessing the role of different ecological processes in determining community structure (McGill et al. 2007).

There are several challenges associated with testing species abundance models. All such models seek to

Manuscript received 3 October 2008; revised 23 January 2009; accepted 13 February 2009. Corresponding Editor: N. J. Gotelli.

<sup>3</sup> E-mail: Sean.Connolly@jcu.edu.au

explain the same, well-documented patterns (i.e., large numbers of rare species and progressively decreasing numbers of more abundant species), and thus they tend to make very similar predictions. The ability to discriminate between the fit of such similar predictions is exacerbated by the fact that conventional goodness-of-fit tests typically have very low power to identify lack of model fit. Consequently, tests of species abundance models are frequently ambiguous, sensitive to how data are represented (Gray et al. 2006), or sensitive to which goodness-of-fit criterion is used (e.g., McGill 2003, Volkov et al. 2003). Goodness-of-fit tests often fail to rule out several alternative species abundance models (McGill et al. 2007), and comparative measures of model fit often yield conflicting or inconclusive results (e.g., McGill 2003, Volkov et al. 2003, Etienne and Olff 2005). Such tests often implicitly assume that species are the units that are sampled, when in reality individuals are sampled: for instance, in chi-squared tests comparing observed and fitted species abundance distributions, the sample size is considered to be the number of species, regardless of how many individuals have been sampled (e.g., McGill 2003, Volkov et al. 2003). In addition, most fits of species abundance models to data implicitly assume that species' abundances are statistically independent of one another (e.g., Pielou 1977), when in reality there will be some statistical interdependence (the abundances of all species must sum to the total number of individuals sampled). Indeed, the validity of parameter estimates and model comparisons based on the independence assumption has been strongly questioned (Etienne and Olff 2004, Etienne and Alonso 2005). Consequently, an improved understanding of the processes that shape patterns of commonness and rarity requires developing methods of testing alternative species abundance models that are more powerful and that better reflect real sampling processes.

In this study, we develop two new approaches to testing the goodness of fit of species abundance models, and we apply those approaches, alongside more established methods, to quantify the fit of three alternative species abundance models to patterns of commonness and rarity in Indo-Pacific tropical reef corals and fishes. The new approaches treat individuals, rather than species, as the units that are sampled in ecological surveys and make different assumptions about the statistical independence of species' abundances. We compare the results of those tests with approaches that treat species as the sampling units (as conventional tests assume) and with conventional model selection. In addition to assessing model fit for the data as a whole, we exploit spatial replication in the data to compare observed site-to-site variability in model fit with that expected due to stochastic sampling effects alone. We also use spatial replication to identify systematic biases in each species abundance model, to estimate the magnitude of those biases, and to determine whether these biases depend upon the assumptions

about the units of sampling. These analyses provide an assessment of species abundance patterns on Indo-Pacific coral reefs that is much more rigorous and comprehensive than any previously conducted. They also make possible an assessment of the potential biases induced by the assumptions typically made in the analysis of species abundances in general and identify methods of assessing model fit that are robust to violations of those assumptions.

We focus on three species abundance models: an analytical version of the "zero-sum multinomial" distribution predicted by Hubbell's (2001) formulation of neutral theory (Etienne and Alonso 2005); the negative binomial distribution, which is predicted by several models (Fisher et al. 1943, Engen and Lande 1996b, Volkov et al. 2007); and the Poisson lognormal distribution, which is also predicted by several species abundance models (May 1975, Engen and Lande 1996a). Each of these models has been proposed to provide good fit to abundance distributions on coral reefs (Hubbell 1997, Connolly et al. 2005, Volkov et al. 2007). Also, for all of these models, it is possible to devise bootstrap analyses that treat species as the units that are sampled (as traditional statistical approaches do), as well as alternative bootstrap analyses that explicitly characterize the sampling of individuals from an underlying species abundance distribution.

## METHODS

### *Sampling design*

We quantified numerical abundances of scleractinian coral colonies and fishes from the family Labridae (i.e., wrasses and parrotfishes) at sites distributed along a 10 000-km transect across the Pacific Ocean. The sampling design was hierarchical, with three habitat types (crest, flat, and slope) sampled at three spatial scales (regions, islands within regions, and sites within islands). For corals, each site consisted of 10 10-m line-intercept transects, along which all colonies were counted and identified to species. Four sites from each habitat type were established on each island, and three islands were sampled within each of five regions (Indonesia, Papua New Guinea, Solomon Islands, American Samoa, and French Polynesia). For fishes, sites consisted of 20-minute belt transects. The sampling design was otherwise identical to that for corals, except that Micronesia replaced the Solomon Islands (civil unrest prevented sampling in the Solomons), and two islands (rather than three) were established in each region. (See Connolly et al. [2005] for further details of the protocol.)

### *Species abundance models*

If the distribution of species abundances in an assemblage follows a lognormal distribution, then a random sample from that community follows a Poisson lognormal distribution:

$$p_r = \frac{1}{r! \sigma \sqrt{2\pi}} \int_{\lambda=0}^{\infty} \lambda^{r-1} \exp \left\{ -\lambda - \frac{1}{2\sigma^2} \left[ \ln \left( \frac{\lambda}{m} \right) \right]^2 \right\} d\lambda \tag{1}$$

(Pielou 1977), where  $p_r$  is the probability that a species has abundance  $r$  in the sample,  $\lambda$  is the Poisson rate parameter, which follows a lognormal distribution, and  $\ln(m)$  and  $\sigma$  are the mean and standard deviation of  $\ln(\text{abundance})$ . If, on the other hand, abundances in the assemblage follow a gamma distribution, then a random sample from that assemblage follows a negative binomial distribution (Fisher et al. 1943, Pielou 1977):

$$p_r = \frac{\Gamma(r+k)}{r! \Gamma(k)} \frac{m^k}{(1+m)^{r+k}} \tag{2}$$

where  $k$  and  $m$  are the shape and scale parameters, respectively, of the gamma distribution. For any given sample of species abundances, estimates of the parameters of either the Poisson lognormal distribution or the negative binomial distribution can be obtained by finding the values that maximize the log-likelihood for the so-called “zero-truncated” forms of Eqs. 1 and 2:

$$LL = \sum_r n_r \log \left( \frac{p_r}{1-p_0} \right) \tag{3}$$

where  $n_r$  is the number of species with abundance  $r$  (Bulmer 1974).

