Density Dependence in Time Series Observations of Natural Populations: Estimation and Testing

Brian Dennis; Mark L. Taper

Stable URL:

http://links.jstor.org/sici?sici=0012-9615%28199405%2964%3A2%3C205%3ADDITSO%3E2.0.CO%3B2-%23

DENSITY DEPENDENCE IN TIME SERIES OBSERVATIONS OF NATURAL POPULATIONS: ESTIMATION AND TESTING[1]

Brian Dennis

Department of Fish and Wildlife Resources[2] and Department of Mathematics and Statistics,
University of Idaho, Moscow, Idaho 83844 USA

Mark L. Taper

Department of Biology, Montana State University, Bozeman, Montana 59717 USA

Abstract. We report on a new statistical test for detecting density dependence in uni-variate time series observations of population abundances. The test is a likelihood ratio test based on a discrete time stochastic logistic model. The null hypothesis is that the population is undergoing stochastic exponential growth, stochastic exponential decline, or random walk. The distribution of the test statistic under both the null and alternate hypotheses is obtained through parametric bootstrapping. We document the power of the test with extensive simulations and show how some previous tests in the literature for density dependence suffer from either excessive Type I or excessive Type II error. The new test appears robust against sampling or measurement error in the observations. In fact, under certain types of error the power of the new test is actually increased. Example analyses of elk (*Cervus elaphus*) and grizzly bear (*Ursus arctos horribilis*) data sets are provided. The model implies that density-dependent populations do not have a point equilibrium, but rather reach a stochastic equilibrium (stationary distribution of population abundance). The model and associated statistical methods have potentially important applications in conservation biology.

Key words: bootstrapping; conservation biology; density dependence; elk; equilibrium; grizzly bear; likelihood ratio; logistic model; nonlinear autoregressive model; population regulation; statistical power; stochastic difference equation; stochastic population model; time series analysis.

Introduction

Whether or not populations in nature tend to have growth rates regulated by their own densities has long been a key but frustrating problem of ecological research (The Biological Laboratory 1957, McLaren 1971, Colinvaux 1973, Kingsland 1985). It is widely acknowledged that possible answers to this question have broad theoretical implications for the structure of communities and the evolution of the species that comprise them (Pianka 1974, Cody and Diamond 1975, May 1976, Roughgarden 1979, Emlen 1984, Diamond and Case 1986). Moreover, assessment of density dependence has attained great practical importance in conservation biology because the strength and form of density dependence have a large influence on the expected survival times of natural populations (Ginzburg et al. 1990, Stacey and Taper 1992).

The availability in recent years of "long term" (≈20 yr) data on population abundances has steered part of the density dependence controversy into a debate about statistical concepts and methods (Eberhardt 1970, Reddingius 1971, 1990, Bulmer 1975, Royama 1977, 1981, Slade 1977, Vickery and Nudds 1984, 1991, Gaston and Lawton 1987, Pollard et al. 1987, den Boer

and Reddingius 1989, Reddingius and den Boer 1989, den Boer 1990, 1991, Solow 1990, 1991, Turchin 1990, Turchin et al. 1991, Crowley 1992, Turchin and Taylor 1992). One conceptual argument, exemplified by den Boer (1991), asserts that if density dependence is to be a cornerstone of ecological theory, a certain burden of proof needs to be satisfied. It is the role of density dependence theorists to demonstrate convincingly that time series abundance data can be distinguished statistically from the trajectories of a density-independent stochastic growth model or even of a random walk model. Another argument discounts the need for statistical detection of density dependence, claiming that population regulation is a purely logical consequence of the ecological processes by which populations grow (Royama 1977, Berryman 1991a). A third argument questions the ecological meaning of existing statistical methods for density dependence testing, because the concept of a population equilibrium is unclear when the environment fluctuates (Wolda 1989). Throughout this conceptual debate, statistical methods for detecting density dependence have been introduced and refined in a steady stream (Reddingius 1971, Bulmer 1975, Slade 1977, Vickery and Nudds 1984, Pollard et al. 1987, Reddingius and den Boer 1989, Turchin 1990, den Boer 1991). The conclusions of various data analysis studies about the prevalence of density dependence seem to vary depending on which methods are used.

One group of studies has used a statistical model of population growth with a density dependence term proportional to the logarithm of population abundance (Reddingius 1971, Bulmer 1975, Gaston and Lawton 1987, den Boer and Reddingius 1989, Reddingius and den Boer 1989, den Boer 1990, Vickery and Nudds 1991, Crowley 1992). While different particular methods for testing whether the density dependence term should be included in this model have different powers (Pollard et al. 1987, Vickery and Nudds 1991), these analyses frequently suggested that density dependence is not as prevalent as expected by ecological theory. The related concepts of "stabilization" (den Boer 1968, 1990) and "density vagueness" (Strong 1986a, b) have been suggested to account for such findings; the concepts essentially take population growth to be density independent (but noisy) over a wide range of densities, with density dependent regulation occurring more or less sharply at very high densities.

However, in contrast to the above studies, Woiwod and Hanski (1992) and Holyoak and Lawton (1992) detected frequent density dependence using tests based on the same logarithmic density dependence model (among other tests). Woiwod and Hanski (1992) analyzed thousands of insect data sets, many of which exceeded 20 observations in length; Holyoak and Lawton (1992) treated 32 insect data sets of 8 or 12 observations. In these studies, longer time series showed increased prevalence of density dependence. Earlier results of Hassell et al. (1989) and Solow and Steele (1990) had also highlighted the importance of sample size to the statistical power of density dependence tests.

Another set of studies employed a model with a density dependence term proportional to population abundance (Turchin 1990, Berryman 1991a, Turchin et al. 1991, Turchin and Taylor 1992). Woiwod and Hanski (1992) and Holyoak and Lawton (1992) used the model as well. These studies found widespread density dependence, sometimes in the form of delayed regulation (second order lags; see Turchin 1990). The statistical methods used to test whether the density dependence term(s) should be included in the model were based on standard results from ordinary regression analysis. Generalization of these analyses to multiple species systems has been reported (Berryman 1991b).

Still other analyses have been based on various statistical properties of random walks (Vickery and Nudds 1984, den Boer 1991, Crowley 1992). The studies regarded a random walk model as the null hypothesis to be rejected by data according to some criterion. While Pollard et al. (1987) suggested that an exponential growth model (containing the random walk model as a special case) makes a more biologically interesting null hypothesis, the possibility that real data sets often cannot be distinguished from random walk trajectories remains unsettling to density dependence proponents. Indeed, den Boer (1991) concludes that Nicholson's (1933) hypothesis that populations "exist in a state of balance because densities fluctuate about a relatively stable norm" is not supported by random walk comparisons or other statistical tests. Though den Boer (1991) does caution that these analyses do not mean that populations obey random walk models, his results should inspire some rereadings of Birch's (1957) and Andrewartha's (1957) earlier density independence arguments.

In this paper, we introduce a new test for density dependence in time series data of population abundances. We propose that a discrete time stochastic logistic model used by Turchin (1990) and Berryman (1991a) can serve as a useful and descriptive model for such testing in a variety of ecological situations. Statistical inference methods for this model, however, have not been well understood in the past. We develop parameter estimation methods and hypothesis testing methods for the model and focus on a likelihood ratio hypothesis test of density-independent vs. density-dependent population growth. Because the distribution of the test statistic is intractable, we show how its critical values can be estimated with a parametric bootstrapping method. The power properties of this new test are documented here with extensive simulations. We illustrate the use of the test with examples. The results of past empirical studies are likely influenced by the statistical testing methods used. In particular, we show that the randomization test of Pollard et al. (1987) has low power (excessive Type II error) compared to the new test. Also, we find that the regression tests of Turchin (1990) and Berryman (1991a) suffer from inflated size (excessive Type I error). The likelihood ratio test proposed here, by contrast, is a size 0.05 test and represents the practical limits of power that can be attained for the stochastic logistic growth model. We discuss the effects of sampling variability, the ecological interpretation of density dependence testing, the concept of a stochastic equilibrium, and the potential use of the new test in population viability analysis.

MODEL DESCRIPTION

Let N_t represent population abundance (as censused, estimated, or indexed) at time t, where $t = 0, 1, 2, \ldots$. The model we present relates N_{t+1} to N_t:

$$N_{t+1} = N_t \exp(a + bN_t + \sigma Z_t). \tag{1}$$

Here a and b are constants, σ is a positive constant, and σZ_t is a random shock to the population growth rate. In this paper, we are mostly concerned with values of b such that $b \leq 0$. We assume that Z_t has a normal distribution with a mean of 0 and a variance of 1 [we write $Z_t \sim$ normal(0, 1)], and that $Z_0, Z_1, Z_2, \ldots$ are uncorrelated. The model involves two essential ideas. First, the per-unit-abundance growth rate is defined in discrete time as $\ln N_{t+1} - \ln N_t$, analogous to $(1/n) \, dn/$

$dt = d \ln n/dt$ in continuous time. Second, that rate so defined is taken to be a linear function of N_t plus noise.

The constant b is the slope of the linear function. If $b = 0$, the per-unit-abundance growth rate does not depend on N_t. When $b < 0$, the per-unit-abundance growth rate decreases as N_t becomes larger. An increasing per-unit-abundance growth rate, or Allee effect, results when $b > 0$ (Dennis 1989a).

The type of variability inherent in the model (Eq. 1) is "environmental" as opposed to "demographic." Models with demographic variability become essentially deterministic as population size becomes large (see discussion by Dennis et al. 1991). A population governed by Eq. 1, however, fluctuates at large as well as small sizes. The distinctions between environmental and demographic variability have ramifications in conservation biology (Leigh 1981, Shaffer 1981, Goodman 1987, Simberloff 1988, Dennis et al. 1991, Wissel and Stöcker 1991).

The population abundances N_0, N_1, N_2, . . . are not independent under this model, even though the random shocks (Z_t) are independent. As we show later in this paper, failure to account for the dependence among the N_t values is the source of flaws in some previous statistical tests for density dependence. The stochastic process N_t defined by Eq. 1 is a Markov process: given that the population has attained some particular size n_t at time t, the future distribution of population sizes depends on n_t, but not on past sizes.

The Markov property is a fairly general assumption applicable in many ecological situations. The deterministic analogue of the Markov property is simply that population abundance can be described by a first-order difference equation. Even in populations with overlapping generations or age structure, some index of population abundance can behave as if governed by a first-order difference equation. For example, Livdahl and Sugihara (1984) and Barlow (1992) document systems in which complex, nonlinear life histories give rise to simple linear dependence of per-unit-abundance growth rate on abundance. Also, Cushing (1989) has provided a theoretical justification of how a simple nonlinear difference equation can emerge from a population projection matrix model (such as a Leslie matrix) in which there is nonlinear dependence of demographic rates on population abundance. We discuss later the evaluation of the model for a given data set by residual analysis and by testing for second-order lags (see *Hypothesis testing* and *Discussion*).

We point out that the mean population abundance at time $t + 1$ under the model is not given by Eq. 1 with $\sigma = 0$. Because $E[\exp(\sigma Z_t)] = \exp(\sigma^2/2)$ and because of the Markov property, the mean population abundance at time $t + 1$ given $N_t = n_t$ is

$$E(N_{t+1} \mid N_t = n_t) = n_t \exp(r + bn_t). \tag{2}$$

Here $r = a + (\sigma^2/2)$.

A deterministic analogue to Eq. 2 is a type of discrete time logistic model that has been analyzed extensively in population ecology (May 1976):

$$n_{t+1} = n_t \exp(r + bn_t). \tag{3}$$

This model is known also as a Ricker equation from its similarity to the Ricker stock-recruitment relationship in fisheries (Ricker 1954). The linear form $r + bn_t$ is a simple way of representing density dependent feedback in the per-unit-abundance growth rate (as defined by $\ln n_{t+1} - \ln n_t$). The deterministic model has a positive point equilibrium at

$$\tilde{n} = -r/b, \tag{4}$$

provided $b < 0$.

Eq. 3 may seem an overly simplified representation of the complex processes of density dependence in natural populations. However, the linear relationship $r + bn_t$ can be regarded as a Taylor series approximation near $\tilde{n}$ of a more biologically detailed rate function (Dennis and Patil 1984, Dennis and Costantino 1988). The stability properties of $\tilde{n}$ and the dynamic behaviors of the deterministic model up to and including chaos are well known (May 1976). The growth model defined by Eq. 1 is a stochastic generalization of Eq. 3 and can be regarded as a type of stochastic, discrete time logistic model.

