LIFTING A VEIL ON DIVERSITY: A BAYESIAN APPROACH TO FITTING RELATIVE-ABUNDANCE MODELS

Duncan J. Golicher,[1,4] Robert B. O'Hara,[2] Lorena Ruíz-Montoya,[1] and Luis Cayuela[3]

[1]*Departamento de Ecología y Sistemática Terrestre, El Colegio de la Frontera Sur, Carretera Panamericana y Periférico Sur s/n, C.P. 29290, San Cristóbal de Las Casas, Chiapas México*
[2]*Department of Mathematics and Statistics, P.O. Box 68, (Gustaf Hällströmin kaku 2b), FIN-00014 University of Helsinki, Helsinki, Finland*
[3]*Departamento de Ecología, Universidad de Alcalá, Carretera de Barcelona km. 33,600, C.P. 28871 Alcalá de Henares, Madrid, Spain*

Abstract. Bayesian methods incorporate prior knowledge into a statistical analysis. This prior knowledge is usually restricted to assumptions regarding the form of probability distributions of the parameters of interest, leaving their values to be determined mainly through the data. Here we show how a Bayesian approach can be applied to the problem of drawing inference regarding species abundance distributions and comparing diversity indices between sites. The classic log series and the lognormal models of relative- abundance distribution are apparently quite different in form. The first is a sampling distribution while the other is a model of abundance of the underlying population. Bayesian methods help unite these two models in a common framework. Markov chain Monte Carlo simulation can be used to fit both distributions as small hierarchical models with shared common assumptions. Sampling error can be assumed to follow a Poisson distribution. Species not found in a sample, but suspected to be present in the region or community of interest, can be given zero abundance. This not only simplifies the process of model fitting, but also provides a convenient way of calculating confidence intervals for diversity indices. The method is especially useful when a comparison of species diversity between sites with different sample sizes is the key motivation behind the research. We illustrate the potential of the approach using data on fruit-feeding butterflies in southern Mexico. We conclude that, once all assumptions have been made transparent, a single data set may provide support for the belief that diversity is negatively affected by anthropogenic forest disturbance. Bayesian methods help to apply theory regarding the distribution of abundance in ecological communities to applied conservation.

Key words: Bayesian approach; butterflies; diversity indices; gamma distribution; hierarchical model; lognormal distribution; Markov-chain Monte Carlo simulation; Montebello National Park, Chiapas, Mexico; Poisson; relative abundance; species abundance distribution.

INTRODUCTION

Ecology is a naturally uncertain discipline in which data sets hold ambiguous messages. Applying appropriate methods of statistical analysis and interpreting their results is a continuing challenge. Statistical artifacts can, at times, lead to lack of consensus over fundamental tenets of ecological theory, such as the relationship between diversity and disturbance (Oksanen 1996). Observational data cause particular difficulties. Recently the traditional reliance on null-hypothesis testing as an automated inferential technique has been questioned (Anderson et al. 2000, Johnson 2003). Alternative forms of drawing inference that focus attention on model comparison, rather than evaluating evidence concerning inconsequential null models, have been advocated (Guthery et al. 2003). The Bayesian

approach to inference is attractive due to its simplicity and consistency (Lindley 2000). Bayesian analyses are set up in terms of evaluating the strength of evidence in favor of meaningful statements given a set of clearly stated prior assumptions. The focus is on providing the distribution for parameters of interest (Ellison 1996), conditional on these assumptions. The Bayesian approach can clarify how inference is being achieved and show how uncertainty associated with predictions enters an analysis. An emphasis on multiple working hypotheses helps to make data analyses more informative with respect to the real ecological questions under consideration (Hilborn and Mangel 1997).

Community ecologists have been slow to explore the practical potential of Bayesian methods, even though they are now routinely used in other disciplines to address delicate inferential problems where objectivity is a strict requirement (Burton et al. 1998, Meyer and Millar 1999). This may in part be attributed to the paucity of published works showing applications in a realistic ecological context. This situation is exacer-

bated by the relative inaccessibility of the specialized statistical literature. To begin to fill this gap we here provide a simple example of how a Bayesian analysis can address one of the commonest inferential problems met by field ecologists, that of comparing species diversity between areas where sample sizes vary.

There is a bewildering array of indices of species diversity. Either implicitly or explicitly they are based on models of the relative abundance of species. The differences between them revolve around the emphasis each places on richness and equitability. The most practical use of diversity indices is to make comparisons between sites. However, in order to do this, sampling effects must be accounted for (Gotelli and Colwell 2001). This can be confusing and easily lead to errors of interpretation. Although statistics cannot remove bias caused by poor sampling design, adopting a Bayesian approach allows the linkage between species-diversity indices and models of relative abundance to be made explicit.

Parametric models of relative abundance are typically used less often than are diversity indices by field ecologists. This seems to be because a conventional parametric form for an underlying model has never been fixed. To explain why, we briefly outline the historical development of two classical models of relative abundance: the log series and the lognormal. The log-series model was published in 1943 (Fisher et al. 1943). Fisher proposed that this could be used to describe the apparent abundance of different species. He derived it as a limiting form of the negative binomial distribution by assuming that the underlying abundances have a gamma distribution. Later, Preston (1948) suggested that the underlying abundance distribution could be lognormally distributed. This distribution has considerable theoretical attraction. If many small effects act multiplicatively, as occurs when populations fluctuate, then a lognormal distribution automatically arises. It has been shown that a few simple assumptions regarding the properties of neutral models of community structure lead to distributions with a similar form to the lognormal (Sugihara 1980, Bell 2000, Hubbell 2001). While these distributions are often subtly different from a lognormal, they are usually similar enough to be indistinguishable in practice (McGill 2003).

