

REVIEW

Bayesian inference in ecology

Aaron M. Ellison

Harvard University, Harvard
Forest, PO Box 68, Petersham,
MA, USA
E-mail: aellison@fas.harvard.edu

Abstract

Bayesian inference is an important statistical tool that is increasingly being used by ecologists. In a Bayesian analysis, information available before a study is conducted is summarized in a quantitative model or hypothesis: the prior probability distribution. Bayes' Theorem uses the prior probability distribution and the likelihood of the data to generate a posterior probability distribution. Posterior probability distributions are an epistemological alternative to *P*-values and provide a direct measure of the degree of belief that can be placed on models, hypotheses, or parameter estimates. Moreover, Bayesian information-theoretic methods provide robust measures of the probability of alternative models, and multiple models can be averaged into a single model that reflects uncertainty in model construction and selection. These methods are demonstrated through a simple worked example. Ecologists are using Bayesian inference in studies that range from predicting single-species population dynamics to understanding ecosystem processes. Not all ecologists, however, appreciate the philosophical underpinnings of Bayesian inference. In particular, Bayesians and frequentists differ in their definition of probability and in their treatment of model parameters as random variables or estimates of true values. These assumptions must be addressed explicitly before deciding whether or not to use Bayesian methods to analyse ecological data.

Keywords

Bayes' Theorem, Bayesian inference, epistemology, information criteria, model averaging, model selection.

Ecology Letters (2004) 7: 509–520

'In the whole round of statistical investigation few principles are so much theoretically neglected at present and yet so largely and unconsciously appealed to as that theorem in probability by which our past experience is made a basis for future conduct.... [T]he extent to which it really lies at the basis of most state, municipal and individual actions is too often disregarded.'

Karl Pearson (1907, p. 365)

INTRODUCTION

Bayesian inference is an alternative method of statistical inference that is frequently being used to evaluate ecological models and hypotheses. Bayesian inference differs from classical, frequentist inference in four ways:

- 1 Frequentist inference estimates the probability of the data having occurred given a particular hypothesis ($P(Y|H)$) whereas Bayesian inference provides a quantitative measure of the probability of a hypothesis being true in light of the available data ($P(H|Y)$);

- 2 Their definitions of probability differ: frequentist inference defines probability in terms of long-run (infinite) relative frequencies of events, whereas Bayesian inference defines probability as a individual's degree of belief in the likelihood of an event.
- 3 Bayesian inference uses prior knowledge along with the sample data whereas frequentist inference uses only the sample data;
- 4 Bayesian inference treats model parameters as random variables whereas frequentist inference considers them to be estimates of fixed, 'true' quantities.

The last three distinctions are epistemic, and one should consider them carefully in choosing whether to use Bayesian or frequentist methods.

This review has three parts. First, I summarize differences between Bayesian and frequentist methods of inference. This section provides the background necessary to decide whether to use Bayesian or frequentist methods. Second, I review briefly the range of ecological problems to which Bayesian inference has been applied. Third, I contrast

frequentist and Bayesian inference in a simple ecological example, using generalized linear models to model species richness across a latitudinal gradient.

SCIENTIFIC INFERENCE

We are all familiar with ‘the’ scientific method of testing statistical null hypotheses (Popper 1959). In brief, we ask what is the probability that we would have obtained our set of data (an independent, random sample of a larger population), *or a more extreme set of data*, if the null hypothesis was true. Generically, we write this as $P(Y|H_0)$, where Y is the data and H_0 is the null hypothesis.¹ Technically, the hypothesis is a model with (known or unknown) parameters. For example, if we are interested in testing if species richness varies with latitude, we compute the probability of obtaining a set of sample data given the null hypothesis of a regression model in which the slope (or unknown parameter) β_1 equals zero. Our data provide an estimate of the mean and variance of the parameter β_1 in our sampled population, and we compute the probability of obtaining these estimates if β_1 equals zero.

If this probability – called the P -value – is ‘small’, we reject the null hypothesis. How small a P -value must be for the null hypothesis to be rejected is a matter of convention: the standard cut-off value is the Neyman–Pearson acceptable probability of committing a Type-I statistical error: $\alpha = 0.05$ (Hubbard & Byarri 2003). It bears repeating that this method of hypothesis testing allows only for falsification or rejection of hypotheses (Popper 1959). The common conclusion drawn from obtaining a small P -value that the alternative hypothesis is true with probability equal to $(1-P)$ is incorrect.² Similarly, a large P -value does not provide evidence in favour of the null hypothesis (Howson & Urbach 1993).

A prerequisite of this method – frequentist statistical inference – is a concept of probability defined as the relative frequency of a particular observation (an *event* or *outcome*). In other words, the probability P of an event A (written as $P(A)$) equals the number of times that event occurs (n_A)

divided by the total number of observed events (n). As n is, in principle, infinite, the *frequency definition* of probability asserts that as $n \rightarrow \infty$, $n_A/n \rightarrow$ the true (population) value of $P(A)$. Under standard design criteria (independent, identically distributed, random samples), the *sample data* provide unbiased estimates of $P(A)$. The interpretation of a frequentist confidence interval follows directly from this definition of probability. A $p\%$ confidence interval calculated from the sample mean and variance asserts that in n hypothetical runs of an experiment, the parameter of interest (e.g. the true population mean μ) is expected to occur in the computed interval in $p\%$ of the experimental runs. Thus, in $(1-p)\%$ of the experimental runs, the computed interval will not include the true value of the parameter (illustrated clearly in the simulations of Blume & Royall 2003). Note that in any particular experiment, the parameter is either in the interval or not, but you never know which. It is incorrect to interpret a confidence interval by asserting that you are $p\%$ sure that the parameter of interest lies in the confidence interval.

Bayesian inference, in contrast, asks what is the probability of our hypothesis (again formulated as a model with known or unknown parameters) being true conditional on the sample data. This probability is found by applying Bayes’ Theorem (Bayes 1763):³

$$P(H|Y) = \frac{f(Y|H)\pi(H)}{P(Y)}. \quad (1)$$

The quantity $P(H|Y)$, or the probability of the hypothesis given the data, is called the *posterior probability distribution*, or simply the *posterior*. The quantity $f(Y|H)$ is the likelihood (Edwards 1992).⁴ The quantity $\pi(H)$ is called the *prior probability distribution*, or just the *prior*, and reflects information available about the hypothesis independent of (and hence prior to) conducting the experiment. The denomin-

¹In this expression, Y can represent either the data and all the more extreme, *but unobserved*, data, or the value of the calculated *test statistic* and all the more extreme, *but unobserved*, test statistics. Note that this leads to a procedure in which ‘a hypothesis which may be true may be rejected because it has not predicted observable results that have not occurred’ (Jeffreys 1961: 385). See Berger & Berry (1988) for a readable critique of using unobserved, and perhaps unobtainable, data to test hypotheses about a given sample.

²In a randomly chosen set of 50 papers published last year (2003) in *Ecology*, 38 of 50 (76%) authors reporting P -values asserted that their data supported or confirmed their ‘hypothesis’. That is, $P(H_A|Y)$ was believed to be large. The prevalence of this conclusion is quite surprising, given the emphasis placed on rejecting null hypotheses in standard undergraduate and graduate statistics courses taken by ecologists.