This stochastic logistic model becomes a first-order nonlinear autoregression model when transformed to a logarithmic scale. By letting $X_t = \ln N_t$, we obtain the following time series model from Eq. 1:

$$X_{t+1} = X_t + a + be^{X_t} + \sigma Z_t. \tag{5}$$

Transforming the model to a logarithmic scale has three main advantages. First, theoretical statistical knowledge about such nonlinear autoregressive models has increased in recent years (Tong 1990). Use of Eq. 5 provides connections between ecological time series data, mathematical population modeling, and established results in mathematical statistics. Second, valid point estimates (but not confidence intervals or hypothesis tests) of the parameters can be obtained with ordinary linear regression packages (see *Parameter estimation*). Third, for some parameter values, one can obtain a diffusion process approximation to X_t. Such an approximation provides simple expressions for long-run statistical properties of X_t, including the stationary distribution and mean first-passage times (see *Discussion*).

The model can be altered to include second- or higher order lags. The model would take the form

$$N_{t+1} = N_t \exp(a + b_1N_t + b_2N_{t-1} + \dots + b_mN_{t-m+1} + \sigma Z_t) \tag{6}$$

for incorporating time lags up to order m. Using this model, Turchin (1990) and Turchin et al. (1991) have argued for the prevalence of second-order lags in ecological populations. We make some preliminary recommendations in this paper (see *Discussion*) about how

testing for second-order lags might be accomplished. A full account of the statistical properties of Eq. 6 and of statistical inference methods for the model must be deferred to a future paper.

An alternative first-order population model was introduced by Reddingius (1971):

$$N_{t+1} = N_t \exp(a + b \ln N_t + \sigma Z_t). \quad (7)$$

Royama (1981) modified this model to incorporate higher order time lags. On a logarithmic scale, Eq. 7 becomes

$$X_{t+1} = X_t + a + bX_t + \sigma Z_t, \quad (8)$$

where, as before, $X_t = \ln N_t$. A statistical motivation for use of this model is that it can be written in the form of a linear, first-order autoregressive model (AR(1)):

$$X_{t+1} - \mu = \beta(X_t - \mu) + \sigma Z_t, \quad (9)$$

with $\mu = -a/b$ and $\beta = 1 + b$. The AR(1) model has well-known statistical properties and established, packaged inference procedures. Written in the form of Eq. 7, the AR(1) model is seen to be a type of discrete time, stochastic Gompertz model [the Gompertz growth equation is $(1/n)dn/dt = a + b \ln n$]. While the statistical convenience of this model is a desirable quality, its biological postulate is that growth rate depends, if at all, only logarithmically on population density. By contrast, the stochastic logistic model (Eq. 1) structurally allows for stronger density dependence.

Computer-generating a time series from the stochastic logistic model using Eq. 5 is a simple procedure. Given numerical values of a, b, and σ^2, and starting at a fixed value $X_0 = x_0$, one can easily calculate X_1, X_2, . . . recursively with the help of a routine for generating standard normal random variables. The simplicity of generating trajectories from the hypothesized stochastic mechanism underlying the data turns out to be a key for convenient and powerful statistical inferences (see *Parameter estimation* and *Hypothesis testing*).

Distinguishing three cases of the model (Eq. 5) is important to density dependence testing. The cases form a series of three nested hypotheses. The simplest is Model 0:

$$H_0: a = 0, b = 0. \quad (10)$$

Model 0 defines X_t as a discrete time Brownian motion process (or Wiener process) with zero drift ($a = 0$). This is the classic "random walk" model; X_t has a normal distribution centered at x_0 with a variance of $\sigma^2 t$. No feedback of population density to the growth rate takes place ($b = 0$).

Model 1 is

$$H_1: a \neq 0, b = 0. \quad (11)$$

Model 1 also defines X_t as a discrete time Brownian motion process, but this time with a positive or negative drift parameter ($a \neq 0$). Under Model 1, X_t has a normal distribution with a mean of $x_0 + at$ and a variance of $\sigma^2 t$. Model 1 in the original population abundance scale (Eq. 1) is a type of stochastic exponential growth or decay model. It is identical to the model described by Dennis et al. (1991) for estimating extinction risks for endangered species (their parameter μ is the same as a in this paper; σ^2 is the same quantity in both papers). In Model 1, no density dependent feedback occurs ($b = 0$).

Finally, it is Model 2 that contains full-fledged density dependence:

$$H_2: a \neq 0, b \neq 0. \quad (12)$$

In many cases, a one-sided variant of Model 2 is of greatest interest, and we can redefine Model 2 as

$$H_2: a \neq 0, b < 0. \quad (13)$$

Testing for density dependence can be regarded as determining whether the added parameter in Model 2 produces noticeably improved description of the data. The first step in such determination is estimating the unknown parameters from the data.

PARAMETER ESTIMATION

Statistically, the problem of connecting the model (Eq. 1) with data amounts to specifying a likelihood function. Let $n_0, n_1, \ldots, n_q$ be the recorded population abundances, so that q is the number of one-step transitions and $q + 1$ is the total number of observations in the time series. Let $x_0 = \ln n_0$, $x_1 = \ln n_1$, . . . , $x_q = \ln n_q$ denote the log-transformed abundances. The likelihood function gives the probability that, under the stochastic mechanism defined by Eq. 1, the outcome of the process N_t would be the observed time series. The likelihood function is defined as the joint probability density function (pdf) for the random variables N_0, N_1, N_2, . . . , N_q evaluated at $n_0, n_1, n_2, \ldots, n_q$. It is more convenient to specify the likelihood function for the log-transformed observations, because of the autoregressive structure indicated by Eq. 5. Given that log-population size is at x_{t-1} at time $t - 1$, the distribution of X_t is, according to Eq. 5, normal with a mean of $x_{t-1} + a + b \exp(x_{t-1})$ and a variance of σ^2. Thus, the pdf for X_t, given $X_{t-1} = x_{t-1}$, is a normal curve:

$$p(x_t | x_{t-1}) = (\sigma^2 2\pi)^{-1/2} \exp[-(x_t - x_{t-1} - a - be^{x_{t-1}})^2/(2\sigma^2)]. \quad (14)$$

Because of the Markov property, the joint likelihood of the data is the likelihood of a transition from x_0 to x_1, multiplied by the likelihood of a transition from x_1 to x_2, etc. This joint likelihood is just a product of normal pdf's of the form given by Eq. 14. It is a function of the data values x_0, x_1, . . . , x_q, and, more importantly, of the unknown model parameters a, b, and σ^2:

TABLE 1. Maximum likelihood estimates for unknown parameter σ_0^2 in Model 0 (random walk), parameters a_1 and σ_1^2 in Model 1 (stochastic exponential growth), and parameters b_2, a_2, and σ_2^2 in Model 2 (stochastic logistic growth). Observed population abundance at time t is n_t; also $x_t = \ln n_t$, $y_t = x_t - x_{t-1}$, $\bar{y} = (y_1 + y_2 + \cdots + y_q)/q$, and $\bar{n} = (n_0 + n_1 + \cdots + n_{q-1})/q$.

Model 0	$\hat{\sigma}_0^2 = \dfrac{1}{q}\sum_{t=1}^{q} y_t^2$		
Model 1	$\hat{a}_1 = \bar{y}$	$\hat{\sigma}_1^2 = \dfrac{1}{q}\sum_{t=1}^{q}(y_t - \bar{y})^2$	
Model 2	$\hat{b}_2 = \dfrac{\sum_{t=1}^{q}(y_t - \bar{y})(n_{t-1} - \bar{n})}{\sum_{t=1}^{q}(n_{t-1} - \bar{n})^2}$	$\hat{a}_2 = \bar{y} - \hat{b}_2\bar{n}$	$\hat{\sigma}_2^2 = \dfrac{1}{q}\sum_{t=1}^{q}(y_t - \hat{a}_2 - \hat{b}_2 n_{t-1})^2$

$$
\begin{aligned}
& L(a, b, \sigma^2) \\
&= \prod_{t=1}^{q} p(x_t \mid x_{t-1}) \\
&= \sigma^2 2\pi^{-q/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{t=1}^{q}(x_t - x_{t-1} - a - be^{x_{t-1}})^2\right].
\end{aligned}
$$
$$(15)$$

This likelihood function $L(a, b, \sigma^2)$ plays a central role in parameter estimation and hypothesis testing.

We must note that Eq. 15, strictly speaking, is not the joint pdf of X_0, X_1, ..., X_q evaluated at the observations (it is not the full likelihood). Rather, it is the joint pdf of X_1, X_2, ..., X_q, conditional on $X_0 = x_0$, and evaluated at the observations. We recommend conditioning on the initial observed population size (and using Eq. 15 as the likelihood) because in practice the probabilistic mechanism producing the observation x_0 is typically unknown.

Maximum likelihood (ML) parameter estimates have numerous desirable statistical properties (Stuart and Ord 1991). ML estimates are defined as the parameter values, denoted $\hat{a}$, $\hat{b}$, and $\hat{\sigma}^2$, that jointly maximize $L(a, b, \sigma^2)$ (or equivalently, $\ln L(a, b, \sigma^2)$). ML estimates are asymptotically efficient (they have the smallest variances in large samples), are consistent (variances approach zero as $q \to \infty$), are asymptotically unbiased (biases approach zero as $q \to \infty$), and have distributions that approach normal distributions for large samples. Standard mathematical statistics books only list these properties for independent, identically distributed observations (Stuart and Ord 1991, Rice 1988). We point out that these desirable properties of ML estimates have been demonstrated for time series models (dependent data) of this type as well (Bhat 1974, Tong 1990).

Obtaining ML estimates for the random walk (Model 0) and exponential growth (Model 1) models is easy. Let $L_0(\sigma^2)$ represent the likelihood function for Model 0, that is, Eq. 15 with $a = 0$ and $b = 0$. Let $L_1(a, \sigma^2)$ represent the likelihood function for Model 1 ($b = 0$). Also, let $y_t = x_t - x_{t-1}$, $t = 1, 2, \ldots, q$. Note that y_t

can be thought of as the per-unit-abundance growth rate observed in the population for transition t. The value of σ^2, denoted $\hat{\sigma}_0^2$, that maximizes $L_0(\sigma^2)$ appears in Table 1. It is easily found by setting $\partial \ln L_0(\sigma^2)/\partial\sigma^2$ equal to zero and solving for σ^2. The values of a and σ^2 that jointly maximize $L_1(a, \sigma^2)$, denoted $\hat{a}_1$ and $\hat{\sigma}_1^2$, also appear in Table 1. They are found by setting $\partial \ln L_1(a, \sigma^2)/\partial a$ and $\partial \ln L_1(a, \sigma^2)/\partial\sigma^2$ equal to zero simultaneously.

The estimates for Models 0 and 1 are familiar ones. Let Y_1, Y_2, ..., Y_q denote the one-step differences in logarithmic population sizes: $Y_t = X_t - X_{t-1}$ (the value realized by Y_t in the data is y_t). Because the Y_t's are increments of Brownian motion under Models 0 or 1, much is known about their statistical properties (see Dennis et al. 1991). Under Model 0, $Y_t \sim$ normal(0, σ^2) and Y_1, Y_2, ..., Y_q are independent. Under Model 1, $Y_t \sim$ normal(a, σ^2) and Y_1, Y_2, ..., Y_q are independent. Both models reduce to simple cases of random sampling from normal distributions. Thus, standard confidence intervals from normal theory for σ^2 (Model 0) or a and σ^2 (Model 1) are valid and exact. Dennis et al. (1991) give further details.

For the full stochastic logistic model (Model 2), obtaining ML estimates is also easy. To keep notation consistent, we will write $L_2(a, b, \sigma^2)$ instead of $L(a, b, \sigma^2)$ for the likelihood function of Model 2 (Eq. 15). The values of b, a, and σ^2 that jointly maximize $L_2(a, b, \sigma^2)$ are listed in Table 1. The Model 2 estimates are also familiar ones. The estimates of a and b are least squares estimates obtained by performing a linear regression of y_t on n_{t-1}, $t = 1, 2, \ldots, q$. Thus, $\hat{a}_2$, $\hat{b}_2$, and $\hat{\sigma}_2^2$ can be calculated with any standard regression package.

Confidence intervals printed by standard regression packages, however, are not valid for Model 2. Printed hypothesis tests for parameter values are not valid either. For the printed intervals and tests to be valid, the Y_t's would have to have independent normal($a + bn_{t-1}$, σ^2) distributions (the standard linear regression model). Under Model 2, though, the Y_t's are not independent, due to the autoregressive structure of the model, nor do they have unconditional normal distributions. The

overlap between the standard linear regression model and Model 2 stops at point estimation; both models happen to have identical ML estimates. The statistical distributions of the ML estimates under the two models are radically different, though, and so interval estimates and hypothesis tests for Model 2 cannot be based on the ordinary regression model.

