AN INTRODUCTION TO BAYESIAN INFERENCE FOR ECOLOGICAL RESEARCH AND ENVIRONMENTAL DECISION-MAKING[1,2]

AARON M. ELLISON

Department of Biological Sciences and Program in Environmental Studies, Mount Holyoke College, South Hadley, Massachusetts 01075-6418 USA

Abstract. In our statistical practice, we ecologists work comfortably within the hypothetico-deductive epistemology of Popper and the frequentist statistical methodology of Fisher. Consequently, our null hypotheses do not often take into account pre-existing data and do not require parameterization, our experiments demand large sample sizes, and we rarely use results from one experiment to predict the outcomes of future experiments. Comparative statistical statements such as "we reject the null hypothesis at the 0.05 level," which reflect the likelihood of our data given our hypothesis, are of little use in communicating our results to nonspecialists or in describing the degree of certitude we have in our conclusions. In contrast, Bayesian statistical inference requires the explicit assignment of prior probabilities, based on existing information, to the outcomes of experiments. Such an assignment forces the parameterization of null and alternative hypotheses. The results of these experiments, regardless of sample size, then can be used to compute posterior probabilities of our hypotheses given the available data. Inferential conclusions in a Bayesian mode also are more meaningful in environmental policy discussions: e.g., "our experiments indicate that there is a 95% probability that acid deposition will affect northeastern conifer forests." Based on comparisons with current statistical practice in ecology, I argue that a "Bayesian ecology" would (a) make better use of pre-existing data; (b) allow stronger conclusions to be drawn from large-scale experiments with few replicates; and (c) be more relevant to environmental decision-making.

Key words: Bayesian inference; decision analysis; environmental decision-making; epistemology; probability; statistical errors; uncertainty.

INTRODUCTION

On the whole, general ecological theory has, so far, been able to provide neither the largely descriptive, scientific conclusions often necessary for conservation decisions, nor the normative basis for policy.— Shrader-Frechette and McCoy 1993

Ecologists are in the midst of an introspective period in which we are questioning intensively the theoretical underpinnings, utility, predictive ability, and epistemological foundations of ecology (e.g., Peters 1991). Serious doubts have been raised about the utility of abstract, general theories that have been shown repeatedly to have little predictive value in field or laboratory situations (Weiner 1995). An alternative "bottom-up" approach, accumulating a succession of case studies from which general conclusions may be derived, has been advocated by Shrader-Frechette and McCoy (1993). The use of meta-analysis to synthesize large numbers of case studies (Gurevitch et al. 1992, Gurevitch and Hedges 1993, Warwick and Clarke 1993, Arnqvist and Wooster 1995) can identify common processes that have occurred under disparate conditions,

but it cannot be used to predict the likelihood of occurrence of those processes in other, unrelated situations. In addition, a bottom-up, inductivist approach may undermine ecology's claim to be a truly scientific endeavor (Popper 1968, Peters 1991).

These doubts are not new to our discipline; critical self-examination has occurred repeatedly since ecology emerged as a science in the late 19th century (e.g., Kingsland 1985, McIntosh 1985). However, the current episode of reflection coincides with a rapidly growing awareness of local, regional, and global environmental problems. Ecologists can and increasingly are expected to estimate the magnitude of responses of populations, communities, and ecosystems to anthropogenic stressors, to formulate experiments to examine potential and actual environmental impacts, and to design and implement strategies to ameliorate these impacts (e.g., Holling 1978, Orians et al. 1986, Lubchenco et al. 1991, Shrader-Frechette and McCoy 1993, Huenneke 1995, Underwood 1995). However, the counsel of ecologists will be considered only if it is perceived by decision-makers as scientifically accurate and legitimate, and if it is communicated intelligibly and meaningfully. Few decision-makers are, or have been, practicing ecologists, and they cannot be expected to interpret conclusions presented in technical jargon.

Bayesian statistical inference can be used to estimate

ecologically meaningful parameters and provides an explicit expression of the amount of uncertainty in these parameter estimates. Since Bayesian inference requires the investigator to use pre-existing data to develop quantitative, probabilistic, and parameterized hypotheses, Bayesian hypothesis testing (decision analysis) inevitably will lead to testable predictions, and permit the rapid development and refinement of applicable theory. Recent discussions of the relative preference for Type I vs. Type II statistical errors in pure and applied ecological contexts (e.g., Peterman 1990, Shrader-Frechette and McCoy 1992, Mapstone 1995) illustrate that conclusions based on falsificationist epistemology (Popper 1968, Strong 1980) are of little use to decision-makers, and allow for little flexibility in dynamic, changeable situations (Holling 1978, Lee 1993). In contrast, conclusions presented in a Bayesian framework can be understood more easily by decision-makers and by the public at large, because these conclusions are presented in clear, familiar language.

This paper is not meant as a complete review either of epistemological issues in ecology or of Bayesian parameter estimation and decision analysis. Rather, I first sketch some difficulties with the way that we infer conclusions from our data. Then, I briefly illustrate that Bayesian inference can provide an alternative statistical framework in which to couch experimental and observational data, synthesize existing information, generate useful (and perhaps novel) ecological theory, and contribute to sound environmental policy. Readers interested in more complete treatments of historical and epistemological issues in ecological thought can begin with recent books by McIntosh (1985), Peters (1991), and Shrader-Frechette and McCoy (1993). Accessible presentations of Bayesian techniques and their applicability to a wide range of "pure" and "applied" situations are provided by Good (1965), Iversen (1984), Howson and Urbach (1991), and Kass and Raftery (1995). Technical introductions of Bayesian inference are given by Lindley (1972), Box and Tiao (1973), Geisser (1993), Bernardo and Smith (1994), and Gelman et al. (1995). The additional papers in this Special Feature provide concrete examples of the use of Bayesian inference in both ecological research and environmental decision-making (Ludwig 1996, Taylor et al. 1996, Ver Hoef 1996, Wolfson et al. 1996). Other examples can be found in Carpenter (1990), Reckhow (1990), Walters and Holling (1990), Solow (1994), Gustafson and Franklin (1995), Nicholson and Barry (1995), and Raftery et al. (1995).

WHAT ECOLOGISTS SHOULD DO WITH STATISTICS, AND WHAT WE'D LIKE TO DO WITH THEM

Ecology is the science which says what everyone knows in language that no one understands.—Elton 1927

As ecologists, we study relationships between organisms and their physical and biological environments (e.g., Pianka 1994), and we describe these relationships using the language of statistics. Anyone who has taught introductory ecology probably would attest that this language can distort and refract Elton's "quantitative natural history" in strange and unpredictable ways. In general, we can identify two fundamental goals of an ecological experiment, whether descriptive or manipulative: *parameter estimation* and *hypothesis testing*. In this and the following section, I discuss frequentist and Bayesian approaches to these goals.