In this study, the zero-sum multinomial (hereafter ZSM) model is fitted in two different ways. First, the genealogical approach developed by Etienne (2005) is applied: this likelihood constrains the total number of individuals in the sample and thus enforces a strict form of statistical interdependence of species abundances (if an individual of one species is encountered on a transect, the total abundance of the remaining species is reduced by one). Model fit statistics (e.g., maximum log-likelihoods) obtained with this approach are not comparable to those obtained above for the Poisson lognormal and negative binomial models, because the latter models assume statistical independence of species abundances (i.e., the number of individuals sampled from one species does not constrain the number of individuals that can be sampled from other species). Therefore, to obtain fits of the ZSM model in a manner comparable to the other models, a pseudo-likelihood was developed, based on the ZSM model’s predicted species abundance distribution. Under this approach, the probability that a species has abundance  $r$  is

$$p_r = \frac{1}{S_{\text{obs}}} \left\{ \frac{\theta}{(I)_N} \binom{N}{r} \int_0^1 (Ix)_r [I(1-x)]_{N-r} \frac{(1-x)^{\theta-1}}{x} dx \right\} \tag{4}$$

(cf. Etienne and Alonso 2005: Eq. 6), where  $S_{\text{obs}}$  is the number of species in the sample,  $\theta$  is the fundamental biodiversity number,  $I = (m/[1-m])(N-1)$ ,  $m$  is the probability of immigration, and  $N$  is the number of individuals in the sample (Etienne and Alonso 2005).

Estimates of the model parameters  $\theta$  and  $m$  can then be obtained by applying Eq. 3, as for the other two models, with one modification: Eq. 4 refers only to species observed in the sample (rather than to all species in the community, as with Eqs. 1 and 2), so, for this model,  $p_0 = 0$ . Many previous uses of neutral models are based on fits to expected abundance distributions such as Eq. 4 (e.g., McGill 2003, Alonso and McKane 2004, Dornelas et al. 2006, Forster and Warton 2007, Volkov et al. 2007), rather than to a likelihood that constrains total sample size. Therefore, in this study, we explicitly compare parameter estimates obtained using the two approaches, and we assess the likely robustness of model fits to assumptions about the statistical interdependence of species abundances.

Both forms of the ZSM model likelihood frequently exhibit two peaks, one of which corresponds to a high immigration rate ( $m$ ) and another to a low immigration rate (Etienne et al. 2006). Because most reef fishes and corals have larvae that spend days to weeks in the plankton and because our sites encompass relatively small areas (a few hundred square meters), the low-immigration peak is biologically implausible (for our data, this latter peak generally implied <1% of recruits spawned outside the sampling area). Therefore, we initialized our parameter searches to find the higher immigration peaks ( $m > 0.05$  in all cases, typically  $> 0.9$ ).

### Analyses

*Goodness-of-fit statistic.*—Our goodness-of-fit statistic (termed  $\hat{c}$ , following White et al. 2001) is based on model deviance, which is a likelihood-based measure of lack of fit (McCullagh and Nelder 1999). Model deviance measures how far away a model is from fitting the data perfectly. Specifically, it is twice the difference between the maximum log-likelihood for the model and the best possible log-likelihood that could have been attained for the data. Some level of imperfection in model fit is expected due to stochastic sampling effects alone. In other words, the expected deviance for any given model is not zero, even under the null hypothesis that the model is correct. The goodness-of-fit statistic  $\hat{c}$  normalizes model deviance relative to this expected level (e.g.,  $\hat{c} = 1.0$  means that the model’s lack of fit is equal to what one would expect, on average, if the model really did generate the data;  $\hat{c} = 1.05$  means the model’s fit is ~5% worse than expected).

We estimate  $\hat{c}$  by means of a parametric bootstrapping procedure (Efron and Tibshirani 1998). Specifically, we simulate a data set according to the assumptions of a species abundance model and quantify the model’s fit to the simulated data. We then estimate  $\hat{c}$  by dividing the deviance for the real data by the deviance for the simulated data. By repeating this procedure 1000 times, we obtain a bootstrap distribution of  $\hat{c}$  values. If the deviance of the empirical data is consistently higher than that of the simulated data, it indicates that lack of fit of the model is greater than would be expected under the null

hypothesis that the data conform to the model's assumptions. For each bootstrap replicate, an aggregate  $\hat{c}$  is calculated for the data as a whole, using total model deviance obtained by summing the deviances across all sites. We also calculate bootstrap distributions of  $\hat{c}$  for each site individually and characterize the site-to-site variation in this statistic in order to assess whether the observed variation among sites is consistent with what one would expect due to effects of stochastic sampling alone.

*Bootstrap simulations: the Poisson lognormal and the negative binomial.*—In this study, we explicitly compare goodness-of-fit statistics obtained by means of three different bootstrap analyses that differ in two important respects: whether individuals, or species, are the units that are sampled and whether species abundances are sampled independently or a fixed sample size (and thus statistical interdependence of abundances) is imposed. These analyses are explained in detail in Appendix A; what follows here is a summary. The “species-based” bootstrap involves independently sampling an abundance value for each species from a species abundance model. This is consistent with previous uses of the parametric bootstrap in species abundance analysis (Diserud and Engen 2000, Connolly et al. 2005) and also with more conventional goodness-of-fit tests, such as chi-squared analyses (e.g., McGill 2003, Volkov et al. 2003), for which sample size is the number of species, not the number of individuals.

For comparison with the species-based bootstrap, we devise two novel, alternative parametric bootstrap analyses, which we term “Poisson” and “hypergeometric.” Both analyses explicitly simulate, first, the distribution of species abundances in the underlying community, and, secondly, the sampling of individuals from that underlying distribution (see Appendix A for details). The analyses differ in how individuals are sampled. In the Poisson bootstrap, a species abundance in the sample is a Poisson random variable that depends upon an overall sampling intensity and on the species relative abundance in the underlying community. Thus, the Poisson bootstrap simulates a probability model in which species abundances are statistically independent of one another (consistent with the species-based bootstrap and with the likelihood functions used here and in most species abundance analyses). The hypergeometric bootstrap also simulates sampling from an underlying community abundance distribution, but with a fixed total number of individuals in the sample. In this case, statistical interdependence is very strong: if one species is unusually abundant, then the places available for other species are proportionately lower. Thus, the hypergeometric bootstrap simulates a probability model in which the species abundance distribution is conditional upon a fixed total sample size (as in the likelihoods for some neutral models).

*Bootstrap simulations: the zero-sum multinomial.*—For the ZSM model, species-based bootstrap simulations are conducted in exactly the same way as for the Poisson

lognormal and negative binomial models: for each species in the sample, a species abundance is drawn at random from the best-fit species abundance distribution. Because sample size is the sum of these simulated species abundances, this approach leads to stochastic variation in the total number of individuals in the sample, as for the species-based bootstraps of the Poisson lognormal and negative binomial models. For the hypergeometric bootstrap, we use Etienne's (2005) Hoppe urn sampling algorithm, which simulates the sampling of individuals from a ZSM distribution. However, because the ZSM model cannot be used to estimate the size of the species pool (the sample is assumed to be a complete census of the community), it is not possible to devise a Poisson bootstrap for this model.