³Bayes’ Theorem originally was expressed as a probability of one event B conditional on another event, A : $P(B|A) = P(A|B)P(B)/P(A)$. In this form, it is uncontroversial, as it is simply a deduction from the axioms of probability. It is only the application of Bayes’ Theorem to *statistical inference*, in which data Y are substituted for event A and the hypothesis of interest H is substituted for event B , that is controversial. The first use of Bayes’ Theorem for statistical inference normally is attributed to Laplace. See Barnett (1999) for further discussion.

⁴Fisher’s likelihood (Fisher 1922) expresses how the probability of the *sample data* Y_{observed} varies with different values of the parameters of the hypothesis or model ($P(Y_{\text{observed}}|H)$). This is neither the same as a P -value ($P(Y|H_0)$) nor the same as the probability of the hypothesis or model for the given data ($P(H_A|Y_{\text{observed}})$). Likelihood inference (Edwards 1992; Hilborn & Mangel 1997) interprets the likelihood as expressing different degrees of support (un-normalized probabilities) for different parameter values given the data ($P(H|Y_{\text{observed}})$). This interpretation is at odds with both the frequency and subjective definitions of probability, and there is no consistent axiomatization of the probability calculus that supports this interpretation of the likelihood (see summary in Barnett 1999: 306 ff).

ator $P(Y)$ is simply a normalizing constant – the marginal probability density of the data across all possible hypotheses – and is equal to $\int_H f(Y|H)\pi(H) dH$.

Bayesian inference is predicated on a different concept of probability: *subjective* probability or an individual's *degree of belief* that a particular event will occur (Howson & Urbach 1993; Barnett 1999). Estimates of degrees of belief may vary from individual to individual, but in all cases are conditional on past experience. Because Bayesian inference has been criticized for its subjectivity and reliance on personal belief, much effort has been dedicated to generating 'objective' measures of degrees of belief (Jeffreys 1961). Most recently, Berger (2003) has proposed a reconciliation of frequentist and Bayesian significance testing, but his approach has been criticized for failure to use the foundation of Bayesian inference: subjective probability (see published comments following Berger 2003). Software used for calculating posterior probability distributions using Bayes' Theorem can accept either informative ('subjective') or non-informative ('objective') priors and the *calculations* proceed independent of one's definition of probability. The *interpretation* of the results, however, does require a definition of probability. A Bayesian posterior is an expression of a degree of belief whereas a frequentist P -value or confidence interval is an expectation of a long-run frequency. In contrast to a frequentist confidence interval, a Bayesian credibility interval is interpreted correctly as one's belief that there is a $p\%$ probability that the parameter of interest lies within the interval.

Bayesian and frequentist inference also differ in their use of prior knowledge. Frequentist testing of statistical null hypotheses assumes that there is no relevant information, such as other observations or experiments, available from past experiences. Computing a P -value is always a *de novo* exercise that begins with the null hypothesis, even if it has been falsified repeatedly in many previous studies. Frequentists view this lack of consideration of prior information positively, as it leads to an *unbiased* assessment of the sample data conditional on the hypothesis.

Bayesians counter that the traditional P -value actually is interpreted subjectively, even if the frequency definition of probability precludes such an interpretation. Further, there is no objective criterion for setting the critical level for rejection of an hypothesis (why not use 0.1 or 0.001 instead of 0.05), and the P -value is based not only on the sample data but also on more extreme data that is not and may never be observed (Jeffreys 1961; Berger & Berry 1988). We almost always have some reasons for conducting a particular experiment, developing a particular hypothesis, or using a particular model to analyse our data. Does it make sense to ignore available data or observations and jump off the shoulders of the giants that have worked before us? Efron (1978, 1986) and Dennis (1996, 2004) provide good

overviews of the arguments for and against subjectivity and objectivity in Bayesian and frequentist inference; see Berger (2003) for an attempt at finding the middle ground.

Bayes' Theorem is also iterative. An investigator may start with little or no information with which to construct the prior, but the posterior derived from the first experiment can then be used as a prior for the next experiment. The iterative nature of Bayesian inference is a central ingredient in the successful implementation of adaptive management (Walters & Holling 1990; Dorazio & Johnson 2003).

Lastly, Bayesian inference treats model parameters as random variables. Thus, not only are the data considered to be samples from a random variable, but also the parameters to be estimated are treated as random variables. This is a very different assumption from that of frequentist (and likelihood) inference, which treats parameters as true, fixed (if unknown) quantities (Fisher 1922; Edwards 1992). The studies by Strong *et al.* (1999) and de Valpine & Hastings (2002) are the only examples I have found where ecologists explicitly rejected a Bayesian method because it considered the parameters to be random variables and not a reflection of a fixed reality. In general, ecologists should consider carefully their epistemological stance when choosing among statistical methods.

HOW DO ECOLOGISTS USE BAYESIAN INFERENCE?

Ecologists have long known of and used Bayes' Theorem. Shortly after Pearson (1907) showed that an approximation of the hypergeometric series could be used to estimate posterior distributions for the condition of multiple events and full prior distributions, Pearl (1917) applied it to estimate the probable error of allelic frequencies in Mendelian populations (see also Karlin 1968; Pollak 1974). This method was elaborated in the 1970s and 1980s to determine the probability of paternity when multiple fertilizations are possible, such as in plants and fruit flies (Levine *et al.* 1980; Adams *et al.* 1992). It continues to be used in population genetic studies, including estimating the probability of introgression into wild populations of genes from genetically modified crops (Cummings *et al.* 2002).

Conditional probabilities calculated using Bayes' Theorem also were used extensively in dynamic models of foraging behaviour (Oster & Heinrich 1976; Clark & Mangel 1984; Valone & Brown 1989) and predator avoidance (Anderson & Hodum 1993). These models explicitly considered that foraging animals used previous experience to modify future foraging activities and take full advantage of the iterative nature of Bayes' Theorem. Although early work on so-called 'Bayesian foragers' used only the expected value (e.g. the mean) of the foragers' probability distributions in their models, current Bayesian models of foraging behaviour use

Table 1 Ecological studies using Bayesian inference published since 1996 in the major ecological journals (*American Naturalist*; *Journal of Ecology*; *Ecology*; *Ecological Monographs*; *Journal of Animal Ecology*; *Oikos*; *Journal of Applied Ecology*; *Oecologia*; *Ecological Applications*; *Conservation Biology*; *Ecology Letters*)

Topic	Number of papers (1996–2003)	Examples
Dynamics of single species		
Population dynamics and estimating extinction risks	10	(Forcada 2000; Drechsler <i>et al.</i> 2003)
Parameter estimation for demographic models	7	(Barrowman <i>et al.</i> 2003; Calder <i>et al.</i> 2003)
Estimating gene frequencies	5	(Cummings <i>et al.</i> 2002; Bolker <i>et al.</i> 2003)
Stock assessment	4	(Pascual & Hilborn 1995; Cooper <i>et al.</i> 2003)
Metapopulation dynamics	3	(O'Hara <i>et al.</i> 2002; Ter Braak & Etienne 2003)
Dispersal models	2	(Clark <i>et al.</i> 1999; Clark <i>et al.</i> 2003a)
Bayesian model averaging	1	(Wintle <i>et al.</i> 2003)
Dynamics of interacting species		
Foraging dynamics	8	(van Gils <i>et al.</i> 2003; Koops & Abrahams 2003)
Predator–prey interactions	5	(Anholt <i>et al.</i> 2000; Schmidt <i>et al.</i> 2001)
Competition and coexistence	3	(Damgaard 1998; Clark <i>et al.</i> 2003b)
Multispecies community ecology		
Detection probability and estimation of species richness	12	(Fleishman <i>et al.</i> 2003; Shen <i>et al.</i> 2003)
Environmental impact assessment	7	(MacNally <i>et al.</i> 2002; Peterson <i>et al.</i> 2003)
Land-use history and community reconstruction	2	(Toivonen <i>et al.</i> 2001; Platt <i>et al.</i> 2002)

Only two examples are given for each type of study; a full bibliography of these studies can be obtained from the author.

the full posterior probability distributions (Olsson & Holmgren 1999; van Gils *et al.* 2003).

