ALTERNATIVES TO STATISTICAL HYPOTHESIS TESTING IN ECOLOGY: A GUIDE TO SELF TEACHING

N. Thompson Hobbs[1,3] and Ray Hilborn[2]

[1]*Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523 USA*
[2]*School of Aquatic and Fisheries Sciences, Box 355020, University of Washington, Seattle, Washington 98195-5020 USA*

Abstract. Statistical methods emphasizing formal hypothesis testing have dominated the analyses used by ecologists to gain insight from data. Here, we review alternatives to hypothesis testing including techniques for parameter estimation and model selection using likelihood and Bayesian techniques. These methods emphasize evaluation of weight of evidence for multiple hypotheses, multimodel inference, and use of prior information in analysis. We provide a tutorial for maximum likelihood estimation of model parameters and model selection using information theoretics, including a brief treatment of procedures for model comparison, model averaging, and use of data from multiple sources. We discuss the advantages of likelihood estimation, Bayesian analysis, and meta-analysis as ways to accumulate understanding across multiple studies. These statistical methods hold promise for new insight in ecology by encouraging thoughtful model building as part of inquiry, providing a unified framework for the empirical analysis of theoretical models, and by facilitating the formal accumulation of evidence bearing on fundamental questions.

Key words: AIC; Bayesian; ecological modeling; information theoretic; likelihood; meta-analysis; model selection.

INTRODUCTION

Ecologists have traditionally relied on a relatively narrow set of statistical techniques to gain insight. This set has typically included analysis of variance, *t* tests, contingency tables, and regression—techniques routinely taught in elementary statistics classes. The preponderance of papers published in contemporary ecological journals use one of these techniques to make inferences from data (Fig. 1). There are, of course, a broad range of problems that yield to such methods, notably problems that are amenable to replicated, manipulative experiments. However, despite the success of these traditional approaches in analyzing data from designed experiments, there is an increasing appreciation among ecologists that a singular focus on manipulative experimentation and associated analyses compresses the range of questions that ecologists can address (Hilborn and Mangel 1997, Anderson et al. 2000, Burnham and Anderson 2002, Hobbs et al. 2006). There is danger that questions are chosen for investigation by ecologists to fit widely sanctioned statistical methods rather than statistical methods being chosen to meet the needs of ecological questions.

During the last decade, dramatic increases in computing power have made it far easier to estimate parameters in ecological models using likelihood and Bayesian techniques. These methods and associated procedures for model selection and evaluation offer new opportunities to gain insight from observational and manipulative studies (Johnson and Omland 2004). To facilitate discussion, we will refer to these methods as the new statistics, acknowledging that these techniques are not new in the sense that they have been recently developed by statisticians. However, they are new to most ecological researchers. Their use has increased rapidly (Fig. 1).

The aim of the new statistics is to evaluate the relative strength of evidence in data for hypotheses represented as models. Traditionally, models used by ecologists for statistical inference have been limited to a relatively small set of linear forms. The functional forms and definitions of parameters in these models were chosen for statistical reasons; that is, they were not constructed to explicitly symbolize biological states and processes. Consequently, composing models to represent hypotheses has traditionally played a relatively minor role in developing testable statements by most ecological researchers. Statistical models were used to represent verbal hypotheses, and little thought was applied to the model building that ultimately supported inference.

The relationship of the new statistics to models is different. Models and modeling come to the fore in the

FIG. 1. Average number of articles per issue in journals of the Ecological Society of America using traditional statistics and what we call the new statistics. Counts were generated for traditional statistics by searching for the strings "analysis of variance," "*t* test," "chi-square," and "linear regression" in the full text of articles published during 1984–2003. Counts for new statistics were based on the search strings "Bayesian," "maximum-likelihood," and "model selection." Counts for each search were then summed by year to generate a total number of articles, which was divided by the number of issues published that year.
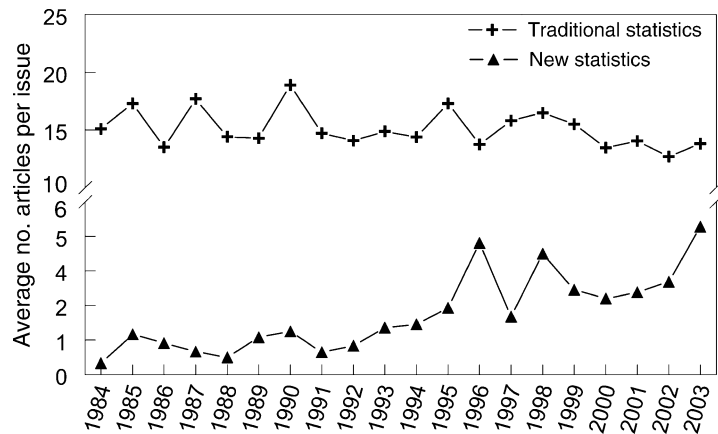
new statistics—the full range of models that can be imagined by the researcher are subject to analysis. As before, these include purely statistical models, whose parameters are not defined biologically, but now it is also possible to analyze biological models, whose variables and parameters explicitly symbolize ecological states and processes. The new statistics require deliberate, thoughtful specification of models to represent competing ecological hypotheses.

We believe that the new statistics offer an important alternative to traditional methods in the analysis of ecological data. We believe these techniques are accessible to the broad range of practicing researchers. In this paper, we provide an entry point for learning the new statistics and their application to ecological problems. We outline the essential features of these approaches, provide supporting text and demonstration material, and offer citations allowing more in-depth study.

This paper will be organized as follows. We begin by describing the role of models in representing ecological hypotheses, an idea that forms the foundation of the new statistics. We then describe the concept of likelihood and sketch its use in estimating values of parameters in biological and statistical models. Likelihood is a central feature of contemporary techniques for model selection and multimodel inference based on information theoretics, which we describe next. We close by schematically treating several related topics, including inferences based on data from diverse sources, Bayesian approaches to parameter estimation, and meta-analysis. It is not our purpose to provide a detailed, comprehensive treatment of these topics, but rather to offer motivation and general concepts needed as a starting point for self-teaching.

MODELS AS HYPOTHESES IN ECOLOGY

We begin with the idea that all statements in science are approximations of a complex truth, and the work of scientists is to evaluate how well or how poorly our statements achieve that approximation (Burnham and Anderson 2002). Models are a particularly precise sci-

entific statement because they avoid the inherent ambiguity of language by using mathematics to symbolize states and processes. Some scientists have distinguished between models and hypotheses (Caswell 1988, Hall 1988, Onstad 1988, Ulanowicz 1988), but we believe that distinction is not useful. Thus, in our view, the starting point for all analyses in the new statistics is the formulation of a set of models representing multiple competing hypotheses about the way ecological systems work. Formulating these alternatives requires the ecologist to bring together his or her knowledge about the system under study, knowledge that can come from scientific literature, previous experiments, or observations and experience. This prior knowledge can enter into the analysis in a formal way, as we will briefly discuss in a later section. However, it is also used informally to develop the candidate set of models to be evaluated. In either case, we cannot overstate the importance of this step: formulating alternative hypotheses expressed mathematically. The ability to gain insight pivots on developing models that make competing predictions and that can be evaluated with data at hand or data that will be collected. These models need to make useful, interpretable statements about processes in nature. Thus, the foundation of the new statistics in ecology is built from the ability of ecologists to represent testable ideas as mathematical models.

One of the most important decisions in building ecological models is how detailed they should be made. Historically, this decision was made subjectively by the modeler. The new statistics allow this decision to be made objectively: the level of detail in the model is decided by the data available to estimate the parameters in the model. All types of models used by ecologists, including dynamic differential and difference equations applied to time-series data (e.g., Forchhammer et al. 1993, Tanaka and Nishii 1997, Grenfell et al. 1998, Mduma et al. 1999, Dennis and Otten 2000, Bjornstad et al. 2002), static linear and nonlinear models of ecological relationships (e.g., Strong et al. 1999,

Vucetich et al. 2002, Hobbs et al. 2003), as well as traditional statistical models (Garrott et al. 2003), can be evaluated by the techniques offered here (also see Hilborn and Mangel 1997, Burnham and Anderson 2002). A discussion of how to formulate models is beyond the scope of this paper, but there are many excellent texts providing a general introduction to ecological modeling (e.g., Edelstein-Keshet 1988, Haefner 1996, Hastings 1997, Gurney and Nisbet 1998, Gershenfeild 1999, Taubes 2001).

USING LIKELIHOOD TO EVALUATE MODELS WITH DATA

Maximum likelihood estimation of model parameters

In this section, we introduce the concept of likelihood, referring the reader to several texts (Azzalini 1996, Royall 1997, Clayton and Hills 1998, Pawitan 2001) for more detailed, in-depth treatments of the topic. The likelihood approach to analysis differs in an important way from traditional statistics (Royall 1997). In the traditional framework, we formulate a null and an alternative hypothesis, design an experiment or sampling protocol to test the null hypothesis, and take observations that are appropriate for our test. We then ask the question, ''What is the probability that we would observe the data if the null hypothesis were true?'' This question is answered by a P value, which is defined as the probability of obtaining a value of the test statistic more extreme than the one observed given that the null hypothesis is true. If the probability of observing the test statistic given the hypothesis is small, then we reject the hypothesis and accept the alternative. Traditional hypothesis testing is based on calculating the probability of observing specific data ($\mathbf{Y}$) or data more extreme than $\mathbf{Y}$ given the null hypothesis, θ_{null}, that is, $P(\mathbf{Y}_{ext} \leq \mathbf{Y} \mid \theta_{null})$ (Edwards 1992, Royall 1997). Note that P values associated with traditional statistical tests do not assess the strength of evidence supporting a hypothesis or model. Rather, they are statements about the probability of extreme events that we have not observed (Royall 1997). In this framework, the hypothesis is fixed and the data are seen as variable.