*Assessing model residuals.*—There are two possible sources of bias in model fit in species abundance analyses. If the true community abundance distribution differs from the distribution predicted by the model, then predicted species abundance patterns will tend to differ systematically from the observed patterns (for instance, by consistently overestimating the number of rare or common species). However, the maximum likelihood method itself can lead to biased parameter estimates and therefore to biases in fitted species abundance distributions. In other words, even if the model is true (e.g., the data really do follow a Poisson lognormal distribution), the maximum likelihood parameter estimates, and thus the best-fit model predictions, may be biased (e.g., Diserud and Engen 2000). The parametric bootstrapping procedure described here can be used to discriminate between these two sources of bias for the species abundance distribution, as follows. For each simulated species abundance distribution, we calculate the difference between the observed and expected fraction of species in each abundance class. If the mean of these residuals differs significantly from zero, it indicates bias due to the method of parameter estimation. One can then compare the distribution of these residuals (across bootstrap simulations) with residuals obtained from model fits to the empirical data. If the empirical residuals lie outside the corresponding bootstrap distributions, it indicates that the empirical species abundance distribution differs systematically from the predictions of the model. By comparing the degree of model bias produced by the Poisson, hypergeometric, and species-based bootstrap analyses, we can also assess whether, and how much, their different assumptions influence the degree of bias in fitted species abundance distributions.

*Comparing alternative models.*—The relative fit of the three alternative models is estimated by calculating model weights according to

$$w_i = \frac{e^{-\Delta\text{MLL}_i}}{\sum_{j \in \text{models}} e^{-\Delta\text{MLL}_j}} = \frac{L_i}{\sum_{j \in \text{models}} L_j} \quad (5)$$

where  $\Delta\text{MLL}_i$  is the difference between maximum log-



TABLE 1. Estimates of goodness of fit ( $\hat{c}$ ) for the Poisson lognormal (PLN), negative binomial (NB), and zero-sum multinomial (ZSM) models fitted to coral and fish data, obtained using the species-based and hypergeometric parametric bootstraps.

Model	Species-based	Hypergeometric	Difference	<i>P</i>
Corals				
PLN	1.04 (1.00, 1.10)	1.07 (1.04, 1.10)	-0.03 (-0.09, 0.03)	0.45
NB	1.19 (1.15, 1.24)	1.17 (1.14, 1.20)	0.02 (-0.04, 0.08)	0.48
ZSM	1.19 (1.14, 1.24)	1.17 (1.14, 1.20)	0.02 (-0.04, 0.08)	0.48
Fishes				
PLN	1.07 (1.00, 1.14)	1.00 (0.96, 1.03)	0.07 (-0.01, 0.15)	0.09
NB	1.04 (0.99, 1.09)	1.00 (0.98, 1.03)	0.04 (-0.02, 0.10)	0.25
ZSM	1.01 (0.96, 1.07)	1.00 (0.97, 1.03)	0.01 (-0.05, 0.07)	0.53

Notes: Values in parentheses are 95% confidence limits, obtained from bootstrap percentiles. "Difference" is the estimated difference between the two  $\hat{c}$  estimates (species-based minus hypergeometric), and the final column is the *P* value testing the null hypothesis that the two estimates are equal.

likelihood of model *i* and the maximum log-likelihood of the best-fitting model and  $L_i$  is the maximum likelihood of model *i*. Because all models have the same number of parameters (two per site), these are equivalent to model weights assigned using Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC) (Boik 2004).

## RESULTS

For the ZSM model, we obtained virtually identical model fits, regardless of whether we used the individual-based, genealogical likelihood (Etienne 2005) or the species-abundance-based likelihood (Eq. 4; Appendix B). Specifically, parameter estimates from the two different likelihood functions were virtually identical; moreover, in the few instances in which differences in best-fit parameters were apparent, this occurred because the log-likelihoods for the two sets of parameters were virtually identical, regardless of which likelihood function was used. Therefore, to allow comparison with the Poisson lognormal and negative binomial models, we use the ZSM model based on the species abundance-based likelihood for all comparisons in this paper.

For the Poisson lognormal and negative binomial models, the two different individual-based bootstrap analyses (hypergeometric and Poisson) yielded virtually identical results, indicating that different assumptions about the statistical independence of species abundances have a negligible impact on our measures of model fit. Consequently, in the following, we present only the results from the hypergeometric analyses. See Appendix C for comparison of results from the hypergeometric and Poisson bootstraps.

*Goodness-of-fit tests.*—Estimates of the magnitude of lack of fit produced by the individual-based and species-based bootstrap analyses were highly consistent. Specifically, for both corals and fishes, the estimated values of  $\hat{c}$  did not differ significantly between individual-based and species-based analyses, for any model (Table 1). However, the individual-based analyses give substantially more precise estimates of model fit: the confidence limits on the goodness-of-fit statistic  $\hat{c}$  were, on average,

about twice as wide for the species-based simulations as for the individual-based simulations. This higher precision of the individual-based approach was highly consistent, producing narrower confidence intervals in every single case: for both taxa, for all models, and for each site considered individually. One consequence of these narrower confidence limits was that the individual-based tests were more likely to detect statistically significant lack of model fit. When goodness-of-fit tests were conducted on each site separately, statistically significant lack-of-model fit ( $\hat{c} > 1$  with >95% confidence) was detected more frequently by the individual-based than the species-based bootstrap analyses for both taxa and all three models: overall, significant lack of fit was found in 23% of cases for individual-based tests vs. 11% of cases for species-based tests.

The individual-based simulations also produced better agreement between observed and predicted among-site variation in model fit, compared with the species-based tests. The species-based simulations typically overestimated the among-site variance in model fit, especially for the Poisson lognormal and negative binomial models, whereas the individual-based simulations typically produced among-site variation in model fit similar to that shown by the empirical data (Figs. 1 and 2).

For corals, goodness of fit of the Poisson lognormal model was significantly better (i.e., closer to 1.0) than the negative binomial or ZSM models, regardless of which bootstrap method was used. In contrast, for fishes, there were no significant differences in goodness of fit between the three models, regardless of bootstrap method (Table 2). For corals, the individual-based bootstrap detected significant lack of model fit ( $\hat{c}$  significantly greater than 1.0) in every case, whereas the species-based analysis detected significant lack of fit only for the negative binomial and ZSM models (Table 1). For fishes, there were no significant differences in goodness of fit between the models, regardless of which bootstrap analysis was used (Table 2), nor did any of the models exhibit statistically significant lack of model fit (Table 1).