Frequentist parameter estimation

Given a random sample of n individuals from a larger population exposed to a treatment (T), we often calculate an average response value of the organism to the treatment. This average response value usually is reported as the mean value of all the observations: $\bar{X} = \Sigma x_i/n$. Assuming that our data (suitably transformed, if necessary) are a random sample, this estimate of the mean is known to be an unbiased estimator of the true population mean μ, no matter what the true value of μ (Efron 1978). More precisely, if we took an infinite number of samples of size n from this population, and calculated the mean of each sample, the expected value (the mean) of these means, $E(\bar{X}) = \mu$. We similarly can calculate the sample variance as $s^2 = \Sigma(x_i - \bar{X})^2/(n - 1)$, the sample standard deviation $s = \sqrt{s^2}$, and the expected standard deviation $E(s) = s/\sqrt{n}$ of the expected distribution of the set of repeatedly determined sample means $\{\bar{X}\}$ (also known as the standard error of the mean, s_X).

A central assumption underlying calculations of these and other commonly used statistical parameters is known as the *frequentist* assumption (e.g., Efron 1978): there is a *true, fixed value,* for each parameter of interest, and the expected value of this parameter is the average value obtained by random sampling repeated ad infinitum (Table 1). From an ecological perspective, there are many difficulties with this assumption. Within experiments, true randomization is difficult, replication is often small, misidentified (Hurlbert 1984), or by virtue of circumstance, nonexistent (e.g., Carpenter 1990, Reckhow 1990). Ecological experiments rarely are repeated independently. No two organisms are exactly alike, and consequently they are unlikely to respond to our treatment in exactly the same way. Evolution virtually guarantees that even if they were alike today, their offspring will be measurably different. Thus, the idea that there is a true, fixed value for any ecologically meaningful statistical parameter is a Platonic phantom. Furthermore, even if there was a true, fixed value for a given statistical parameter, frequentist statistical theory states that we can never know it. Although we implicitly hope that we have sampled appropriately and obtained a realistic estimator of our population mean, it is likely that the best we can hope

TABLE 1. Some fundamental differences between frequentist and Bayesian statistical inference
in their uses and interpretations of statistical concepts and terms.

Concept or term	Frequentist interpretation	Bayesian interpretation
Probability	Result of an infinite series of trials conducted under identical conditions	The observer's degree of belief, or the organized appraisal in light of the data
Data	Random (representative) sample	Fixed (all there is)
Parameters	Fixed	Random
$k\%$ confidence interval	This interval will include the true value of a given parameter in $k\%$ of all possible samples	$k\%$ of the possible parameter values will fall within the confidence (credibility) interval
Treatment of nuisance parameters	Conditions on sufficient statistics or maximum likelihood estimate	Integrates over all possible values
Conclusion	$P(x \mid H)$	$P(H \mid x)$

to do is to state the probability distribution in which the average response value, and unobserved values, are likely to occur (Simberloff 1980).

The standard way we get around this dilemma is to compute the reliability of our estimated statistical parameter by constructing a "$k\%$ confidence interval" around $\bar{X}$

$$P(\bar{X} - t_{\alpha[n-1]} \cdot s_{\bar{X}} \leq \bar{X} \leq \bar{X} + t_{\alpha[n-1]} \cdot s_{\bar{X}}) = (1 - \alpha) = k \tag{1}$$

where $t_{\alpha[n-1]}$ = the percent (α) of area in both tails of a t distribution beyond the indicated value of t for $n-1$ degrees of freedom. By convention, we most commonly compute a 95% confidence interval based on a normal distribution, where $t_{0.05[\infty]} = 1.96$. Unfortunately, Eq. 1 tells us only that the interval $(\bar{X} - 1.96 \cdot s_{\bar{X}}, \bar{X} + 1.96 \cdot s_{\bar{X}})$ will include the true value of μ in 95% ($P = 0.95$) of the infinitely many repeated samples from our population (Table 1). We cannot use Eq. 1 to state that there is a $k\%$ probability that the true mean does in fact occur within the confidence interval we created from our experiment, and Eq. 1 is not a probability distribution in which we expect μ to occur (Howson and Urbach 1991). Since μ is fixed, it is either inside a confidence interval or outside it. It makes no sense to assert that a fixed parameter would occur in a fixed confidence interval only $k\%$ of the time (Sokal and Rohlf 1995). In addition, for most parameterized, nonnormal distributions, the actual probability that Eq. 1 includes μ is substantially ($\geq 60\%$) less than the expected coverage k (e.g., Robinson 1975).

Frequentist hypothesis testing

Once we have an estimate of the parameter of interest, we normally test hypotheses regarding that parameter; e.g., was the parameter affected by the applied treatment? Assume that we have collected a set of random, independent observations $x = \{x_1, \ldots, x_n\}$ from a normal population to which an experimental treatment has been applied, and that the sample variance approximates the population variance. If we substitute the expected mean value (μ_0) under a point null statistical hypothesis (e.g., H_0: $\mu = \mu_0$, where μ_0 may be

obtained from a set of controls) for $\bar{X}$ in Eq. 1, and use $t_{\alpha[n-1]}$, we obtain the upper and lower bounds of the "rejection region" for this null hypothesis at the α level. In other words, for the two-tailed test H_1: $\mu \neq \mu_0$, if $\bar{X} <$ the lower bound or $\bar{X} >$ the upper bound of Eq. 1, it is unlikely that our sample comes from the population where the expected value of $\bar{X}$, $E(\bar{X}) = \mu_0$. In standard statistical jargon, we would "reject the null hypothesis" (H_0) that $\mu = \mu_0$ at the α level of significance, or we would state that our results are "significant" with probability (P value) $P \leq \alpha$.

The meaning of this result is far from obvious; Sokal and Rohlf (1995) devote nearly 12 pages to explaining it. Most importantly, supporting H_1 or inferring cause and effect from P values is not possible (see also Lewontin 1974). Precisely, if we reject H_0 at the α level, we are asserting that we will incorrectly reject a true null hypothesis (i.e., commit a Type I statistical error) with frequency α, if we repeated this experiment an infinite number of times. The converse error (Type II), accepting a false null hypothesis, depends on the specified level of α, the sample size, the magnitude of the true effect, and varies with the underlying (and often unknown) population distribution.