The new statistics depart from this approach. In evaluating a set of models, we presume that we have data in hand. This doesn't mean that all analyses need to be post hoc; the approaches we describe are equally suitable to designed investigations. However, it does mean that when the analysis starts, we have data and we wish to understand the strength of support in the data for competing models (i.e., hypotheses). So, in contrast to the probability statement above, we are now interested in the likelihood ($\mathcal{L}$) of competing hypotheses given the data, which is proportionate to the probability of the data given the hypothesis (Edwards 1992):

$$\mathcal{L}(\theta \mid \mathbf{Y}) \propto P(\mathbf{Y} \mid \theta). \qquad (1)$$

It is important to understand that θ represents a set of parameter values specifying a particular model or models and that we are interested in the probability of observing the data $\mathbf{Y}$ given those parameter values. So, competing models can have the same set of parameters with different values for them, or more importantly, can have different sets of parameters specifying different model formulations. In this framework, the data are fixed and the hypothesis is variable. Estimating the likelihood of different models given the data (Eq. 1) allows us to evaluate the relative support in the data for each model. It also allows us to evaluate the strength of evidence supporting each model relative to competitors. The remainder of this section will schematically describe how to do that.

Ecological models make predictions about states and processes of interest as functions of parameters (θ). Thus, in this context, ''model,'' ''parameters,'' and ''hypothesis'' can be used more or less interchangeably. We can gain insight from ecological models by comparing their predictions to observations. For a single observation, the likelihood of the prediction of a model is proportionate to the probability of making that observation conditional on the model's parameters. More formally, presume we have a model $f(\theta)$ that makes predictions on a variable of interest, for example, population density, carbon flux, or metabolic rate. We have a data set $\mathbf{Y}$ composed of n individual observations on that variable, $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$. It follows from Eq. 1 that likelihood of the value of the parameter (or parameters) in our model given a single observation y_i is

$$\mathcal{L}(\theta \mid y_i = g(y_i \mid \theta) \qquad (2)$$

where the function $g(\ )$ is a probability function (if the y_i are discrete) or a probability density function (if the y_i are continuous). Because we are interested in making statements about the *relative* support in data for alternative models, we can assume that the constant of proportionality equals 1. Eq. 2 gives the likelihood of the model parameters conditional on a single observation, but to evaluate models based on full data sets, we must be able to maximize likelihoods by comparing multiple observations with multiple predictions. When we can assume that the deviations are independent of one another, then

$$\mathcal{L}(\theta \mid \mathbf{Y}) = \prod_{i=1}^{n} g(y_i \mid \theta). \qquad (3)$$

Eq. 3 simply means that the likelihood of the parameters given the full data set is the product of the probabilities of the individual likelihoods given the model parameters. Again, we seek to find a set of model parameters that maximizes the likelihoods. For computational simplicity, as well as for more fundamental reasons, it is often more useful to maximize the log likelihoods, in which case we have the following:

$$\ln[\mathcal{L}(\theta \mid \mathbf{Y})] = \sum_{i=1}^{n} \ln[g(y_i \mid \theta)]. \qquad (4)$$

Computing the likelihood or the log likelihood as a function of parameter values or model predictions provides a likelihood profile, which allows us to see how the model's likelihood changes as parameter values are changed (Appendix A). Likelihood profiles can be used to calculate confidence intervals on model parameters (Hilborn and Mangel 1997, Clayton and Hills 1998) as we will illustrate subsequently.

In our view, the concept of likelihood is revealed more clearly by example than statistical formalism, so we offer a simple illustration in Appendix A. Of course, the starting point for any analysis based on likelihood is choosing the correct likelihood function, $g(\ )$ (Eq. 2). In Appendix B, we outline how that choice is made.

Model selection using information theoretics

We argued above that formulating competing models of ecological systems represents the starting point of the new statistics. Evaluating the evidence in data for these competitors is a process known formally as model selection. Although model selection can be accomplished using several approaches (Johnson and Omland 2004), we will focus on techniques using information theoretics, touching on Bayesian methods in a later section. There are two reasons for our emphasis. First, information theoretics enjoy a fundamental, theoretical basis for separating the response of models to noise in data from their response to information. We believe, philosophically, that this separation is a basic aim of science. Second, these techniques are quite accessible to ecologists; self teaching can be pursued in depth using the readable and comprehensive reference of Burnham and Anderson (2002). We highly recommend this book to those who wish to learn how to evaluate evidence supporting models. In this section, we introduce some of the central ideas developed more fully in Burnham and Anderson (2002).

It is important to understand the philosophical basis for model selection using information theoretics. Imagine that we have a set of ecological models, and we want to compare that set to a perfect model—one that represents reality without error—a model that contains all of the information about the process or system of interest. We will call this the true model (recognizing the non sequitur) or more simply, the truth. We would like to know how well, or how poorly, the individual models in our set approximate the truth.

We can answer that question by examining the consequences of using a model to represent truth; in particular, we are interested in how much information we lose when we use a model to portray the "true" system. This quantity is given by the Kullback-Leibler (hereafter, KL) distance (also called the KL information discrepancy), which measures the amount of information

that is lost when a given model is used to represent truth (Burnham and Anderson 2002). As such, it represents a fundamental quantity in science.

Although the mathematical details of the calculation of the KL distance are beyond the scope of this paper, these calculations explicitly contain a term for truth, a term that is vital to the interpretation of the KL distance, but makes it difficult to use in a practical, statistical sense. Scientists never know "truth." Thus, an enormous advance in science was made when Hirotugu (Akaike 1973) linked fundamental concepts in information theory with fundamental concepts in mathematical statistics by deriving a way to estimate the expected value of the relative KL distance without explicitly knowing truth. This linkage provides a rigorous way to evaluate the strength of evidence in data for alternative models. Reiterating, Akaike's Information Criterion, or AIC, is an estimate of the expected value of the KL distance:

$$\text{AIC} = -2 \ln[\mathcal{L}(\hat{\theta} \mid \mathbf{Y})] + 2K \qquad (5)$$

where $\mathcal{L}(\hat{\theta} \mid \mathbf{Y})$ is the likelihood associated with the maximum likelihood estimates of the model parameters, $\hat{\theta}$, given the data set $\mathbf{Y}$, and K is the number of parameters in the model plus any variance terms needed in the likelihood function that are estimated from the data. There are variations of AIC suitable for small samples (i.e., AIC_c) and for overdispersed data (QAIC), which are discussed more fully elsewhere (Burnham and Anderson 2002). Here, it suffices to understand that AIC provides a way to evaluate ecological models using data. Models can be ranked using AIC where the best model, the one that sacrifices the least information when it is used to approximate truth, has the lowest AIC value. The relative support in the data for these models can be quantified using techniques described in the next section.

Model selection uncertainty, strength of evidence, and multimodel inference

Inference in the new statistics can be viewed as a contest among competing models, a contest arbitrated by data. Although AIC provides a basis for choosing the best model and ranking its competitors, there is some uncertainty that the chosen "best" KL model would emerge as superior given a different dataset. Akaike weights provide insight into that uncertainty. They are calculated as follows. Presume we have formulated a set of $r = 1 \cdots R$ candidate models. We will define the difference in AIC values between the best model and model r as $\Delta_r = \text{AIC}_r - \min(\text{AIC})$. Thus, the Δ_r for the best model, the one with the lowest AIC, is 0, and other models are ranked in descending order of the Δ_r relative to the best one. The likelihood of model r given the data is

$$\mathcal{L}(\theta_r \mid \mathbf{Y}) = e^{-2\Delta_r} \qquad (6)$$

and the Akaike weight, w_r, is calculated for each model as

$$w_r = \frac{e^{-2\Delta_r}}{\sum_{i=1}^{R} e^{-2\Delta_i}}. \tag{7}$$

The Akaike weights are very useful quantities, providing an estimate of model selection uncertainty, allowing assessment of relative strength of evidence in data for alternative models, and forming a basis for multimodel inference. They take values from 0 to 1. To quantify model selection uncertainty, the w_r can be interpreted as the ''probability'' that model r would emerge as the best model given many repetitions of the model selection exercise. Thus, the best models will have Akaike weights closer to 1 than the poor models have. This is not a probability in the Bayesian sense of a posterior distribution (more about this idea later), but rather is surmised from empirical, bootstrapped simulations where the proportion of times that a model is chosen as the best model is well approximated by the w_r.

Ratios of the w_r provide estimates of the relative support in data for two models. Suppose we have two models with Akaike weights, w_i and w_j, $w_i > w_j$, $\rho = w_i/w_j$. Calculating the ratio of the Akaike weights (ρ) with the weight for the better model in the numerator permits statements like ''Model i has ρ times more support in the data than model j.'' The precise interpretation of likelihood ratios and the technical meaning of the term support is covered nicely by Royall (1997). The summed likelihoods cancel in ratios of Akaike weights, leaving a ratio of likelihoods adjusted to account for differences in the number of parameters in models. Likelihood ratios illustrate a fundamental tenant of likelihood that all evidence is relative (Royall 1997). The central idea here is that it is not possible to evaluate a model without comparing its ability to represent the data relative to another model. It is important to understand that ''another model'' may be a model with different parameter values and the same structure, or it may be a model with a different structure, that is, with different sets of parameters and/or different mathematical operations.