Despite the virtues of simplicity, communicability, and theoretical backing, the lognormal distribution has a most unfortunate weakness. The empirical abundance distributions of samples of individuals rarely look anything like it (Dewdney 1998). Instead of the expected unimodal shape of the histogram of logged abundances, relative abundances in samples tend to form a monotonically declining inverse ''J'' curve. This consistent empirical lack of fit would be expected to rule out a model. However in this particular case it does not. Lack of fit is almost inevitable due to sampling effects. Under the lognormal most species are comparatively rare. As

ecological samples are not usually large enough to include more than half the potential species, the rest of the distribution is hidden in the non-observed species. Although rare species are arguably the most important components of diversity, their abundances are inevitably poorly typified (Pielou 1977). The difficulty becomes even more acute as species richness increases. This is one of the greatest challenges that any study of diversity has to confront. Sound inference seems hardest to achieve in exactly the circumstances in which it is most desirable. Those (of us) whose field work is in the poorly studied, yet species-rich, tropics face by far the most difficult situation.

Inference is impossible without assumptions. Preston (1948) did suggest a reasonable one. The lognormal distribution could be truncated using a ''veil line.'' In other words, Preston assumed that species with an expected abundance lower than 1 would not be found in a sample. This straightforward idea was a sensible short cut for its time. However it is not mathematically justifiable (Dewdney 1998). The abundances of the rarer species in the sample inevitably cross the line in both directions. A superior solution is to return to Fisher's original derivation of the log series while assuming that the Poisson model of sampling also applies to the underlying lognormal (Cassie 1962). Once this is done the two seemingly competing models become rather similar in their underlying structure (Dewdney 1998). The so-called ''Poisson lognormal distribution'' is very well known to statisticians, but has been underused by ecologists, because of the technical nature of model fitting using maximum likelihood (Kempton and Taylor 1974, Hayek and Buzas 1997).

The development of the Markov-chain Monte Carlo (MCMC) simulation removes many of the technical problems involved in fitting models involving distributions with more than one level of stochastic variability. Modern Bayesian methods rely on using MCMC to simulate the posterior distribution. An accessible introduction is provided by Link et al. (2002), and details of the basis of MCMC and techniques for its implementation are described in Gilks et al. (1996). The output of an MCMC simulation is a long list of parameter estimates. If the MCMC algorithm has been designed correctly, then the distribution of these simulated parameters will be very close to that of the actual posterior distribution. A useful result is that it is easy with MCMC to obtain credible intervals for the posterior abundances of all species, conditional on the data and distributional assumptions. MCMC is in one sense producing simulated communities, which links the technique to a common theme in theoretical ecology.

Implementation of MCMC simulation of Bayesian posterior distributions has been made easy by the development of the BUGS program (Spiegelhalter et al. 1999), which automates the choice and coding of the algorithms for the estimation, allowing the user to concentrate on developing the statistical models. BUGS is

freely available on the Internet.[5] Fitting most compound distributions using MCMC is simple. The inputs to models such as the Poisson gamma or Poisson lognormal are vectors of the number of individuals for each species. No further mathematical assumptions are required, beyond ''uninformative'' priors for the hyper parameters.

The MCMC solution to the problem of model fitting is, however, only partial. The rare-species effect remains to be settled. The usual practice is to ignore species that are not found in a sample of individuals. This implies that they are not part of the community that the sample has been drawn from, even though this assumption is not stated and perhaps not deliberately made. Some researchers do attempt to estimate the number of species from a sample (e.g., Colwell and Coddington 1994). Unfortunately these methods appear to be unreliable unless the underlying abundance distribution is known (O'Hara 2005).

Knowledge about the species that are to be found in a region can be used to place a limit on the number of species that may be expected to be found in a sample. In most practical conservation settings, even in the tropics, some relevant knowledge of a species pool is available. It can come from replication within a study, reference collections, previous studies, or, in the very worst case, informal natural-history observations made in the surrounding area. This provides an upper bound on the number of species. Absence of a species in a sample of individuals can then be represented explicitly by zero abundance rather than a missing value. Of course adding zeros to the data instead of assuming complete absence adds an assumption, although it can usually be justified. Many ecologists naturally assume that small samples have been drawn from open communities. In other words, as samples become larger they eventually include a representative from most species present in a regional species pool. Whether very rare species are considered accidentals or genuine members of the community is an ecological, rather than a mathematical, issue. It seems preferable to make the issue explicit and thus open it to scientific, rather than statistical, criticism. We will return to this issue in the *Discussion*, below.

We illustrate the power of this approach, and the associated MCMC fitting routine with a simple example involving the fruit-feeding butterflies of Montebello National Park, Mexico. The example uses the sort of data that is typically produced by comparatively short, although nevertheless costly, field campaigns in the tropics. The researchers were particularly interested in the hypothesis that diversity differed among four forest sites in the study area. The first of the four areas was in comparatively undisturbed forest while the other three had suffered varying amounts of anthropogenic disturbance. Our goal was to provide an informative

and accurate means of comparing diversity for data sets differing in their sample size while fully quantifying the uncertainty involved in the inferential process and keeping the assumptions transparent. We also wanted to emphasize the incorporation of expert knowledge concerning the presence of rare species in the region in order to help solve a delicate inferential problem. The hope is that the method outlined in this study may be considered by field ecologists as an alternative to traditional methods for estimating diversity when the main goal is to make comparisons between sites.

MATERIAL AND METHODS

The data

The data for the example come from a one-year study on diversity of fruit-feeding butterflies in Montebello National Park, Chiapas, Mexico. This is an area of mountain rainforest (sensu Breedlove 1981) that has been affected by a range of anthropogenic activities. In 1998 the park suffered a severe fire. From June 1997 to May 1998 sampling of fruit-feeding butterfly diversity was undertaken in four patches with differing vegetation types ranging from undisturbed species-rich forest to highly disturbed secondary areas. The first area was undisturbed forest. The a priori hypothesis of the researchers was that butterfly diversity in this area might, in some respect, be either higher or lower than in other areas. Full details of the sampling scheme used are provided by Ruiz-Montoya et al. (2000). Traps and trapping occasions were replicated. However many traps were stolen or vandalized leading to less replication that would have been ideal. Crucially, there were marked differences in sampling intensities between sites. All the essential information in the data, apart from the identities of the individual species, can be condensed into four vectors of abundances of individuals, one for each site. These data are shown concisely in Fig. 1 as a series of empirical rank-abundance plots with the number of individuals per species included. There were a total of 50 species found in the pooled data.