Formally, a P value is the probability of observing our results conditional on H_0 (i.e., $\alpha = P(x|H_0)$ *and* the probability of observing all results less likely than the observed result. In other words, the P value overstates how unlikely the data really are, because it is weighted by additional, unlikely, and unobserved results (e.g., Berger 1985, Berger and Berry 1988). Note also what a P value does not tell us. It does not tell us how probable our null hypothesis actually is given the data (i.e., $P(H_0|x)$), nor does it tell us the probability of our alternative hypothesis (i.e., $P(H_1|x)$). Based on the statement that observed data are unlikely given a null hypothesis ($P(x|H_0)$ is small), one cannot conclude that the alternative hypothesis is likely (or "true") given the data ($P(H_1|x)$ is large); this commits the logical fallacy of affirming the consequent (e.g., Howson and Urbach 1991). That is, if a null hypothesis implies a particular set of data ($p \rightarrow q$), then although observing different data implies that the null hypothesis is false

(not q → not p, or symbolically $\neg q \rightarrow \neg p$), it does not necessarily imply that the alternative hypothesis produced the observed data ($\neg p \not\rightarrow \neg q$). In fact, even if $P(x|H_0)$ is small, $P(H_0|x)$ may be quite large (and $P(H_1|x)$ consequently small), with frequency as high as an order of magnitude $>\alpha$ (Lindley 1957, Berger and Delampady 1987, Berger and Sellke 1987). Similarly, failure to reject H_0 does not indicate the evidence in favor of it (Kass and Raftery 1995). In point of fact, $P(H_1|x)$, the probability that the treatment "caused" the results (or the likelihood of the alternative hypothesis, given the data), is what most scientists and policy-makers actually want to know.

It is arguable that most ecologists are testing an implicit *statistical* null hypothesis of the form H_0: $\mu = 0$, and that they are using their statistical analysis as a benchmark to indicate that they in fact had obtained an adequate sample to observe any trend. Put another way, the point of statistical analyses in ecological research is the testing of a *scientific* hypothesis that the imposed treatment had the hypothesized effect (and that the sample size was large enough to detect it). This difference between statistical and scientific hypotheses also illustrates the importance of reporting statistical power (= 1 − the probability of committing a Type II statistical error). If sample size is small, power will be low, and the *statistical* null hypothesis may be falsely accepted. However, if sample size is large, power may be so high that it may be very unlikely to ever accept a statistical null hypothesis, even if the true effect is small and biologically irrelevant (see also Raftery 1995). An example illustrates this distinction between statistical and biological significance: for even modest sample size ($n = 50$) and one independent variable, a correlation coefficient need only explain 7% of the variance in the data to be significant at the 0.05 level (Rohlf and Sokal 1995). For larger sample sizes, it may be virtually impossible to ever accept the statistical null hypothesis, even if the observed treatment effect is ecologically irrelevant. One way around this difficulty is to set a substantially lower α level for rejection of the statistical null hypothesis when sample size and statistical power are high (e.g., Raftery 1995). There are no frequentist guidelines for such reductions in α for given sample sizes, however.

A randomly chosen set of 50 papers published in *Ecology* in 1994 bears out the assertion that ecologists principally are testing scientific, not statistical hypotheses. Forty-nine of the papers examined at least one scientific hypothesis using standard statistical techniques, such as *t* tests, regression, ANOVA, etc., and regularly reported *P* values (however, only 1 of 50 reported statistical power). All but one found statistically "significant" results. Despite the oft-repeated statement that the use of explicit statistical null hypotheses or multiple working (scientific) hypotheses are needed if ecology is to mature as a science (e.g., Strong 1980, Peters 1991) only 10 of these papers ex-

plicitly presented null hypotheses, while 14 (including 5 of the 10 with null models) used multiple hypotheses. Thus, in 61% of the sample (30 of 49 papers), it was not clear what statistical or scientific hypothesis was being rejected by the reported *P* values. Similar results have been found in literature surveys of aquatic sciences (Bourget and Fortin 1995), psychology, and medicine (Sterling et al. 1995).

In general, however, classical hypothesis testing and rejection appeared unimportant, as the authors of 47 of these papers asserted that their results (i.e., data) supported or confirmed their hypothesis. In other words, 96% of the authors considered $P(H_1|x)$ to be high, despite the fact that the authors actually tested $P(x|H_0)$. If we were really serious about testing plausible ecological null hypotheses (i.e., those that we expect to accept with high frequency), there would not be a dearth of papers reporting "nonsignificant" results. In fact, we normally expect to accept a single ecological alternative hypothesis by rejecting a statistical H_0 (and consequently rarely state it); otherwise we would not have done the experiment in the first place. On the other hand, if we expect to reject an ecological H_0 in the first place, then it would be more informative to know the likelihood of both H_0 and H_1 once the experiment is complete, in light of the data obtained. Additional data could then be used to revise the likelihood of these, and other, alternative hypotheses. Frequentist hypothesis testing and *P* values do not provide this type of information. Construction of sound, predictive ecological theory that can both advance ecological understanding and contribute to environmental policy decisions requires that we precisely state how likely a particular ecological process is to affect variables of interest.

BAYESIAN INFERENCE FOR ECOLOGISTS

When we make a scientific generalization we do not assert the generalization and its consequences with certainty; we assert that they have a high degree of probability on the knowledge available to us at the time, but that this probability may be modified by additional knowledge.—Jeffreys 1931

Bayesian inference provides a mechanism, based on the probability calculus, to quantify the uncertainty in parameter estimates, and to determine the probability that an explicit scientific hypothesis is true given ("conditional on") a set of data. Bayesian inference treats statistical parameters as random variables (Table 1), and uses a likelihood function to express the relative plausibility of obtaining different values of this parameter when particular data have been observed (McCullagh and Nelder 1991).

Bayesian parameter estimation

Bayesian parameter estimation begins with the observation that the joint probability of two events, $P(\theta$

x), equals the product of the probability of one of the events and the conditional probability of the second event given the first one.

$$P(x) \cdot P(\theta|x) = P(\theta x) = P(\theta) \cdot P(x|\theta). \quad (2)$$

Rearranging terms in Eq. 2 yields an expression for $P(\theta|x)$, or the *posterior probability* of obtaining the parameter θ given the data at hand

$$P(\theta|x) = \frac{P(x|\theta) \cdot P(\theta)}{P(x)}. \quad (3)$$

This expression is known as Bayes' theorem (Bayes 1763). In this expression, $P(\theta)$ is the *prior probability* of obtaining the specified parameter. In other words, $P(\theta)$ is the probability of observing θ that is expected by the investigator *before* the experiment is conducted.

There are three interpretations of the prior probability $P(\theta)$ (Cox and Hinkley 1974): (1) a frequency distribution whose parameters reflect analysis and synthesis of existing data (see Taylor et al. 1996 for an example); (2) an "objective" statement of what is rational to believe about the initial parameter or distribution, given initial ignorance of this parameter (see Ludwig 1996 and Ver Hoef 1996 for examples); or (3) a subjective measure of what the investigator actually believes. (Wolfson 1996, and references therein, discusses how to elicit such statements of belief.)