The w_r also allow us to use more than one model in making predictions and inferences about parameters. Presume we want to compose a set of models allowing us to be $\alpha \times 100\%$ certain that the set contains the best KL model. That set includes all of the models for which $w_r \geq \alpha$ and we call it an α confidence set. Given that all of the models in this set are supported by the data, it is nonsensical to use the best one alone to make inferences. The Akaike weights allow us to average model predictions and estimate parameters (contained in all models in the confidence set). To do so, we renormalize the weights such that the denominator of Eq. 7 is summed over the models in the set. In a process known as model averaging, we then obtain a weighted average of the models' predictions, $\hat{y}$, by summing the predictions of the individual models multiplied by their weights:

$$\hat{y} = \sum_{r=1}^{R'} w_r f_r(x, \hat{\theta}) \tag{8}$$

where R' is the number of models in the confidence set and $f_r(x, \hat{\theta})$ gives the unweighted prediction of the rth model in the set. Model averaged estimates of parameter values can be obtained similarly for parameters that occur in all models in the confidence set.

Although these procedures identify the best models from a set of candidates, it remains possible that all of the models identified have low predictive power. Plots and regressions of observations on predictions (Haefner 1996:159), as well as confidence intervals on model predictions (Burnham and Anderson 2002:164), can inform the question of the ability of models to make predictions.

As before (Appendix A), we believe that these concepts become most transparent in a worked example. To that end, Appendix C illustrates the computations involved in obtaining maximum likelihood estimates of model parameters and selecting best approximating models using information theoretics.

The new statistics and designed experiments

Clearly, manipulative experiments with treatments and controls can be analyzed reliably with traditional statistics, and in this context, traditional statistics have served ecologists well. However, the approaches we describe here can also be used for manipulative experiments (e.g., Hilborn 1997:177–179, Hobbs 2003). As a simple example, consider a one-factor experiment where we measure the response of a variable to a treatment with two levels and a control. This can be cast as a problem in parameter estimation and model selection by estimating a model with a single parameter (μ_0, averaged across all treatments), two parameters (μ_c for the control and μ_t for the treatments), or three parameters (μ_c for the control, μ_1 for the first level of the treatment, and μ_2 for the second level). Following this logic, one could easily imagine how the analysis could be structured to handle all of the usual contrasts between means by evaluating the strength of evidence for alternative models. However, the interpretation of this analysis differs from the interpretation of ANOVA. Rather than trying to establish significant differences among means, we are able to compare the relative support in the data for several competing hypotheses about the effects of treatment. The P values calculated in analysis of variance do not estimate evidence (Royall 1997).

Using more than one source of data in parameter estimation

Many statistical tools (regression, analysis of variance) assume we observe a single dependent variable,

but biological models often make predictions about several observed variables. For instance, age and sex structured population models predict the total population size as might be observed in a survey, the age composition as might be observed in a sample from the population, and the sex ratio as might be observed in either a survey or a sample (Fournier and Archibald 1982, Cooper et al. 2003). A model of brown bear predation on salmon predicted both the number of fish eaten by bears on each day, and the time the fish had spent in the spawning stream (Gende et al. 2004).

Likelihood theory can easily incorporate multiple types of observations like these (Aguiar and Sala 1999, Pawitan 2001; Appendix D). Assuming the observations are independent, then the likelihood of the model parameters conditional on all types of observations is the product of the likelihoods for each type of observation or the sum of the sum of the log likelihoods. For example, if we have three types of observations, say x, y, and z, then the log likelihood of the three types of data given model parameters θ is simply

$$\ln[\mathcal{L}(\theta \,|\, x, y, z)] = \ln[\mathcal{L}(\theta \,|\, x)] + \ln[\mathcal{L}(\theta \,|\, y)] + \ln[\mathcal{L}(\theta \,|\, z)] \tag{9}$$

where $[\mathcal{L}(\theta \,|\, x, y, z)]$ is the log likelihood of model parameters θ given observations x, y, and z and the right-hand side of the equation is the individual log likelihoods. This simple approach provides enormous power to estimation of parameters or comparison of models that are biologically realistic (Schnute 1994, Punt and Hilborn 1997, Lubow et al. 2002, White and Lubow 2002, Cooper et al. 2003).

In the new statistics, we write down the competing biological models and then consider what predictions these models make that can be compared to observations (Appendix D). For each observation, we consider the appropriate likelihood, and each of these likelihoods becomes a component of the total likelihood. In this approach to modeling, estimation and model comparison is conveniently considered in the framework of state space models (Aoki 1987, Schnute 1994, also see Newman et al. 2006):

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \theta) \qquad \mathbf{y}_t = g(\mathbf{x}_t, \phi) \tag{10}$$

where $\mathbf{x}_t$ is the vector of state variables that completely describe the system at time t, $\mathbf{y}_t$ is the vector of observations of the system, θ are the parameters of the dynamic model relating the state variables from one time to the next, and ϕ are the parameters of the observation model that relate the observations $\mathbf{y}$ to the state variables $\mathbf{x}$.

The biological model is used to predict the x values, and the observation model is used to predict the observations (y). The likelihoods for each observation are then used to compute the total likelihood. For example, in an age-structured model, the state variables are nor-

mally the number of individuals alive by age and sex (Appendix D). The observations might be the total population size and the number of individuals of each age in a sample or samples. The parameters of the dynamic model might be birth and death rates while the parameters of the observation model might be the fraction of individuals seen in a survey method and any age or sex selectivity in sampling methods. Newman et al. (2006) give a detailed treatment of these kinds of models. The point here is that a likelihood framework provides a simple and intuitive way to estimate their parameters.

Using existing knowledge and data

In traditional statistics, each experiment is assumed to be independent of previous work, and all evaluation of hypotheses is related only to the data collected in the individual experiment. These methods never ask how to combine independent experiments to provide a summary of total knowledge. For example, consider again the age-structured population models which customarily have the natural mortality rate as a parameter distinct from a parameter that represents the harvest induced mortality. We often find that it is difficult to estimate the natural mortality rate. As a result of this difficulty, it is common practice (Hilborn and Walters 1992, Punt and Hilborn 1997) to treat natural mortality as known without error and simply input plausible estimates of mortality as constants into the models. The estimated values are frequently taken from another experiment or study, perhaps a tagging study, or an age sample obtained in a place or time that had no harvest. The problem with this approach is that assuming that natural mortality is known without error is clearly flawed, and results in model predictions with less uncertainty than they should have. Models need to use what is known about the uncertainty in previous parameter estimates to assure that this uncertainty is included in model predictions.

Using the new statistics, we would consider the natural mortality rate an unknown parameter of the age structured model and include a likelihood component for the natural mortality rate derived from the previous study. For example, if an earlier study had estimated the natural mortality rate (m) to be normally distributed with a mean μ and a standard deviation σ, then the total log likelihood for the age structured model would be

$$\ln(\mathcal{L}_{\text{total}}) = \ln[\mathcal{L}(\,|\,\theta\,|\,\mathbf{Y})]$$
$$+ \ln\left[\left(\frac{1}{\sigma\sqrt{2\pi}}\right)\exp\left(-\frac{(m - \mu)^2}{2\sigma^2}\right)\right] \tag{11}$$

where $\mathbf{Y}$ is a time series of data available for estimating the parameters of the age structured model; θ is the vector of parameters of the age structured model including m; $\mathcal{L}_{\text{total}}$ is the total likelihood for the model;

$\ln[\mathcal{L}(\theta\,|\,\mathbf{Y})]$ is the log likelihood of the model's parameters (including m) conditional on the current data, and

$$\ln\left\{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)\exp\left[-\frac{(m-\mu)^2}{2\sigma^2}\right]\right\}$$

is the log likelihood of the estimate of m given data from the previous study. The interplay of previous knowledge and current knowledge can be seen in Eq. 11. If the estimated likelihood for the parameter departs from the previous estimate (μ), then the total likelihood gets smaller. The strength of the effect of the previous estimate on the current estimate depends on the variance of the previous one (σ^2). Small variance means that the previous estimate will have a strong effect on the total likelihood; large variance weakens this effect.

Likelihood provides a convenient way to summarize our knowledge from a series of experiments (e.g., Cooper et al. 2003); so as long as these experiments are independent, then the likelihoods will multiply together, or as above, the log likelihoods will be added. When likelihoods are not independent, then the variance/covariance structure of the likelihoods needs to be included in the analysis. Treatment of variance–covariance structure is beyond the scope of this paper, but is covered nicely by Mood et al. (1963).

Bayesian analysis

The basic logic of accumulation of knowledge is as follows: what you know after a new set of data are analyzed is what you knew before the new data were available modified by what you learned from the new data. Consider a simple model with a single parameter θ. For our example here, we will assume that the θ_i are discrete values. We can write this as a likelihood statement as a variant on Eq. 11

$$\mathcal{L}(\theta_i\,|\,\text{prior},\,\mathbf{Y}) = \mathcal{L}\theta_i\,|\,\text{prior})\,\mathcal{L}(\theta_i\,|\,\mathbf{Y}) \qquad (12)$$

where $\mathbf{Y}$ is the new data and prior is all the information available before the new data were collected. In likelihood theory, the likelihoods are relative and are not to be interpreted at probability per se, but rather are proportionate to probabilities.

It is a natural transition to move from using pre-existing knowledge with likelihood to Bayesian analysis. Rather than treating pre-existing knowledge as optional (as in the likelihood framework), Bayesian analysis requires that pre-existing knowledge of all model parameters be included in any current estimates of those parameters (Gelman et al. 1995). As such, Bayesian analysis is the natural framework to consider in the process of accumulation of knowledge in scientific endeavors because it requires you to summarize what you know before the new data is analyzed. There is not enough space in this paper to offer even an elementary tutorial on Bayesian methods. Other sources provide this introduction (Hilborn 1997, Punt and Hilborn 1997, 2002, Link et al. 2002, Calder et al. 2003).