*Assessing model residuals.*—Residuals of fitted species abundance distributions illustrate systematic differences

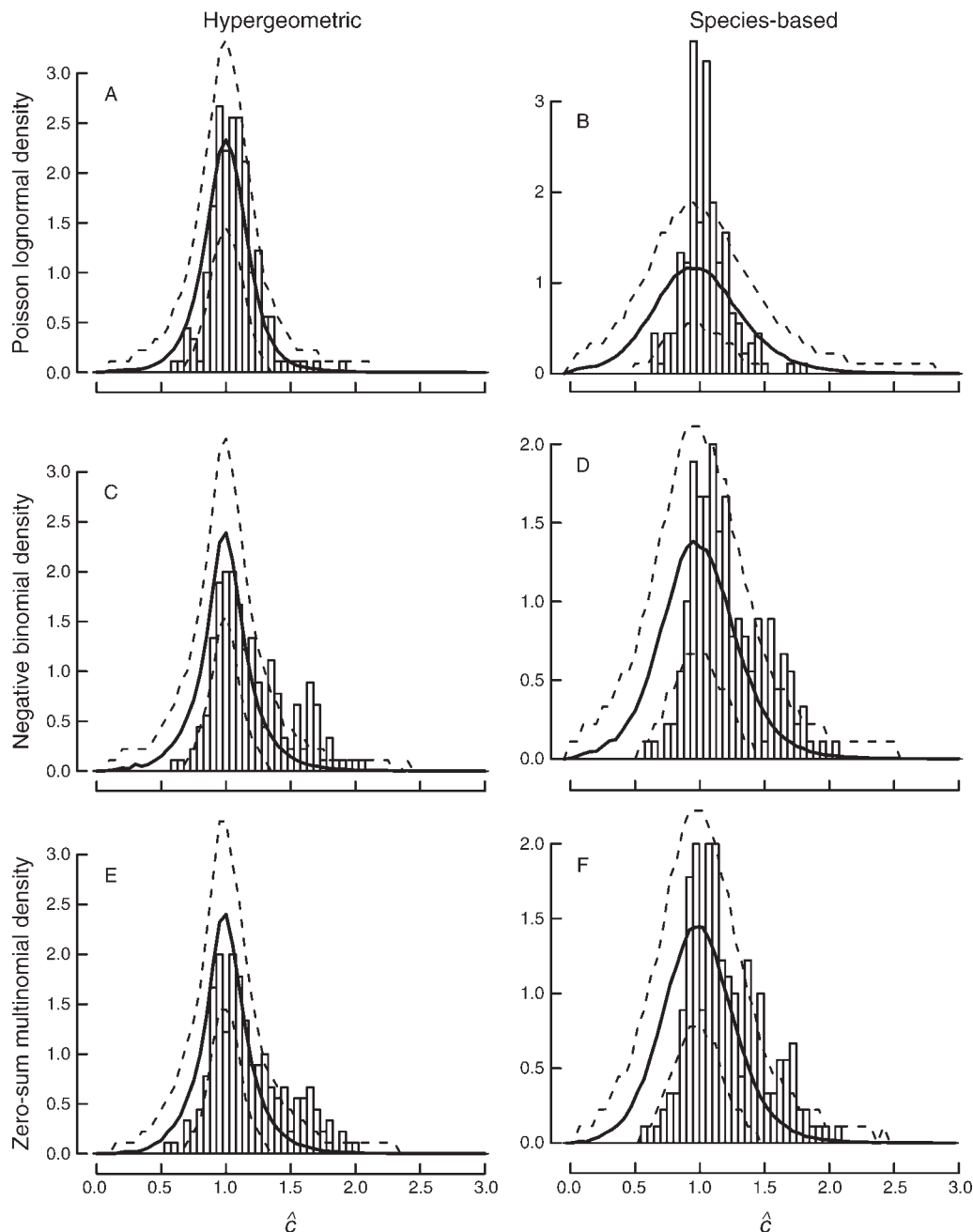


FIG. 1. Frequency distributions of the goodness-of-fit statistic  $\hat{\epsilon}$  across sites for (A, C, E) hypergeometric and (B, D, F) species-based parametric bootstrapping of the (A, B) Poisson lognormal, (C, D) negative binomial, and (E, F) zero-sum multinomial (ZSM) models, for corals. Histograms show the frequency distribution of  $\hat{\epsilon}$  across sites for the actual data. The solid line approximates the expected distribution under the assumptions of the (A, C, E) hypergeometric and (B, D, E) species-based bootstraps (obtained by averaging the frequency distributions from the bootstrap simulations). Dashed lines trace 95% confidence limits, also obtained from the bootstrap simulations.

between the empirical data, the simulated data, and the fitted species abundance models (Fig. 3). These differences were virtually identical for the different bootstrap methods, so we show here only the species-based bootstrap (Fig. 3; see Appendix C for the equivalent results for the individual-based bootstraps). The differences between the observed and predicted species

abundances were calculated for each site, and these residuals were averaged across sites to yield the value plotted as a triangle. Sideways histograms represent the expected distribution of that residual value, obtained from the parametric bootstrap. In other words, the histograms encompass the range of values we would expect, if the data really did conform to the fitted