Ecologists and others trained in frequentist statistics are most likely to use the first or second interpretation of $P(\theta)$ when designing and analyzing an experiment. The information required to construct $P(\theta)$ as a frequency distribution normally is reported qualitatively in the introduction to a research paper or in descriptions of study sites and study species, but these data rarely are used quantitatively, even in review papers (Peters 1991). Objective statements of $P(\theta)$ reflecting total ignorance can be expressed with *noninformative* priors, such as $P(\theta) =$ a uniform distribution, where all values are equally likely (Jeffreys 1961; see Ludwig 1996 for an example). However, since scientific research is an incremental process, in which new hypotheses are based on preexisting data, it is unlikely that we would ever be in a situation of complete ignorance regarding possible values of hypothetical parameters (Wolfson 1996).

The other term in the numerator, $P(x|\theta)$ is Fisher's *likelihood function* for the parameter (Box and Tiao 1973, Reckhow 1990). The denominator is the expected value of the likelihood function, and acts as a scaling constant that normalizes the sum or integral of the area under the posterior probability distribution. Because the denominator in Eq. 3 is a constant, this equation has the form of:

posterior probability $\propto$ likelihood $\times$ prior probability.

This statement indicates that the likelihood function is what modifies prior knowledge into posterior expectations (Box and Tiao 1973).

A simple example of parameter estimation illustrates the use of Eq. 3 and the interaction between the investigators' prior probability distributions and the likelihood function. Consider two ecologists interested in estimating the fraction of foliar area (here denoted by β) of red spruce (*Picea rubens* Sarg.) affected by a predefined concentration of acid deposition in a previously unstudied location in Vermont. The first ecologist (A) has extensive experience with red spruce at other sites in Vermont, and before visiting the new study site, she estimates β to be 0.4 $\pm$ 0.05 (mean ± 1 SD), and asserts that population values of β form a normal distribution; in statistical notation, $\beta \sim N$ (0.4, 0.05²). Therefore, ecologist A's *prior probability* function (solid black line in Fig. 1) is:

$$P_A(\beta) = \frac{1}{\sqrt{2\pi} \cdot 0.05} \exp\left[-\frac{1}{2}\left(\frac{\beta - 0.4}{0.05}\right)^2\right]. \quad (4)$$

In addition to being an expression of possible values for β, Eq. 4 is also a testable hypothesis about the distribution of possible β values. The second ecologist (B) has much less experience with red spruce in Vermont, and estimates $\beta \sim N(0.2, 0.1^2)$. Therefore,

$$P_B(\beta) = \frac{1}{\sqrt{2\pi} \cdot 0.1} \exp\left[-\frac{1}{2}\left(\frac{\beta - 0.2}{0.1}\right)^2\right] \quad (5)$$

(dotted line in Fig. 1).

Subsequent to obtaining adequate funding, these two ecologists measure β by growing 10 red spruce seedlings in a controlled environment and applying nitric acid at a specified concentration. The results are $b = \{b_1, \ldots, b_{10}\}$ values that are normally distributed with mean $\bar{b} = 0.3$ and standard deviation 0.075. The standardized likelihood function (shaded curve in Fig. 1) for this result is $\sim N(\bar{b}, [s/\sqrt{n}]^2)$, where the variance = the squared standard error of the mean. Following Box and Tiao (1973), the posterior distribution of β given $\{b_i\}$, $P(\beta|b)$, $\sim N(\bar{\beta}_n, \bar{\sigma}_n^2)$, where

$$\bar{\beta}_n = \frac{1}{w_0 + w_n} \cdot (w_0\beta_0 + w_n\bar{b}) \quad (6)$$

and

$$\frac{1}{\bar{\sigma}_n^2} = w_0 + w_n \quad (7)$$

$n =$ the sample size, β_0 is the prior mean, $w_0 = 1/\sigma_0^2$ (the reciprocal of the prior variance) and $w_n = n/s^2$. In this example, ecologist A reports a posterior estimate of foliar loss $\bar{\beta}_n = 0.318 \pm 0.021$ (mean ± 1 SD), while ecologist B reports $\bar{\beta}_n = 0.295 \pm 0.023$.

Despite the large initial discrepancies between the two ecologists' estimates of β and their differing degrees of certitude in their estimates (expressed as each ecologist's prior σ_0^2), the application of Bayesian inference to their data following a single experiment leads to close agreement of their subsequent parameter
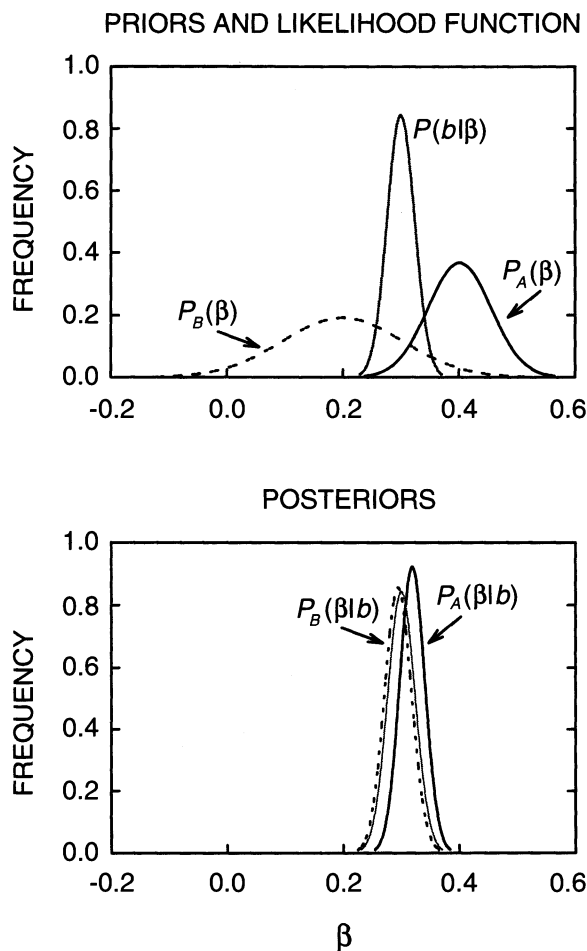
PRIORS AND LIKELIHOOD FUNCTION

POSTERIORS

FIG. 1. Top: illustration of the relationship between the two hypothetical prior probability distributions described by Eqs. 6 and 7, and the likelihood function based on the hypothetical experimental data. Bottom: Posterior probability distributions resulting from applying Bayes' theorem to the top panel.

estimates (Fig. 1). Since Bayesian inference is an iterative process, the posterior probability distribution of β obtained by this experiment can now be used as the prior probability distribution for a new set of experiments, should further refinement of this estimate or additional hypothesis testing be required.