Instead, we will touch on a few key points with the aim of motivating self-teaching.

Bayes Law for discrete hypotheses is

$$P(\theta_i\,|\,\mathbf{Y}) = \frac{P(\theta_i\,|\,\text{prior})\mathcal{L}(\theta_i\,|\,\mathbf{Y})}{\sum_j P(\theta_j\,|\,\text{prior})\mathcal{L}(\theta_j\,|\,\mathbf{Y})}. \qquad (13)$$

The two obvious differences between writing Bayes Law this way and the earlier likelihood statement (Eq. 12) is that now the pre-existing information about θ_i is stated as probabilities rather than likelihoods, and the denominator, which simply scales the numerators, assures that the probability across all hypotheses adds to 1.0.

Twenty years ago the use of Bayesian statistics in science was widely debated (Efron 1986, Howson and Urbach 1991, Jeffreys and Berger 1992, Dennis 1996), but now, the preponderance of statisticians use Bayesian statistics to some extent, and Bayesian methods are widely accepted. Much of the earlier controversy revolved around the use of subjective ''opinion'' in formulating the prior distributions. In modern applications, it is far more common to use pre-existing data sets to formulate priors, or to use ''diffuse'' or uninformative priors that do not give strong prior probability to any hypotheses.

Most people continue to believe that the distinguishing character of Bayesian statistics is the use of prior information, yet as we saw in the previous section, prior information can be used in a strictly likelihood framework. There is nothing uniquely Bayesian about accumulating knowledge from one experiment to another. However, Bayesian statistics are unique in two very important, but unappreciated, ways. First, the application of Bayes law produces probabilities of hypotheses; no other statistical approach does this. Bayes Law is simply a restatement of the laws of conditional probability and is not challenged by anyone as a mathematical truism. If you want to produce probabilities of hypotheses, there is no alternative to Bayes Law and Bayesian analysis (Gelman et al. 1995). Frequentist statistics consider frequency distributions of estimators and probabilities of observing data, but deny that probability of hypotheses is an important subject for discussion. Many students continue to believe that the 95% confidence interval from a traditional statistical analysis says that there is a 95% probability the real parameter value is within that range, instead of the proper frequentist statement that if the experiment was repeated many times, the estimate of the parameter would fall within the 95% confidence interval 95% of the time (Gelman et al. 1995). Similarly, likelihoods are proportionate to probabilities, but are not equal to probabilities.

The distinction between probabilities and relative strength of evidence is not particularly important when estimating a single parameter, and the likelihood profile

conveys the information needed to understand what we have learned about a parameter from a particular data set. However, if the purpose of the data analysis is to provide input to decision makers to help them weigh the odds of alternative outcomes, then the formal tool is statistical decision analysis (Berger 1985, Clemen 1996) where probabilities of hypotheses are needed in order to make probability statements about outcomes. If you wish to ask any questions about expected value of alternative decisions, or the frequency distribution of outcomes, you must assign probabilities to alternative hypotheses, which only Bayesian statistics will calculate.

For example, assume that you have data ($\mathbf{Y}$) on the mortality rate as a function of exposure to a toxic substance (pcbs, tobacco smoke, etc.), and wish to determine how many individuals will die as a result of this exposure. The expected value of the number of deaths per thousand, m, is

$$E(m\,|\,\mathbf{Y}) = \sum_i P(m_i\,|\,\mathbf{Y})m_i. \qquad (14)$$

Although likelihoods are proportionate to probabilities, they are not equal to probabilities—you need Bayes law to make statements about the probability of a hypothesis. It can easily be seen from Eq. 13 that if the prior probability of all hypotheses is equal, then the probabilities are just the likelihoods normalized to sum to one. This assumes that we had no prior information about the competing hypotheses before we collected data ($\mathbf{Y}$). More importantly, this simple result is only true in one dimension. Once we deal with more than one parameter at a time, major differences emerge between likelihood and Bayesian analysis. Perhaps because Bayesian analysis is the mathematical language of decision making, its earliest use was the scientific analysis of gambling and decision making in business schools. Many scientists adopt a naïve view that ''I am a scientist, not a decision maker,'' but they fail to recognize that scientists make decisions every day about what experiments to perform (Hilborn 1997) based on the probabilities they intuitively assign to competing hypotheses.

The second fundamental difference between Bayesian analysis and traditional statistics and likelihood methods is that estimating parameters in the Bayesian framework requires integrating across the other parameters, while in likelihood, you maximize across those parameters. Assume you are calculating the likelihood profile for one parameter θ_1 in a model that contains a second parameter θ_2. You find the likelihood profile for θ_1 by iterating across values of θ_1, and for each value finding the value of θ_2 that maximizes the likelihood. In Bayesian analysis you integrate the likelihood of all possible values of θ_2 for each value of θ_1. This distinction often makes little difference for many simple problems, but we often find with more complex and realistic biological models that the point in parameter space with the highest likelihood is outside the 95% probability bounds in a Bayesian analysis (Schnute and Kronlund 2002).

Although Bayesian analysis has become very popular and is now extensively used, we find two ongoing difficulties. The first is simple computer implementation. The integration required can often take enormous computer time, often 24 to 48 hours on modern desktop computers, whereas maximum likelihood estimates and likelihood profiles can be obtained in a matter of minutes. However, this disadvantage must be weighted against the fact that likelihood methods that mix process and observation errors cannot be solved by traditional maximum likelihood methods (De Valpine and Hastings 2002), and presently, only Bayesian methods are computationally possible for such mixed models.

More challenging is identifying appropriate prior distributions for parameters when no specific studies are available to provide guidance. The first approach is to try to define ''uninformative'' priors. In a single dimension, a uniform prior over the range of plausible parameter values is uninformative, but as soon as you move beyond one dimension, what is uninformative for one parameter may provide information about another (Walters and Ludwig 1994). In some cases, log-uniform priors for one parameter may be non-informative for another. The alternative to uninformative priors is to summarize pre-existing knowledge in a process known as meta-analysis.

Hierarchical models, which are analogous to traditional random effects models, are growing in popularity (Link et al. 2002, Sauer and Link 2002, Calder et al. 2003, Clark 2003), and provide a framework for describing the variability among individuals or populations. Hierarchical models are an essential element of most modern meta-analysis discussed in the next section.

Meta-analysis

Perhaps the greatest challenge facing ecologists and statisticians is how to make collective sense of individual scientific findings, to sort through everything that has been done, and come up with a summary of ''what does this all mean.'' One of the emerging tools in statistics is meta-analysis, an approach to combine multiple experimental results into a statistical statement of cumulative knowledge (Hedges and Olkin 1985, Hunt 1997, Gurevitch et al. 2001).

Meta-analysis is the combined analysis and quantitative synthesis of a collection of multiple studies generally using a set of summary statistics for each study (Hedges and Olkin 1985). In recent applications in ecology, meta-analysis has been extended to include the combined analysis of complete data sets from the individual studies, not only their summary statistics. The original motivation for meta-analysis was the syn-

thesis of results from controlled experiments (Mann 1990). This remains its primary use in medicine, where it has seen the most development, and also in ecology (Gurevitch et al. 2001). Several experiments designed to study the same phenomenon may produce discrepant, or even contradictory, estimates of the effect of interests by chance alone or due to logistic differences. A very large study that will guarantee an unbiased and precise estimate of the effect may not be possible. The summary of results from multiple studies by meta-analysis is crucial in order to show the big picture and to allow for accurate estimation of the underlying effect. In ecology, analogs of small experiments are the short and noisy time series that may not, in isolation, present enough information to estimate parameters of interest. The simplest meta-analysis attempts to estimate the value of a single parameter from a series of studies.

A second form of meta-analysis (known as hierarchic meta-analysis) deals with parameters whose values differ among populations. Meta-analysis attempts to estimate the distribution of the parameter. The simplest form of hierarchic meta-analysis is drawing a histogram of the parameter as estimated from studies of multiple populations. A distribution can be fit to the histogram and can be used as a summary of the understanding of the parameter values among populations. Formal hierarchic meta-analysis moves beyond histograms by considering the statistical uncertainty in the estimation of the parameter of interest in each population and correcting for such measurement error.

Hierarchic meta-analysis has seen its most intensive use in ecological fields in the analysis of fisheries data, presumably because of the demands of decision makers for statements of cumulative knowledge. For example, it is well established that power plants that use pass-through cooling inflict mortality on the fish eggs and larvae that are entrained in the water as it enters the power plants. In the 1970s, enormous scientific energy was devoted to determining the impact of a series of Hudson River power plants on striped bass (and other species) (Barnthouse and Van Winkle 1988). A key issue was the intensity of density dependence in the egg and larval stages; the more density dependence, the less impact egg and larval mortality had on the adult population size. The result of all the studies in the 1970s was that the intensity of density dependence in Hudson River striped bass was very difficult to estimate given the data available at that time. In the last few years, the same issues have come up with respect to other power plants, where again, the data available for the species of interest in the area of interest provide little information about density dependence. However, meta-analysis of density dependence in fishes has provided estimates of the range of density dependence that has been seen in other fish stocks (Myers et al. 2002). Such meta-analysis was used to evaluate the risk the power plants posed to the affected fishes.