The models

The classic detailed investigation of the properties of the lognormal and log-series models is given by Kempton and Taylor (1974). Both models are identical at their lowest level. It is assumed that the number of individuals of species i, n_i, that are found in a sample come from a Poisson distribution with mean μ_i:

$$\Pr(N = n) = \frac{\mu_i^n e^{-\mu_i}}{n!}.$$

The μ_i's vary between species, and can be described by some distribution. There are many candidates for this distribution, but here we shall only examine two. The lognormal is the following:
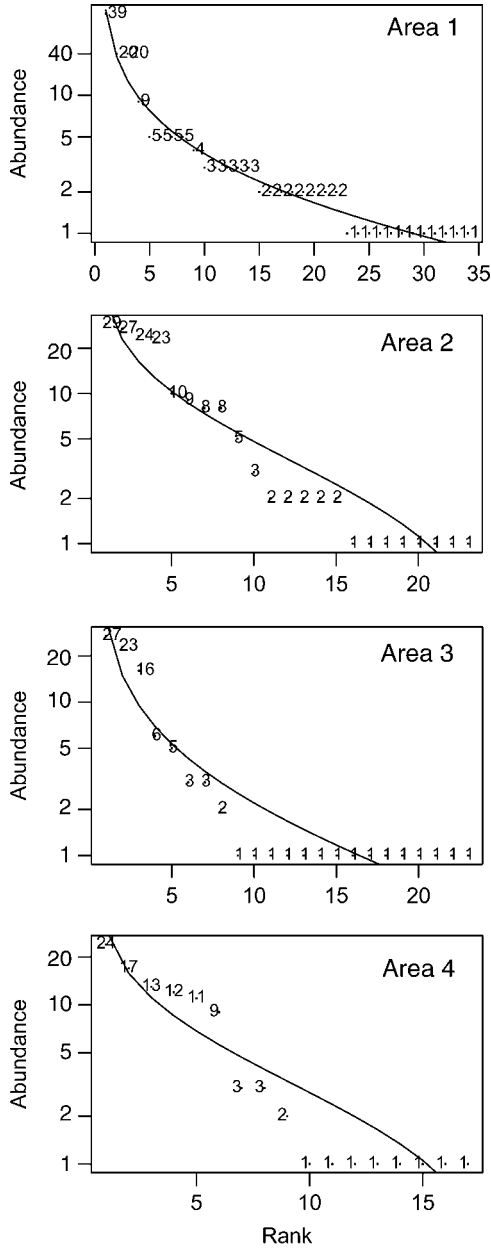
FIG. 1. Rank–abundance plots with fitted veiled lognormal curves. Numbers refer to number of individuals of each butterfly species caught over the space of a year in each of the four survey areas (Montebello National Park, Chiapas, Mexico). The dots represent the underlying points upon which the numbers have been superimposed. Note the *y*-axis log scale.

$$f(\mu_i) = \frac{1}{\mu_0\sqrt{2\pi\sigma^2}}e^{-(1/2\sigma^2)(\log \mu_i - \mu_0)^2}$$

where μ_0 and σ^2 are the two hyperparameters. In other words, they are the parameters that describe the distribution of the μ_i's. The mean of the distribution is $\exp(\mu_0 + \tfrac{1}{2}\sigma^2)$ and the variance is $(\exp(\sigma^2) - 1)$

$\exp(2\mu_0 + \sigma^2)$. An alternative underlying distribution could be a gamma distribution:

$$f(\mu_i) = \frac{\beta^\alpha}{\Gamma(\alpha)}\mu_i^{\alpha-1}e^{-\beta\mu_i}$$

where α and β are the hyperparameters of the distribution. The gamma distribution has a mean of α/β, and a variance of α/β^2. This distribution is much more flexible than the lognormal, making weaker assertions about the allowable shape of the underlying distribution. If $\alpha > 1$, the distribution looks similar to the lognormal distribution. At the extreme, when α tends to 0, the log-series is produced (Fisher et al. 1943). The difficulty with the gamma distribution from a general ecological perspective is that fitting it tells us less about community properties than does the lognormal. The two parameters are more difficult to interpret, as they are less well known and are not as well linked to theory. On the other hand the σ of the lognormal is a very well understood measure of community equitability in a theoretical context.

The models were fitted using a Bayesian approach. For this, prior distributions need to be specified. Here we used vague, non-informative priors. The prior used for the σ of the lognormal was a gamma distribution with shape = 1 and scale = 0.01. The prior for the α of the gamma distribution was an exponential with rate = 1. For β we used a gamma with shape = 0.1 and scale = 1. These priors are flat over the region of interest. Trials showed that the small changes in the form of the priors had little or no practical influence on the outcome. In some circumstances informative priors could be useful if it is desired to strengthen inference by combining data with the result from previous studies.

The simplicity of these particular models does not demonstrate the full power of the Bayesian approach for solving otherwise intractable problems. In a pragmatic sense, so long as uninformative priors are used, a Bayesian approach can be seen as another way of reproducing the results of maximum-likelihood distribution fitting. However there are some very convenient properties that can be used to great advantage in this setting, as we will show. The posterior distribution of any secondarily derived parameters can be calculated within the model. We can therefore simultaneously calculate any additional parameters that we could be interested in together with their distributions under the sampling assumptions made. Obvious candidates are Shannon's or Simpson's indices. These can be based on abundances that are simulated by MCMC (Markov-chain Monte Carlo). This provides an easy way of providing confidence intervals for diversity indices while at the same time fitting distributions. It can also provide credible intervals for the posterior abundances of all species, including those not actually found, although the interpretation of the latter result must be placed very firmly in its mathematical context.

Model comparisons

We compared the fit of the models using the deviance information criterion (DIC) (Spiegelhalter et al. 2002). This is a Bayesian equivalent of AIC (the Akaike information criterion) and can be seen as a measure of optimality of a model, trading off complexity, and fit to the data. In principle, lower values of DIC suggest that a model should be preferred. However, the properties of this statistic are still being investigated, so the results should be viewed with some caution.