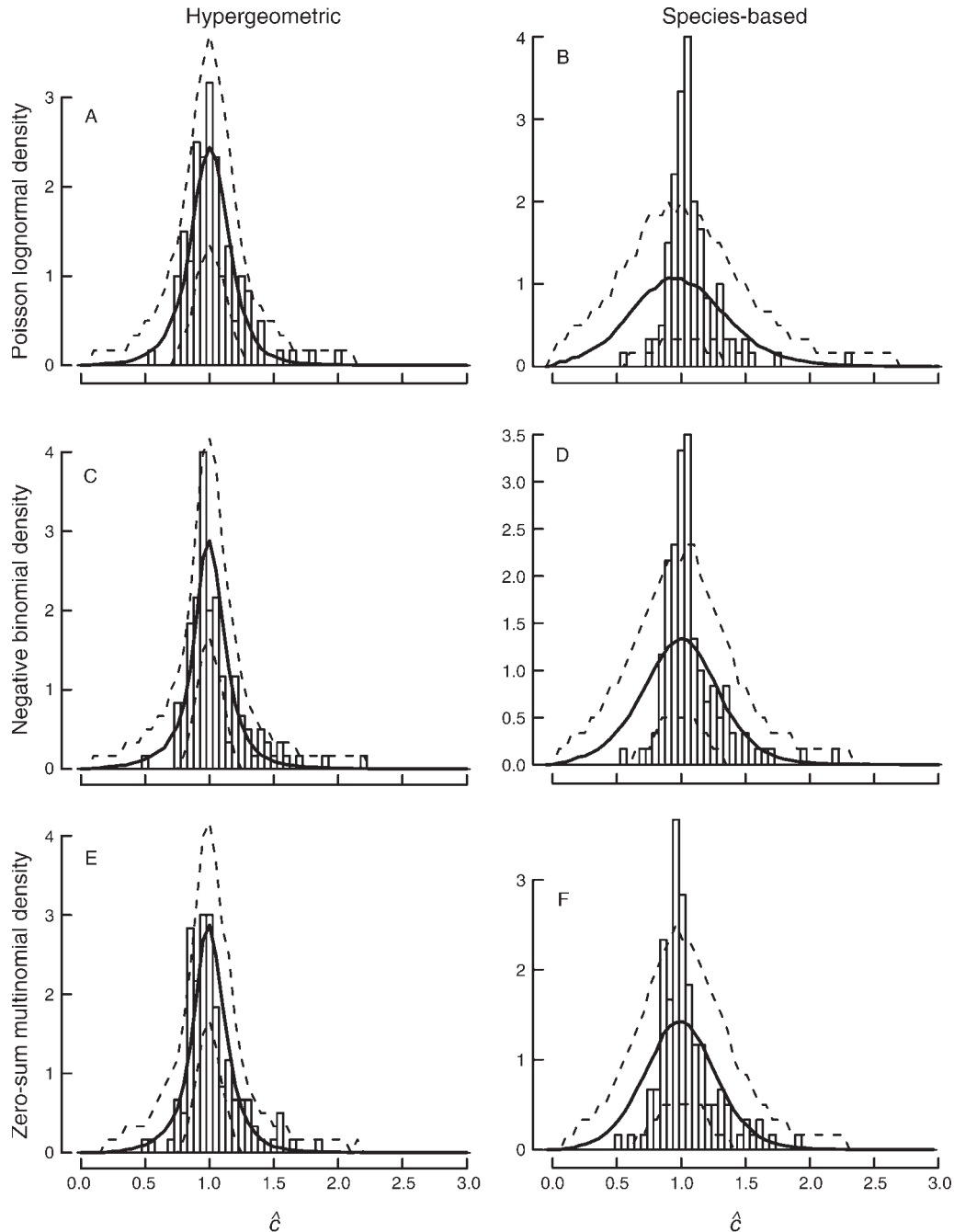


FIG. 2. Frequency distributions of the goodness-of-fit statistic  $\hat{\chi}$  across sites for (A, C, E) hypergeometric and (B, D, F) species-based parametric bootstrapping of the (A, B) Poisson lognormal, (C, D) negative binomial, and (E, F) zero-sum multinomial (ZSM) models, for fishes. Histograms show the frequency distribution of  $\hat{\chi}$  across sites for the actual data. The solid line approximates the expected distribution under the assumptions of the (A, C, E) hypergeometric and (B, D, E) species-based bootstraps (obtained by averaging the frequency distributions from the bootstrap simulations). Dashed lines trace 95% confidence limits, also obtained from the bootstrap simulations.

species abundance model. Thus, where a triangle lies above (below) the corresponding histogram, it indicates that the number of species in that abundance class is greater (less) than predicted by the model. Triangles within the range of the corresponding histograms

indicate that the data are consistent with the species abundance model.

Often, the expected distributions of residuals are not centered on zero, indicating that maximum likelihood estimates of species abundances are somewhat biased,

TABLE 2. Comparison of goodness of fit ( $\hat{c}$ ) of the Poisson lognormal (PLN), negative binomial (NB), and zero-sum multinomial (ZSM) models reported in Table 1.

Comparison	Species-based		Hypergeometric	
	Difference	<i>P</i>	Difference	<i>P</i>
Corals				
PLN vs. NB	-0.15 (-0.22, -0.08)	<0.001	-0.10 (-0.14, -0.06)	<0.001
PLN vs. ZSM	-0.15 (-0.22, -0.08)	<0.001	-0.10 (-0.14, -0.06)	<0.001
NB vs. ZSM	0.00 (-0.07, 0.07)	0.98	0.00 (-0.04, 0.04)	0.89
Fishes				
PLN vs. NB	0.03 (-0.06, 0.12)	0.54	0.00 (-0.04, 0.04)	0.76
PLN vs. ZSM	0.06 (-0.03, 0.15)	0.23	0.00 (-0.05, 0.05)	0.97
NB vs. ZSM	0.03 (-0.05, 0.11)	0.50	0.00 (-0.04, 0.04)	0.78

Notes: Values are estimates of the difference in  $\hat{c}$  values (with 95% CLs): positive values indicate that the first model fits worse than the second model; negative values indicate the converse. The *P* values are for tests of the null hypothesis that the two estimates are equal.

particularly for the Poisson lognormal and negative binomial models (see also Diserud and Engen 2000). However, bias in the fitted species abundance values tends to be small for all models (means of bootstrap distributions, in most cases, differ from zero by ~1% or less). We also find that this bias is virtually identical, regardless of which parametric bootstrap is used to generate the expected distributions (Fig. 3, Appendix C). In other words, bias is not sensitive to the different assumptions of the alternative bootstrap methods.

A comparison of the empirical residuals with their corresponding bootstrap distributions suggests that the Poisson lognormal model exhibits better fit to the empirical abundance distributions overall, for both corals and fishes (Fig. 3). Specifically, for the Poisson lognormal model, observed residuals generally fall within their expected distributions, although there is some suggestion of lack of fit, particularly for corals (triangles near the outer tails of the corresponding bootstrap distributions: Fig. 3A, B). However, the magnitude of lack of fit is exceedingly small: observed residuals deviate from the means of their corresponding bootstrap distributions by ~1% or less. This is consistent with the goodness-of-fit tests, which indicated that lack of model fit was small in magnitude and statistically significant only for corals (Table 1). In contrast, the data deviate much more substantially from the corresponding expected distributions for the negative binomial and ZSM distributions, most strikingly through an excess of singletons (i.e., species represented by a single individual) and a paucity of species with intermediate abundances (Fig. 3). For corals, these results are consistent with the goodness-of-fit tests, in which lack of model fit for the negative binomial model and the ZSM model was statistically significant and substantially greater in magnitude than for the Poisson lognormal model (Tables 1 and 2). For fishes, however, the goodness-of-fit tests do not appear to be sufficiently powerful to detect the lack of model fit apparent from the species abundance residuals.

*Model selection.*—Model selection results are consistent with bootstrap goodness-of-fit tests and the

assessment of species abundance residuals, but model selection distinguishes differences in model fit more strongly than goodness-of-fit tests. For corals, the Poisson lognormal distribution is the best model with nearly 100% confidence ( $\Delta\text{MLL}_{\text{PLN}} = 0$ ,  $w_{\text{PLN}} \approx 1.000$ ;  $\Delta\text{MLL}_{\text{NB}} = 197$ ,  $w_{\text{NB}} \approx 0.000$ ;  $\Delta\text{MLL}_{\text{ZSM}} = 224$ ,  $w_{\text{ZSM}} \approx 0.000$ ), consistent with the finding that its  $\hat{c}$  is significantly and substantially closer to 1.0 than the other two models (Tables 1 and 2). However, model selection also favors the negative binomial model over the ZSM model, whereas neither goodness-of-fit tests nor species abundance residuals clearly distinguish between these two models (Table 1, Fig. 3C, E). For fishes, model selection favored the Poisson lognormal model ( $\Delta\text{MLL}_{\text{PLN}} = 0$ ,  $w_{\text{PLN}} \approx 1.000$ ;  $\Delta\text{MLL}_{\text{NB}} = 43$ ,  $w_{\text{NB}} \approx 0.000$ ;  $\Delta\text{MLL}_{\text{ZSM}} = 48$ ,  $w_{\text{ZSM}} \approx 0.000$ ), consistent with the larger discrepancies between observed and predicted species abundances observed for the negative binomial model and the ZSM model (Fig. 3). As for corals, model selection favored the negative binomial model over the ZSM model, whereas discrepancies between observed and fitted species abundances were similar for those two models.