Myers et al. (2002) used hierarchical meta-analysis in which one does not assume that there is a single value of density for all populations, but rather that there is a distribution of density dependence among all populations of specific species of life history groups. Individual populations of fish within these groupings have a density dependence parameter that is drawn from this distribution. Fig. 2 shows the distribution of density dependence parameters (known as steepness) for four different groups of fish. Myers et al. (2002) showed that the existing information about these life history types provided strong information about density dependence, even in the total absence of data for the species in the area of concern.

In the absence of formal meta-analysis, scientists must either say they know nothing other than what the data they have in hand tell them, or they must find a way to summarize their professional opinion. Formal methods for meta-analysis are easily learned by anyone familiar with likelihood (Gelman et al. 1995), and have recently been extensively used in fisheries (Hilborn and Liermann 1998, Myers 2001, Myers et al. 2001, 2002, Dorn 2002) and ecology (Bender et al. 1998, Connor et al. 2000, Schmitz et al. 2000, Chalfoun et al. 2002, Fahrig 2002, Hillebrand and Cardinale 2004).

Meta-analysis is not without its critics (Wachter 1988). The biggest problem in meta-analysis is ensuring that one has a random sample from the population one wishes to describe. For example, in the meta-analysis of Myers et al. (2002) described above, we must assume that the fish stocks for which we have data are random samples from the population of stocks we wish to make inferences about. This is likely not true. Larger, more productive stocks are more likely to be studied and have longer time series of data available for analysis than smaller less productive stocks. Even within the data sets available, the choice of which to use and not to use is problematic. These problems are not unique for formal meta-analysis, as anyone hoping to summarize what has been learned from the studies available would be forced into the same questions.

DISCUSSION

An alternative to ''strong inference''

The highly influential paper of Platt (1964) described a method for achieving rapid progress in science; a method known as ''strong inference.'' In this method, knowledge accumulates by critical tests of hypotheses conducted sequentially—at each step in the sequence, one hypothesis is conclusively discarded in favor of another. Metaphorically, the process resembles climbing a tree where each bifurcation in branches represents two competing hypotheses. Platt's approach was motivated by the history of science in physics, chemistry, and molecular biology where progress in the middle of the 20th century was particularly swift. However, ''strong inference'' advocated by Platt has proven less
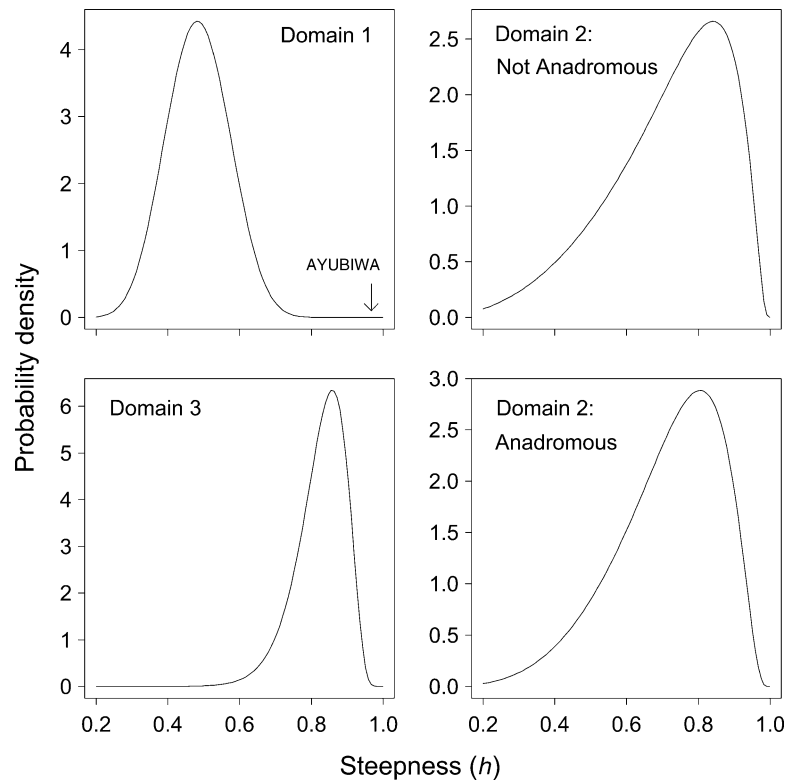
FIG. 2. Myers et al. (2001) used meta-analysis to construct likelihood profiles of steepness parameter for Hudson River fish species grouped by life history characteristics. The steepness parameter indicates the strength of density dependence, where a value of 0.2 indicates strong density dependence with recruitment proportional to spawning stock. A value of 1.0 indicates that recruitment is independent of density. The figure is redrawn from Myers et al. (2001).

successful in ecology, largely because ecological experiments often produce partial support for competing views rather than unambiguous rejection of one over another. This is because ecology deals with problems at the system level where interactions are complex, composite effects are common, true controls are rare, replicates are often difficult to obtain, and experiments often take many years to complete. Whereas in physics, chemistry, and molecular biology, researchers can find model systems that have little natural variability, in ecology, natural variability is an important part of the questions we ask. Moreover, physical models have a direct, unambiguous analog in nature, unlike the micro- and mesocosms used in ecology, which may poorly reflect the systems they are designed to represent. Thus, there is reason to doubt that Platt's approach applies without modification to ecological research.

In the physical sciences, statistics are often superfluous, an idea made plain by the eminent physicist Earnest Rutherford who said ''If your experiment needs statistics, you ought to have done a better experiment'' (Bailey 1967). In Platt's and Rutherford's world, you learned by sequential discarding of hypotheses, whereas in ecological research, we more frequently learn by accumulation of evidence—a single experiment rarely

provides overwhelming evidence in favor of one competing hypotheses. Examples of this process of accumulation of evidence in ecology include understanding the relationship between diversity and productivity (Schlapfer and Schmid 1999, Bullock et al. 2001, Sala 2001, Downing and Leibold 2002, Worm et al. 2002, Worm and Duffy 2003), portraying the operation of trophic interactions in structuring communities (Leibold et al. 1997, McCann et al. 1998, Huxel 1999, Polis et al. 2000, Schmitz et al. 2000, Bell et al. 2003, Nystrom et al. 2003), and resolving effects of herbivory on ecosystem net primary production (McNaughton 1979, Belsky 1986, Detling 1988, Georgiadis et al. 1989, Deangelis and Huston 1993, Noy-Meir 1993, Frank et al. 1998, 2002, Anderson and Frank 2003). In all of these cases, our understanding of nature emerges from the accumulation of evidence from multiple studies rather than crisp rejections of hypotheses from a few definitive experiments.

Closely related to the need to accumulate evidence over time is the idea that, at any one time, ecological phenomena can usually be explained by multiple hypotheses. Sometimes it is possible to soundly reject all but one hypothesis (e.g., Hobbs et al. 2003). In other cases, some models can be rejected, but more than one

will have support in the data (e.g., Vucetich et al. 2002). This is the way it should be. Many processes in ecology represent the action of a composite of forces (e.g., Mc-Naughton 1983, Callaway 1995, Steinauer and Collins 2001, Allcock and Hik 2003). When this is the case, several hypotheses, each representing a portion of the composite, will be more or less true rather than absolutely true or false. As a result, it should not be surprising that experiments often do not conclusively eliminate a hypothesis, but instead identify more than one with support in data. The new statistics are able to deal with more than one model. Likelihoods, likelihood profiles, model averaging, and Bayesian probabilities allow us to evaluate multiple hypotheses simultaneously and do not force us to conclude that only one of them is supported by the data.

In contrast, the emphasis on testing a single hypothesis against a null hypothesis is poorly suited to ecological questions where we are unable to conclusively eliminate all hypotheses and are interested in the relative support for competing ideas. The use of P values as the standard statistic for reporting the results of statistical analysis has been under strong criticism for a number of years (Berger and Sellke 1987, Hilborn 1997, Goodman 1999, Johnson 1999, Anderson et al. 2000), and a number of journals have policies either discouraging or prohibiting papers presenting P values (Fidler et al. 2004). Therefore, it is remarkable that almost all statistics courses taught to ecologists feature P values as the primary result of statistical analysis and that ecological journals are peppered with them.

Empiricism and theory in ecology

Historically, ecologists chose between two career paths, one of which emphasized empiricism, the other, theory (e.g., Fretwell 1972). Empirically inclined ecologists tended to gain insight using linear statistical models and hypothesis tests tightly tied to data from manipulative experiments. However, their statistical analysis was only loosely related to biology; that is, the models that supported analysis were not composed to represent ecological processes. Theorists, on the other hand, have emphasized development of nonlinear models tightly tied to biology; these models were explicitly motivated by ecology. These models portrayed states and processes in nature (May 1981, Gotelli 1998, Case 2000). However, theoretical models were often loosely tied to data (Hall 1988) because there was no widely accepted statistical framework for testing and comparing them. The new statistics encourage a tight linkage between theory and data in ecology by offering a unified approach to the evaluation of virtually all types of models—statistical and biological, linear and nonlinear.

The emphasis on parameter estimation and model evaluation inherent in the new statistics encourages the accumulation of understanding in ecology rather than the accumulation of disconnected facts. Understanding can accumulate from models in two ways. First, parameters that are appropriately estimated in one study can be used to enhance subsequent estimates, allowing knowledge gained in the past to contribute to future knowledge, as described above. Second, deep knowledge of ecology, or any topic for that matter, has a hierarchical structure—an understanding of lower levels of ecological organization provides a mechanistic understanding of phenomena operating at higher levels. The emergence of sets of tested models provides a way to quantitatively link understanding across levels in this hierarchy. An excellent example of this linkage can be seen in the paper of Canham and Uriarte (2006), who show how evaluating models of processes operating at the level of individual trees allows insight at the level of forested landscapes.