Instead, approximate confidence intervals for parameters in Model 2 can be found by either bootstrapping or jackknifing. Bootstrapping involves estimating the distributions of $\hat{a}_2$, $\hat{b}_2$, and $\hat{\sigma}_2{}^2$ in some fashion using the data. The following parametric bootstrap method makes efficient use of the information present in modest-sized samples ($q \approx 20$). The procedure is simple but requires some computer programming. First, from the data calculate the ML parameter estimates for Model 2 (Table 1). Second, repeatedly (say, 2000 times) computer-generate time series of the same length as the original data from the *estimated* model:

$$X_t = X_{t-1} + \hat{a}_2 + \hat{b}_2 \exp(X_{t-1}) + \hat{\sigma} Z_t. \qquad (16)$$

A given series generated from the estimated model ("bootstrapped" series) will be denoted with asterisks: $x_0{}^*, x_1{}^*, \ldots, x_q{}^*$. Each series should be started at the observed initial value: $X_0 = x_0{}^* = x_0$. To each such bootstrapped time series, fit Model 2 by calculating bootstrapped ML estimates, denoted $\hat{a}_2{}^*$, $\hat{b}_2{}^*$, and $\hat{\sigma}_2{}^{2*}$, using the expressions in Table 1. Third, for an approximate 95% confidence interval for a, take the 2.5[th] and the 97.5[th] sample percentiles of the 2000 $\hat{a}_2{}^*$ values. Likewise, use the $\hat{b}_2{}^*$ and the $\hat{\sigma}_2{}^{2*}$ values to construct confidence intervals for b and σ^2. An adjustment known as the bias-corrected percentile method (see Efron and Gong 1983) might make the coverage rate of the bootstrap confidence interval closer to 95%.

Jackknifing also involves using the data to estimate the distributions of $\hat{a}_2$, $\hat{b}_2$, and $\hat{\sigma}_2{}^2$, and it typically requires less computing time. Lele (1991) describes a jackknifing technique for dependent data. For Model 2, Lele's technique entails dropping transitions from the data one by one and refitting the model each time to the remaining transitions. A transition here is a change in the data over one time step, that is, from x_{t-1} to x_t. One performs the regressions on the pairs $(y_1, n_0), (y_2, n_1), \ldots, (y_q, n_{q-1})$, each time leaving out (y_t, n_{t-1}), where $t = 1, 2, \ldots, q$. Lele (1991) provides expressions for consistent estimates of the variances and covariances of $\hat{a}_2$, $\hat{b}_2$, and $\hat{\sigma}_2{}^2$.

We have focused our computer simulations in this paper on evaluating hypothesis testing methods and cannot make any recommendations at this time concerning which type of confidence intervals for Model 2 parameters have superior properties. A large-scale evaluation of the coverage probabilities for bootstrapped and jackknifed confidence intervals is a topic for future research.

Missing data can be handled in the ML estimates by simply incorporating in the analysis all the one-step transitions present in the data. Thus, if the jth year (or whatever time period) population size, n_j, was missing, one would perform the regression calculations using $(y_1, n_0), (y_2, n_1), \ldots, (y_{j-1}, n_{j-2}), (y_{j+2}, n_{j+1}), \ldots, (y_q, n_{q-1})$. One missing year means that there are $q - 2$ one-step transitions present in the data (two transitions missing). The ML formulas (Table 1) would have $q - 2$ instead of q as a divisor, and $\bar{n}$ (Table 1) would include $n_0, n_1, \ldots, n_{j-2}, n_{j+1}, n_{j+2}, \ldots, n_{q-1}$ in the sum (but not n_{j-1} or n_q).

HYPOTHESIS TESTING

Statistical theory draws a careful distinction between a statistical hypothesis and a scientific hypothesis (for instance, see Stuart and Ord 1991). A statistical hypothesis is an assumption about the form of a probability model, and a statistical hypothesis test is the use of data to make a decision between two probability models. A scientific hypothesis, on the other hand, is an explanatory assertion about some aspect of nature.

For density dependence studies, a general scientific hypothesis of interest is the assertion that a population's abundance produces a negative feedback effect on its growth rate (Berryman 1991a). From this assertion, we expect that time series observations of a density-dependent population would lead us to favor Model 2 over Model 1 as a model of the population's abundance. However, investigators should be aware that other stochastic mechanisms besides ecological feedback can produce observations that pass statistical density dependence tests, including the test described here (see *Discussion*).

When deciding between two models, the likelihood ratio (LR) test originating with Neyman and Pearson provides the benchmark for test power (Neyman and Pearson 1933, Stuart and Ord 1991). If the two probability models are completely specified (no unknown parameters), the LR test has power greater than or equal to any other size α test, according to the Neyman–Pearson Lemma (Stuart and Ord 1991). Here, however, Models 0, 1, and 2 are not completely specified, that is, they contain unknown parameters. LR tests that are modified to accommodate unknown parameters are sometimes called "generalized LR tests." The statistical criteria for choosing test methods are more complex when one or more of the models is not completely specified. In some simple textbook cases (for example, a one-sided t test), the LR test is the uniformly most powerful test. In other cases, the power of the LR test tends to compare quite favorably to other tests according to various definitions of "asymptotic relative efficiency" (these criteria are reviewed by Serfling 1980 and Stuart and Ord 1991). Thus, LR tests for comparing Models 0 and 1, Models 1 and 2, or Models 0 and 2, if feasible to construct, would likely offer desirable power properties.

In general, the LR test for comparing two models, i and j say, is constructed as follows. The model with

the smallest number of unknown parameters typically forms the null hypothesis, while the more complex model becomes the alternate hypothesis. Let $\hat{L}_i$ denote the likelihood function for the null hypothesis (Model i), evaluated at the ML estimates of all unknown parameters in the model. Essentially, $\hat{L}_i$ represents the estimated joint probability density of the data (or estimated likelihood that the data would have arisen) under Model i. Likewise, let $\hat{L}_j$ denote the alternate hypothesis (Model j) likelihood function, maximized over all parameter values permissible in the model. The LR test is to choose between Model i and Model j on the basis of the value of the LR test statistic:

$$\Lambda_{ij} = \hat{L}_i/\hat{L}_j. \tag{17}$$

The decision is made in favor of Model i if $\Lambda_{ij} > c$, where c is some constant cutoff point selected by the investigator and is made in favor of Model j if $\Lambda_{ij} \leq c$. The value of c is selected so that the probability of wrongly choosing Model j when the data in fact arise from Model i (that is, the probability of a Type I error) is fixed at some small number, α (the size of the test). The study of such LR tests occupies a prominent portion of any mathematical statistics text (e.g., Bain and Englehardt 1987, Rice 1988).

The essential problem in constructing an LR test is finding the value of c corresponding to the desired test size, α. In the normal-based linear models of analysis of variance and regression, the test statistic Λ_{ij} is a monotone function of the more familiar variance ratio statistic. Under the null hypothesis, the variance ratio statistic has an F distribution. The value of c then is calculated by transforming the $100(1 - \alpha)^{\text{th}}$ percentile of the appropriate F distribution. In a broad class of other models, including many nonlinear regression models, time series models, and loglinear models, the statistic $G_{ij}^2 = -2 \ln \Lambda_{ij}$ has, under the null hypothesis, a distribution that converges to a chi-square distribution as sample sizes increase. The value of c is obtained (approximately) from the $100(1 - \alpha)^{\text{th}}$ percentile of the chi-square distribution. Unfortunately, for testing among Models 0, 1, and 2, blind application of these traditional results can lead to erroneous inferences.

In the case of testing Model 0 (random walk) vs. Model 1 (exponential growth), the LR statistic does reduce to a statistic with an F distribution (with 1 and $q - 1$ degrees of freedom), or equivalently, a Student's t distribution ($q - 1$ degrees of freedom). Under Model 0, $\hat{L}_0 = L_0(\hat{\sigma}_0^2)$ is the likelihood function evaluated at the ML estimate of σ^2 (Table 1), and $\hat{L}_1 = L_1(\hat{a}_1, \hat{\sigma}_1^2)$ is the likelihood function evaluated at the ML estimates of a and σ^2 (Table 1). The LR test statistic can be written in several algebraically equivalent forms:

$$\Lambda_{01} = L_0(\hat{\sigma}_0^2)/L_1(\hat{a}_1, \hat{\sigma}_1^2) = (\hat{\sigma}_0^2/\hat{\sigma}_1^2)^{-q/2}$$
$$= \{1 + [T_{01}^2/(q - 1)]\}^{-q/2}. \tag{18}$$

Here,

$$T_{01} = \hat{a}_1/[q\hat{\sigma}_1^2/(q - 1)]^{1/2} \tag{19}$$

is the familiar t statistic for testing whether the mean is zero or not for independent normal(a, σ^2) random variables. Under Model 0, T_{01} has a Student's t distribution with $q - 1$ degrees of freedom. Alternatively, $T_{01}^2 = F_{01}$ has an F distribution with 1 and $q - 1$ degrees of freedom. Thus, the cutoff point c for Λ_{01} is a simple function of an F or a Student's t percentile. Model 1 is favored over Model 0 if $|T_{01}| \geq t_{\alpha/2}$, where $t_{\alpha/2}$ is the $100[1 - (\alpha/2)]^{\text{th}}$ percentile of the appropriate Student's t distribution.

In some circumstances, a one-sided hypothesis about a might be of interest. For instance, a test of H_0: $a = 0$ vs. H_1: $a < 0$ could be used to determine if an endangered species is in decline or not, provided the species is not abundant enough for density dependence effects to be important. The one-sided test in this case would reject H_0 if $T_{01} \leq t_{1-\alpha}$.

Unfortunately, for testing Model 1 vs. Model 2, the distribution of the LR statistic is unknown. The likelihood function maximized under Model 2 (Eq. 15) is given by $\hat{L}_2 = L_2(\hat{a}_2, \hat{b}_2, \hat{\sigma}_2^2)$. The LR test statistic can be written in the following forms:

$$\Lambda_{12} = L_1(\hat{a}_1, \hat{\sigma}_1^2)/L_2(\hat{a}_2, \hat{b}_2, \hat{\sigma}_2^2)$$
$$= (\hat{\sigma}_1^2/\hat{\sigma}_2^2)^{-q/2} = \exp(-G_{12}^2/2)$$
$$= \{1 + [T_{12}^2/(q - 2)]\}^{-q/2}. \tag{20}$$

Here $G_{12}^2 = -2 \ln \Lambda_{12}$, and

$$T_{12} = \hat{b}_2[(q - 2) \sum_{t=1}^{q} (n_{t-1} - \bar{n})^2/(q\hat{\sigma}_2^2)]^{1/2}. \tag{21}$$

The statistic T_{12} is identical to the familiar t statistic used for testing whether the slope parameter is nonzero in a linear regression. However, T_{12} (Eq. 21) does not have a Student's t distribution, not even approximately, due to the time dependence of the observations. Testing for density dependence based on an assumed Student's t distribution for T_{12} produces unacceptably inflated Type I error rates (see *Discussion*).

One should not even be lulled into using the traditional chi-square approximation for the distribution of G_{12}^2. Under the exponential growth model of the null hypothesis, the population is not ergodic (N_t does not probabilistically tend to return to any given abundance level). The value $b = 0$ is at the edge of the set of values ($b < 0$) for which the stochastic process N_t is ergodic. Without ergodicity, the theorems of mathematical statistics that give the chi-square approximation for G_{12}^2 do not apply. Simulations (not reported here) indicated that the use of the chi-square approximation produces inflated Type I error rates. As can be seen from Eq. 9, the situation is akin to testing whether $\beta = 1$ in an AR(1) process, a well-known case in which the chi-square approximation for G_{ij}^2 fails (Dickey and Fuller 1981).

Instead, the distribution of Λ_{12} (or T_{12}, or $G_{12}{}^2$) can be estimated from the data through parametric bootstrapping. The critical percentile, c, is an unknown function of the two unknown parameters, a and σ^2, in the null hypothesis (Model 1). If Model 1 did indeed give rise to the data, then the ML estimates (Table 1) are in principle quite good estimates of a and σ^2. In fact, the ML estimate $\hat{a}_1$ and the bias-corrected estimate of σ^2 given by

$$\tilde{\sigma}_2{}^2 = q\hat{\sigma}_1{}^2/(q - 1) \tag{22}$$

are the uniformly minimum variance unbiased estimates. We would expect therefore that the time series model given by

$$X_t = X_{t-1} + \hat{a}_1 + \tilde{\sigma}_1 Z_t, \tag{23}$$

where $Z_1, Z_2, \ldots$ are independent normal(0,1) random variables, represents a reasonably good estimate of the mechanism that produced the data under the null hypothesis. We have found in our simulations a slight but detectable advantage to using the unbiased estimate, $\tilde{\sigma}_1{}^2$, in place of $\hat{\sigma}_1{}^2$ (see *Test validation*).