Calculating diversity indices

Models were analyzed within BUGS and by using the coda package (Plummer et al. 2004) in the R statistical environment (R Development Core Team 2004). The BUGS code, the data used, R code, and a TclTk graphical interface to the models are available in the online Supplement to this paper. The models were set up to provide confidence intervals for many commonly used diversity indices and dissimilarity measures. Here we demonstrate their utility by showing only two standard diversity indices that we calculate using posterior abundances under the two models: (a) Shannon's index, H; and (b) the inverse of Simpson's index, $1/D$.

RESULTS

Rank-abundance curves are a commonly used means of visualizing relative abundances (Wilson 1991). The fitted curves in Fig. 1 were produced using the R package ''vegan'' (J. Oksanen, *unpublished software*), which provides useful complementary diversity analyses to the Bayesian method we present. In particular the lognormal can be treated as a generalized linear model with a logarithmic link function. A veiled lognormal can be analyzed by fitting a nonlinear model for one parameter while estimating other parameters and the log-likelihood within a generalized linear-model framework (J. Oksanen, *unpublished software* [2004]).[6] While this can be helpful in some contexts the procedure is difficult to follow and is known to fail with small data sets. It does not provide measures of fit that are clearly interpretable with respect to a model of the data-generating process. As implemented in R, confidence intervals are not available for the key parameters of interest. It is typically used to compare fits with alternative models such as the Zipf–Mandelbrot rather than drawing comparisons between sites directly.

We fitted the Poisson gamma and the Poisson lognormal distributions after adding zeros to represent species that were not caught in every area but found in the pooled data and thus assumed to be potentially present. All four vectors had a length of 50. Markov-chain Monte Carlo (MCMC) can sometimes fail to converge. This makes diagnostics vital when new MCMC models are being developed and tested. Few problems were expected given the comparative simplicity of these models. Nevertheless in order to check conver-gence we ran two chains for each model for 10 000 iterations using over-dispersed initial values. The Gelman diagnostic was used on key variables to check for convergence. In the case of the Poisson lognormal we applied it to σs. In the case of the Poisson gamma we looked at α. Values of the diagnostic were not substantially above 1, clearly showing that there were no problems with convergence. Fig. 2 shows the trace plots for the σ of the lognormal. The trace demonstrates good mixing. Similarly satisfactory results were observed for the Poisson-gamma model. We can therefore assume that these models can usually be safely run using a single chain. A Heidelberger and Welch (HW) test (Heidelberger and Welch 1983) can then be used to establish the length of the chain needed for reliability. Based on the HW diagnostic we report the results from 6000 iterations after discarding the first 1000 as a burn in. The chain can be thinned by a factor of 3, although with a small data set thinning is unnecessary given the capacity of contemporary computers.

Table 1 shows a breakdown of the deviance information criteria (DIC). This suggests that an underlying gamma provides a better fit, particularly for areas 2, 3, and 4. As further analysis will clarify, these are the least diverse assemblages. The smallest DIC suggests the model that makes the best short-term predictions and should be interpreted in the same way as AIC. However, if the difference between DIC is, small (<5), models provide similar levels of fit. In this case the choice of model should take into account other factors, such as interpretability and connection to ecological theory. The lognormal thus might still be considered. It has the useful property of interpretability even if it does not fit the data quite as well as the gamma. Lack of fit can arise from several causes. Even though sampling stochasticity was included in the models, extra Poisson variability in the data may still arise through the effects of overdispersal due to sampling artifacts rather than as genuine features of the underlying distribution.

Fig. 3 shows the posterior distributions for σ, the key parameter of the lognormal distribution. Results of MCMC simulations can be summarized in many forms, emphasizing that parameters in a Bayesian analysis are treated as random variables. Here the mode for the 2000 simulations is given together with 95% higher posterior density (HPD) confidence (or credible) intervals calculated within R (Martyn Plummer, *personal communication*). HPD intervals are similar to the usual confidence intervals except that instead of having the same proportion in each tail the tails are defined so that the density within the interval is higher than the density outside the interval on either side.

The overall median for all four areas can be extracted and used as a reference baseline. This lies outside the 95% HPD interval for the σ of area 1. When two parameters are of interest they form a cloud of points, the density of which can be represented as contour
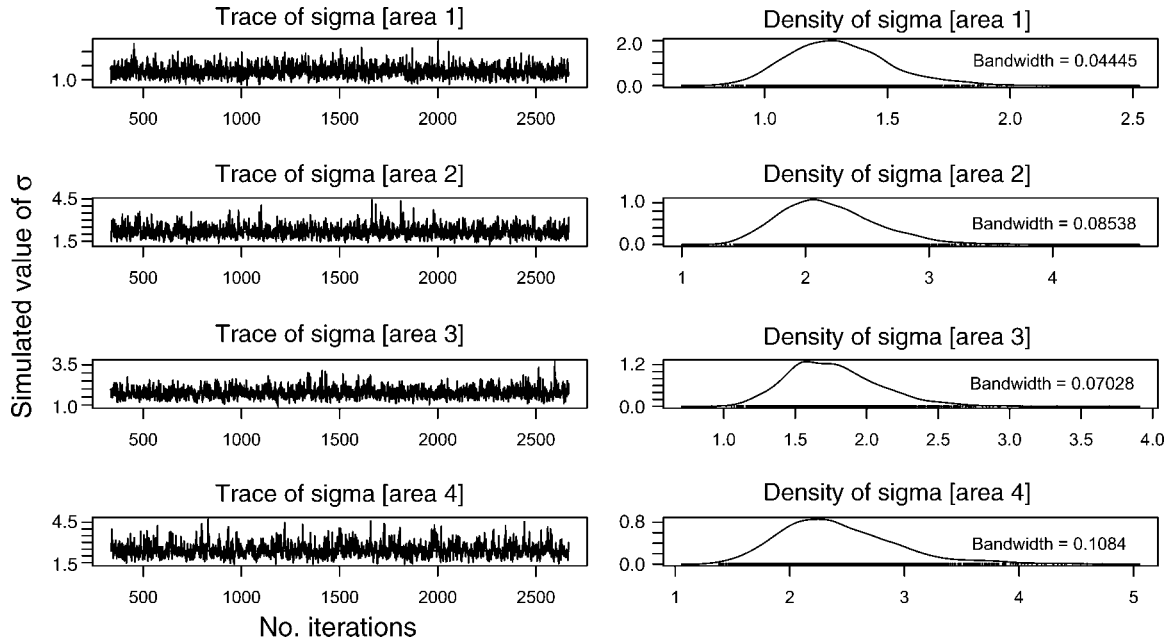
FIG. 2. Trace and density plots for σ, the standard deviation of the lognormal relative-abundance model, produced using the coda package in R (Plummer et al. 2004). For all density plots, $N = 2333$ simulations; bandwidths are given on the right side of each density plot.