Applying ecological knowledge to policy and management

The difference in the old and new statistics has fundamental implications for the application of ecological knowledge to management and policy. These implications are well illustrated in the debate on the precautionary principal and the burden of proof. In many public policy settings, the burden of proof has traditionally been on those who wanted to argue for regulation—someone would be allowed to pollute, for instance, unless there was ''proof'' of damage. The precautionary principle says that the burden of proof should be reversed—pollution would not be allowed until it was proven to have no impact (Cameron and Abouchar 1991, Peterman and M'Gonigle 1992, M'Gonigle et al. 1994). Traditional statistics are well suited to these alternatives, with traditional approaches to regulation allowing pollution if you could not conclusively demonstrate an effect and the precautionary principle forbidding pollution unless you could conclusively demonstrate that effects do not exist. However, the question can be reframed in a more useful way: how large must an effect be to justify regulation? Would a totally trivial effect mean the pollution would not be allowed (Holm and Harris 1999)?

An alternative to hypothesis testing in this context is to concentrate on estimating the size of the effect and the risk it entails. This alternative supports a variant on the precautionary principle, known as the precautionary approach, suggesting that decisions should be made based on weighing the relative risks of actions (Foster et al. 2000, Hilborn et al. 2001). Formulated by the Food and Agriculture Organization of the United Nations (FAO) for fisheries, part of the precautionary approach includes ''the standard of proof to be used in decisions regarding authorization of fishing activities should be commensurate with the potential risk to the resource while also taking into account the expected benefits of the activities'' (Food and Agriculture Or-

ganization of the United Nations 1996). This requires that we assess the probability of competing hypotheses and estimate the size of effects. The new statistics are well suited to this decision making approach.

A second example of opportunities for applying ecological knowledge to decisions is created by the interplay of models and data in adaptive management. Adaptive management of natural resources (Walters and Hilborn 1978, Walters 1986) is an approach that uses management actions to help discriminate among competing models of resource dynamics and the effect of harvest on those dynamics. Adaptive management has at its core the recognition that there are alternative models about the underlying system dynamics. Statistics are used to estimate the support that existing data provide for these models and to plan management actions needed to evaluate them with improved, future data. These different levels of support, in turn, are used to ''weight'' the expected value of different management actions (Walters 1986). As originally formulated, adaptive management *requires* the use of Bayesian statistics. Although adaptive management has been described (Lee and Lawrence 1986, Halbert and Lee 1991) as being no different from the strong inference advocated by Platt (1964), except perhaps in the scale of inquiry, we believe (also see McLain and Lee 1996) that this description is fundamentally inaccurate and represents a dramatic departure from the original ideas advocated by Walters (1986) and Walters and Hilborn (1978). The original vision of adaptive management embraced multiple competing models and acknowledged that support in data for more than one model is likely (Walters 1986, Walters and Holling 1990, Walters and Green 1997). Walters (1986:185) specifically avoided any treatment of traditional statistical methods (contrast this with Lee 1993:74–75). Contemporary adaptive management can benefit from using statistical tools that are well-aligned with its original formulation (Walters 1986). We believe that one of the primary reasons that adaptive management has not been implemented as it was designed is simply that researchers and managers in natural resources have not been trained in the likelihood and Bayesian methods that are critical to understanding how adaptive management is supposed to work.

Training for the new statistics

The new statistics encourages us to evaluate the evidence in data for biologically meaningful hypotheses expressed as models. However, applying these methods requires training. In our experience, students often emerge from the typical sequence of courses in statistical methods without an integrated, intuitive understanding of the relationship between models and data. Instead, they know a series of techniques and some rules for when they should be applied. The more methods courses they take, the more techniques and rules

they accumulate, but they fail to see a connection among these techniques at a deep level.

We believe that a new approach to statistical training would provide that connection by focusing on the general relationship among scientific models (hypotheses), probability models, and data rather than emphasizing specific testing procedures. Although a first course in mathematical statistics goes a long way toward providing this understanding, such courses are usually structured to meet the needs and abilities of students in statistics rather than other areas of science. In our opinion, large dividends can emerge for ecology by formal training (and self-teaching) in contemporary methods of parameter estimation and model evaluation using likelihood, information theoretics, and Bayesian approaches. There are increasing numbers of graduate programs, courses, and workshops that treat these methods in ways that are accessible to ecologists, for example, the Program for Interdisciplinary Ecology Mathematics and Statistics[4] and Systems Ecology[5] at Colorado State University, Ecological Models and Data at the University of Florida,[6] and Modeling for Conservation of Populations at the University of Washington.[7] We see a rich opportunity for ecologists and statisticians to work together designing curricula that will prepare students to evaluate the strength of evidence in data for competing models of ecological processes.

Literature Cited

Aguiar, M. R., and O. E. Sala. 1999. Patch structure, dynamics and implications for the functioning of arid ecosystems. Trends in Ecology and Evolution **14**:273–277.

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petrov and F. F. Csaki, editors. Second International Symposium on Information Theory. Akademiai Kiado, Budapest, Hungary.

Allcock, K. G., and D. S. Hik. 2003. What determines disturbance–productivity–diversity relationships? The effect of scale, species and environment on richness patterns in an Australian woodland. Oikos **102**:173–185.

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

Anderson, M. T., and D. A. Frank. 2003. Defoliation effects on reproductive biomass: importance of scale and timing. Journal of Range Management **56**:501–516.

---

[4] ⟨http://www.primes.colostate.edu/⟩
[5] ⟨http://www.cnr.colostate.edu/class_info/nr575/⟩
[6] ⟨http://www.zoo.ufl.edu/bolker/emd/⟩
[7] ⟨http://www.fish.washington.edu/classes/fish458/index.htm⟩

Aoki, M. 1987. State space modeling of time series. Springer-Verlag, New York, New York, USA.

Azzalini, A. 1996. Statistical inference based on the likelihood. Chapman and Hall, London, UK.

Bailey, N. T. J. 1967. The mathematical approach to biology and medicine. Wiley, New York, New York, USA.

Barnthouse, L. W., and W. Van Winkle. 1988. Analysis of impingement impacts on Hudson River fish populations. Pages 182–190 *in* Science, law and Hudson River power plants: a case study in environmental impact assessment. American Fisheries Society, Bethesda, Maryland, USA.

Bell, T., W. E. Neill, and D. Schluter. 2003. The effect of temporal scale on the outcome of trophic cascade experiments. Oecologia **134**:578–586.

Belsky, A. J. 1986. Does herbivory benefit plants—a review of the evidence. American Naturalist **127**:870–892.

Bender, D. J., T. A. Contreras, and L. Fahrig. 1998. Habitat loss and population decline: a meta-analysis of the patch size effect. Ecology **79**:517–533.

Berger, J. O. 1985. Statistical decision theory and Bayesian analysis. Springer-Verlag, New York, New York, USA.

Berger, J. O., and T. Sellke. 1987. Testing a point null hypothesis: the irreconcilability of *P* values and evidence. Journal of the American Statistician Association **82**:112–122.

Bjornstad, O. N., B. F. Finkenstadt, and B. T. Grenfell. 2002. Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. Ecological Monographs **72**:169–184.

Bullock, J. M., R. F. Pywell, M. J. W. Burke, and K. J. Walker. 2001. Restoration of biodiversity enhances agricultural production. Ecology Letters **4**:185–189.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multi-model inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.

Calder, C., M. Lavine, P. Muller, and J. S. Clark. 2003. Incorporating multiple sources of stochasticity into dynamic population models. Ecology **84**:1395–1402.

Callaway, R. M. 1995. Positive interactions among plants. Botanical Review **61**:306–349.

Cameron, J., and J. Abouchar. 1991. The precautionary principle: a fundamental principle of law and policy for the protection of the global environment. Boston College International and Comparative Law Review **14**:1–27.

Canham, C. D., and M. Uriarte. 2006. Analysis of neighborhood dynamics of forest ecosystems using likelihood methods and modeling. Ecological Applications **16**:62–73.

Case, T. J. 2000. An illustrated guide to theoretical ecology. Oxford University Press, Oxford UK.

Caswell, H. 1988. Matrix population models. Sinauer, Sunderland, Massachusetts, USA.

Chalfoun, A. D., F. R. Thompson, and M. J. Ratnaswamy. 2002. Nest predators and fragmentation: a review and meta-analysis. Conservation Biology **16**:306–318.

Clark, J. S. 2003. Uncertainty and variability in demography and population growth: a hierarchical approach. Ecology **84**:1370–1381.

Clayton, D., and M. Hills. 1998. Statistical models in epidemiology. Oxford Science Publications, Oxford, UK.

Clemen, R. T. 1996. Making hard decisions: an introduction to decision analysis. Duxbury Press, Belmont, California, USA.

Connor, E. F., A. C. Courtney, and J. M. Yoder. 2000. Individuals–area relationships: the relationship between animal population density and area. Ecology **81**:734–748.

Cooper, A. B., R. Hilborn, and J. W. Unsworth. 2003. An approach for population assessment in the absence of abundance indices. Ecological Applications **13**:814–828.

Deangelis, D. L., and M. A. Huston. 1993. Further considerations on the debate over herbivore optimization theory. Ecological Applications **3**:30–31.

Dennis, B. 1996. Discussion: should ecologists become Bayesians? Ecological Applications **6**:1095–1103.