The bootstrap idea is straightforward. Generate data sets repeatedly from the *estimated* null hypothesis model (Eq. 23). For each of these "bootstrap" data sets, fit Models 1 and 2 and calculate an LR statistic (Λ_{12}, T_{12}, or $G_{12}{}^2$). The resulting 2000 or so LR statistic values constitute a random sample from the *estimated* distribution of the test statistic under the null hypothesis. The appropriate sample percentile of those values becomes the estimated critical value for the test. We use the term "parametric bootstrap test" instead of "Monte Carlo test" to emphasize the fact that the model under H_1 is being estimated (Beran 1986, Efron 1986). The terminology is common in the statistics literature (e.g., Schork 1992).

The parametric bootstrap likelihood ratio (PBLR) test is quite simple to conduct using the following steps. The two-sided test of H_1: $b = 0$ vs. H_2: $b \neq 0$ is described first. (1) Obtain ML estimates for all parameters in Models 1 and 2 using the expressions in Table 1. (2) Calculate $T_{12}{}^2$ (or $G_{12}{}^2$, or Λ_{12}) as in Eq. 21. (3) Generate 2000 or more data sets in the form $x_0{}^*, x_1{}^*, \ldots, x_q{}^*$ from the estimated null model (Eq. 23). Each of these bootstrap data sets should start at $x_0{}^* = x_0$ and be the same length as the original set. (4) Calculate for each bootstrap data set the parameter estimates for Models 1 and 2 (Table 1), obtaining $\hat{a}_1{}^*, \hat{\sigma}_1{}^{2*}, \hat{a}_2{}^*, \hat{b}_2{}^*, \hat{\sigma}_2{}^{2*}$. (5) Obtain in this fashion 2000 (or so) bootstrap values of the LR test statistic, $T_{12}{}^{2*}$ (Eq. 21; or $\Lambda_{12}{}^*$ or $G_{12}{}^{2*}$). Each of these $T_{12}{}^{2*}$ values represents an independent observation from the *estimated* distribution of $T_{12}{}^2$ (likewise Λ_{12}, $G_{12}{}^2$). (6) Take the $100(1 - \alpha)^{th}$ sample percentile, $\hat{f}_\alpha$, as the estimate of the critical percentile of the distribution of $T_{12}{}^2$. (7) Reject H_1: $b = 0$ in favor of H_2: $b \neq 0$ if the original value of $T_{12}{}^2$ is greater than $\hat{f}_\alpha$.

Alternatively, at step 6 one can estimate a P value

for the test with the proportion of $T_{12}{}^{2*}$ values that are greater than or equal to $T_{12}{}^2$. The null hypothesis would be rejected if $\hat{P} \leq \alpha$, where $\hat{P}$ is the estimated P value.

Note that in each bootstrap cycle of the calculations, only the value of $T_{12}{}^{2*}$, the original data, the original ML estimates, and the original test statistic need to be retained; the values $\hat{a}_1{}^*, \hat{b}_1{}^*, \hat{a}_2{}^*, \hat{b}_2{}^*, \hat{\sigma}_2{}^{2*}$, and $x_0{}^*$, $x_1{}^*, \ldots, x_q{}^*$ do not need to be stored. The procedure is easily programmed and is acceptably fast: our program, written in the GAUSS matrix language (Aptech Systems 1991), completes the calculations, using 8000 bootstrap samples, for one moderately sized ($q \approx 16$) data set in ≈ 11 min on an old IBM AT/286. (A short SAS program written by the authors for conducting the PBLR test takes ≈ 60 min to run on a 286 machine.)

The one-sided PBLR test of H_1: $b = 0$ vs. H_2: $b < 0$ is straightforward. In step 1 above, reject H_2 outright if $\hat{b}_2 \geq 0$; otherwise continue. In step 5 above, calculate $T_{12}{}^*$ using Eq. 21 for each of the bootstrap samples, instead of $T_{12}{}^{2*}$. In step 6, estimate the critical percentile of the distribution of T_{12} with the $100\alpha^{th}$ sample percentile, $\hat{f}_{1-\alpha}$, of the $T_{12}{}^*$ values. Alternatively, estimate a P value by the proportion of $T_{12}{}^*$ values that are less than T_{12}. In step 7, reject H_1 if $T_{12} \leq \hat{f}_{1-\alpha}$, or if $\hat{P} < \alpha$.

We point out that the distribution of T_{12} is not symmetric, nor is it centered at zero. The two-sided test conducted with the $100(\alpha/2)^{th}$ and $100[1 - (\alpha/2)]^{th}$ sample percentiles of the $T_{12}{}^*$ values is different from the previously described two-sided test that uses the $100(1 - \alpha)^{th}$ percentile of the $T_{12}{}^{2*}$ values. The power properties of both tests have not been compared.

A PBLR test of Model 0 against Model 2 might be of interest in some studies. Other methods for distinguishing a drift-free random walk from a density-dependent process (e.g., den Boer 1991) are likely not as powerful. The one-sided test of H_0: $a = 0$, $b = 0$ vs. H_2: $a \neq 0$, $b < 0$ would be conducted as follows. (1) Obtain ML estimates for all parameters in Models 0 and 2 using Table 1. Reject H_2 outright if $\hat{b}_2 \geq 0$; otherwise continue. (2) Calculate the LR statistic as

$$\begin{aligned}\Lambda_{02} &= L_0(\hat{\sigma}_0{}^2)/L_2(\hat{a}_2, \hat{b}_2, \hat{\sigma}_2{}^2) \\ &= (\hat{\sigma}_0{}^2/\hat{\sigma}_2{}^2)^{-q/2}. \end{aligned} \tag{24}$$

(3) Generate bootstrap data sets in the form $x_0{}^*, x_1{}^*, \ldots, x_q{}^*$ from the estimated random walk model,

$$X_t = X_{t-1} + \hat{\sigma}_0 Z_t, \tag{25}$$

starting at $x_0{}^* = x_0$. (4) For each bootstrap data set, calculate the parameter estimates for Models 0 and 2: $\hat{\sigma}_0{}^{2*}, \hat{a}_2{}^*, \hat{b}_2{}^*, \hat{\sigma}_2{}^{2*}$. From these parameter estimates, calculate a value of the LR statistic as

$$\Lambda_{02}{}^* = \begin{cases} 1, & \hat{b}_2{}^* \geq 0; \\ (\hat{\sigma}_0{}^{2*}/\hat{\sigma}_2{}^{2*})^{-q/2}, & \hat{b}_2{}^* < 0. \end{cases} \tag{26}$$

(5) Obtain in this fashion 2000 (or so) values of $\Lambda_{02}{}^*$.

The provision of setting $\Lambda_{02}{}^*$ equal to 1 when $\hat{b}_2{}^* \geq 0$ will result in a one-sided test. (6) Take the $100\alpha^{\text{th}}$ percentile, $\hat{\lambda}_\alpha$, as the estimate of the critical percentile of Λ_{02}. Alternatively, estimate a P value by the proportion of $\Lambda_{02}{}^*$ values that are less than Λ_{02}. (7) Reject H_0 if $\Lambda_{02} \leq \hat{\lambda}_\alpha$, or if $\hat{P} < \alpha$.

Of course, Model 1 contains Model 0 as a special case. The previously described test of Model 1 against Model 2 thus implicitly includes Model 0 in the null hypothesis. Normally, the test of Model 1 against Model 2 should be used, unless there are specific reasons for restricting the null hypothesis to a pure random walk.

The PBLR tests can accommodate missing data. The fundamental "observation" in the likelihood function (Eq. 15) is not a (logarithmic) population size, x_t, but a *transition* from x_{t-1} to x_t. If x_j is missing from the data, it means that two transitions are missing: x_{j-1} to x_j, and x_j to x_{j+1}. The likelihood function (Eq. 15) with $p(x_j | x_{j-1})$ and $p(x_{j+1} | x_j)$ omitted from the product is the joint pdf of $X_1, X_2, \ldots, X_{j-1}$ given x_0 and X_{j+2}, $X_{j+3}, \ldots, X_q$ given x_{j+1}. This likelihood is used as the fundamental building block for parameter estimates and hypothesis tests.

When observations are missing, the ML parameter estimates for Model 2 are easiest to calculate with a least squares approach. The formulae (Table 1) otherwise must be altered in messy ways. Take as the fundamental data set the logarithmic differences $y_1, y_2,$ $\ldots, y_{j-1}, y_{j+2}, y_{j+3}, \ldots, y_q$ (y_j, y_{j+1} missing). The sample mean of the $y_t{}^2$ values provides the ML estimate of σ^2 for Model 0. The sample mean of the differences, $\bar{y}$, and the sample mean of the $(y_t - \bar{y})^2$ values provide respectively the ML estimates of a and σ^2 for Model 1. The linear regression of y_t on n_{t-1} (using all transitions present in the data) provides ML estimates of a and b for Model 2. The ML estimate for σ^2 in Model 2 would be the sum of squared errors (from the regression) divided by the number of transitions present ($q - 2$, if just one population size is missing). The bootstrap data sets are generated (Eq. 23 or 25) as a series of one-step transitions starting at x_0 ($x_0, X_1{}^*,$ $X_2{}^*, \ldots, X_{j-1}{}^*$), followed by a series starting at x_{j+1} ($x_{j+1}, X_{j+2}{}^*, \ldots, X_q{}^*$). All parameter estimates and tests are thus conditioned on the observed starting values, x_0 and x_{j+1}, of series of one-step transitions.

An alternative approach to testing with missing data is to condition only on x_0. Parameter estimates and the test statistic are computed as described above using all one-step transitions present in the data. However, bootstrap data sets are generated starting at x_0 for all times ($x_0, X_1{}^*, X_2{}^*, \ldots, X_q{}^*$), including missing times. The bootstrap values of the test statistic are then calculated after omitting from the bootstrap data sets the generated observations occurring at missing times (that is, omit $X_j{}^*$ before calculating $T_{12}{}^*$).

The two approaches to handling missing data are subtly different. The first treats the uninterrupted time series essentially as separate series, but assumes the series are governed by the same (density-independent or -dependent) model. The first approach would be preferred if, for instance, the population was restarted at size x_{j+1} after some drastic change (a harvest or catastrophe). The second approach treats the uninterrupted series as one single series with some observations (the missing ones) simply unknown. Which approach is most appropriate will be case specific, although the resulting tests are not likely to differ much unless the number of missing transitions is large. The statistical properties of the two approaches have not yet been compared.

Because the PBLR test is a parametric test, some additional model checking is advised in any application. Judicious use of model diagnostic techniques will help minimize problems associated with "Type III error" (fitting the wrong model). In particular, serious departures of the data from the Markov property or from the model could likely be detected through some form of residual analysis. For the stochastic logistic model, diagnostic techniques would center around the conditional residuals: $\hat{e}_t = x_t - x_{t-1} - \hat{a} - \hat{b}\exp(x_{t-1})$, $t = 1, 2, \ldots, q$. Under the model assumptions, these residuals should be approximately normal white noise. The residuals can be subjected to the customary analysis techniques of linear time series modeling (see Tong 1990:322 for discussion). We often use the Lin-Mudholkar test for normality against asymmetric alternatives (Lin and Mudholkar 1980, Tong 1990:324) in addition to the standard normal plots and autocorrelation tests. There is a caveat, however: in the nonlinear setting, the adequacy of the normal-white noise approximation is unknown and varies from model to model. The properties of these diagnostic techniques for the stochastic logistic model would be a worthwhile topic for future study.

According to statistical principles of asymptotic relative efficiency, the PBLR tests represent approximate size α tests having power functions that cannot be exceeded by much. The principles rest on large-sample theorems of statistics (Serfling 1980). But do these asymptotic assurances apply to the data sets of moderate lengths likely to be encountered in ecological practice? A large-scale, Monte Carlo power assessment of the PBLR test of Model 1 against Model 2, along with comparative studies of other available tests, provides some answers.