The application of Bayesian inference to ecological questions has blossomed since the publication in 1996 of a series of papers on Bayesian inference for ecological research and environmental decision making (Dixon & Ellison 1996). Bayesian methods have been used most widely in population and community ecology (Table 1), in which there are many competing models to explain ecological phenomena (Hilborn & Mangel 1997), the parameter values of the models have high levels of uncertainty, and the reporting of this uncertainty (as standard errors or confidence intervals) is common. Bayesian inference is used extensively to model dynamics of single species, forecast population dispersal, growth, and extinction, and predict changes in meta-population structure on fragmented landscapes (Table 1). Foraging dynamics and predator–prey interactions continue to benefit from Bayesian methods, but they are used rarely in studies of competition; there has been a parallel 20-year decline in studies that estimate niche breadth and associated competition coefficients (Chase & Liebold 2003). Among community ecologists, Bayesian inference has been used most frequently for estimating species occurrences and species richness from geographically or logistically constrained samples, or in response to expected environmental change (He *et al.* 2003). A promising new avenue for research is the use of Bayesian methods to reconstruct palaeocommunity structure and to place estimates of

uncertainty on those reconstructions (Toivonen *et al.* 2001; Platt *et al.* 2002). In marked contrast, ecosystem studies have applied Bayesian inference only rarely (but see Carpenter *et al.* 1996; Cottingham & Carpenter 1998).

Bayesian inference is central component of formal decision analysis (Berger 1985), and has been used to assess environmental impacts (Reckhow 1990), to decide among alternative management regimes (Raftery *et al.* 1995; Layton & Levine 2003), and to structure adaptive management programs (Dorazio & Johnson 2003). Nonetheless, despite its utility for expressing uncertainty of predictions made by conservation biologists (Wade 2001) and environmental managers (Ellison 1996), Bayesian methods have not been adopted broadly by these groups. I suspect this is due to computational difficulties, lack of user-friendly software, and the requirements for precise quantification of management options and their associated utilities or outcomes.

A WORKED EXAMPLE

In this section, I use a simple example to contrast three aspects of Bayesian and frequentist inference: parameter estimation and hypothesis testing; model selection; and model averaging. Although a single example is useful to illustrate some general principles of statistical inference, it is unlikely to be satisfying to many readers. Particular models of interest to any individual are unlikely to be included in the example, and there is the danger that the example will be reified to represent all the positive or negative characteristics

of both Bayesian and frequentist inference. Similarly, because of the general lack of familiarity of Bayesian methods, a single example could result in Bayesian inference being applied to only a single class of models. Either of these outcomes would be unfortunate, and it is not my intent to capture the wide range of models that are addressable using Bayesian inference (see Box & Tiao 1992; Gelman *et al.* 1995; Sivia 1996; Hilborn & Mangel 1997; Carlin & Louis 2000; Congdon 2001 for compendia of examples).

The data: a latitudinal gradient of species richness

Latitudinal and elevational changes in species richness are well known and intensively studied ecological patterns (Huston 1994). In this example (data in Table 2), the response variable is the number of species of ants found in 64-m² sampling grids at each of 22 bogs and in the forests surrounding them. The sites span a scant 3° of latitude in New England, USA (Gotelli & Ellison 2002b). Here, I use a

Table 2 Data used for the worked example (from Gotelli & Ellison 2002b)

Site	Species richness		Latitude	Elevation
	Forest	Bog		
TPB	6	5	41.97	389
HBC	16	6	42.00	8
CKB	18	14	42.03	152
SKP	17	7	42.05	1
CB	9	4	42.05	210
RP	15	8	42.17	78
PK	7	2	42.19	47
OB	12	3	42.23	491
SWR	14	4	42.27	121
ARC	9	8	42.31	95
BH	10	8	42.56	274
QP	10	4	42.57	335
HAW	4	2	42.58	543
WIN	5	7	42.69	323
SPR	7	2	43.33	158
SNA	7	3	44.06	313
PEA	4	3	44.29	468
CHI	6	2	44.33	362
MOL	6	3	44.50	236
COL	8	2	44.55	30
MOO	6	5	44.76	353
CAR	6	5	44.95	133

Ant species richness was sampled in bogs and surrounding forests at 22 sites in Connecticut, Massachusetts, and Vermont (USA). Environmental variables used in the analysis are habitat (forest or bog), latitude (decimal degrees), and elevation (metres above sea level).

generalized linear model (McCullagh & Nelder 1989) to relate a discrete (Poisson) random variable (ant species richness) to two continuous predictor variables (degrees north latitude and elevation in metres above sea level) and one categorical predictor variable (habitat type – forest or bog).

Classical inference on the ant data

I examined simple additive models (richness S as a function of habitat type, latitude, elevation) and models that included all possible interaction terms. The ‘best’ model was chosen from the set of candidate models by minimizing Akaike’s information criterion (AIC; Burnham & Anderson 2002):

$$\text{AIC} = -2 \log(L(\hat{\beta}|\text{data})) + 2k. \quad (2)$$

In eqn 2, $L(\hat{\beta}|\text{data})$ is the likelihood of the model (which has parameters β) conditional on the data (see Footnote 4), and k is the number of parameters in the model. The model for which AIC was minimized was a simple additive model with an intercept (β_0) and all three main effects (β_1 , β_2 , β_3), but no interaction terms (Table 3):

$$\log(\hat{S}_i) = \hat{\beta}_0 + \hat{\beta}_1 \text{latitude}_i + \hat{\beta}_2 \text{elevation}_i + \hat{\beta}_3 \text{habitat}_i. \quad (3)$$

The fit of this model to the data is illustrated in Fig. 1 (top). The maximum likelihood estimates of the parameters, their standard deviations, and 95% confidence intervals are presented in the first column of Table 4. The null hypothesis that β equals zero is rejected for each β_i .

A key assumption of this model is that the observations are independent random samples, and in particular, that the forest and bog observations at a given site are independent of each other. This independence is observed in two ways. First, the deviance residuals are uncorrelated (Fig. 1b), which supports the *statistical* criterion for independence. Second, the bog and forest samples are *biologically* independent, as they are separated by hundreds of metres (far greater than the foraging distance of a single ant colony) and bog and forest ant assemblages share few species in common (Gotelli & Ellison 2002a,b).