lines. This is a convenient means of investigating the relationship between them. In this case it can be seen that the two diversity indices, 1-D (the inverse of Simpson's) and H (Shannon's), are quite closely correlated, and again area 1 stands out clearly. Differences between sites can also be extracted and comparisons can be made. Formal hypothesis testing is not usually conducted by Bayesian analysts. However if over 95% of the posterior density of the differences lies above or below zero this can be taken as evidence of a notable difference between sites, providing the assumptions of the model have been met.

Fig. 4 shows a similar breakdown of the Poisson gamma model. The shape parameter, α, now becomes our guide to equitability. If α is larger than 1 the gamma appears closer in form to a lognormal. Again the model shows the difference between area 1 and the other three areas.

The confidence regions for the Shannon's and Simpson's indices are almost identical under either model.

We conclude that the intervals provided by the MCMC simulation for these indices are therefore mainly a result of the assumption of Poisson variability in the sampling process. An alternative technique that does a similar task is bootstrap resampling. We found that it provided similar results and confidence intervals. We stress, however, the convenience of combining estimation of these metrics in the same analysis that fits the underlying distributions rather than using a combination of separate techniques.

A final feature of the MCMC procedure is the ability to pool the simulated abundances of all the species in order to produce a density plot for the posterior relative abundances (Fig. 5). Notice that this is not quite the same as the lognormal model of abundance, which is derived from the posterior abundances. The differences between the two density plots help to reveal where the data dominate the underlying model. In this case it is clear that the "less diverse" plots all had more individuals for the common species than the model ex-

TABLE 1. Breakdown of the deviance information criterion (DIC) for both models calculated for all four areas.

Model	Area 1			Area 2			Area 3			Area 4		
	DIC	$\bar{D}$	p_D	DIC	$\bar{D}$	p_D	DIC	$\bar{D}$	p_D	DIC	$\bar{D}$	p_D
Lognormal	77.1	49.7	27.4	62.7	39.9	22.8	60.9	39.7	21.2	50.7	32.6	18.1
Gamma	74.2	45.3	28.9	56.3	34.0	22.3	56.4	34.9	21.5	44.2	27.2	17.0
Difference	2.9			6.4			4.5			6.5		

Notes: D is the deviance; $\bar{D}$ is the posterior mean of $-2\log L$; p_D is the estimated number of parameters. Note that DIC = $\bar{D} + p_D$. Data are from a one-year study of diversity of fruit-feeding butterflies in four areas of Montebello National Park (Chiapas, Mexico) that differed in vegetation type and history (ranging from undisturbed species-rich forest [area 1] to highly disturbed secondary forest [area 4]).
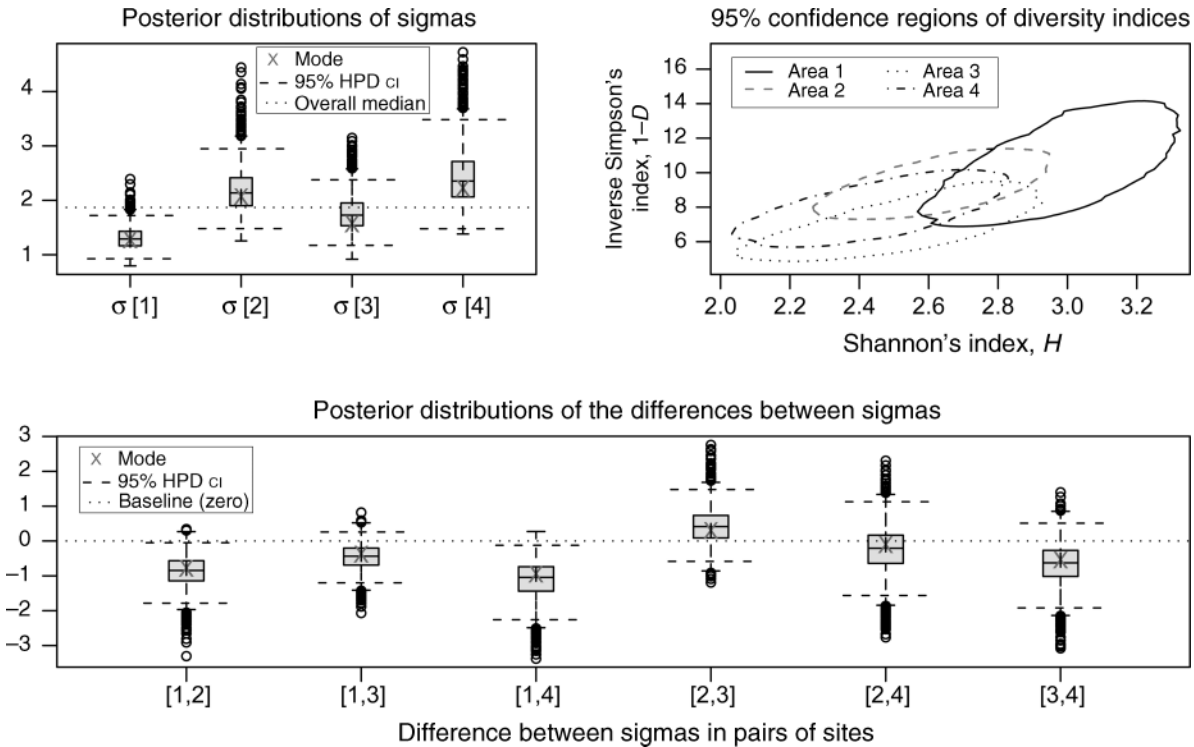
Posterior distributions of sigmas

95% confidence regions of diversity indices

Posterior distributions of the differences between sigmas

FIG. 3. Posterior distributions for the σ parameter of the lognormal, confidence regions for the two diversity indices, and posterior distributions for the difference in σ between sites. Boxes represent the interquartile range. Whiskers with short bars extend to 1.5 times the length of the box. Whiskers with long bars show the 95% higher posterior density (HPD) intervals.

pected. With such small data sets the conclusions drawn have to be approached with caution, but the ability of the MCMC procedure to clarify features in the data while maintaining the abstraction needed for evaluating theory is demonstrated.