This example also illustrates a general principle of Bayesian inference. For parametric distributions, as long as the likelihood function "dominates" the prior distributions (i.e., $w_0 \ll w_n$), the data will have a much greater effect on the posterior probability function than will the prior (Box and Tiao 1973). They also point out that this conforms to standard scientific practice. If the prior dominates the likelihood, then the experiment is probably irrelevant, since that implies the existence of more prior information than the subsequent experiment can supply to influence posterior estimates. This is an important result, as one of the common criticisms of Bayesian inference is that specifying a prior

probability distribution is a subjective process that is inappropriate to scientific research, wherein the data are expected to speak for themselves. Recall, however, the subjective nature of what scientists regularly do. We rarely, if ever, test all possible hypotheses, and most of us use substantial prior knowledge about the behavior of a system in designing our experiments (e.g., Berger and Berry 1988, Johnson 1990, Lewontin 1991). Unlike classical frequentist statistical practice, Bayesian inference requires the investigator to state assumptions explicitly and use pre-existing information quantitatively in order to define the prior distribution or hypothesis.

Each ecologist can place a 95% Bayesian credibility interval on her estimate of β:

$$P(\bar{\beta}_n - 2\sqrt{D} \le \beta \le \bar{\beta}_n + 2\sqrt{D}) = 0.95 \qquad (8)$$

where $D = \bar{\sigma}_n^2$ from Eq. 7 (Efron 1978). While the Bayesian credibility interval (or bounds on the posterior probability) is calculated analogously to a frequentist confidence interval (Eq. 1), the interpretations are very different (Table 1). Because Bayesian inference treats statistical estimates (here, β) as random variables, in this example the interpretation of Eq. 8 is that 95% of the potential values of β will fall within the boundaries of the credibility interval. If the investigator uses a noninformative prior, the bounds of Eq. 8 will equal the bounds of Eq. 1. However, given some prior information, Eq. 8 normally will be narrower than Eq. 1 (e.g., Deely and Zimmer 1969).

Bayesian hypothesis testing

Alternative prior probability distributions for specific parameters can also be viewed as multiple working hypotheses. Bayes' theorem (Eq. 3) can be extended to assess the relative probabilities of such alternative, quantitative hypotheses given the available data (e.g., Jeffreys 1961, Iversen 1984, Reckhow 1990, Bernardo and Smith 1994). Although a frequentist analysis attempts to "reject" an hypothesis, and a Bayesian speaks in terms of its "likelihood," both probably would agree that statistical hypothesis testing should be used to assess the evidence in favor of the null hypothesis.

For multiple hypotheses H_i, Eq. 3 can be generalized:

$$P(H_i \mid x) = \frac{P(x \mid H_i) \cdot P(H_i)}{P(x)} \qquad (9)$$

where

$$P(x) = \begin{cases} \sum P(x \mid H_j) \cdot P(H_j) & H \text{ discrete} \\ \int P(x \mid H) \cdot P(H) \, dH & H \text{ continuous} \end{cases} \qquad (10)$$

again is a scaling constant equal to the sum of the conditional probabilities $P(x \mid H_j)$ weighted by their prior probabilities $p(H_j)$. The ratio of the posterior probabilities of two alternative hypotheses (often called the

TABLE 2. Interpretation of Bayes factors suggested by Jeffreys (1961) and Kass and Raftery (1995). In this table, the value given is based on B_{10}, the evidence *against* H_0, where the likelihood of H_0 would be the denominator and the likelihood of H_1 would be the numerator in Eq. 11. This reciprocal of the likelihoods given in Eq. 11 is how most ecologists would be likely to use Bayes factors in day-to-day research.

Jeffreys (1961)			Kass and Raftery (1995)		
$\log_{10}(B_{10})$	B_{10}	Evidence against H_0	$2 \times \log_e(B_{10})$	B_{10}	Evidence against H_0
0–0.5	1–3.2	Not worth more than a bare mention	0–2	1–3	Not worth more than a bare mention
0.5–1	3.2–10	Substantial	2–6	3–20	Positive
1–2	10–100	Strong	6–10	20–150	Strong
>2	>100	Decisive	>10	>150	Very strong

"odds ratio") gives the relative evidence for one hypothesis over another:

$$\frac{P(H_0 \mid x)}{P(H_1 \mid x)} = \frac{P(H_0)}{P(H_1)} \cdot \frac{P(x \mid H_0)}{P(x \mid H_1)}. \quad (11)$$

Eq. 11 is often stated as:

posterior odds = prior odds $\times B_{01}$

and B_{01} is called the "Bayes factor" (reviewed by Kass and Raftery 1995). Jeffreys (1961) and Kass and Raftery (1995) proposed scales using Bayes factors with which to decide whether or not data fail to support H_0 or favor H_1 (Table 2). Note that if two hypotheses are considered to be equally likely a priori ($P(H_0) = P(H_1) = 0.5$), then the Bayes factor = the posterior odds in favor of H_0. If two hypotheses are simple distributions (all parameters specified), and H_0 is nested within H_1, then B_{01} is simply the standard likelihood ratio (Reckhow 1990, Kass and Raftery 1995). For nested hypotheses, if any parameter (e.g., expected mean, variance) is unknown for either of the hypotheses (so-called *nuisance parameters*), then the values for the terms of B_{01} must be obtained by integrating (or summing, in the discrete case), *not maximizing,* the likelihood over the parameter space (see Eq. 10; Kass and Raftery 1995). If H_0 and H_1 are not nested, then standard frequentist likelihood ratio tests are inapplicable (Kass and Raftery 1995). Use of Bayes factors requires parameterization of the prior. If the prior probability distributions are noninformative (e.g., the uniform distribution), then B_{01} is also undefined.

As an example, consider a simple comparison between two means, where the means are each the amount of red spruce foliage lost to two different applications of nitric acid ($n = 10$ plants per treatment). The null hypothesis H_0 is $\mu_1 = \mu_2$ (or, equivalently, $\mu_1 - \mu_2 = 0$), and the alternative H_1 is $\mu_1 \neq \mu_2$. Assume that the mean responses are $\bar{X}_1 = 0.31$ and $\bar{X}_2 = 0.30$, and the underlying populations are known to be normal, with $\sigma^2 = 0.01$. A standard (frequentist) t test on simulated data (drawn from $N(0.31, 0.01)$ and $N(0.30, 0.01)$) finds that these two samples are "significantly different" at $P = 0.016$. For this simple case, where the populations are both normal and their variances are known, the posterior odds in favor of H_0 can be calculated using methods outlined by Cox and Hinkley (1974: Chapter

10). If the prior probabilities of H_0 and H_1 both equal 0.5, then the posterior odds in favor of H_0 in this example = 0.06. Equivalently, this means that the odds are $\approx 17:1$ in favor of the alternative hypothesis that the two means are different. This provides "positive" (Kass and Raftery 1995) to "strong" (Jeffreys 1961) evidence in favor of H_1 (Table 2). Note that if there were a priori evidence that H_0 was more likely than H_1, then the posterior odds in favor of H_0 would increase as a function of $P(H_0)/P(H_1)$. In addition, this process and attendant results are substantially more informative about the quantitative importance of acid deposition than a statement that simply rejects a null hypothesis that acid deposition has no effects on spruce foliage. Reckhow (1990) similarly showed that effective use of prior probabilities allowed for biologically meaningful resolution among statistically contradictory outcomes of studies examining effects of acidification on sulfate concentrations, acid neutralizing capacity, and total base cations in Adirondack lakes.