From the frequentist analysis, the inferences are:

- 1 The data are improbable given the null hypothesis that the parameters of the model (i.e. the regression coefficients β) equal zero. In other words, because $P(\text{data} | H_0) < 0.05$, the null hypothesis is rejected.
- 2 The model fitting procedure provides maximum likelihood estimates of the parameters. We can use the standard errors of these estimates to construct confidence intervals on these parameters (Table 4). The conclusion is that in repeated sampling (which is unlikely, as collecting

Model	AIC	DIC	p_D
$S = H$	77.08	237.43	1.98
$S = L$	83.97	243.29	1.48
$S = E$	90.33	250.67	1.99
$S = H + L$	56.29	216.32	2.81
$S = H + E$	62.65	223.01	2.99
$S = L + E$	76.37	236.61	2.83
$S = H + L + E$	48.68	208.76	3.85
$S = H + L + E + H \times E$	50.27	210.17	4.75
$S = H + L + E + L \times E$	50.32	211.16	5.02
$S = H + L + E + H \times L$	50.64	210.39	4.29
$S = H + L + E + H \times E + L \times E$	51.90	211.18	5.26
$S = H + L + E + H \times E + H \times L$	52.26	210.98	4.86
$S = H + L + E + L \times E + H \times L$	53.30	211.38	5.16
$S = H + L + E + H \times E + L \times E + H \times L$	53.90	217.18	7.93
$S = H + L + E + H \times E + L \times E + H \times L + H \times L \times E$	55.76	215.66	7.92

The best fitting model (bold) included all three main effects: $\log(S) = 11.95 - 0.24L - 0.001E + 0.64H$, and had a residual deviance of 40.68 with 40 residual degrees of freedom. Models fit with S-Plus 6.1 (Insightful Corp., Seattle, WA) using the glm function; AICs calculated using the stepAIC function; DICs and the effective number of parameters (p_D) calculated in WinBUGS version 1.4 using 25 000 iterations after a burn-in of 25 000 iterations.

this single sample required >3000 person-hours), 95% of the time the true values of the parameters will fall within the estimated confidence intervals.

- 3 The model fit is reasonable. A linear regression of the observed data on the predicted values illustrates that the model accounts for 55% of the variance in the data.

Bayesian inference on the ant data

Bayesian inference uses not only the sample data but also any available prior information. Using Bayes' Theorem to calculate the posterior probability of the model conditional on the data requires explicit specification of the prior probability of the model – i.e. prior probability distributions for each of the model's parameters. Thus, we use eqn 1 to estimate the posterior:

$$P(\beta|\text{data}) \propto f(\text{data}|\beta) \times \pi(\beta). \quad (4)$$

The term $f(\text{data}|\beta)$ in eqn 4 is the likelihood of the data (the same as $L(\beta|\text{data})$ in eqn 2). As in the classical model, the likelihood is modelled as a Poisson random variable. The term $\pi(\beta)$ in eqn 4 is the prior. Many investigators choose to use non-informative normal (Gaussian) priors that reflect prior 'ignorance' (e.g. distributions of each of the parameters are centred on zero with very large variances so that the prior is integrable but is essentially uniform over the range of the data). Alternatively, priors can be gleaned from the literature or constructed using techniques developed for eliciting expert opinion (see Wolfson *et al.* 1996 for

an ecological example). Initially, I used uninformative, Gaussian priors on each of the β_i terms in eqn 3 ($\beta_i \sim N(0, 1000)$).

The computation of the posterior $P(\beta|\text{data})$ using Bayes' Theorem often involves numerical approximations of solutions of integrals (Gelman *et al.* 1995; Carlin & Louis 2000), especially when priors and likelihoods are not 'conjugate' (i.e. are of different functional forms: Gelman *et al.* 1995), as in this example. Available software most frequently uses Markov chain Monte Carlo (MCMC) methods (Gilks *et al.* 1996). For the ant example, I used WinBUGS version 1.4 (Spiegelhalter *et al.* 2003), which implements MCMC methods using a Gibbs sampler (Chib & Greenberg 1995). Posterior probability distributions on the regression parameters β were sampled from normal distributions. The most credible estimates of the parameters (Table 4) using the uninformative priors and the simple additive model (eqn 3) were nearly identical to the maximum likelihood estimates (Fig. 2).

Diversity patterns of ants have been documented around the world, and so it is reasonable to use the published literature to generate more informative priors for the model parameters. I derived priors for latitudinal gradients in temperate ants from Gotelli & Arnett (2000): $\beta_1 \sim N(-0.017, 0.04)$; for effects of elevation from Gotelli & Arnett (2000) and Brühl *et al.* (1999): $\beta_2 \sim N(-0.002, 0.0003)$; and for differences between 'open' habitats such as bogs and 'closed' habitats such as forests from Jeanne (1979), Gotelli & Arnett (2000) and Kaspari *et al.* (2000):

Table 3 Results of model selection for possible log-linear models relating species richness (S) to habitat type (H: forest or bog), latitude (L: decimal degrees), and elevation (E: metres above sea level)

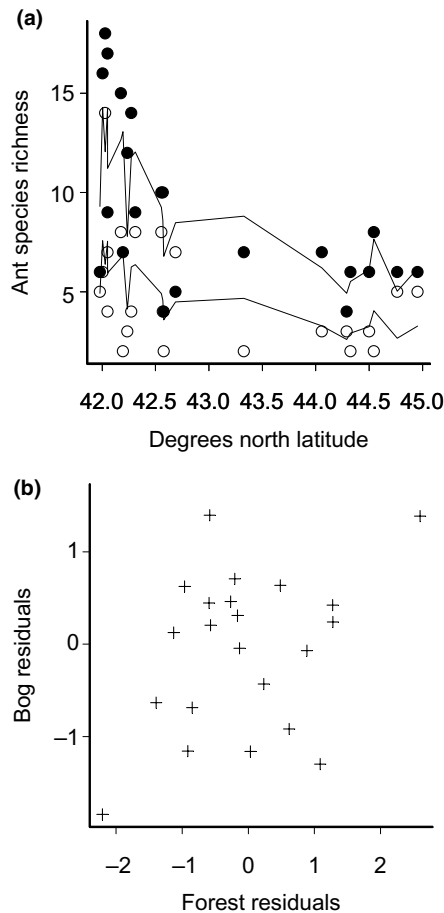


Figure 1 (a) Observed (points) and maximum likelihood (lines) estimates of the ant species richness at 22 bog (open circles) and surrounding forest (closed circles) sites in New England. (b) Paired deviance residuals of bog and forest habitats at each of the 22 sites.

$\beta_3 \sim N(0.37, 1)$. These priors are illustrated in the last column of Fig. 2, along with the posteriors estimated from these priors. Because the likelihood (i.e. the information in the data) had much smaller variance than these informative priors, the posteriors estimated from the informative priors differed only slightly from those with uninformative priors.

I also compared all the models listed in Table 3. One method of choosing among competing Bayesian models is the deviance information criterion (DIC) (Spiegelhalter *et al.* 2002):

$$\text{DIC} = \overline{D(\theta)} + p_D. \quad (5)$$

In words, DIC equals the posterior mean of the deviance of all the candidate models $\overline{D(\theta)}$ plus the effective number of parameters in the model (p_D). The posterior mean of the deviance $\overline{D(\theta)}$ itself equals -2 times the log of the likelihood, and the effective number of parameters in the

model is estimated as the posterior mean of the deviance minus the deviance of the posterior means ($\overline{D(\theta)} - D(\bar{\theta})$). In the absence of any prior information, $\text{DIC} = \text{AIC}$ (eqn 3), but the inclusion of prior information results in increases in both $\overline{D(\theta)}$ and the effective number of parameters (Spiegelhalter *et al.* 2002).