DISCUSSION

The best solution to the problems of inference that arise from small sets of data is to obtain more data. However this is not always possible. It is the role of statistics to ensure that the available data are properly interpreted rather than discarded as completely uninformative. Bayesian analyses are ideal for making use of provisional information as the interpretation of the results includes uncertainty. The main finding we draw from this study is that anthropogenic disturbance of the forest of Montebello National Park (Chiapas, Mexico) is believed to result in a shift toward a less diverse assemblage of fruit-feeding butterflies. It was possible to derive this tentative statement from a relatively short-term study. Considerable temporal and spatial variability is to be expected in butterfly communities (DeVries and Walla 2003). When the results of fitting models such as these are reported, the assumptions involved must be made very clear. In our examples, if confidence limits for the diversity statistics are cited it should be mentioned that they are derived under the assumption of simple Poisson sampling. The limited

temporal and spatial extent of the survey must also be mentioned. If width parameters, or their derivatives, are used it should be stated that they are obtained from assuming Poisson sampling from an underlying lognormal, or gamma, distributed species pool consisting of 50 species.

The overall findings derived from this method do not necessarily differ from those obtained through alternative approaches. The key virtue is found in the simplicity and transparency of the procedure. As MCMC (Markov-chain Monte Carlo) has made the Bayesian approach extremely accessible and easy to use, we suggest that its power should be more widely exploited. An interesting question is whether the conclusions drawn from parametric model fitting would be expected to differ from those drawn from analysis of diversity indices alone. We feel that diversity indices that combine both species richness and equitability apply to the sample and are not truly inferential in the sense that an underlying population model is not involved. Our results confirmed this by showing that confidence regions for the diversity statistics were independent of the assumptions made regarding the prior form of the underlying distribution. Thus whether or not this particular model-fitting technique is used, we suggest that analyses of diversity should always include at least one explicitly parametric index together with an estimate of the associated uncertainty regarding its value.
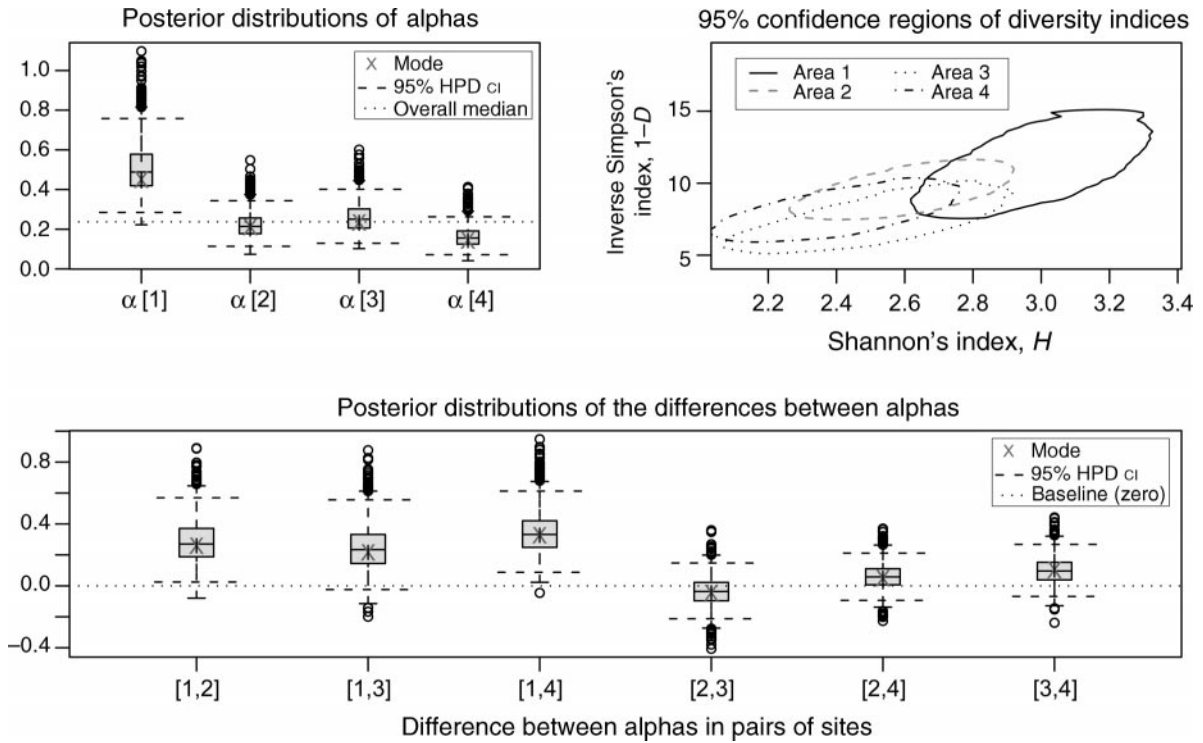
FIG. 4. Posterior distributions for the α parameter of the gamma, posterior confidence regions for the two diversity indices, and posterior distributions for the difference in α between sites. Boxes represent the interquartile range. Whiskers with short bars extend to 1.5 times the length of the box. Whiskers with long bars show the 95% higher posterior density (HPD) intervals.

A further strength of the Bayesian framework is the facility with which models can be extended, to provide both greater predictive power and greater insight into processes operating in ecological communities. More complex models can be built that reflect the full complexity of both study design and ecological phenomena themselves (e.g., Gelfand et al. 2005). Hierarchical statistical models are natural complements for the hierarchical structure of communities themselves (Whittaker et al. 2002). More complex formulations are possible where the structure of data sets allows. For larger data sets, overdispersal terms can easily be added in order to model the effects of extra Poisson variation (e.g., O'Hara 2005), but for small data sets like those here it is not possible to separate out the variation due to overdispersal from the variation in the actual abundances.