TEST VALIDATION

It is desirable that statistical tests have several properties. First, their size should be close to their nominal size, that is, if the null hypothesis is true the probability of rejecting the null hypothesis should be close to what the investigator thinks it is. Second, the test should be powerful enough to detect scientifically interesting deviations from the null hypothesis. Third, the test should be robust to measurement error. This last property is
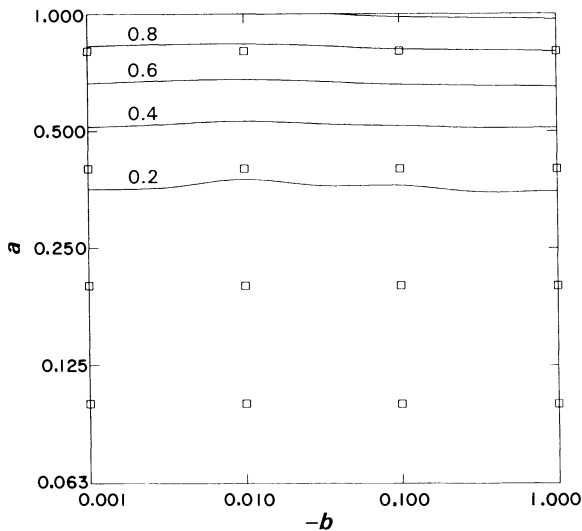
FIG. 1. A contour plot showing estimated power of the parametric bootstrap likelihood ratio test of density dependence as a function of model parameters a and b. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis is Model 2 ($b < 0$). Each □ indicates location of 1000 simulated tests; each test used 200 bootstrap samples and the values $q = 19$, $n_0 = -a/b$, and $\sigma = 0.2$. The power surface was estimated using a distance weighted least squares algorithm (McLain 1974).

particularly important when dealing with population abundance data, which commonly contain substantial uncertainty.

We have investigated the qualities of the PBLR test of Model 1 against Model 2 (both one- and two-sided) with Monte Carlo simulation. Using a known set of parameters we generate a time series of simulated population densities according to Eq. 1. This simulated time series is then subjected to the PBLR test exactly as if it were data from field observations. For each set of parameters (a, b, σ, q, and n_0: we have couched the simulations in terms of σ rather than σ^2) chosen for study, this process was repeated a large number of times, usually 1000. All tests were conducted at a nominal 0.05 level. The proportion of times the null hypothesis was rejected was recorded for each parameter set. If the parameter b was zero, that is if the null hypothesis of no density dependence was true, this proportion represents an estimate of the size of the test. If b was not zero then the proportion of rejections is an estimate of the power of the test under the set of parameters. We denote by $\phi(a, b, \sigma, q, n_0)$ the probability of rejection of the null hypothesis as a function of model parameters (power function), and by $\hat{\phi}(a, b, \sigma, q, n_0)$ its estimate from simulations.

Test size

We examined the size of the PBLR test (one- and two-sided) under a broad range of parameters. The parameters a and σ ranged from 0.05 to 1.6, while q

ranged from 4 to 64. In all, 64 different sets of parameters were tested each with two separate simulations of 1000 trials. The sample mean of these 128 values of $\hat{\phi}$ for the one-sided tests was 0.0504 with a sample variance of 0.0000470. The underlying population mean of the $\hat{\phi}$ values is thus not significantly different from 0.05 ($Z = 0.66$, $N = 128$, $P = .25$). The sample variance was close to the variance expected under binomial sampling, $0.0000475 (= 0.05(0.95)/1000)$. Further, despite the wide variety of parameters used, the distribution of $\hat{\phi}$ values observed was not significantly different from a normal distribution with a mean and variance of 0.05 and 0.0000475, according to a Kolmogorov–Smirnov test ($D = 0.0611$, $N = 128$, $P > .5$). Results for the two-sided tests (based on $T_{12}{}^2$) were similar. Thus there is no reason to believe that the true size of the PBLR test is different from its nominal size. If any deviations do exist, they are of insignificant magnitude.

The size results reported above were obtained for the PBLR test that uses the unbiased estimate, $\tilde{\sigma}_1{}^2$, instead of the ML estimate, $\hat{\sigma}_1{}^2$, in the estimated null hypothesis model (Eq. 23). We detected through simulations a slight but noticeable increase in the test size over the nominal size of 0.05 when the ML estimate is used. While the increase is small enough to be of little practical importance, it is easily corrected simply by using $\tilde{\sigma}_1{}^2$.

Test power

As with the size of the test, we investigated power extensively with simulations. The power of a test depends in general on the specific true values of the parameters. We have assessed the influence of b, a, σ, q, and n_0 on power. A number of our results contradict unreflective intuition.

First, the probability, $\phi(a, b, \sigma, q, n_0)$, of rejecting the null hypothesis is nearly independent of the parameter b as long as b is not zero (Fig. 1). Thus, the influence of b on power is not continuous; instead, b acts as a switch to change the qualitative behavior of the model. This discontinuity in the power function is, from the standpoint of statistical theory, unusual (e.g., Bain and Engelhardt 1987:373). One would normally expect the power function to increase smoothly, starting from a level of α, as the parameter in question becomes farther from its hypothesized null value.

However, such smooth textbook dependence of power on b would in fact be an undesirable property. The parameter b is related to the level around which N_t is fluctuating according to the density-dependent model (Eq. 1). While the concept of point equilibrium (carrying capacity) is of questionable meaning in a stochastic model (Dennis and Costantino 1988, Wolda 1989), we can see that the level $-a/b$ (Eq. 5) represents a center for the return tendencies of N_t. If $N_t > -a/b$, then $\ln N_{t+1}$ is expected to decrease (Eq. 5), while if $N_t < -a/b$, $\ln N_{t+1}$ is expected to increase. One presum-
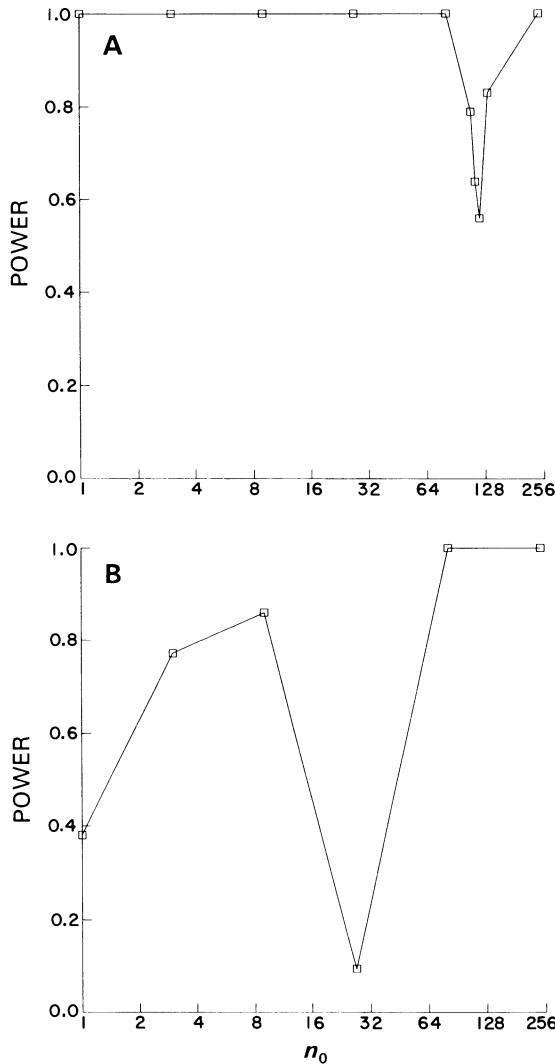
FIG. 2. Estimated power of the parametric bootstrap likelihood ratio test of density dependence as a function of initial population density, n_0. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis is Model 2 ($b < 0$). Each square represents 1000 simulated tests; each test used 200 bootstrap samples and the values $q = 9$, $b = -0.01$, $\sigma = 0.05$. (a) $a = 1.2$. (b) $a = 0.3$.

ably would not want the power of a density dependence test to depend on whether the population was fluctuating around 1000 or 100, but rather only on whether the population was fluctuating around (that is, was showing some tendency to return to) some unspecified level. It is thus entirely reasonable that the numerical value of b does not affect the power of the test under Model 2, because b is simply a reflection of the units in which population size is measured.

The interaction of initial population size, n_0, and power is also interesting. Power is quite low when n_0 is near $-a/b$. As n_0 deviates from this value, power increases (Fig. 2a). What distinguishes the density dependent model from the null model is the tendency for

the population density to move toward a central value when displaced from it. It makes sense that a deviation in the initial population size would increase the test power. However, if the parameter a is low, then power decreases again if n_0 is too far below the value of $-a/b$ (Fig. 2b). With a low a and a low initial size, the population will tend to increase for a number of time steps, making it difficult to distinguish the time series from one that would be produced by a population undergoing exponential growth.

We turn now to the effect of environmental variation in growth rate on the power of the PBLR test. The parameter that measures environmental variation is σ, the standard deviation of $Y_t = \ln (N_t/N_{t-1})$ conditional on $N_{t-1} = n_{t-1}$. (The standard deviation of Y_t conditional on $N_0 = n_0$ is an unknown, increasing function of σ that depends on t as well.) Recall that σ is estimated by the root-mean-squared error in a linear regression of the y_t values on the n_{t-1} values (Table 1). The influence of σ on power is as intriguing as the influence of n_0. As σ increases so does the power of the PBLR test, although the effect is minimal until σ is around the magnitude of a (Fig. 3). This increase in power is counterintuitive to investigators accustomed to thinking about "error" in standard regressions. However, the above discussion of n_0 and power resolves the apparent contradiction. In a nutshell, the test works by detecting return toward an abundance level from deviations away from that level. Stochastic fluctuations provide some of these deviations.

As would be expected, increasing the length of the time series, $q + 1$, increases power. Fig. 4 shows $\hat{\phi}$ as

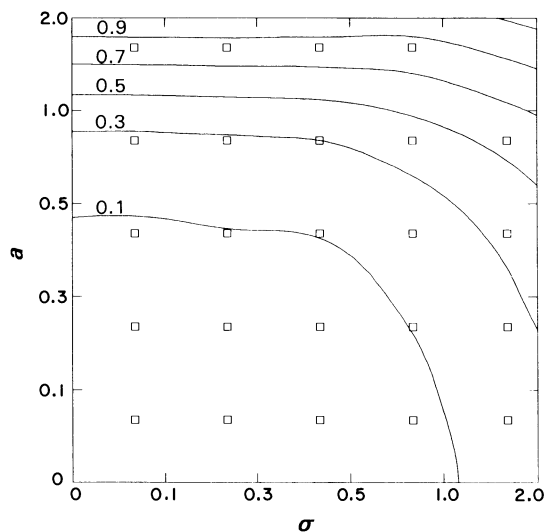FIG. 3. Estimated power of the parametric bootstrap likelihood ratio test of density dependence as a function of model parameters a and σ. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis is Model 2 ($b < 0$). Each square represents 1000 simulated tests; each test used 200 bootstrap samples and the values $q = 9$, $b = -0.01$, $n_0 = -a/b$. Contours were drawn as in Fig. 1.
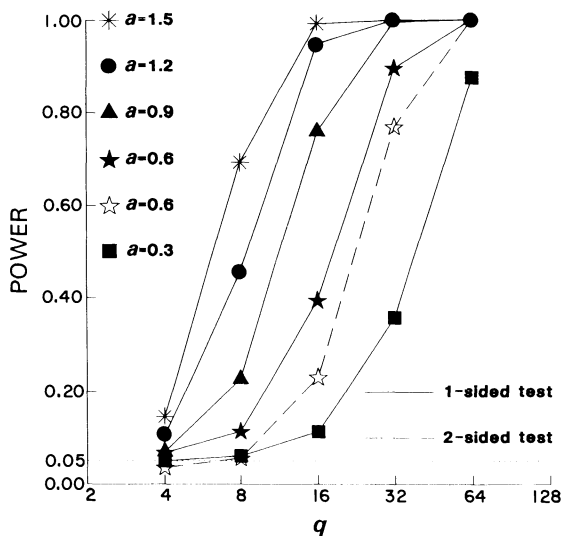
FIG. 4. Estimated power of the parametric bootstrap likelihood ratio test of density dependence as a function of sample size, q (number of one-step transitions in the time series). The null hypothesis for the test is Model 1 ($b = 0$). Each symbol represents 1000 simulated tests; each test used 200 bootstrap samples and the values $b = -0.01$, $\sigma = 0.05$, $n_0 = -a/b$.

a function of q for a spectrum of a values ranging from 0.3 to 1.2. All simulations in this figure had initial sizes fixed at $n_0 = -a/b$, and thus power is at a local minimum. Nonetheless, power becomes quite reasonable by the time q is 16. Even time series as short as eight transitions can have a nontrivial probability of rejecting the false null hypothesis if the rate parameter a is not small. In the figure, the two curves marked with stars portray power for a series of runs with a equal to 0.6. The solid line is for one-sided tests and the dashed line represents two-sided tests. The difference between the two types of tests in the probability of correctly rejecting the null hypothesis may be important, particularly when power is modest. In the figure when q is 16, the power of the one-sided test is almost twice the power of the two-sided test. Many of the abundance records for natural populations are only 10–30 yr long. Power is expected to only be moderate. Thus we strongly recommend the use of the one-sided test.