As with AIC, the model with the smallest DIC is selected to be the 'best' model. For the ant data set, applying the DIC to the set of models estimated with uninformative priors yielded the same result as applying AIC to the maximum likelihood models: the simple, additive model provided the best-fit with the fewest effective parameters (Table 3).

From the Bayesian analysis, the inferences are:

- 1 The additive model is a believable description of how latitude, elevation, and habitat can be used to predict species richness of ants in New England, USA.
- 2 There is a 95% probability that the estimated values of the model parameters in fact fall within the calculated credible sets (Table 3).
- 3 The model provides a good fit to the data. Figure 3 illustrates expected values and associated 95% credible sets for species richness in each habitat at each sampled site. For 73% (16 of 22) of the forests and 55% (12 of 22) of the bogs sampled, the probability is at least 95% that the model accurately predicts the observed value.

Uncertainty in model selection

There is recognized uncertainty in the parameter estimates of both classical and Bayesian models. Less often appreciated is the uncertainty involved in selecting a particular model relative to other plausible models (Chatfield 1995; Draper 1995). Yet, the incorrect specification or choice of a statistical model can result in faulty inferences or predictions. Automated tools for model selection such as the stepAIC function in S (Venables & Ripley 2002), the MARK software (White & Burnham 1999), or the DIC function in WinBUGS (Spiegelhalter *et al.* 2002) may have the unintended consequence of discouraging scientists from thinking about uncertainty in model selection. Recognizing uncertainty in parameter estimates and predictions of ecological models (e.g. IPCC 2001) and communicating the uncertainty in the range of ecological models considered (Wintle *et al.* 2003) can lead to better understanding by ecologists of the power and limitations of statistical inference and prediction.

I considered 15 models to 'explain' the species richness of ants in New England (Table 3). This example is relatively simple, as many complex ecological models include dozens of factors and the number of candidate models increases exponentially with the number of predictor variables. In this example, the same model was selected using AIC and DIC, but these values differed only by a few percent among

Table 4 Parameter estimates for the additive model (eqn 3) predicting ant species richness from habitat, elevation, and latitude

	Classical model (maximum likelihood estimate)	Bayesian models		
		Posterior mode, non-informative prior	Posterior mode, informative prior	Averaged model, non-informative prior
$\hat{\beta}_0$	11.95 (2.65) [6.81,17.73]	11.49 (1.87) [7.89, 15.32]	12.18 (2.22) [6.89, 16.33]	12.03 (2.65)
$\hat{\beta}_1$	-0.24 (0.06) [-0.36, -0.11]	-0.23 (0.04) [-0.31, -0.14]	-0.24 (0.05) [-0.33, -0.12]	-0.24 (0.06)
$\hat{\beta}_2$	-0.001 (0.0003) [-0.002, -0.0004]	-0.001 (0.0004) [-0.002, -0.0004]	-0.001 (0.0004) [-0.002, -0.0004]	-0.001 (0.0004)
$\hat{\beta}_3$	0.64 (0.06) [0.44, 0.75]	0.64 (0.12) [0.40, 0.88]	0.63 (0.12) [0.40, 0.84]	0.64 (0.12)

Values in parentheses are standard deviations of the parameter estimate; values in brackets are 95% confidence intervals (maximum likelihood estimates) or credible sets (Bayesian estimates). Bayesian posteriors calculated using WinBUGS version 1.4 using 25 000 iterations after a burn-in of 25 000 iterations. The last column gives the parameter estimates and standard errors for the averaged model that combines both the complete additive model (eqn 3) and the additive model without elevation. Model averaging done in S-Plus 6.1 using the bic.glm function.