Dennis, B., and M. R. M. Otten. 2000. Joint effects of density dependence and rainfall on abundance of San Joaquin kit fox. Journal of Wildlife Management **64**:388–400.

Detling, J. K. 1988. Grasslands and savannas—regulation of energy flow and nutrient cycling by herbivores. Pages 131–148 *in* L. R. Pomeroy and J. J. Alberts, editors. Concepts in ecosystem ecology: a comparative view. Sprinter-Verlag, Berlin, Germany.

De Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs **72**:57–76.

Dorn, M. W. 2002. Advice on West Coast rockfish harvest rates from Bayesian meta-analysis of stock-recruit relationships. North American Journal of Fisheries Management **22**:280–300.

Downing, A. L., and M. A. Leibold. 2002. Ecosystem consequences of species richness and composition in pond food webs. Nature **416**:837–841.

Edelstein-Keshet, L. 1988. Mathematical models in biology. McGraw Hill, New York, New York, USA.

Edwards, E. W. F. 1992. Likelihood. Johns Hopkins University Press, Baltimore, Maryland, USA.

Efron, B. 1986. Why isn't everyone a Bayesian? American Statistician **40**:1–11.

Fahrig, L. 2002. Effect of habitat fragmentation on the extinction threshold: a synthesis. Ecological Applications **12**:346–353.

Fidler, F., N. Thomason, G. Cumming, S. Finch, and J. Leeman. 2004. Editors can lead researchers to confidence intervals, but can't make them think—statistical reform lessons from medicine. Psychological Science **15**:119–126.

Food and Agriculture Organization of the United Nations (FAO). 1996. Precautionary approach to capture fisheries and species introductions. Food and Agriculture Organization of the United Nations, Rome, Italy.

Forchhammer, M. C., N. C. Stenseth, E. Post, and R. Langvatn. 1993. Population dynamics of Norwegian red deer: density-dependence and climatic variation. Proceedings of the Royal Society of London Series B: Biological Sciences **265**:341–350.

Foster, K. R., P. Vecchia, and M. H. Repacholi. 2000. Risk management—science and the precautionary principle. Science **288**:979.

Fournier, D. A., and C. Archibald. 1982. A general theory for analyzing catch at age data. Canadian Journal of Fisheries and Aquatic Sciences **39**:1195–1207.

Frank, D. A., M. M. Kuns, and D. R. Guido. 2002. Consumer control of grassland plant production. Ecology **83**:602–606.

Frank, D. A., S. J. McNaughton, and B. F. Tracy. 1998. The ecology of the Earth's grazing ecosystems. BioScience **48**:513–521.

Fretwell, S. 1972. Populations in a seasonal environment. Princeton University Press, Princeton, New Jersey, USA.

Garrott, R. A., L. L. Eberhardt, P. J. White, and J. Rotella. 2003. Climate-induced variation in vital rates of an unharvested large–herbivore population. Canadian Journal of Zoology-Revue Canadienne De Zoologie **81**:33–45.

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

Gende, S. M., T. P. Quinn, R. Hilborn, A. P. Hendry, and B. Dickerson. 2004. Brown bears selectively kill salmon with higher energy content but only in habitats that facilitate choice. Oikos **104**:518–528.

Georgiadis, N. J., R. W. Ruess, S. J. McNaughton, and D. Western. 1989. Ecological conditions that determine when

grazing stimulates grass production. Oecologia **81**:316–322.

Gershenfeild, N. 1999. The nature of mathematical modeling. Cambridge University Press, London, UK.

Goodman, S. N. 1999. Toward evidence-based medical statistics. 1: the *P* value fallacy. Annals of Internal Medicine **130**:995–1004.

Gotelli, N. J. 1998. A primer of ecology. Sinauer, Sunderland, Massachusetts, USA.

Grenfell, B. T., K. Wilson, B. F. Finkenstadt, T. N. Coulson, S. Murray, S. D. Albon, J. M. Pemberton, T. H. Clutton-Brock, and M. J. Crawley. 1998. Noise and determinism in synchronized sheep dynamics. Nature **394**:674–677.

Gurevitch, J., P. S. Curtis, and M. H. Jones. 2001. Meta-analysis in ecology. Advances in Ecological Research **32**:199–247.

Gurney, W. S. C., and R. M. Nisbet. 1998. Ecological dynamics. Oxford University Press, New York, New York, USA.

Haefner, J. W. 1996. Modeling biological systems: principles and applications. Chapman and Hall, London, UK.

Halbert, C. L., and K. N. Lee. 1991. Implementing adaptive management—a conversation with Lee, Kai N. Northwest Environmental Journal **7**:136–150.

Hall, C. A. S. 1988. An assessment of several of the historically most influential theoretical models used in ecology and the data provided in their support. Ecological Modeling **43**:5–31.

Hastings, A. 1997. Population biology: concepts and models. Springer-Verlag, New York, New York, USA.

Hedges, L. V., and I. Olkin. 1985. Statistical methods for meta-analysis. Academic Press, Orlando, Florida, USA.

Hilborn, R. 1997. Statistical hypothesis testing and decision theory in fisheries science. Fisheries **20**:19–20.

Hilborn, R., and M. Liermann. 1998. Standing on the shoulders of giants: learning from experience in fisheries. Reviews in Fish Biology and Fisheries **8**:273–283.

Hilborn, R., J. J. Maguire, A. M. Parma, and A. A. Rosenberg. 2001. The precautionary approach and risk management: can they increase the probability of successes in fishery management? Canadian Journal of Fisheries and Aquatic Sciences **58**:99–107.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, New York, New York, USA.

Hillebrand, H., and B. J. Cardinale. 2004. Consumer effects decline with prey diversity. Ecology Letters **7**:192–201.

Hobbs, N. T., J. E. Gross, L. A. Shipley, D. E. Spalinger, and B. A. Wunder. 2003. Herbivore functional response in heterogeneous environments: a contest among models. Ecology **84**:666–681.

Hobbs, N. T., S. Twombly, and D. Schimel. 2006. Deepening ecological insights using contemporary statistics. Ecological Applications **16**:3–4.

Holm, S., and J. Harris. 1999. Precautionary principle stifles discovery. Nature **400**:398.

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

Hunt, M. 1997. How science takes stocks. A history of meta-analysis. Russell Sage Foundation, New York, New York, USA.

Huxel, G. R. 1999. On the influence of food quality in consumer–resource interactions. Ecology Letters **2**:256–261.

Jeffreys, W. H., and J. O. Berger. 1992. Ockham's razor and Bayesian analysis. American Scientist **80**:64–72.

Johnson, D. H. 1999. The insignificance of statistical significance testing. Journal of Wildlife Management **63**:763–772.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution **19**:101–108.

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

Lee, K. N., and J. Lawrence. 1986. Adaptive management: learning from the Columbia River basin fish and wildlife program. Environmental Law **16**:431–460.

Leibold, M. A., J. M. Chase, J. B. Shurin, and A. L. Downing. 1997. Species turnover and the regulation of trophic structure. Annual Review of Ecology and Systematics **28**:467–494.

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

Lubow, B. C., F. J. Singer, T. L. Johnson, and D. C. Bowden. 2002. Dynamics of interacting elk populations within and adjacent to Rocky Mountain National Park. Journal of Wildlife Management **66**:757–775.

Mann, C. 1990. Meta-analysis in the breech. Science **249**:479–480.

May, R. M., editor. 1981. Theoretical ecology: principles and applications. Sinauer, Sunderland, Massachusetts, USA.

McCann, K. S., A. Hastings, and D. R. Strong. 1998. Trophic cascades and trophic trickles in pelagic food webs. Proceedings of the Royal Society of London Series B-Biological Sciences **265**:205–209.

McLain, R. J., and R. G. Lee. 1996. Adaptive management: promises and pitfalls. Environmental Management **20**:437–448.

McNaughton, S. J. 1979. Grazing as an optimization process—grass ungulate relationships in the Serengeti. American Naturalist **113**:691–703.

McNaughton, S. J. 1983. Serengeti grassland ecology: the role of composite environmental factors and contingency in community organization. Ecological Monographs **53**:291–320.

Mduma, S. A. R., A. R. E. Sinclair, and R. Hilborn. 1999. Food regulates the Serengeti wildebeest: a 40-year record. Journal of Animal Ecology **68**:1101–1122.

M'Gonigle, R. M., T. L. Jameison, M. K. McAllister, and R. M. Peterman. 1994. Taking uncertainty seriously: from permissive regulation to preventative design in environmental decision making. Osgoode Hall Law Journal **32**:99–169.

Mood, A. M., F. A. Graybill, and D. C. Boes. 1963. Introduction to the theory of statistics. Third edition. McGraw-Hill, New York, New York, USA.

Myers, R. A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science **58**:937–951.

Myers, R. A., N. J. Barrowman, R. Hilborn, and D. G. Kehler. 2002. Inferring Bayesian priors with limited direct data: applications to risk analysis. North American Journal of Fisheries Management **22**:351–364.

Myers, R. A., B. R. MacKenzie, K. G. Bowen, and N. J. Barrowman. 2001. What is the carrying capacity for fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod. Canadian Journal of Fisheries and Aquatic Sciences **58**:1464–1476.

Newman, K. B., S. T. Buckland, S. T. Lindley, L. Thomas, and C. Fernández. 2006. Hidden process models for animal population dynamics. Ecological Applications **16**:74–86.

Noy-Meir, I. 1993. Compensating growth of grazed plants and its relevance to the use of rangelands. Ecological Applications **3**:17–20.

Nystrom, P., A. R. McIntosh, and M. J. Winterbourn. 2003. Top-down and bottom-up processes in grassland and forested streams. Oecologia **136**:596–608.