I illustrated above (*Frequentist hypothesis testing*) that as sample size increases, it becomes increasingly unlikely to accept H_0 for a fixed significance level (e.g., $\alpha = 0.05$), and that to distinguish biological significance from statistical significance, the α level should be reduced as sample size increases (Raftery 1995). Although there are no frequentist guidelines for such reductions, Cox and Hinkley (1974) illustrate that the posterior odds in favor of H_0 scale directly with sample size for a given significance level. However, they find that, when contrasting means from normal distributions where the variances between treatments are approximately equal, the posterior odds in favor of H_0 will be roughly constant if the α is reduced as a function of $n \log(n)$, where n is sample size. Thus, the frequentist and Bayesian approaches to hypothesis testing give comparable results with respect to the "truth" of H_0 if α is reduced as n increases (Cox and Hinkley 1974). Note, however, that a frequentist would not assess directly the probability *in favor of* either the null or the alternative hypothesis.

Bayesian inference is not limited to parameter estimation or simple hypothesis testing. There are Bayesian procedures to perform post-hoc multiple comparisons among means (Waller and Duncan 1969); to con-

struct and test alternative regression models (e.g., Hartigan 1969, Hoadley 1970, Lindley and Smith 1972, Box and Tiao 1973, Halpern 1973, Dempster et al. 1977); to estimate the relative importance of variance components in common ANOVA models (fixed effects models, random effects models, mixed models, block designs); to analyze time-series models (e.g., Spall 1988, Pole et al. 1994); and to conduct sensitivity analyses of simulation models whose parameterization is uncertain (Raftery et al. 1995).

BAYESIAN INFERENCE, ADAPTIVE MANAGEMENT, AND ENVIRONMENTAL DECISION-MAKING

Ecological understanding of complex phenomena is essential if society is to anticipate and ameliorate the environmental effects of human activities.—Lubchenco et al. 1991

Ecologists cannot ignore anthropogenic impacts on the biosphere, and many of us want to contribute substantively to decision-making processes surrounding environmental problems (e.g., Orians et al. 1986, Lubchenco et al. 1991, Shrader-Frechette and McCoy 1993, Vitousek 1994, Huenneke 1995, Underwood 1995). Despite Vitousek's (1994) assertion that we are certain that there are a number of components of global environmental change that are occurring and are driven by human activities, a quantitative expression of that certitude, stated in a way that is meaningful to decision-makers, normally is absent from ecological publications. The lack of quantifiable uncertainty often is used by ecologists to justify their lack of involvement with the decision-making process, and by some decision analysts as a vehicle to avoid using scientific information in the process (Underwood 1995:242). However, uncertainty is fundamental to *all* scientific activities, and people regularly make decisions based on uncertain data (e.g., weather forecasts). Ludwig et al. (1993) clearly illustrated that ignoring ecological uncertainty has led repeatedly to environmental catastrophes.

Most formal frameworks for making decisions explicitly incorporate uncertainty by using odds ratios (Eq. 11) to decide among alternative courses of action (e.g., Chernoff and Moses 1959, Lindley 1971, Berger 1985, Smith 1988, Chechile and Carlisle 1991). For any situation in which a decision must be made, there is a problem to be solved, a set of possible actions, a set of uncertain events associated with each action, and a set of consequences that can occur subsequent to the events (Bernardo and Smith 1994). The degree of uncertainty that a particular event will occur can be expressed quantitatively as a prior probability, the decision can be seen as an experiment, and potential consequences can be estimated as posterior probabilities. Each possible consequence also needs to be assigned a quantitative value (usually referred to as its *utility*; Lindley 1971). Although most often expressed in monetary terms, the utility need only express some standardized value that individuals or society place on each hypothesized outcome. For example, the utilities of different emission-control policies could be expressed as expected changes in net photosynthetic rates of affected trees.

Bayes' theorem is used in decision analysis to estimate the consequences of a decision (as posterior probabilities) based on uncertainty (prior probability) and events (likelihood functions). Bayesian decision theory demonstrates that the optimal decision is the one that maximizes the product of the utility and the posterior probability of the consequence of the given decision (e.g., Lindley 1971, Berger 1985, Chechile and Carlisle 1991). Alternatively, Shrader-Frechette and McCoy (1993) suggest that a *minimax* decision rule (choose the decision for which the maximum possible loss or risk is minimized) is preferable in environmental decision-making, as it guards against choosing a course of action leading to the most egregious environmental impacts (recall Peterman 1990). However, Bernardo and Smith (1994) point out that the minimax rule (which does not depend on a prior probability distribution) is equivalent to the Bayes decision rule that uses the prior probability distribution associated with the highest expected risk. In other words, the minimax decision rule is appropriate in environmental situations only when prior data provide strong evidence for substantial negative effects (and dominate the likelihood functions). This is likely to be the case only when there is substantial prior information and little uncertainty about possible outcomes.

However, environmental decisions rarely are made in light of complete and certain data, so decisions should be made in ways that reflect the uncertainty and that can be modified when new data become available. "Adaptive management" (e.g., Holling 1978, Orians et al. 1986, Lee 1993) incorporates initial uncertainty, treats decisions as hypotheses to be tested, and demands that managers learn from the consequences of their decisions and alter their decisions (and implement new decisions) accordingly. Adaptive management is precisely analogous to an iterative Bayesian learning and decision process. Prior information is specified, decisions are made, and consequences are observed. The consequences are treated not as final events, but as new sources of information (new prior probability functions) for subsequent "experiments" (events, likelihood functions) that lead to modifications in management practices (new decisions). As suggested by Holling (1978), Orians et al. (1986), and Underwood (1995), ecologists can contribute to this process at all stages. We can gather, synthesize, and meta-analyze information to construct prior hypotheses, we can treat decisions as experiments, and we can analyze results as posterior hypotheses to be tested rather than as done deals. Bayesian inference and decision theory provide a quantitative framework and intelligible language in

which to analyze and express adaptive management procedures.