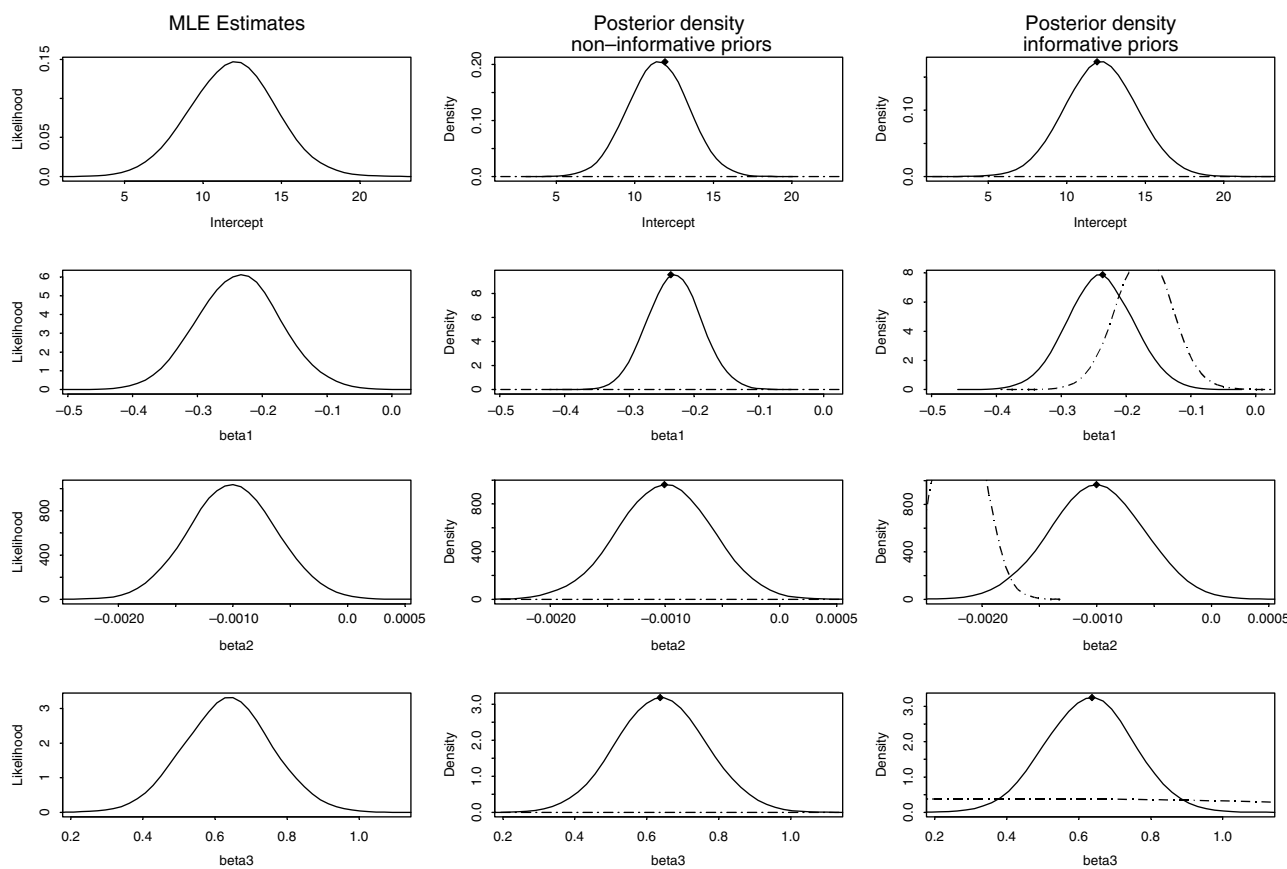


Figure 2 Maximum likelihood and Bayesian estimates for the parameters of the additive model: $\log(\hat{S}_i) = \hat{\beta}_0 + \hat{\beta}_1 \text{latitude}_i + \hat{\beta}_2 \text{elevation}_i + \hat{\beta}_3 \text{habitat}_i$. The first column of plots illustrates the likelihoods of the parameters. The second column illustrates Bayesian posteriors (solid lines) using non-informative priors (dotted lines), and the third column illustrates Bayesian posteriors (solid lines) using informative priors (dotted lines). The maximum likelihood estimate is shown as a solid diamond on the corresponding Bayesian posterior. In a given row, all plots have the same x-axis scaling to allow for comparisons between inferences. The y-axis scaling varies within rows, however.

several models and it is possible that one of the other models actually may be the ‘true’ model. One way to account for uncertainty in model construction and selection

is to create and use an ‘average’ model. The contribution of each individual model to the averaged model is weighted by its plausibility or posterior weight of evidence.

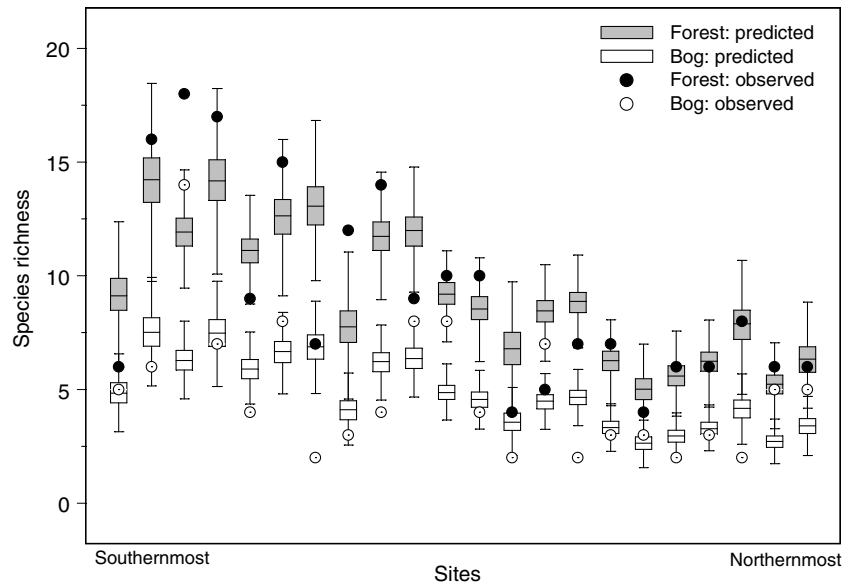


Figure 3 Bayesian estimates of species richness in each habitat at each of the 22 sampled sites. Box-plots illustrate expected median, quartiles, and 95% credible sets for species richness. Circles are observed values.

Frequentist model averaging is a nascent and promising area of statistical research (Claeskens & Hjort 2003; Hjort & Claeskens 2003), but it has not developed yet to the extent that it can be applied to even basic ecological problems. In contrast, Bayesian model averaging (reviewed by Hoeting *et al.* 1999) is an established method for combining models that has been applied only recently to ecological questions (Wintle *et al.* 2003). In the combined or averaged model, the individual models are weighted by their degree of plausibility. Normally, all possible individual models are not included in the averaged model. Rather, only those that meet a defined selection criterion are used. Madigan & Raftery (1994) suggested two criteria for inclusion: Occam's Window, which excludes models that predict the data 'far less well' (e.g. when the Bayes factor, the ratio of the posterior probabilities of the candidate model to the best model, is less than 0.05); and Occam's Razor, which excludes any complex model (e.g. more terms, more interactions) that receives less support from the data than simpler models.

Averaging generalized linear models, such as those used in the ant example, is relatively straightforward (Hoeting *et al.* 1999) and can be accomplished with freely available software.⁵ Using the S-function, `bic.glm` (Volinsky *et al.* 1997), and assigning equal prior weights to all the models, two plausible models were included in the averaged model: the additive model with all three predictors identified as previously as the 'best' model, and an additive model that included only habitat and latitude (Table 4). The parameter estimates were similar to the individual models, but the standard errors were larger, reflecting the uncertainty inherent in model averaging.

⁵<http://www.research.att.com/~volinsky/bma.html>.

THE FUTURE OF BAYESIAN ECOLOGY

Bayesian inference is fast becoming an accepted statistical tool among ecologists. Bayes' Theorem provides an intuitively clear alternative method for estimating parameters and expressing the degree of confidence or uncertainty in those estimates. Bayesian methods allow for the explicit incorporation of as much or as little existing data or prior knowledge that is available, and provides a direct measure of the probability of one or more hypotheses of interest.

The analysis of designed experiments can be approached either with frequentist or Bayesian methods. Information-theoretic and likelihood methods for choosing among multiple candidate models (Burnham & Anderson 2002) are better understood than Bayesian model selection methods, but both Bayesian and likelihood approaches can be used for model-based analyses (Hilborn & Mangel 1997). Finally, deciding whether to use Bayesian or frequentist inference demands an understanding of their differing epistemological assumptions. Strong statistical inference demands that ecologists not only confront models with data but also confront their own assumptions about how the world is structured.

ACKNOWLEDGMENTS

I thank Leon Blaustein and Nick Gotelli for suggesting and inviting this review; Brian Beckage, Nick Gotelli, Jacqueline Mohan, Michael Papaik, Mark Velland and participants in the Harvard Forest lab group for comments on the manuscript; Philip Dixon for advice about Poisson model structure in WinBUGS; Brian Dennis for re-awakening my interest in philosophy and epistemology; and members of

the likelihood short-course at the Institute for Ecosystem Studies, three anonymous reviewers, and Fangliang He for sharpening the presentation. The Harvard Forest supports my research on Bayesian inference.