Onstad, D. W. 1988. Population-dynamic theory: the roles of analytical, simulation, and super computer models. Ecological Modeling **43**:111–124.

Pawitan, Y. 2001. In all likelihood: statistical modeling and inference using likelihood. Oxford Scientific Publications, Oxford, UK.

Peterman, R. M., and M. M'Gonigle. 1992. Statistical power analysis and the precautionary principle. Marine Pollution Bulletin **24**:231–234.

Platt, J. R. 1964. Strong inference—certain systematic methods of scientific thinking may produce much more rapid progress than others. Science **146**:347–353.

Polis, G. A., A. L. W. Sears, G. R. Huxel, D. R. Strong, and J. Maron. 2000. When is a trophic cascade a trophic cascade? Trends in Ecology and Evolution **15**:473–475.

Punt, A. E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries **7**:35–63.

Punt, A. E., and R. Hilborn. 2002. Bayesian stock assessment methods in fisheries. Food and Agriculture Organization of the United Nations, Rome, Italy.

Royall, R. 1997. Statistical evidence: a likelihood paradigm. Chapman and Hall/CRC, New York, New York, USA.

Sala, O. E. 2001. Ecology—price put on biodiversity. Nature **412**:34–36.

Sauer, J. R., and W. A. Link. 2002. Hierarchical modeling of population stability and species group attributes from survey data. Ecology **83**:1743–1751.

Schlapfer, F., and B. Schmid. 1999. Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. Ecological Applications **9**:893–912.

Schmitz, O. J., P. A. Hamback, and A. P. Beckerman. 2000. Trophic cascades in terrestrial systems: a review of the effects of carnivore removals on plants. American Naturalist **155**:141–153.

Schnute, J. T. 1994. A general framework for developing sequential fisheries models. Canadian Journal of Fisheries and Aquatic Sciences **51**:1676–1688.

Schnute, J. T., and A. R. Kronlund. 2002. Estimating salmon stock-recruitment relationships from catch and escapement data. Canadian Journal of Fisheries and Aquatic Sciences **59**:433–449.

Steinauer, E. M., and S. L. Collins. 2001. Feedback loops in ecological hierarchies following urine deposition in tallgrass prairie. Ecology **82**:1319–1329.

Strong, D. R., A. V. Whipple, A. L. Child, and B. Dennis. 1999. Model selection for a subterranean trophic cascade: root-feeding caterpillars and entomopathogenic nematodes. Ecology **80**:2750–2761.

Tanaka, S., and R. Nishii. 1997. Model of deforestation by human population interactions. Environmental and Ecological Statistics **4**:83–91.

Taubes, C. H. 2001. Modeling differential equations in biology. Prentice Hall, Upper Saddle River, New Jersey, USA.

Ulanowicz, R. E. 1988. On the importance of higher-level models in ecology. Ecological Modeling **43**:45–56.

Vucetich, J. A., R. O. Peterson, and C. L. Schaefer. 2002. The effect of prey and predator densities on wolf predation. Ecology **83**:3003–3013.

Wachter, K. W. 1988. Disturbed by meta-analysis? Science **241**:1407–1408.

Walters, C., and D. Ludwig. 1994. Calculations of Bayes posterior probability distributions for key population parameters. Canadian Journal of Fisheries and Aquatic Sciences **51**:713–722.

Walters, C. J. 1986. Adaptive management of renewable resources. Macmillan, New York, New York, USA.

Walters, C. J., and R. Green. 1997. Valuation of experimental management options for ecological systems. Journal of Wildlife Management **61**:987–1006.

Walters, C. J., and R. Hilborn. 1978. Ecological optimization and adaptive management. Marine Environmental Research **9**:157–188.

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

White, G. C., and B. C. Lubow. 2002. Fitting population models to multiple sources of observed data. Journal of Wildlife Management **66**:300–309.

Worm, B., and J. E. Duffy. 2003. Biodiversity, productivity and stability in real food webs. Trends in Ecology and Evolution **18**:628–632.

Worm, B., H. K. Lotze, H. Hillebrand, and U. Sommer. 2002. Consumer versus resource control of species diversity and ecosystem functioning. Nature **417**:848–851.

APPENDIX A

An example calculation of likelihood (*Ecological Archives* A016-001-A1).

APPENDIX B

A discussion regarding what likelihood to use (*Ecological Archives* A016-001-A2).

APPENDIX C

A discussion on how to use likelihood and information theoretics to evaluate support in data for competing models (*Ecological Archives* A016-001-A3).

APPENDIX D

A further discussion on the concept of likelihood (*Ecological Archives* A016-001-A4).

SUPPLEMENT

Calculations shown in Appendix C (*Ecological Archives* A016-001-S1).

Alldredge, M. W., K. H. Pollock, T. R. Simons, J. Collazo, and S. A. Shriner. 2007. Time of detection method for estimating abundance from point count surveys. Auk, *in press*.

Best, L. B. 1981. Seasonal changes in detection of individual bird species. Studies in Avian Biology 6:252–261.

Both, C., S. Bouwhuis, C. M. Lessells, and M. E. Visser. 2006. Climate change and population declines in a long-distance migratory bird. Nature 441:81–83.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, New York, New York, USA.

Burnham, K. P. 1981. Summarizing remarks: environmental influences. Studies in Avian Biology 6:324–325.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.

DeJong, M. J., and J. T. Emlen. 1985. The shape of the auditory detection function and its implication for songbird censusing. Journal of Field Ornithology 56:213–223.

Diefenbach, D. R., D. W. Brauning, and J. A. Mattice. 2003. Variability in grassland bird counts related to observer differences and species detection rates. Auk 120:1168–1179.

Diehl, B. 1981. Bird populations consist of individuals differing in many respects. Studies in Avian Biology 6:225–229.

Emlen, J. T., and M. J. DeJong. 1981. The application of song detection threshold distance to census operations. Studies in Avian Biology 6:346–352.

Emlen, J. T., and M. J. DeJong. 1992. Counting birds: the problem of variable hearing abilities. Journal of Field Ornithology 63:26–31.

Faanes, C. A., and D. Bystrak. 1981. The role of observer bias in the North American breeding bird survey. Studies in Avian Biology 6:353–359.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414–425.

Kendall, W. L., B. G. Peterjohn, and J. R. Sauer. 1996. First-time observer effects in the North American Breeding Bird Survey. Auk 113:823–829.

Mayfield, H. F. 1981. Problems in estimating population size through counts of singing males. Studies of Avian Biology 6: 220–224.

McShea, W. J., and J. H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology. 68:367–375.

Morton, E. S. 1975. Ecological sources of selection on avian sounds. American Naturalist 109:17–34.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393–408.

Norvell, R. E., F. P. Howe, and J. R. Parrish. 2003. A seven-year comparison of relative-abundance and distance-sampling methods. Auk 120:1013–1028.

Ralph, J. C., S. Droege, and J. R. Sauer. 1995. Managing and monitoring birds using point counts: standards and applications. Pages 161–168 *in* J. C. Ralph, J. R. Sauer, and S. Droege, editors. Monitoring bird populations by point counts (U.S. Forest Service General Technical Report PSW-GTR-149.

Read, A. F., and D. M. Weary. 1992. The evolution of bird song: comparative analyses. Philosophical Transactions of the Royal Society of London B 338:165–187.

Richards, D. G. 1981. Environmental acoustics and censuses of singing birds. Studies in Avian Biology 6:297–300.

Robbins, C. S. 1981. Effect of time of day on bird activity. Studies in Avian Biology 6:275–286.

Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter. 2002. Landbird counting techniques: current practices and an alternative. Auk 119:46–53.

Sauer, J. R., J. E. Hines, and J. Fallon. 2003. The North American Breeding Bird Survey, results and analysis 1966–2002. Version 2003.1. USGS Patuxent Wildlife Research Center, Laurel, Maryland, USA.

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. Auk 111:50–62.

Scott, J. M., F. L. Ramsey, and C. B. Kepler. 1981. Distance estimation as a variable in estimating bird numbers from vocalizations. Studies in Avian Biology 6:334–340.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. Auk, *in press*.

Skirvin, A. A. 1981. Effect of time of day and time of season on the number of observations and density estimates of breeding birds. Studies in Avian Biology 6:271–274.

Thompson, W. L. 2002. Towards reliable bird surveys: accounting for individuals present but not detected. Auk 119:18–25.

Wiley, R. H., and D. G. Richards. 1982. Adaptations for acoustic communication in birds: sound transmission and signal detection. Pages 131–181 *in* D. E. Kroodsma, E. H. Miller, and H. Ouellet, editors. Acoustic communication in birds. Volume 1. Academic Press, New York, New York, USA.

Wilson, D. M., and J. Bart. 1985. Reliability of singing bird surveys: effects of song phenology during the breeding season. Condor 87:69–73.

ERRATA

In the paper by N. Thompson Hobbs and Ray Hilborn (2006) published as part of the "Contemporary Statistics and Ecology" Invited Feature, *Ecological Applications* 16(1):5–19, the equations published for Eqs. 6 and 7 were incorrect. The text spanning pp. 8–9 should read as follows:

... The likelihood of model *i* given the data is

$$\mathcal{L}(g_i|\mathbf{Y}) = e^{-\frac{1}{2}\Delta_i} \tag{6}$$

and the Akaike weight, w_i, is calculated for each model as

$$w_i = \frac{e^{-\frac{1}{2}\Delta_i}}{\sum_{r=1}^{R} e^{-\frac{1}{2}\Delta_r}}. \tag{7}$$