CONCLUDING REMARKS

What is unknown and of interest? H. What is known? X. Calculate P(H | X). How? The probability calculus is the only tool. The Bayesian argument provides a recipe.—Lindley, in Efron 1986

Theory and methods of Bayesian inference have developed rapidly in the last 50 yr, but they are still discussed rarely, if at all, where ecologists encounter and learn statistics: in introductory statistics classes and biometry textbooks. Efron (1986) argued that Bayesian inference is applied rarely to scientific data analysis because it is more difficult to solve for Bayesian posterior probabilities than it is to compute P values. For some ecological data, Eq. 9 would require integration of functions that may be solvable only by numerical approximation. As yet, there is little easily used software for conducting Bayesian analyses, but this is changing rapidly (e.g., Smith and Gelfand 1992, Albert 1993, Cook and Broemeling 1995, Kass and Raftery 1995, Raftery et al. 1995, Ver Hoef 1996). Many ecological datasets and hypotheses, however, could be analyzed using Bayesian techniques without undue difficulty. Bernardo and Smith (1994) provide formulae, likelihood functions, and methods to calculate prior and posterior distributions for 8 univariate and 2 multivariate parameterized discrete distributions, and 16 univariate and 8 multivariate parameterized continuous distributions. The use of Bayesian inference to analyze nonparametric data (i.e., where the distribution functions are unknown) has to date received less attention (Ferguson 1973, Leonard 1978, Bernardo and Smith 1994).

Conceptually, Bayesian inference is the most straightforward way of analyzing and interpreting our hypotheses in light of our data. As ecologists, we have a vast store of natural history and experimental data with which to address uncertain hypotheses. New observations, experiments, and statistical analyses are the only tools we have to evaluate critically these hypotheses. If we are seriously interested in testing our hypotheses in light of our data, we should use tools appropriate to the task.

LITERATURE CITED

Albert, J. H. 1993. Teaching Bayesian statistics using sampling methods and MINITAB. American Statistician **47**: 182–191.

Arnqvist, G., and D. Wooster. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends in Ecology and Evolution **10**:236–240.

Bayes, T. 1763. An essay towards solving a problem in the doctrine of chances. Philosophical Transactions of the Royal Society of London **53**:370–418.

Berger, J. O. 1985. Statistical decision theory and Bayesian analysis. Springer-Verlag, Berlin, Germany.

Berger, J. O., and D. A. Berry. 1988. Statistical analysis and the illusion of objectivity. American Scientist **76**:159–165.

Berger, J. O., and M. Delampady. 1987. Testing precise hypotheses (with comments and rejoinder). Statistical Science **2**:317–352.

Berger, J. O., and T. Sellke. 1987. Testing a point null hypothesis: the irreconcilability of P values and evidence (with comments and rejoinder). Journal of the American Statistical Association **82**:112–139.

Bernardo, J. M., and A. F. M. Smith. 1994. Bayesian theory. John Wiley and Sons, New York, New York, USA.

Bourget, E., and M.-J. Fortin. 1995. A commentary on current approaches in the aquatic sciences. Hydrobiologia **300/301**:1–16.

Box, G. E. P., and G. C. Tiao. 1973. Bayesian inference in statistical analysis. Addison-Wesley, Reading, Massachusetts, USA.

Carpenter, S. R. 1990. Large-scale perturbations: opportunities for innovation. Ecology **71**:2038–2043.

Chechile, R. A., and S. Carlisle, editors. 1991. Environmental decision making: a multidisciplinary perspective. Van Nostrand Reinhold, New York, New York, USA.

Chernoff, H., and L. E. Moses. 1959. Elementary decision theory. John Wiley and Sons, New York, New York, USA.

Cook, P., and L. D. Broemeling. 1995. Bayesian statistics using Mathematica. American Statistician **49**:70–76.

Cox, D. R., and D. V. Hinkley. 1974. Theoretical statistics. Chapman and Hall, London, UK.

Deely, J. J., and W. J. Zimmer. 1969. Shorter confidence intervals using prior observations. Journal of the American Statistical Association **64**:378–386.

Dempster, A. P., M. Schatzoff, and N. Wermuth. 1977. A simulation study of alternatives to ordinary least squares (with comments and rejoinder). Journal of the American Statistical Association **72**:77–106.

Efron, B. 1978. Controversies in the foundations of statistics. American Mathematical Monthly **85**:231–246.

———. 1986. Why isn't everyone a Bayesian? (with comments and reply). American Statistician **40**:1–11.

Elton, C. S. 1927. Animal ecology. Macmillan, New York, New York, USA.

Ferguson, T. S. 1973. A Bayesian analysis of some nonparametric problems. Annals of Statistics **1**:209–230.

Geisser, S. 1993. Predictive inference: an introduction. Chapman and Hall, New York, New York, USA.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, London, UK.

Good, I. J. 1965. The estimation of probabilities: an essay on modern Bayesian methods. M.I.T. Press, Cambridge, Massachusetts, USA.

Gurevitch, J., and L. V. Hedges. 1993. Meta-analysis: combining the results of independent experiments. Pages 378–398 *in* S. M. Scheiner and J. Gurevitch, editors. Design and analysis of ecological experiments. Chapman and Hall, New York, New York, USA.

Gurevitch, J., L. Morrow, A. Wallace, and J. Walsh. 1992. The meta-analysis of competition in field experiments. American Naturalist **140**:539–572.

Gustafson, L., and W. Franklin. 1995. Predictors of neonatal mortality in the guanacos (*Lama guanaco*) of Torres del Paine, Chile: a pilot study in adaptive modeling of wildlife population health. Pages 285–289 *in* T. B. Herman, S. Bon-

drup-Nielson, J. H. M. Willison, and N. P. Munro, editors. Ecosystem monitoring and protected areas. Science and Management of Protected Areas Association, Wolfville, Nova Scotia, Canada.

Halpern, E. F. 1973. Polynomial regression from a Bayesian approach. Journal of the American Statistical Association **68**:137–143.

Hartigan, J. A. 1969. Linear Bayesian methods. Journal of the Royal Statistical Society, Series B **31**:446–454.

Hoadley, B. 1970. A Bayesian look at inverse linear regression. Journal of the American Statistical Association **65**:356–369.

Holling, C. S., editor. 1978. Adaptive environmental assessment and management. John Wiley and Sons, Chichester, UK.

Howson, C., and P. Urbach. 1991. Bayesian reasoning in science. Nature **350**:371–374.

Huenneke, L. F. 1995. Involving academic scientists in conservation research: perspectives of a plant ecologist. Ecological Applications **5**:209–214.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs **54**: 187–211.

Iversen, G. R. 1984. Bayesian statistical inference. Sage, Beverly Hills, California, USA.

Jeffreys, H. 1931. Scientific inference. Cambridge University Press, Cambridge, UK.

―――. 1961. Theory of probability. Third edition. Oxford University Press, Oxford, UK.

Johnson, J. G. 1990. Method of multiple working hypotheses: a chimera. Geology **18**:44–45.