Modifying our optimization approach for the ZSM model to allow for either high immigration or low immigration peaks in the likelihood function alters the relative performance of the ZSM and the negative binomial models, but the Poisson lognormal model remains the selected best model with ~100% confidence (corals,  $\Delta\text{MLL}_{\text{PLN}} = 0$ ,  $\Delta\text{MLL}_{\text{ZSM}} = 172$ ,  $\Delta\text{MLL}_{\text{NB}} = 197$ ; fishes,  $\Delta\text{MLL}_{\text{PLN}} = 0$ ,  $\Delta\text{MLL}_{\text{ZSM}} = 12$ ,  $\Delta\text{MLL}_{\text{NB}} = 43$ ).

DISCUSSION

*Commonness and rarity on coral reefs*

The combination of approaches used here (individual-based and species-based goodness-of-fit tests, analysis of species abundance residuals, and model selection) strongly indicates that the Poisson lognormal model provides a better fit to local species abundance patterns than the negative binomial model or the ZSM model.



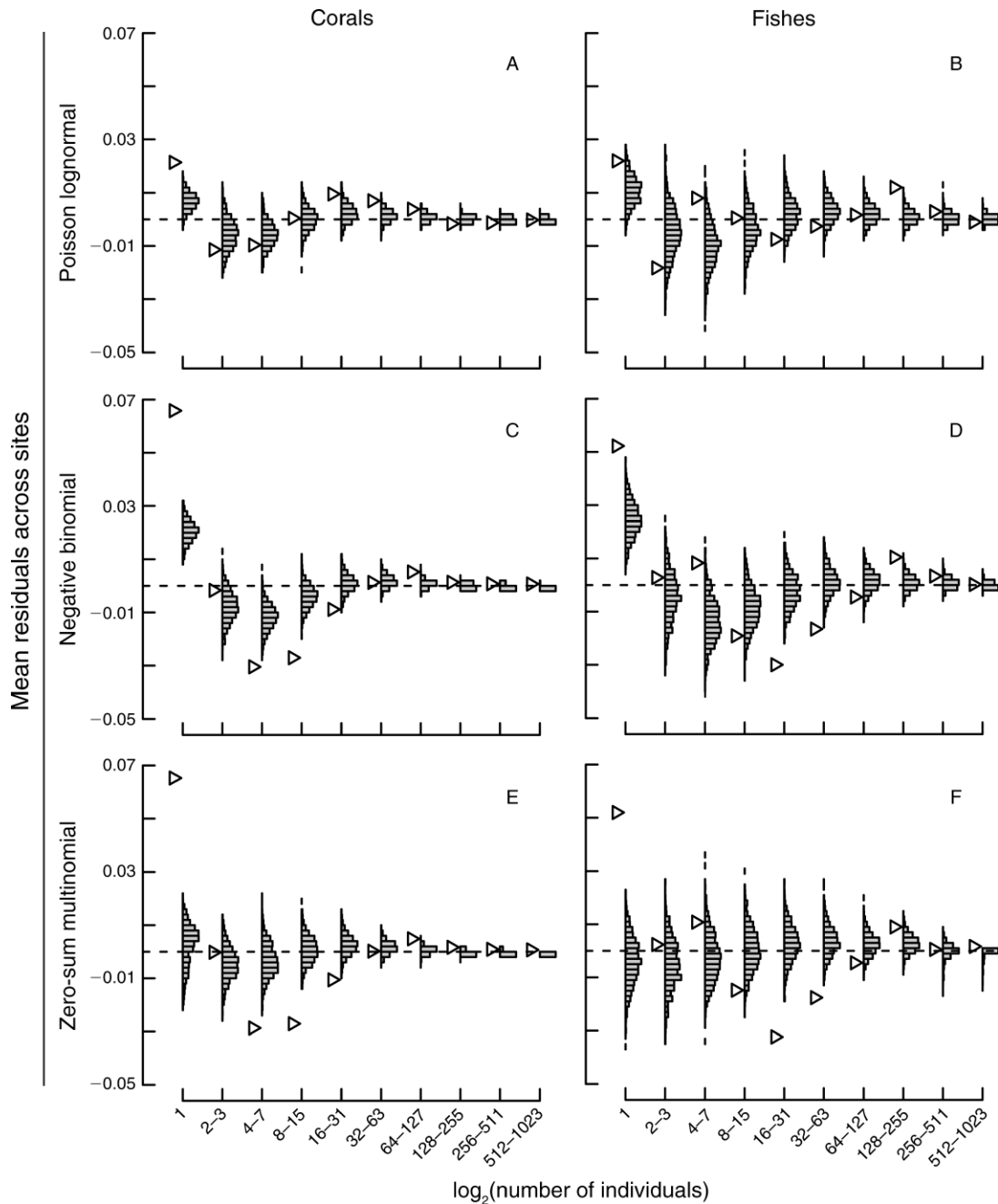


FIG. 3. Comparison of observed and expected species abundance distributions for the (A, B) Poisson lognormal, (C, D) negative binomial, and (E, F) zero-sum multinomial models for (A, C, E) corals and (B, D, F) fishes, for the species-based bootstrap. Open triangles represent mean residuals across sites (observed – expected) for each octave (i.e., the difference between observed and fitted frequency of singletons is calculated for each site and then averaged across all sites to obtain the value plotted as an open triangle). The sideways histograms show the expected distribution of residuals, obtained from fits to data simulated using the species-based bootstrap.

For both corals and fishes, likelihood-based model selection strongly favored the Poisson lognormal model over the negative binomial model and the ZSM model. Similarly, application of parametric bootstrapping to identify systematic patterns in species abundance residuals revealed greater discrepancies between fitted models

and data for the negative binomial model and the ZSM model, compared to the Poisson lognormal model. Likewise, for corals, when bootstrap analyses were used to calculate goodness-of-fit statistics for the data considered as a whole, our estimates of lack of model fit were approximately three-fold greater (i.e.,  $\hat{c}$  three

times farther away from 1.0) for the negative binomial model and the ZSM model than for the Poisson lognormal model. However, for fishes, our goodness-of-fit tests were not sufficiently powerful to detect significant lack of model fit for any of the models, even though the species abundance residuals suggested that lack of fit was present, at least for the negative binomial and the ZSM models.

Inspection of species abundance residuals suggests that the superior fit of the Poisson lognormal model is due to its thicker tail of highly abundant species, relative to the negative binomial and the ZSM models. For this reason, it accurately captures the combination of many rare species and a few highly abundant species that characterize our coral reef communities. Comparatively thin-tailed distributions, such as the negative binomial and the ZSM models, are unable to do this as accurately. To capture the thick tail of highly abundant species, they systematically underpredict the number of very rare species and overpredict the number of moderately rare species (Fig. 3C–F). Because neutral models, such as those generating negative binomial or ZSM abundance distributions, specifically assume demographic equivalence of species, there is likely to be a limit to which such models can capture the simultaneous presence of many very rare species and a few highly abundant species (although this is likely to depend to some extent on ancillary assumptions, such as the speciation process [Etienne et al. 2007]). Consequently, explaining species abundance patterns with high heterogeneity, such as observed in this study, may well require the incorporation of differences in species' niche space or demographic rates.