Knowledge of a test's power is extremely useful when interpreting results. If power is low then failure to reject the null hypothesis contributes only weak evidence in favor of the null model. True power can only be quantified if the real parameters are known. For the density dependence test, the power can be estimated in a statistically consistent fashion by substituting the empirically estimated parameters for the true parameters and conducting a Monte Carlo simulation such as described above. However, we suggest some caution in that such power estimates will only approximate the true power in small samples. The maximum likelihood parameter estimates under the density dependent model have a finite sample bias. While this has no influence on the

size of the PBLR test for significance testing, it will affect power estimates.

Measurement error

The sensitivity of regression analysis to measurement errors in the predictor variables is well known (Fuller 1987). We were particularly concerned about the impact of measurement error on the PBLR test because the "predictor" variable in the regression is population density. Populations are frequently estimated rather than censused, and in some cases the time series consist of relative population indices (light trap counts, redd counts, etc.). Sampling error could potentially be an important source of variability in the time series in these situations. What happens if the PBLR test is applied directly to such time series data?

We have investigated the consequences of sampling error with simulations similar to those discussed above in estimating size and power. As above, we generated a time series of simulated population densities according to Eq. 1 using a known set of parameters. This simulated time series was then contaminated with a noise variable representing measurement error and subjected to the PBLR test exactly as if it were data from field observations. Since many population density estimates or relative indices are based in some fashion on binomial or Poisson sampling, the variance of the noise contaminating N_t was itself made proportional to N_t (the observations entering the data set had zero-truncated normal(N_t, cN_t) distributions). The results of this study are summarized in Table 2. Remarkably, the size of the test is hardly influenced even by massive amounts of sampling error. Even more remarkably, power is somewhat increased by this common type of sampling error. We have noted earlier that

TABLE 2. Estimated power of the parametric bootstrap likelihood ratio test of density dependence in the presence of sampling or measurement error. Each observation, n_t, simulated from the stochastic logistic model was replaced by an observation generated from a zero-truncated normal(n_t, cn_t) distribution before testing. Power estimates are based on 1000 trials. In each trial, $q = 9$, $a = 0.5$, $\sigma = 0.1$, and $n_0 = 50$.

b	$c^{1/2}$	Rejection probability
0	0.00	0.051
0	0.05	0.055
0	0.10	0.038
0	0.20	0.052
0	0.30	0.046
0	0.40	0.040
0	0.50	0.037
−0.01	0.00	0.102
−0.01	0.05	0.098
−0.01	0.10	0.099
−0.01	0.20	0.129
−0.01	0.30	0.137
−0.01	0.40	0.147
−0.01	0.50	0.180

TABLE 3. Parametric bootstrap likelihood ratio test of density dependence for the grizzly bear population in the Yellowstone region. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis is Model 2 ($b < 0$). Number of bootstrap samples = 8000.

Population size data, N_t*	33	36	34	39	35	34	38	36	37	41	39	51
	47	57	48	59	64							

Maximum likelihood parameter estimates under Model 1 (density independence)	$\hat{a}_1 = 4.14 \times 10^{-2}$; $\hat{\sigma}_1^2 = 1.49 \times 10^{-2}$
Maximum likelihood parameter estimates under Model 2 (density dependence)	$\hat{a}_2 = 1.41 \times 10^{-1}$; $\hat{b}_2 = -2.41 \times 10^{-3}$; $\hat{\sigma}_2^2 = 1.45 \times 10^{-2}$
Likelihood ratio statistics§	$T_{12} = -0.60$; $t_{0.95} = -2.9$; $P = 0.72$

* Original data consist of the yearly counts of adult females seen with cubs, from 1973–1991 (Eberhardt et al. 1986; R. R. Knight, *personal communication*). Values listed here for N_t are calculated from the original data as a 3-yr moving sum (sum of 1973, 1974, and 1975 counts, sum of 1974, 1975, and 1976 counts, etc.).

§ Likelihood ratio test statistic (Eq. 21), estimated fifth percentile of the test statistic distribution under Model 1, and estimated P value for the test.

deviations increase the power of this test. Apparently, these deviations do not even have to be entirely real.

EXAMPLES

We present, in this section, several worked examples of the PBLR test of density dependence. Numerous examples of density dependence testing in the literature have involved insect populations. To the scarce supply of large mammal examples in the density dependence debate, we add a few more (Tables 3–5). Fowler (1984, 1987) has given additional information and insights about density dependence in large mammals. In this section we also analyze 16 insect data sets assembled by den Boer and Reddingius (1989) in order to compare results of the PBLR test to earlier published results.

The grizzly bear (*Ursus arctos horribilis*) population of the greater Yellowstone ecosystem shows no evidence of density dependence in time series abundance data (Table 3). The data (Table 3) consist of a 3-yr running sum of adult females seen with cubs. An adult female produces cubs on average every 3 yr, so the 3-yr running sum of this relatively visible component of the population represents an estimate of the minimum number of adult females in the population (see Knight and Eberhardt 1984, 1985, Eberhardt et al. 1986, and Dennis et al. 1991 for discussion). Table 3 reflects the counts from 1973 through 1991 (Eberhardt et al. 1986, R. R. Knight, *personal communication*). Dennis et al. (1991) found increased variability in the data after 1971, possibly related to the garbage dump closures in 1970–1971 or to institution of new aerial survey methods. The outcome of the density dependence test suggests that the population has not yet reached carrying capacity, that is, the time series gives no reason to favor Model 2 over Model 1.

Results of model diagnostic procedures for the grizzly data are mixed. The residuals from both models are normally distributed, according to the Lin-Mudholkar test (Model 1: LM = -0.28, $P = .78$; Model 2: LM = -0.78, $P = -.43$; see Tong 1990:324). However, the residuals from both models have some autocorrelation, according to standard tests with the first- and second-order sample autocorrelation statistics

(Model 1: $\sqrt{q}\hat{\rho}_1 = -2.73$, $P = .0064$; $\sqrt{q}\hat{\rho}_2 = 1.77$, $P = .077$; Model 2: $\sqrt{q}\hat{\rho}_1 = -2.21$, $P = .027$; $\sqrt{q}\hat{\rho}_2 = 1.71$, $P = .088$; see Tong 1990:324). While the properties of these or other white noise tests have not been investigated for the residuals of Model 2, the results suggest that the grizzly female population has a higher order autocorrelation structure not accounted for by either Model 1 or Model 2. Oscillations from year-class imbalances in the population could cause such autocorrelation. The large variability of the population, though, gives reason for concern about its long-term viability (Dennis et al. 1991).

Two elk (*Cervus elaphus*) populations in the greater Yellowstone ecosystem have noticeable density dependence (Tables 4 and 5). The data on the northern Yellowstone population (Table 4) are winter census records from Houston (1982:17); the data on the central valley population in Grand Teton National Park (Table 5) are from Boyce (1989) and represent summer mark–recapture estimates.

The northern Yellowstone population increased rapidly after artificial removals from the park were ended in 1969. The population appears to have subsequently attained a stochastic equilibrium. The power of the density dependence test was enhanced because the initial population was far from equilibrium (see *Test validation*). Residual plots and tests show no outliers, no significant first- or second-order autocorrelation, and no significant departures from normality (Model 1: LM = 0.91, $P = .36$; $\sqrt{q}\hat{\rho}_1 = 0.89$, $P = .37$; $\sqrt{q}\hat{\rho}_2 = 1.26$, $P = .26$; Model 2: LM = -1.08, $P = .28$; $\sqrt{q}\hat{\rho}_1 = -1.53$, $P = .13$; $\sqrt{q}\hat{\rho}_2 = 0.21$, $P = .83$; see Tong 1990:324).

The central valley population in Grand Teton National Park fluctuates substantially, and the estimated P value for the test is just under .05. The Grand Teton population has a missing observation in year 1983, and so the test was conditioned on n_{21} (= 1453) in addition to n_0 (= 1627) (see *Hypothesis testing*). The second transition (1527 to 824) is a possible outlier for both Model 1 and Model 2, with a standardized residual of ≈ 2.7 for Model 2. No significant first- or second-order autocorrelation is evident in the first 19 consecutive

TABLE 4. Parametric bootstrap likelihood ratio test of density dependence for the northern Yellowstone elk population. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis is Model 2 ($b < 0$). Number of bootstrap samples = 8000.

Population size data, N_t*	3172	4305	5543	7281	8215	9981	10 529
	12 607	10 807	10 741	11 855	10 768		
Maximum likelihood parameter estimates under Model 1 (density independence)	$\hat{a}_1 = 1.11 \times 10^{-1}$; $\hat{\sigma}_1^2 = 2.06 \times 10^{-2}$						
Maximum likelihood parameter estimates under Model 2 (density dependence)	$\hat{a}_2 = 4.68 \times 10^{-1}$; $\hat{b}_2 = -4.14 \times 10^{-5}$; $\hat{\sigma}_2^2 = 4.89 \times 10^{-3}$						
Likelihood ratio statistics§	$T_{12} = -5.37$; $t_{0.95} = -2.8$; $P = 0.0025$						

  * Data are winter census records from 1968–1979 listed by Houston (1982). The 1977 value is an adjusted value given by Houston (1982:23).
  § Likelihood ratio test statistic (Eq. 21), estimated fifth percentile of the test statistic distribution under Model 1, and estimated P value for the test.

residuals (Model 1: $\sqrt{19}\hat{\rho}_1 = -0.69$, $P = .49$; $\sqrt{19}\hat{\rho}_2 = -0.95$, $P = .34$; Model 2: $\sqrt{19}\hat{\rho}_1 = 0.48$, $P = .63$; $\sqrt{19}\hat{\rho}_2 = -0.32$, $P = .75$). With the outlier removed, the residuals are acceptably normal (Model 1: LM = 1.95, $P = .051$; Model 2: LM = 1.91, $P = .056$).

den Boer and Reddingius (1989) used the Pollard et al. (1987) randomization test to look for density dependence in 16 insect populations. Their paper provides a table with the original data. The randomization test did not flag a single population as density dependent. By contrast, the PBLR test rejects density independence for two of the populations at the .05 significance level (Table 6). If the regression test for $b < 0$ based on the Student's t distribution is used, the density dependent count jumps to eight (Table 6).

The den Boer and Reddingius (1989) data illustrate the role of test power. The regression test is obviously more powerful but is inappropriate because it is not a size 0.05 test (see *Discussion*). Both the randomization and the PBLR are close to size 0.05 tests. Because of the asymptotic relative efficiency of LR tests, the PBLR test probably represents the practical limit of power for testing $b < 0$ in the stochastic logistic model. Even though the power can exceed that of the randomization test by 50% (see *Discussion*), the basic thrust of den Boer and Reddingius' results remains intact. If there is density dependence in these 16 populations, it is difficult to detect from time series data alone.

Could the two density dependent cases have oc-

curred simply by chance? If one assumes that all 16 populations follow the null hypothesis model (stochastic exponential growth), then two or more "successes" out of 16 trials is a plausible outcome when the success probability is .05 (if $W \sim$ binomial (16, 0.05), then $P[W \geq 2] \approx .19$). However, there is more information present in the collection of P values than simply the number of them $< .05$. If all the populations were realizations of the null hypothesis model, the P values would represent 16 independent observations from a uniform(0, 1) distribution. Then $-2 \ln P_i$ would be an observation from a chi-square(2) distribution, and the sum of k such values would be an observation from a chi-square($2k$) distribution (Fisher's test; see Fisher 1958). From the P values in Table 6, we find that $\Sigma - 2 \ln P_i = 46.1$, a value that is just below the 95th percentile of the chi-square(32) distribution ($P' \approx .051$). Enough of the P values in this meta-analysis are "leaning" toward the alternate hypothesis end so as to cast doubt upon the assumption that all 16 populations are realizations of the density-independent model, though one would not reject that assumption at a strict .05 significance level.

DISCUSSION

The PBLR test can easily be adapted, if desired, to testing for density dependence under the Gompertz model (Eqs. 7 and 8). The modification would use x_t values in place of n_t values in the parameter estimates

TABLE 5. Parametric bootstrap likelihood ratio test of density dependence for the elk population in the central valley of Grand Teton National Park. The null hypothesis for the test is Model 1 ($b = 0$), and the alternate hypothesis for the test is Model 2 ($b < 0$). Number of bootstrap samples = 8000.