REFERENCES

- Adams, W.T., Griffin, A.R. & Moran, G.F. (1992). Using paternity analysis to measure effective pollen dispersal in plant populations. *Am. Nat.*, 140, 762–780.
- Anderson, D.J. & Hodum, P.J. (1993). Predator behavior favors clumped nesting in an oceanic seabird. *Ecology*, 74, 2462–2464.
- Anholt, B.R., Werner, E. & Skelly, D.K. (2000). Effect of food and predators on the activity of four larval ranid frogs. *Ecology*, 81, 3059–3521.
- Barnett, V. (1999). *Comparative Statistical Inference*, 3rd edn. John Wiley & Sons, Ltd, Chichester, UK.
- Barrowman, N.J., Myers, R.A., Hilborn, R., Kehler, D.G. & Field, C.A. (2003). The variability among populations of coho salmon in the maximum reproductive rate and depensation. *Ecol. Appl.*, 13, 784–793.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philos. Trans.*, 53, 370–418.
- Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd edn. Springer-Verlag, New York, NY, USA.
- Berger, J.O. (2003). Could Fisher, Jeffreys and Neyman have agreed on testing (with comments and rejoinder). *Stat. Sci.*, 18, 1–32.
- Berger, J.O. & Berry, D.A. (1988). Statistical analysis and the illustration of objectivity. *Am. Sci.*, 76, 159–165.
- Blume, J.D. & Royall, R.M. (2003). Illustrating the law of large numbers (and confidence intervals). *Am. Stat.*, 57, 51–57.
- Bolker, B.M., Pacala, S.W. & Neuhauser, C. (2003). Spatial dynamics in model plant communities: what do we really know? *Am. Nat.*, 162, 135–148.
- Box, G.E.P. & Tiao, G.C. (1992). *Bayesian Inference in Statistical Analysis*. John Wiley & Sons, Inc., New York.
- Brühl, C.A., Mohamed, V. & Linsenmair, K.E. (1999). Altitudinal distribution of leaf litter ants along a transect in primary forests on Mount Kinabalu, Sabah, Malaysia. *J. Trop. Ecol.*, 15, 265–277.
- Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Inference: A Practical Information-theoretic Approach*, 2nd edn. Springer-Verlag, New York.
- Calder, C., Lavine, M., Muller, P. & Clark, J.S. (2003). Incorporating multiple sources of stochasticity into dynamic population models. *Ecology*, 84, 1395–1402.
- Carlin, B.P. & Louis, T.A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd edn. Chapman & Hall/CRC Press, Boca Raton, FL.
- Carpenter, S.R., Kitchell, J.F., Cottingham, K.L., Schindler, D.E., Christensen, D.L., Post, D.M. *et al.* (1996). Chlorophyll variability, nutrient input, and grazing: evidence from whole-lake experiments. *Ecology*, 77, 725–735.
- Chase, J.M. & Liebold, M.A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago, IL.
- Chatfield, C. (1995). Model uncertainty, data mining and statistical inference (with discussion). *J. R. Stat. Soc.*, A158, 419–466.
- Chib, S. & Greenberg, E. (1995). Understanding the Metropolis–Hastings algorithm. *Am. Stat.*, 49, 327–335.
- Claeskens, G. & Hjort, N.L. (2003). The focused information criterion. *J. Am. Stat. Assoc.*, 98, 900–945.
- Clark, C.W. & Mangel, M. (1984). Foraging and flocking strategies: information in an uncertain environment. *Am. Nat.*, 123, 626–641.
- Clark, J.S., Silman, M., Kern, R., Macklin, E. & HilleRisLambers, J. (1999). Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology*, 80, 1475–1494.
- Clark, J.S., Lewis, M., McLachlan, J.S. & HilleRisLambers, J. (2003a). Estimating population spread: what can we forecast and how well? *Ecology*, 84, 1979–1988.
- Clark, J.S., Mohan, J., Dietze, M. & Ibanez, I. (2003b). Coexistence: how to identify trophic trade-offs. *Ecology*, 84, 17–31.
- Congdon, P. (2001). *Bayesian Statistical Modeling*. John Wiley & Sons, Ltd, Chichester, UK.
- Cooper, A.B., Hilborn, R. & Unsworth, J.W. (2003). An approach for population assessment in the absence of abundance indices. *Ecol. Appl.*, 13, 814–828.
- Cottingham, K.L. & Carpenter, S.R. (1998). Population, community, and ecosystem variates as ecological indicators: phytoplankton responses to whole-lake enrichment. *Ecol. Appl.*, 8, 508–530.
- Cummings, C.L., Alexander, H.M., Snow, A.A., Rieseberg, L.H., Kim, M.J. & Culley, T.M. (2002). Fecundity selection in a sunflower crop-wild study: can ecological data predict crop allele changes? *Ecol. Appl.*, 12, 1661–1671.
- Damgaard, C. (1998). Plant competition experiments: testing hypotheses and estimating the probability of coexistence. *Ecology*, 79, 1760–1767.
- Dennis, B. (1996). Discussion: should ecologists become Bayesians? *Ecol. Appl.*, 6, 1095–1103.
- Dennis, B. (2004). Statistics and the scientific method in ecology (with commentaries and rejoinder). In: *The Nature of Scientific Evidence* (eds Taper, M.L. & Lele, S.R.). University of Chicago Press, Chicago, IL (in press).
- Dixon, P. & Ellison, A.M. (1996). Introduction: ecological applications of Bayesian inference. *Ecol. Appl.*, 6, 1034–1035.
- Dorazio, R.M. & Johnson, F.A. (2003). Bayesian inference and decision theory – a framework for decision making in natural resource management. *Ecol. Appl.*, 13, 556–563.
- Draper, D. (1995). Assessment and propagation of model uncertainty (with discussion). *J. R. Stat. Soc.*, B57, 45–97.
- Drechsler, M., Frank, K., Hanski, I., O'Hara, R.B. & Wissel, C. (2003). Ranking metapopulation extinction risk: from patterns in data to conservation management decisions. *Ecol. Appl.*, 13, 990–998.
- Edwards, A.W.F. (1992). *Likelihood*. Johns Hopkins University Press, Baltimore, MD.
- Efron, B. (1978). Controversies in the foundations of statistics. *Am. Math. Monthly*, 85, 231–246.
- Efron, B. (1986). Why isn't everyone a Bayesian? *Am. Stat.*, 40, 1–11.
- Ellison, A.M. (1996). An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecol. Appl.*, 6, 1036–1046.
- Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. London*, 222, 309–368.
- Fleishman, E., Nally, R.M. & Fay, J.P. (2003). Validation tests of predictive models of butterfly occurrence based on environmental variables. *Conserv. Biol.*, 17, 806–817.