Despite the very close agreement between the Poisson lognormal model and the coral and fish data, there were statistically significant differences between the data and the model for corals. There are two potential reasons for this, related, respectively, to each of the two core assumptions of the Poisson lognormal model: lognormally distributed abundances in the community and Poisson sampling of individuals from that distribution (i.e., all individuals sampled independently and at random with respect to species). Non-Poisson sampling of individuals can arise due to phenomena such as local aggregation effects (e.g., Karlson et al. 2007, Dornelas and Connolly 2008) or interspecific variation in detectability (due, for instance, to differences in colony size [corals] or behavioral responses to observers [fishes]). Such effects are likely to be present to some degree in most, if not all, species abundance data.

Another possibility is that the significant lack of model fit occurs because the underlying distribution of abundances in the community deviates to some degree from the lognormal. Such small discrepancies are consistent with the two most general biological explanations offered for lognormal species abundance distributions. Specifically, lognormal abundance distributions arise as a limiting case when abundance is determined by the multiplicative interaction of many stochastic factors,

such as annual population growth rates or niche widths along multiple niche axes (May 1975). Similarly, the lognormal distribution may arise as a second-order mathematical approximation for the action of an arbitrary number of “non-neutral” factors such as partial niche overlap, environmental stochasticity, or interspecific differences in demographic rates (Pueyo 2006). Under both explanations, the lognormal should provide a good approximation for most species abundance distributions (coral reef and otherwise), but small departures from the lognormal should be detectable given a sufficiently powerful test. Alternatively, the discrepancies we found may be caused by samples from communities that have multi-modal abundance distributions. For instance, a very large survey (~40 000 colonies) of a single coral community at Lizard Island, Great Barrier Reef, is best fit by a mixture of multiple Poisson lognormal distributions (Dornelas and Connolly 2008). A sample of several hundred individuals (typical for our sites) from such a community would be dominated by the distribution with highest mean abundance and thus might well be closely fit by a single Poisson lognormal model. However, some species from the rarer distributions would still be sampled occasionally, and this could potentially cause an excess of singletons in the data, compared to a single Poisson lognormal model.

#### *Approaches to testing species abundance models*

There is now a broad consensus that prevailing approaches to testing species abundance models are weak and prone to ambiguous results (Alonso et al. 2006, Gray et al. 2006, McGill et al. 2007). Like the species-based bootstrap test used here, conventional tests for species abundance distributions, such as chi-squared statistics, implicitly assume that species are the units that are sampled; they treat species' abundances as statistically independent of one another; and they typically also require binning of data into categories of abundance to satisfy sample size requirements (e.g., McGill 2003, Volkov et al. 2003). Intuitively, one might expect that such assumptions about sampling of species will overestimate the variability in model fit due to stochastic sampling effects, because there are typically substantially fewer species than individuals in ecological samples. This intuition is borne out by the bootstrap analyses conducted here. The individual-based hypergeometric and Poisson bootstraps explicitly characterize the sampling of individuals from an underlying abundance distribution. Data simulated using this approach exhibit site-to-site variability in model fit that is in relatively close accord with that of the real data (Figs. 1 and 2A, C, E). In contrast, the species-based bootstrap simulations generate site-to-site variability in model fit that is typically substantially greater than that of the empirical data, especially for the Poisson lognormal and negative binomial models (Figs. 1 and 2B, D). This greater variability leads to less-precise goodness-of-fit tests: the confidence limits on  $\hat{c}$  are

substantially and very consistently (i.e., for corals, fishes, all models, and even all sites individually) broader for the species-based bootstrap compared to the hypergeometric and Poisson bootstraps.

The greater precision with which the individual-based bootstraps estimate goodness of fit implies that they are likely to provide more powerful tests of model fit than the species-based approach, at least for species abundance patterns similar to our data. For our goodness-of-fit statistic, statistically significant lack of fit is detected when 95% confidence limits on  $\hat{c}$  do not overlap 1. Our analyses indicate that  $\hat{c}$  does not differ between the individual-based and species-based approaches (consistent with the assumption that the two approaches are estimating the same quantity) and that the individual-based estimate has narrower confidence limits. If so, the individual-based approach should identify lack of model fit more frequently than the species-based approach, as our analyses of the coral and fish data show.

Our results also indicate that the traditional approach to fitting species abundance models is probably robust to statistical interdependence of species abundances. The overwhelming majority of fits of species abundance models in the literature make the implicit assumption that species abundances are sampled independently of one another; indeed, the prevalence of this assumption has led some workers to formulate alternative likelihoods, which are explicitly conditional on the total number of individuals sampled (e.g., Dewdney 1998, Etienne 2005, Etienne and Olff 2005). Here, rather than formulating alternative likelihoods, we focused on quantitatively assessing whether and how much the independence assumption of the more conventional fitting approach might bias fits of species abundance models to data, when the independence assumption is violated in the data. Specifically, the Poisson bootstrap simulates each species abundance independently, consistent with the assumptions of the likelihoods used in our model fitting. Conversely, the hypergeometric bootstrap enforces a fixed total sample size, so every time an individual of one species is recorded, this reduces, by one individual, the number of individuals that can belong to one of the other species. The fact that virtually identical results were obtained from these two different bootstraps (Appendix C) indicates that the classical likelihood-based approach to fitting species abundance models is unlikely to be biased by statistical interdependence of species abundances.

In reality, typical ecological sampling probably falls between the extremes of Poisson and hypergeometric sampling. For instance, space occupied by one coral species reduces space available to colonies of other species, but, because space is typically far from saturated by corals, an increase in larval supply can lead to more colonies of one species without displacing other species' colonies. We know of no statistical models for species abundance distributions that impose such an intermediate form of statistical dependence, so the apparent robustness

of standard model-fitting approaches to the extremes of hypergeometric and Poisson sampling is encouraging.

### Conclusions

The analyses in this paper confirm that the Poisson lognormal model provides better fit to patterns of commonness and rarity of corals and reef fishes than either the negative binomial model or ZSM. This better fit appears to be due to the fact that the comparatively thin tails of the latter models limit their ability to capture the combination of many very rare and few highly abundant species, a pattern that characterizes patterns of commonness and rarity for both corals and fishes. Moreover, assessments of model fit were robust to two important kinds of assumptions about sampling that are typical of species abundance models, but likely to be unrealistic: species, not individuals, are the units that are sampled; and species' abundances are either statistically independent of one another or strictly interdependent. Nevertheless, we found one important difference between species-based and individual-based model tests: the latter predict less stochastic variability in model fit than the former. This lower level of variability was more consistent with empirical site-to-site variation in model fit and yielded substantially more powerful tests of model fit, in individual-based tests.