Kass, R. E., and A. E. Raftery. 1995. Bayes factors. Journal of the American Statistical Association **90**:773–795.

Kingsland, S. E. 1985. Modeling nature: episodes in the history of population ecology. University of Chicago Press, Chicago, Illinois, USA.

Lee, K. N. 1993. Compass and gyroscope: integrating science and politics for the environment. Island Press, Washington, DC, USA.

Leonard, T. 1978. Density estimation, stochastic processes and prior information (with discussion). Journal of the Royal Statistical Society, Series B **40**:113–146.

Lewontin, R. C. 1974. The analysis of variance and the analysis of causes. American Journal of Human Genetics **26**: 400–411.

―――. 1991. Facts and the factitious in natural sciences. Critical Inquiry **18**:140–153.

Lindley, D. V. 1957. A statistical paradox. Biometrika **44**: 187–192.

―――. 1971. Making decisions. John Wiley and Sons, London, UK.

―――. 1972. Bayesian statistics, a review. Conference Board of the Mathematical Sciences—National Science Foundation Regional Conference Series in Applied Mathematics, Number 2. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, USA.

Lindley, D. V., and A. F. M. Smith. 1972. Bayes estimates for the linear model (with discussion). Journal of the Royal Statistical Society, Series B **34**:1–41.

Lubchenco, J., A. M. Olson, L. B. Brubaker, S. R. Carpenter, M. M. Holland, S. P. Hubbell, S. A. Levin, J. A. MacMahon, P. A. Matson, J. M. Melillo, H. A. Mooney, C. H. Peterson, H. R. Pulliam, L. A. Real, P. J. Regal, and P. G. Risser. 1991. The sustainable biosphere initiative: an ecological research agenda. Ecology **72**:371–412.

Ludwig, D. 1996. Uncertainty and the assessment of extinction possibilities. Ecological Applications **6**:1067–1076.

Ludwig, D., R. Hilborn, and C. Walters. 1993. Uncertainty,

resource exploitation, and conservation: lessons from history. Science **260**:17, 36.

Mapstone, B. D. 1995. Scalable decision rules for environmental impact studies: effect size, Type I, and Type II errors. Ecological Applications **5**:401–410.

McCullagh, P., and J. A. Nelder. 1991. Generalized linear models. Second edition. Chapman and Hall, London, UK.

McIntosh, R. P. 1985. The background of ecology: concept and theory. Cambridge University Press, Cambridge, UK.

Nicholson, M., and J. Barry. 1995. Inferences from spatial surveys about the presence of an unobserved species. Oikos **72**:74–78.

Orians, G. H., J. Buckley, W. Clark, M. Gilpin, C. Jordan, J. Lehman, R. May, G. Robilliard, D. Simberloff, W. Erchmann, D. Policansky, and N. Grossblatt. 1986. Ecological knowledge and environmental problem solving. National Academy Press, Washington, D.C., USA.

Peterman, R. M. 1990. The importance of reporting statistical power: the forest decline and acidic deposition example. Ecology **71**:2024–2027.

Peters, R. H. 1991. A critique for ecology. Cambridge University Press, Cambridge, UK.

Pianka, E. R. 1994. Evolutionary ecology. Harper Collins, New York, New York, USA.

Pole, A., M. West, and J. Harrison. 1994. Applied Bayesian forecasting and time series analysis. Chapman and Hall, New York, New York, USA.

Popper, K. R. 1968. The logic of scientific discovery. Harper and Row, New York, New York, USA.

Raftery, A. E. 1995. Bayesian model selection in social research. Pages 111–196 *in* P. V. Marsden, editor. Sociological methodology 1995. Blackwell Scientific, Oxford, UK.

Raftery, A. E., G. H. Givens, and J. E. Zeh. 1995. Inference from a deterministic population dynamics model for bowhead whales (with comments and rejoinder). Journal of the American Statistical Association **90**:402–430.

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

Robinson, G. K. 1975. Some counterexamples to the theory of confidence intervals. Biometrika **62**:155–161.

Rohlf, F. J., and R. R. Sokal. 1995. Statistical tables. Third edition. W. H. Freeman, New York, New York, USA.

Shrader-Frechette, K. S., and E. D. McCoy. 1992. Statistics, costs and rationality in ecological inference. Trends in Ecology and Evolution **7**:96–99.

Shrader-Frechette, K. S., and E. D. McCoy. 1993. Method in ecology: strategies for conservation. Cambridge University Press, Cambridge, UK.

Simberloff, D. 1980. A succession of paradigms in ecology: essentialism to materialism and probabilism. Synthèse **43**: 3–39.

Smith, A. F. M., and A. E. Gelfand. 1992. Bayesian statistics without tears: a sampling—resampling perspective. American Statistician **46**:84–88.

Smith, J. Q. 1988. Decision analysis: a Bayesian approach. Chapman and Hall, London, UK.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. Third edition. W. H. Freeman, New York, New York, USA.

Solow, A. R. 1994. On the Bayesian estimation of the number of species in a community. Ecology **75**:2139–2142.

Spall, J. C., editor. 1988. Bayesian analysis of time series and dynamic models. Marcel Dekker, New York, New York, USA.

Sterling, T. D., W. L. Rosenbaum, and J. J. Weinkam. 1995. Publication decisions revisited: the effect of the outcome of statistical tests on the decision to publish and vice versa. American Statistician **49**:108–112.

Strong, D. R., Jr. 1980. Null hypotheses in ecology. Synthèse **43**:271–285.

Taylor, B. L., P. R. Wade, R. A. Stehn, and J. F. Cochrane.

1996. A Bayesian approach to classification criteria for Spectacled Eiders. Ecological Applications **6**:1077–1089.

Underwood, A. J. 1995. Ecological research and (and research into) environmental management. Ecological Applications **5**:232–247.

Ver Hoef, J. M. 1996. Parametric empirical Bayes methods for ecological applications. Ecological Applications **6**: 1047–1055.

Vitousek, P. M. 1994. Beyond global warming: ecology and global change. Ecology **73**:1861–1876.

Waller, R. A., and D. B. Duncan. 1969. A Bayes rule for the symmetric multiple comparisons problem. Journal of the American Statistical Association **64**:1484–1503.

Walters, C. J., and C. S. Holling. 1990. Large-scale management experiments and learning by doing. Ecology **71**: 2060–2068.

Warwick, R. M., and K. R. Clarke. 1993. Comparing the severity of disturbance: a meta-analysis of marine macrobenthic community data. Marine Ecology Progress Series **92**:221–231.

Weiner, J. 1995. On the practice of ecology. Journal of Ecology **83**:153–158.

Wolfson, L. J., J. B. Kadane, and M. J. Small. 1996. Bayesian environmental policy decisions: two case studies. Ecological Applications **6**:1056–1066.