Population size data, N_t*	1627	1527	824	891	1140	1322	1431	1733	1131
	1611	1644	1991	1762	1076	1442	1800	1667	1558
	1396	1753	$\cdots$	1453	1804				
Maximum likelihood parameter estimates under Model 1 (density independence)	$\hat{a}_1 = 1.45 \times 10^{-2}$; $\hat{\sigma}_1^2 = 6.85 \times 10^{-2}$								
Maximum likelihood parameter estimates under Model 2 (density dependence)	$\hat{a}_2 = 7.31 \times 10^{-1}$; $\hat{b}_2 = -4.93 \times 10^{-4}$; $\hat{\sigma}_2^2 = 4.65 \times 10^{-2}$								
Likelihood ratio statistics§	$T_{12} = -2.92$; $t_{0.95} = -2.88$; $P = 0.044$								

  * Data are summer mark–recapture population estimates from 1963–1985 (1983 missing) listed by Boyce (1989).
  § Likelihood ratio test statistic (Eq. 21), estimated fifth percentile of the test statistic distribution under Model 1, and estimated P value for the test.

(Table 1) for the alternate hypothesis model. The resulting test is essentially the modification of Vickery and Nudd's (1984) "simulation" test that was suggested by Pollard et al. (1987). The test procedure is otherwise identical to the test we have proposed. Mountford (1988) employed such a Gompertz-based PBLR test using Model 0 (random walk) as the null hypothesis.

Reddingius (1971) recognized the theoretical desirability of LR tests, and he developed a small table of critical values for the Gompertz-based LR statistic using Monte Carlo simulation. In practice, the investigator must estimate parameters from the data before referring to the table. Use of Reddingius' table thus is quite similar to conducting a PBLR test for Gompertz-type density dependence.

Many subsequent investigators have taken (implicitly or explicitly) the Gompertz model as the alternate hypothesis in density dependence tests, but have used different test statistics (Bulmer 1975, Royama 1977, Slade 1977, Vickery and Nudds 1984, 1991, Gaston and Lawton 1987, Pollard et al. 1987, den Boer and Reddingius 1989, Reddingius and den Boer 1989, den Boer 1990, Holyoak and Lawton 1992, Woiwod and Hanski 1992). Crowley (1992) modified the Gompertz to include sampling variability. The statistics include the sample correlation of the y_t's and x_{t-1}'s (Pollard et al. 1987), the slopes of principal and reduced major axes (Slade 1977), the reciprocal of von Neumann's ratio (Bulmer 1975), and the number of times that the one-step transitions have moved toward (or away) from a given abundance level (Crowley 1992).

Of the tests studied by these investigators, the randomization test of Pollard et al. (1987) based on the sample correlation coefficient appears the most powerful (Vickery and Nudds 1991, Crowley 1992). This is not surprising; the LR statistic for testing whether $b = 0$ in the Gompertz model (or in the logistic model) is a monotone function of the squared sample correlation coefficient:

$$\Lambda_{12} = [1 - R^2]^{q/2}. \qquad (27)$$

Here R is the sample correlation coefficient of the y_t's and x_{t-1}'s (or n_{t-1}'s if the test is adapted to the logistic model given by Eq. 1). The randomization procedure proposed by Pollard et al. (1987) estimates the distribution of R or R^2 by taking random permutations of the y_t's to construct new time series data sets. A new value of R^2 is obtained from each set. This is essentially a form of nonparametric bootstrapping. However, the resulting estimate of the distribution of R^2 (or R, or Λ) under the null hypothesis does not make use of the sufficient statistics $\hat{a}$ and $\hat{\sigma}^2$, and therefore does not make the most efficient use of the data ("sufficient statistics" contain all the information about model parameters that is present in the data; see Rice 1988).

Our simulations reveal that the randomization test of Pollard et al. (1987) is considerably less powerful

TABLE 6. Results of two density dependence tests performed on 16 insect data sets listed by den Boer and Reddingius (1989). Shown are the values of the likelihood ratio test statistic (T_{12}), the number of one-step transitions (q), the P values estimated by parametric bootstrapping $(\hat{P})$, and the P values resulting from a Student's t distribution with $q - 2$ degrees of freedom (P_{student}). The null hypothesis is Model 1 $(b = 0)$, and the alternate hypothesis is Model 2 $(b < 0)$. Order of entry corresponds to order in den Boer and Reddingius' (1989) Table 1.

T_{12}	q	$\hat{P}$	P_{student}
−1.84	18	.243	.042
−2.54	18	.089	.011
−3.14	14	.033	.004
−3.07	14	.033	.005
−1.51	13	.436	.080
−2.15	13	.154	.027
−2.52	12	.095	.015
−0.91	28	.776	.185
−1.45	18	.349	.083
−1.29	19	.457	.107
−1.61	14	.378	.067
−2.07	13	.198	.031
+2.17	11	.999	.971
−0.62	11	.759	.275
−1.86	11	.228	.048
−1.66	10	.397	.067

than the PBLR test when the data are generated from the stochastic logistic model given by Eq. 1 (Table 7). The parametric test attains as much as a 50% increase in power over a range of parameter values. Even if the randomization test is adapted to the logistic model, by using n_{t-1}'s instead of x_{t-1}'s in Eq. 27, the power levels of the parametric test are not approached (Table 7).

There might be some nonparametric benefits in terms of robustness of the randomization test, though such benefits have not been assessed. The stochastic logistic (Eq. 1) that we have assumed as the alternate hypothesis is a fairly general model that can describe many situations. Possible modifications of the model for studying robustness of tests might include use of a heavy-tailed distribution for the Z_t's instead of the normal distribution, or allowing the Z_t's to be autocorrelated. Few abundance data sets have more than 30 observations, however, and it is likely that the full benefits of nonparametric approaches (which depend heavily on large-sample consistency theorems) to density dependence testing will not be realized in practice.

Use of the Gompertz stochastic model in and of itself could involve a loss of power. Detecting growth-rate feedback that is proportional to $\ln N_t$ instead of N_t will likely require data covering wider ranges of abundance. As noted before, a PBLR test can be constructed for the Gompertz model; such a test would represent the practical limit of power that can be attained under that model. If this PBLR-Gompertz test is performed on data arising from a stochastic logistic, the size of the test remains ≈ 0.05, but a detectable loss of power ensues (Table 7). The difference in power between the PBLR-logistic and the PBLR-Gompertz quantifies the

TABLE 7. Estimated sizes and powers of various one-sided density dependence tests, when data arise from the stochastic logistic model.* Tests are the parametric bootstrap likelihood ratio (PBLR) tests based on the stochastic logistic and the stochastic Gompertz and randomization tests based on the stochastic logistic and stochastic Gompertz. All tests were conducted at a nominal significance level of .05 and used 200 bootstrap or randomized samples where appropriate. Each estimate was obtained from 1000 trials. For all tests, $q = 9$.

b	n_0	σ	a	PBLR (logistic)	PBLR (Gompertz)	Randomization (logistic)	Randomization (Gompertz)
0	100	0.25	0.5	0.039	0.053	0.041	0.039
0	100	0.25	1.0	0.051	0.039	0.052	0.052
0	100	0.25	1.5	0.045	0.060	0.050	0.051
0	100	0.50	0.5	0.052	0.029	0.038	0.050
0	100	0.50	1.0	0.050	0.048	0.048	0.044
0	100	0.50	1.5	0.046	0.045	0.046	0.039
0	100	0.75	0.5	0.041	0.045	0.035	0.042
0	100	0.75	1.0	0.049	0.045	0.063	0.059
0	100	0.75	1.5	0.056	0.047	0.052	0.044
0	100	1.00	0.5	0.046	0.055	0.040	0.049
0	100	1.00	1.0	0.065	0.050	0.043	0.042
0	100	1.00	1.5	0.054	0.040	0.039	0.052
−0.01	−a/b	0.25	0.5	0.104	0.101	0.086	0.084
−0.01	−a/b	0.25	1.0	0.362	0.354	0.304	0.290
−0.01	−a/b	0.25	1.5	0.759	0.742	0.699	0.692
−0.01	−a/b	0.50	0.5	0.152	0.112	0.104	0.092
−0.01	−a/b	0.50	1.0	0.424	0.399	0.322	0.270
−0.01	−a/b	0.50	1.5	0.807	0.608	0.682	0.665
−0.01	−a/b	0.75	0.5	0.191	0.121	0.106	0.103
−0.01	−a/b	0.75	1.0	0.494	0.269	0.333	0.250
−0.01	−a/b	0.75	1.5	0.775	0.458	0.675	0.421
−0.01	−a/b	1.00	0.5	0.237	0.131	0.107	0.095
−0.01	−a/b	1.00	1.0	0.562	0.220	0.336	0.188
−0.01	−a/b	1.00	1.5	0.788	0.298	0.629	0.308

* The model is $Y_t = a + bN_{t-1} + \sigma Z_t$, where N_t is population abundance at time t, $Y_t = \ln(N_t/N_{t-1})$, $Z_t \sim$ normal(0,1), and $t = 1, 2, \ldots, q$.

intrinsic effect of looking for density dependence with a model of log-density dependence.

Of course, one could just as easily take the Gompertz as the "true" model in the simulations. The logistic, though, is fundamentally a nonlinear dynamic model and possesses a wider range of dynamic behaviors (such as limit cycles and chaos). The Gompertz is fundamentally a linear dynamic model (see Eq. 9) and therefore has a restricted repertoire of dynamic behaviors. The logistic would seem a more flexible choice for modeling the dynamic behavior of natural populations, and we have therefore centered our investigations around statistical properties arising from the logistic.

Some investigators using Gompertz-based methods have concluded that density dependence is frequently weak and not as widely prevalent as theoretical ecologists might expect (den Boer and Reddingius 1989, den Boer 1990). Without extensive logistic-Gompertz model evaluations and analyses using many data sets, it is not clear whether their results arise from nature or from increased Type II error rates inherent in the Gompertz-based statistical methods. The low power of such methods has been acknowledged (Vickery and Nudds 1991), though not explained. We speculate that searching for density dependence with a model of log-density dependence might be akin to trying to photograph a distant bird with a wide-angle instead of a telephoto lens.

On the other hand, investigators using the stochastic logistic model, particularly with modifications for second-order lags (Eq. 6), have reported pervasive evidence of density dependence (Turchin 1990, Turchin et al. 1991, Berryman 1991a). Some of these results are probably influenced by the properties of the regression-based statistical testing methods they used. As we have noted, the t and F statistics for testing whether slope parameters are zero in Eq. 6 do not have t or F distributions. Our simulations indicate that the Type I error rates in one-sided tests of first- or second-order lags using Student's t distributions are markedly inflated over the nominal rate of 0.05 (Table 8). Users of such tests will find density dependence too often when it in fact is weak or absent. Extensions of the t and F tests to multispecies versions of the logistic (Berryman 1991b) must be called into question as well, until such tests receive further study. In addition, the regression-based, two-sided test for a first-order lag has inflated Type I error rates for some parameter values (Table 8). The regression-based, two-sided test for a second-order lag appears to possess reasonable Type I error rates (Table 8). It is noteworthy that the investigators have tended to use the two-sided regression test for second-order lags. Thus, their results concerning the prevalence of second-order lags might hold up when the data are analyzed with techniques that have received thorough evaluation.

The PBLR procedure can be adapted to test for second-order lags. Such a test provides a check on the

Markov assumption implicit in the stochastic logistic model. One would use Model 2 (Eq. 1) as the null hypothesis and the second-order lag model (Eq. 6 with b_3, b_4, ... set to zero) as the alternate hypothesis. The test is conditioned on n_0 and n_1, so the ML estimates for the null hypothesis are based on the time series starting with n_1. The likelihood function for the alternate hypothesis is a product of conditional normal pdf's of the form $p(x_t | x_{t-1}, x_{t-2})$, because X_t conditioned on x_{t-1} and x_{t-2} has a normal $(x_{t-1} + a + b_1 \exp(x_{t-1}) + b_2 \exp(x_{t-2}), \sigma^2)$ distribution. It can be shown that the ML estimates of a, b_1, and b_2 are the least squares estimates obtained by performing a multiple linear regression of y_t on n_{t-1} and n_{t-2}. The bootstrap data sets would be obtained from the estimated null hypothesis model, and bootstrap values of the LR statistic would be obtained by fitting both models to each bootstrap data set. A study of the power properties of this test is in progress.

What is the ecological interpretation of rejecting Model 1 in favor of Model 2? Essentially, the outcome results when the data contain sufficient information to estimate an additional parameter. That parameter, b, imparts an ergodic behavior to the model (when b is negative): large populations tend to decline, and small populations tend to increase. Model 2 thus quantifies a *return tendency* in the data. The return point of the population is $-a/b$; this represents the population abundance at which the *average* change in ln N_t, conditional on N_{t-1}, is zero (Eq. 5). Failure to reject Model 1 can occur when the return point, if one exists, is simply too large or too small to be estimated (out of the range of the data). According to Model 1, a "density-vagueness" (Strong 1986a, b) prevails over the range of the data. Note that for a growing population, it might only be a matter of time before the return point can be estimated. Similarly, a population declining toward the return point at first also resembles Model 1.