More broadly, our individual-based bootstrap analyses can be adapted for any species abundance model that characterizes sampling from an underlying community distribution of species abundances, as the Poisson lognormal and negative binomial models do. Where there is bias in parameter estimates and predicted species abundances (as is likely to be the case for most species abundance models, given their highly asymmetric shapes), residuals from such bootstrap analyses can estimate the magnitudes of such biases (e.g., sideways histograms in Fig. 3), allowing an assessment of the extent to which discrepancies between model fits and data can be attributed to such statistical artifacts and how much to real departures from a species abundance model. Moreover, the approach makes possible a goodness-of-fit test that appears to be substantially more powerful than tests that treat species as the units that are sampled. Thus, this approach may yield new insights where previous tests of model fit have been inconclusive and, more broadly, help to resolve the problem of low statistical power that has plagued the analysis of patterns of commonness and rarity for decades (McGill et al. 2007).

### ACKNOWLEDGMENTS

We thank numerous staff, students, and volunteers at the Australian Research Council Centre of Excellence for Coral Reef Studies for assistance with data collection; M. Hisano for mathematical programming; R. Karlson, K. Roy, R. Etienne, and an anonymous reviewer for helpful comments on various drafts; and R. Etienne for MATLAB and PARI-GP code to fit the genealogical ZSM. This work was supported by the Australian Research Council, the National Geographic Society, and James Cook University.

## LITERATURE CITED

- Alonso, D., R. S. Etienne, and A. J. McKane. 2006. The merits of neutral theory. *Trends in Ecology and Evolution* 21:451–457.
- Alonso, D., and A. J. McKane. 2004. Sampling Hubbell's neutral theory of biodiversity. *Ecology Letters* 7:901–910.
- Boik, R. J. 2004. Why likelihood? Commentary. Pages 167–180 in M. L. Taper and S. R. Lele, editors. *The nature of scientific evidence: statistical, philosophical, and empirical considerations*. University of Chicago Press, Chicago, Illinois, USA.
- Bulmer, M. G. 1974. On fitting the poisson lognormal distribution to species-abundance data. *Biometrics* 30:101–110.
- Connolly, S. R., T. P. Hughes, D. R. Bellwood, and R. H. Karlson. 2005. Community structure of corals and reef fishes at multiple scales. *Science* 309:1363–1365.
- Dewdney, A. K. 1998. A general theory of the sampling process with applications to the “veil line.” *Theoretical Population Biology* 54:294–302.
- Diserud, O. H., and S. Engen. 2000. A general and dynamic species abundance model, embracing the lognormal and the gamma models. *American Naturalist* 155:497–511.
- Dornelas, M., and S. R. Connolly. 2008. Multiple modes in a coral species-abundance distribution. *Ecology Letters* 11:1008–1016.
- Dornelas, M., S. R. Connolly, and T. P. Hughes. 2006. Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440:80–82.
- Efron, B., and R. J. Tibshirani. 1998. *An introduction to the bootstrap*. Chapman and Hall, Boca Raton, Florida, USA.
- Engen, S., and R. Lande. 1996a. Population dynamics models generating the lognormal species abundance distribution. *Mathematical Biosciences* 132:169–183.
- Engen, S., and R. Lande. 1996b. Population dynamics models generating species abundance distributions of the gamma type. *Journal of Theoretical Biology* 178:325–331.
- Etienne, R. S. 2005. A new sampling formula for neutral biodiversity. *Ecology Letters* 8:253–260.
- Etienne, R. S., and D. Alonso. 2005. A dispersal-limited sampling theory for species and alleles. *Ecology Letters* 8:493–504.
- Etienne, R. S., M. E. F. Apol, H. Olf, and F. J. Weissing. 2007. Modes of speciation and the neutral theory of biodiversity. *Oikos* 116:241–258.
- Etienne, R. S., A. M. Latimer, J. A. Silander, Jr., and R. M. Cowling. 2006. Comment on “Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot.” *Science* 311:610b.
- Etienne, R. S., and H. Olf. 2004. A novel genealogical approach to neutral biodiversity theory. *Ecology Letters* 7:170–175.
- Etienne, R. S., and H. Olf. 2005. Confronting different models of community structure to species-abundance data: a Bayesian model comparison. *Ecology Letters* 8:493–504.
- Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42–58.
- Forster, M. A., and D. I. Warton. 2007. A metacommunity-scale comparison of species-abundance distribution models for plant communities of eastern Australia. *Ecography* 30:449–458.
- Gray, J. S., A. Bjørgesæter, and K. I. Ugland. 2006. On plotting species abundance distributions. *Journal of Animal Ecology* 75:752–756.
- Hubbell, S. W. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* 16:S9–S21.
- Hubbell, S. W. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, New Jersey, USA.
- Karlson, R. H., H. V. Cornell, and T. P. Hughes. 2007. Aggregation influences coral species richness at multiple spatial scales. *Ecology* 88:170–177.
- Magurran, A. E. 2004. *Measuring biological diversity*. Blackwell, Oxford, UK.
- May, R. M. 1975. Patterns of species abundance and diversity. Pages 81–120 in M. L. Cody and J. M. Diamond, editors. *Ecology and evolution of communities*. Harvard University Press, Cambridge, Massachusetts, USA.
- McCullagh, P., and J. A. Nelder. 1999. *Generalized linear models*. Second edition. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- McGill, B. J. 2003. A test of the unified neutral theory of biodiversity. *Nature* 422:881–885.
- McGill, B. J., et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10:995–1015.
- Motomura, I. 1932. On the statistical treatment of communities. *Zoological Magazine, Tokyo* 44:379–383.
- Pielou, E. C. 1977. *Mathematical ecology*. John Wiley and Sons, New York, New York, USA.
- Pueyo, S. 2006. Diversity: between neutrality and structure. *Oikos* 112:392–405.
- Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences (USA)* 101:10854–10861.
- Tokeshi, M. 1999. *Species coexistence: ecological and evolutionary perspectives*. Blackwell, Oxford, UK.
- Volkov, I. J., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. *Nature* 424:1035–1037.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2007. Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450:45–49.
- White, G. C., K. P. Burnham, and D. R. Anderson. 2001. Advanced features of Program MARK. Pages 368–377 in R. Field, R. J. Warren, H. Okarma, and P. R. Sievert, editors. *Wildlife, land, and people: priorities for the 21st century. Proceedings of the Second International Wildlife Management Congress*. The Wildlife Society, Bethesda, Maryland, USA.

## APPENDIX A

Parametric bootstrap algorithms (*Ecological Archives* E090-225-A1).

## APPENDIX B

Comparison of alternative zero-sum multinomial fits (*Ecological Archives* E090-225-A2).

## APPENDIX C

Poisson vs. hypergeometric bootstrap results (*Ecological Archives* E090-225-A3).