- Forcada, J. (2000). Can population surveys show if the Mediterranean monk seal colony at Cap Blanc is declining in abundance? *J. Appl. Ecol.*, 37, 171–181.
- Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.
- Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC Press, Boca Raton, FL.
- van Gils, J.A., Schenk, I.W., Bos, O. & Piersma, T. (2003). Incompletely informed shorebirds that face a digestive constraint maximize net energy gain when exploiting patches. *Am. Nat.*, 161, 777–793.
- Gotelli, N.J. & Arnett, A.E. (2000). Biogeographic effects of red fire ant invasion. *Ecol. Lett.*, 3, 257–261.
- Gotelli, N.J. & Ellison, A.M. (2002a). Assembly rules for New England ant assemblages. *Oikos*, 99, 591–599.
- Gotelli, N.J. & Ellison, A.M. (2002b). Biogeography at a regional scale: determinants of ant species density in bogs and forests of New England. *Ecology*, 83, 1604–1609.
- He, F., Zhou, J. & Zhu, H. (2003). Autologistic regression model for the distribution of vegetation. *J. Agric. Biol. Environ. Stat.*, 8, 205–222.
- Hilborn, R. & Mangel, M. (1997). *The Ecological Detective: Confronting Models with Data*. Princeton University Press, Princeton, NJ.
- Hjort, N.L. & Claeskens, G. (2003). Frequentist model average estimators. *J. Am. Stat. Assoc.*, 98, 879–889.
- Hoeting, J.A., Madigan, D., Raftery, A.E. & Volinsky, C.T. (1999). Bayesian model averaging: a tutorial. *Stat. Sci.*, 14, 382–401.
- Howson, C. & Urbach, P. (1993). *Scientific Reasoning: The Bayesian Approach*, 2nd edn. Open Court Publishing Company, Peru, IL.
- Hubbard, R. & Byarri, M.J. (2003). Confusion over measures of evidence (P 's) vs. errors (α 's) in classical statistical testing (with discussion and rejoinder). *Am. Stat.*, 57, 171–182.
- Huston, M.A. (1994). *Biological Diversity: The Coexistence of Species on Changing Landscapes*. Cambridge University Press, New York.
- IPCC (2001). *Climate Change 2001: Synthesis Report*. Intergovernmental Panel on Climate Change, Geneva, Switzerland.
- Jeanne, R.L. (1979). A latitudinal gradient in rates of ant predation. *Ecology*, 60, 1211–1224.
- Jeffreys, H. (1961). *Theory of Probability*, 3rd edn. Oxford University Press, Oxford, UK.
- Karlin, S. (1968). Rate of approach to homozygosity for finite stochastic models with variable population size. *Am. Nat.*, 102, 443–455.
- Kaspari, M., O'Donnell, S. & Kercher, J.R. (2000). Energy, density, and constraints to species richness: ant assemblages along a productivity gradient. *Am. Nat.*, 155, 280–293.
- Koops, M.A. & Abrahams, M.V. (2003). Integrating the roles of information and competitive ability on the spatial distribution of social foragers. *Am. Nat.*, 161, 586–600.
- Layton, D.F. & Levine, R.A. (2003). How much does the far future matter? A hierarchical Bayesian analysis of the public's willingness to mitigate ecological impacts of climate change. *J. Am. Stat. Assoc.*, 98, 533–544.
- Levine, L., Asmussen, M., Olvera, O., Powell, J.R., de la Rosa, M.E., Salceda, V.M. *et al.* (1980). Population genetics of Mexican *Drosophila*. V. A high rate of multiple insemination in a natural population of *Drosophila pseudoobscura*. *Am. Nat.*, 116, 493–503.
- MacNally, R., Horrocks, G. & Pettifer, L. (2002). Experimental evidence for potential beneficial effects of fallen timber in forests. *Ecol. Appl.*, 12, 1588–1594.
- Madigan, D. & Raftery, A.E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *J. Am. Stat. Assoc.*, 89, 1535–1546.
- McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*, 2nd edn. Chapman & Hall, London.
- O'Hara, R.B., Arjas, E., Toivonen, H. & Hanski, I. (2002). Bayesian analysis of metapopulation data. *Ecology*, 83, 2408–2415.
- Olsson, O. & Holmgren, N.M.A. (1999). Gaining ecological information about Bayesian foragers through their behaviour. I. Models with predictions. *Oikos*, 87, 251–263.
- Oster, G. & Heinrich, B. (1976). Why do bumblebees major? A mathematical model. *Ecol. Monogr.*, 46, 129–133.
- Pascual, M.A. & Hilborn, R. (1995). Conservation of harvested populations in fluctuating environments: the case of the Serengeti wildebeest. *J. Appl. Ecol.*, 32, 468–480.
- Pearl, R. (1917). The probable error of a Mendelian class frequency. *Am. Nat.*, 51, 144–156.
- Pearson, K. (1907). On the influence of past experience on future expectation. *London, Edinburgh and Dublin Philos. Mag. J. Sci.*, 13 - Sixth Series, 365–378.
- Peterson, G.D., Carpenter, S.R. & Brock, W.A. (2003). Uncertainty and the management of multistate ecosystems: an apparently rational route to collapse. *Ecology*, 84, 1403–1411.
- Platt, W.J., Beckage, B., Doren, R.F. & Slater, H.H. (2002). Interactions of large-scale disturbances: prior fire regimes and hurricane mortality of savannah pines. *Ecology*, 83, 1566–1572.
- Pollak, E. (1974). The survival of a mutant gene and the maintenance of polymorphism in subdivided populations. *Am. Nat.*, 108, 20–28.
- Popper, K.R. (1959). *The Logic of Scientific Discovery*. Hutchinson, London.
- Raftery, A.E., Givens, G.H. & Zeh, J.E. (1995). Inference from a deterministic population dynamics model for bowhead whales (with comments and rejoinder). *J. Am. Stat. Assoc.*, 90, 402–430.
- Reckhow, K.H. (1990). Bayesian inference in non-replicated ecological studies. *Ecology*, 71, 253–259.
- Schmidt, K.A., Goheen, J.R., Naumann, R., Ostfeld, R.S., Schaubert, E.M. & Berkowitz, A. (2001). Experimental removal of strong and weak predators: mice and chipmunks preying on songbird nests. *Ecology*, 82, 2927–2936.
- Shen, T.J., Chao, A. & Lin, C.F. (2003). Predicting the number of new species in further taxonomic sampling. *Ecology*, 84, 798–804.
- Sivia, D.S. (1996). *Data Analysis: A Bayesian Tutorial*. Oxford University Press, Oxford, UK.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *J. R. Stat. Soc. B*, 64, 583–639.
- Spiegelhalter, D., Thomas, A., Best, N. & Lunn, D. (2003). *WinBUGS version 1.4 User Manual*. <http://www.mrc-bsu.cam.ac.uk/bug> (last accessed 10 April 2004).
- Strong, D.R., Whipple, A.V., Child, A.L. & Dennis, B. (1999). Model selection for a subterranean trophic cascade: root-feeding caterpillars and entomopathogenic nematodes. *Ecology*, 80, 2750–2761.
- Ter Braak, C.J.F. & Etienne, R.S. (2003). Improved Bayesian analysis of metapopulation data with an application to a tree frog metapopulation. *Ecology*, 84, 231–241.

- Toivonen, H.T.T., Mannila, H., Korhola, A. & Olander, H. (2001). Applying Bayesian statistics to organism-based environmental reconstruction. *Ecol. Appl.*, 11, 618–630.
- Valone, T.J. & Brown, J.S. (1989). Measuring patch assessment abilities of desert granivores. *Ecology*, 70, 1800–1810.
- de Valpine, P. & Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecol. Monogr.*, 72, 57–76.
- Venables, W.N. & Ripley, B.D. (2002). *Modern Applied Statistics with S-Plus*, 4th edn. Springer-Verlag, New York.
- Volinsky, C.T., Madigan, D., Raftery, A.E. & Kronmal, R.A. (1997). Bayesian model averaging in proportional hazard models: predicting the risk of a stroke. *Appl. Stat.*, 46, 433–448.
- Wade, P.R. (2001). Bayesian methods in conservation biology. *Conserv. Biol.*, 14, 1308–1316.
- Walters, C.J. & Holling, C.S. (1990). Large-scale management experiments and learning by doing. *Ecology*, 71, 2060–2068.
- White, G.C. & Burnham, K.P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study*, 46 (Suppl.), 120–138.
- Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003). The use of Bayesian model averaging to better represent uncertainty in ecological models. *Conserv. Biol.*, 17, 1579–1590.
- Wolfson, L.J., Kadane, J.B. & Small, M.J. (1996). Bayesian environmental policy decisions: two case studies. *Ecol. Appl.*, 6, 1056–1066.

Editor, Fangliang He

Manuscript received 17 December 2003

First decision made 5 February 2004

Manuscript accepted 1 April 2004