Kempton and Taylor (1974) stressed that very little information regarding an underlying relative-abundance distribution can be derived from a sample without adding assumptions regarding the maximum number of species. Inferential diversity indices, such as the σ of the lognormal in our analyses rely on this. Our approach was based on the work of Kempton and Taylor, and the results we obtained support this interpretation. Under some forms of noninferential analyses, species-poor areas that support a small number of equally abundant species can appear more equitable than species-rich areas with many rare species (Wilson et al. 1999). This is a particular problem with Simpson's index. Treating rarity explicitly as nonoccurrence ensures that the problem is at least handled consistently. Either of the two width parameters from the underlying models can be used as diversity indices. The familiarity of the lognormal makes σ the most easily interpretable measure of equitability. Species-poor samples occur either because they contain a low total number of individuals, in which case the uncertainty regarding the value of the width parameter is naturally increased, or because the probability of the capture of rarer species is extremely low due to a genuinely wide distribution of abundances. These alternative explanations are made very explicit under our analysis.

A reviewer of an earlier draft of this paper suggested that fixing the total number of species through adding zeros to the data was in some respects un-Bayesian. The veil line should be estimated as a parameter. In fact either model presented in the paper can be easily extended within BUGS to accommodate the possibility that species given zero abundance in a sample are genuinely not present in the community rather than merely being present, but at a low abundance. It can be done through fitting mixture models. These assume that there is some additional probability, p, that a species is pre-
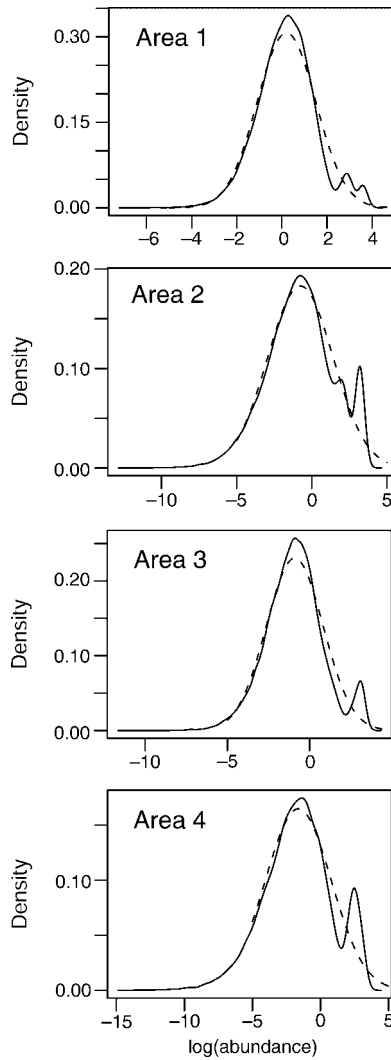
FIG. 5. Posterior relative-abundance distributions (solid lines) and the fitted lognormal model (dashed lines).

sent that is independent of Poisson sampling. If the species is present then the number of individuals sampled is as described in the original models. If the species is genuinely not part of the community, then we also will have sampled zero individuals, but the information will not be used when fitting a relative-abundance distribution. The differences between the two situations can in some sense be teased apart by using the zero inflated Poisson model which models underlying presence or absence as a binomial process:

$$\Pr(N_i = n) = \begin{cases} p.e^{-\mu_i} + (1 - p) & n = 0 \\ p\dfrac{\mu^n e^{-\mu_i}}{n!} & n > 0. \end{cases}$$

When implementing these models we found that the result was to shift uncertainty away from the values for the parameters of the underlying models and toward

uncertainty regarding the number of species. As there is usually much better information available regarding the number of species than the shape of a distribution, we concluded that for practical purposes it was preferable, if not ideal, to use a fixed number of species. The method can be extended easily to include sensitivity analysis regarding this assumption.

We emphasize that given the complexity of ecological communities we do not expect the assumptions used in building any of these models to hold in their entirety. This is precisely why statistical modeling should be transparent and open to criticism. We agree with Box (1979:202) in believing that "all models are wrong but some are useful." It is impossible to rigorously test all the assumptions of a model using a single data set. To suggest otherwise would be very misleading. DIC (deviance information criterion) did produce an assessment of the models fit. The difference between competing models based on this criterion alone was not great. This was also expected with such small data sets. Many similar models could fit equally well. The usefulness of these models is derived from the insight they provide regarding the true question of interest, which in this case concerned the difference between sites, not between models.

Samples taken from ecological communities are always small when placed in their true temporal and spatial context. Thus they can hardly be expected to reveal a complete picture unaided. Classical independent replication is often impossible in observational data sets (Reckhow 1990, Oksanen 2001). The Bayesian attitude to the problem of non-independence is not to ignore it, but to make it explicit and draw on other sources to support inference where necessary. In some respects this provides a bridge between theoretical and empirical ecology (Holling and Allen 2003). A model that suggests a prior distribution is an explicit requirement for Bayesian analysis. Ecological theory is one source for priors (Wade 2000, Ludwig et al. 2001), and theories which are able to quantify uncertainty provide useful priors. Theories that suggest overoptimistic levels of precision are less helpful. In our example we could have used Preston's canonical hypothesis to provide a more informative prior distribution for the variance of the lognormal model. We did not do this because although the canonical hypothesis suggests the most probable values for σ we did not, at this stage, have a suitable model for its probability distribution. Meta-analyses could be used to find such a model. A great strength of Bayesian analysis is that new evidence can be added to our models if it becomes available, making meta-analysis particularly straightforward.

Conclusion

We conclude, in view of our results, that a relatively small data set can provide support for the belief that butterfly diversity is negatively affected by anthropogenic forest disturbance. Edwards (1996) criticizes

Bayesian methods on the grounds that data analysis should be ''journalistic'' rather than ''editorial.'' We prefer the transparent statement of assumptions that a Bayesian analysis provides. Such methods would seem particularly valuable for workers in as yet poorly studied tropical regions such as ours where conservation concerns motivate ecological research. Under the Bayesian paradigm, editorials can always be written afresh as research proceeds. Bayesian ecologists adapt and change their views in response to new evidence.

Literature Cited

Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. Journal of Wildlife Management **64**:912–923.

Bell, G. 2000. The distribution of abundance in neutral communities. American Naturalist **155**:606–617.

Box, G. E. P. 1979. Robustness in scientific model building. Pages 201–236 *in* R. L. Launer and G. P. Wilkinson, editors. Robustness in statistics. Academic Press, New York, New York, USA.

Breedlove, D. E. 1981. Introduction to the flora of Chiapas. Flora of Chiapas. Volume 1. California Academy of Sciences, San Francisco, California, USA.

Burton, P. R., L. C. Gurrin, and M. J. Campbell. 1998. Clinical significance not statistical significance: a simple Bayesian alternative to *P* values. Journal of Epidemiology and Community Health **52**:318–323.

Cassie, R. M. 1962. Frequency distribution models in the ecology of plankton and other organisms. Journal of Animal Ecology **31**:65–92.

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.

DeVries, P. J., and T. R. Walla. 2003. Species diversity and community structure in neotropical fruit-feeding butterflies. Biological Journal of the Linnean Society **74**:1–15.

Dewdney, A. K. 1998. A general theory of the sampling process with applications to the ''veil line.'' Theoretical Population Biology **54**:294–302.

Edwards, D. 1996. Comment: the first data analysis should be journalistic. Ecological Applications **6**:1090–1094.

Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications **6**:1036–1046.

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.

Gelfand, A. E., A. M. Schmidt, S. Wu, J. A. Silander, Jr., A. Latimer, and A. G. Rebelo. 2005. Modelling species diversity through species level hierarchical modelling. Journal of the Royal Statistical Society Series C **54**:1–20.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, editors. 1996. Markov chain Monte Carlo in practice. Chapman and Hall, London, UK.

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

Guthery, F. S., J. J. Lusk, and M. J. Peterson. 2003. The fall of the null hypothesis: liabilities and opportunities. Journal of Wildlife Management **65**:379–384.

Hayek, L. C., and M. A. Buzas. 1997. Surveying natural populations. Columbia University Press, New York, New York, USA.

Heidelberger, P., and P. D. Welch. 1983. Simulation run length control in the presence of an initial transient. Operations Research **31**:1109–1144.

Hilborn, R., and M. Mangel. 1997. The ecological detective. Princeton University Press, Princeton, New Jersey, USA.

Holling, C. S., and C. R. Allen. 2003. Adaptive inference for distinguishing credible from incredible patterns in nature. Ecosystems **5**:319–328.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.

Johnson, D. H. 2003. The role of hypothesis testing in wildlife science. Journal of Wildlife Management **66**:272–276.

Kempton, R. A., and L. R. Taylor. 1974. Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. Journal of Animal Ecology **43**:381–399.

Lindley, D. V. 2000. The philosophy of statistics. Journal of the Royal Statistical Society Series D–The Statistician **49**:293–319.

Link, W. A., E. Cam, J. D. Nichols, and E. G. Cooch. 2002. Of BUGS and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. Journal of Wildlife Management **66**:277–291.

Ludwig, D., M. Mangel, and B. Haddad. 2001. Ecology, conservation and public policy. Annual Review of Ecology and Systematics **32**:481–517.

McGill, B. 2003. Strong and weak tests of macroecological theory. Oikos **102**(3):679–685.

Meyer, R., and R. B. Millar. 1999. BUGS in Bayesian stock assessments. Canadian Journal of Fisheries and Aquatic Sciences **56**:1078–1087.

O'Hara, R. B. 2005. Species richness estimators: how many species can dance on the head of a pin? Journal of Animal Ecology **74**:375–386.

Oksanen, J. 1996. Is the humped relationship between species richness and biomass an artefact due to plot size? Journal of Ecology **84**:293–295.

Oksanen, L. 2001. Logic of experiments in ecology: is pseudoreplication a pseudoissue? Oikos **94**:27–38.

Pielou, E. C. 1977. Mathematical ecology. John Wiley and Sons, New York, New York, USA.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2004. coda: output analysis and diagnostics for MCMC. R package version 0.8–3. ⟨http://www-fis.iarc.fr/coda/⟩

Preston, F. W. 1948. The commonness and rarity of species. Ecology **29**:254–283.

R Development Core Team. 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org./⟩

Reckhow, K. H. 1990. Bayesian inference in non-replicated ecological studies. Ecology **71**:2053–2059.

Ruiz-Montoya, L., A. Horvath, R. Martinez-Castellanos, P. Enriquez-Rocha, and J. L. Rangel-Salazar. 2000. Situaciónactual de la Fauna Silvestre en el Parque Nacional Lago de MonteBello. Informe Final Sistema di Investigación Fegional Benito Juarez-Conacyt. El Colegio de la Frontera Sur, San Cristóbal de Las Casas, Chiapas, Mexico.

Speigelhalter, D. J., G. B. Nicola, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society B **64**:583–639.

Spiegelhalter, D. J., A. Thomas, and N. G. Best. 1999. WinBUGS version 1.2 user manual. MRC Biostatistics Unit, Cambridge, UK.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. American Naturalist **116**(6):770–787.

Wade, P. R. 2000. Bayesian methods in conservation biology. Conservation Biology **14**:1308–1316.

Whittaker, R. J., K. J. Willis, and R. Field. 2002. Scale and species richness: towards a general, hierarchical theory of species diversity. Journal of Biogeography **28**:453–470.

Wilson, J. B. 1991. Methods for fitting dominance/diversity curves. Journal of Vegetation Science **2**:35–46.

Wilson, J. B., J. B. Steel, W. M. King, and H. Gitay. 1999. The effect of spatial scale on evenness. Journal of Vegetation Science **10**:463–468.

SUPPLEMENT

The code for WinBUGS models, an example data set (Mexican butterfly diversity), and an R interface that will allow the user to fit the compound distributions described in the paper to small data sets are available in this supplement (*Ecological Archives* A016-012-S1).