The return point, $-a/b$, is not an equilibrium. The "equilibrium" of the discrete time stochastic logistic model is not a point; rather, it is a long-term stationary probability distribution of population sizes. Wolda (1989, 1991) has questioned the meaning of density dependence tests that rely on high densities being above and low densities being below an "equilibrium," because of the impossibility of separating "fluctuating equilibrium values" from "fluctuating deviations from those equilibrium values" (Wolda 1991). Indeed, it is not likely that any statistical method will be able to distinguish these mechanisms of fluctuation from time series data alone. The model (Eq. 1) instead accommodates both of these mechanisms; the noise represented by Z_t describes in a phenomenological fashion a population's growth rate fluctuating for whatever reasons. Once noise is admitted, an ecological Rubicon of sorts is crossed: there no longer is a point equilibrium, conceptually, mathematically, or empirically. It is not correct to claim that a point equilibrium may

TABLE 8. Estimated sizes of density dependence tests in which the critical percentiles of the test statistics are taken from Student's t distributions used in regression analysis. Each estimate was obtained from 1000 trials; all tests were conducted at a nominal significance level of .05.

σ	a	First-order lag*		Second-order lag†	
		One-sided	Two-sided	One-sided	Two-sided
0.25	0.5	0.110	0.076	0.091	0.054
0.25	1.0	0.064	0.047	0.077	0.035
0.25	1.5	0.052	0.051	0.073	0.046
0.50	0.5	0.158	0.107	0.088	0.052
0.50	1.0	0.105	0.075	0.080	0.034
0.50	1.5	0.077	0.064	0.063	0.032
0.75	0.5	0.227	0.141	0.088	0.041
0.75	1.0	0.136	0.094	0.086	0.049
0.75	1.5	0.095	0.069	0.063	0.035
1.00	0.5	0.258	0.162	0.078	0.041
1.00	1.0	0.173	0.122	0.076	0.048
1.00	1.5	0.115	0.088	0.057	0.031

* The base model is $Y_t = a + bN_{t-1} + \sigma Z_t$, where N_t is population abundance at time t, $Y_t = \ln(N_t/N_{t-1})$, $Z_t \sim$ normal(0,1), and $t = 1, 2, \ldots, q$. Null hypothesis is H_1: $b = 0$, one-sided alternate hypothesis is H_2: $b < 0$, and two-sided alternate hypothesis is H_2: $b \neq 0$. For all simulations, $b = 0$, $n_0 = 100$, and $q = 9$. Critical percentiles were obtained from a Student's t distribution with $q - 2$ df.

† The base model is $Y_t = a + b_1 N_{t-1} + b_2 N_{t-2} + \sigma Z_t$, where $t = 2, 3, \ldots, q + 1$. Null hypothesis is H_0: $b_2 = 0$, one-sided alternate hypothesis is H_1: $b_2 < 0$, and two-sided alternate hypothesis is H_1: $b_2 \neq 0$. For all simulations, $b_2 = 0$, $b_1 = -0.01$, $n_0 = -a/b$, and $q = 15$. Critical percentiles were obtained from a Student's t distribution with $q - 3$ df.

emerge as a result of density dependence analyses (Berryman 1991a). Wolda (1989) has stated it well: "Equilibrium is not a point but a cloud of points." The stationary distribution of the stochastic logistic model can be approximated by a positively skewed distribution known as a gamma distribution (Dennis and Patil 1984, Dennis 1989b). Some ecological ramifications of this concept of a stochastic equilibrium have been discussed elsewhere (May 1974, Dennis and Patil 1984, Dennis and Costantino 1988, Dennis 1989b, Desharnais et al. 1990, Costantino and Desharnais 1991, Kemp and Dennis 1993).

Investigators must carefully distinguish the statistical hypothesis of density dependence (as exemplified by Model 2) from the ecological hypothesis (biological mechanism of negative feedback on growth rate). Analyses of time series data have all the pitfalls of any observational studies. Other stochastic mechanisms with stationary distributions, such as a series of independent, identically distributed population sizes drawn from some statistical distribution, are better described by Model 2 (and similar models such as the stochastic Gompertz) than Model 1 or Model 0 (Wolda and Dennis 1993). Indeed, if $b = -1$ in the stochastic Gompertz (Eq. 8), then population sizes *are* independent, identically distributed lognormal random variables. Quantities such as annual rainfall and spring snowpack levels thus qualify as "density dependent" under statistical tests (see Wolda and Dennis 1993),

but it is questionable whether ecologists would consider such quantities to be density dependent in an ecological sense. We believe the PBLR test can be a useful component of a case for ecological density dependence, but should not be the sole component.

One important application of density dependence testing is in conservation biology. A critical question is how to estimate population trends and properties of the first-passage distribution from time series data (e.g., Dennis et al. 1991). The first-passage distribution is the probability distribution of the time it will take for a stochastic process (population size) to first attain some lower (or higher) value. Properties of interest include the mean time to reach a lower value and the probability of reaching it before reaching a given higher value. Preliminary evidence indicates that estimates of first-passage properties can vary substantially depending on whether or not a density dependent model is used (Ginzburg et al. 1990, Stacey and Taper 1992). The PBLR test represents a potentially valuable aid to deciding whether or not to account for density dependence in population viability analysis.

LITERATURE CITED

Andrewartha, H. G. 1957. The use of conceptual models in population ecology. Cold Spring Harbor Symposia on Quantitative Biology 22:219–232.

Aptech Systems. 1991. GAUSS. Aptech Systems, Kent, Washington, USA.

Bain, L. J., and M. Engelhardt. 1987. Introduction to probability and mathematical statistics. PWS Publishers, Boston, Massachusetts, USA.

Barlow, J. 1992. Nonlinear and logistic growth in experimental populations of guppies. Ecology 73:941–950.

Beran, R. 1986. Discussion of the paper by C. F. J. Wu. Annals of Statistics 14:1295–1298.

Berryman, A. A. 1991a. Stabilization or regulation: what it all means! Oecologia 86:140–143.

———. 1991b. Can economic forces cause ecological chaos? The case of the northern California Dungeness crab fishery. Oikos 62:106–109.

Bhat, B. R. 1974. On the method of maximum-likelihood for dependent observations. Journal of the Royal Statistical Society Series B 36:48–53.

The Biological Laboratory. 1957. Population studies: animal ecology and demography. Cold Spring Harbor Symposia on Quantitative Biology. Volume 22. Long Island Biological Association, New York, New York, USA.

Birch, L. C. 1957. The role of weather in determining the distribution and abundance of animals. Pages 203–215 in The Biological Laboratory. Population studies: animal ecology and demography. Cold Spring Harbor Symposia on Quantitative Biology. Volume 22. Long Island Biological Association, New York, New York, USA.

Boyce, M. S. 1989. The Jackson elk herd: intensive management in North America. Cambridge University Press, Cambridge, England.

Bulmer, M. G. 1975. The statistical analysis of density dependence. Biometrics 31:901–911.

Cody, M. L., and J. M. Diamond, editors. 1975. Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA.

Colinvaux, P. 1973. Introduction to ecology. John Wiley & Sons, New York, New York, USA.

Costantino, R. F., and R. A. Desharnais. 1991. Population dynamics and the Tribolium model: genetics and demography. Monographs on Theoretical and Applied Genetics 13. Springer-Verlag, New York, New York, USA.

Crowley, P. H. 1992. Density dependence, boundedness, and attraction: detecting stability in stochastic systems. Oecologia 90:246–254.

Cushing, J. M. 1989. A strong ergodic theorem for some nonlinear matrix models for the dynamics of structured populations. Natural Resource Modeling 3:331–357.

den Boer, P. J. 1968. Spreading of risk and stabilization of animal numbers. Acta Biotheoretica 18:165–194.

———. 1990. On the stabilization of animal numbers. Problems of testing. 3. What do we conclude from significant test results? Oecologia 83:38–46.

———. 1991. Seeing the trees for the wood: random walks or bounded fluctuations of population size? Oecologia 86:484–491.

den Boer, P. J., and J. Reddingius. 1989. On the stabilization of animal numbers. Problems of testing. 2. Confrontation with data from the field. Oecologia 79:143–149.

Dennis, B. 1989a. Allee effects: population growth, critical density, and the chance of extinction. Natural Resource Modeling 3:481–537.

———. 1989b. Stochastic differential equations as insect population models. In L. McDonald, B. Manly, J. Lockwood, and J. Logan, editors. Estimation and analysis of insect populations. Lecture Notes in Statistics 55:219–238.

Dennis, B., and R. F. Costantino. 1988. Analysis of steady-state populations with the gamma abundance model: application to Tribolium. Ecology 69:1200–1213.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. Ecological Monographs 61:115–143.

Dennis, B., and G. P. Patil. 1984. The gamma distribution and weighted multimodal gamma distributions as models of population abundance. Mathematical Biosciences 68:187–212.

Desharnais, R. A., B. Dennis, and R. F. Costantino. 1990. Genetic analysis of a population of Tribolium. IX. Maximization of population size and the concept of a stochastic equilibrium. Genome 33:571–580.

Diamond, J., and T. J. Case, editors. 1986. Community ecology. Harper and Row, New York, New York, USA.

Dickey, D. A., and W. A. Fuller. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49:1057–1072.

Eberhardt, L. L. 1970. Correlation, regression, and density dependence. Ecology 51:306–310.

Eberhardt, L. L., R. R. Knight, and B. M. Blanchard. 1986. Monitoring grizzly bear population trends. Journal of Wildlife Management 50:613–618.

Efron, B. 1986. Discussion of the paper by C. F. J. Wu. Annals of Statistics 14:1301–1304.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician 37:36–48.

Emlen, J. M. 1984. Population biology. Macmillan, New York, New York, USA.

Fisher, R. A. 1958. Statistical methods for research workers. Hafner, New York, New York, USA.

Fowler, C. W. 1984. Density dependence in cetacean populations. Reports of the International Whaling Commission Special Issue 6:373–379.

———. 1987. A review of density dependence in populations of large mammals. Pages 401–441 in H. H. Genoways, editor. Current mammalogy. Volume 1. Plenum, New York, New York, USA.

Fuller, W. A. 1987. Measurement error models. John Wiley & Sons, New York, New York, USA.

Gaston, K. J., and J. H. Lawton. 1987. A test of statistical techniques for detecting density dependence in sequential censuses of animal populations. Oecologia 74:404–410.

Ginzburg, L. R., S. Ferson, and H. R. Akçakaya. 1990. Reconstructibility of density dependence and the conservative assessment of extinction risks. Conservation Biology 4:63–70.

Goodman, D. 1987. Consideration of stochastic demography in the design and management of biological reserves. Natural Resource Modeling 1:205–234.

Hassell, M. P., J. Latto, and R. M. May. 1989. Seeing the wood for the trees: detecting density dependence from existing life-table studies. Journal of Animal Ecology 58:883–892.

Holyoak, M., and J. H. Lawton. 1992. Detection of density dependence from annual censuses of bracken-feeding insects. Oecologia 91:425–430.

Houston, D. B. 1982. The northern Yellowstone elk: ecology and management. Macmillan, New York, New York, USA.

Kemp, W. P., and B. Dennis. 1993. Density dependence in rangeland grasshoppers (Orthoptera: Acrididae). Oecologia 96:1–8.

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

Knight, R. R., and L. L. Eberhardt. 1984. Projected future abundance of the Yellowstone grizzly bear. Journal of Wildlife Management 48:1434–1438.

Knight, R. R., and L. L. Eberhardt. 1985. Population dynamics of Yellowstone grizzly bears. Ecology 66:323–334.

Leigh, E. G. 1981. The average lifetime of a population in a varying environment. Journal of Theoretical Biology 90:213–239.

Lele, S. 1991. Jackknifing linear estimating equations: asymptotic theory and applications in stochastic processes. Journal of the Royal Statistical Society Series B 53:253–267.

Lin, C.-C., and G. S. Mudholkar. 1980. A simple test for normality against asymmetric alternatives. Biometrika 67:455–461.

Livdahl, J. P., and G. Sugihara. 1984. Non-linear interactions of populations and the importance of estimating per capita rates of change. Journal of Animal Ecology 53:573–580.

May, R. M. 1974. Stability and complexity in model ecosystems. Second edition. Princeton University Press, Princeton, New Jersey, USA.

———, editor. 1976. Theoretical ecology: principles and applications. W. B. Saunders, Philadelphia, Pennsylvania, USA.

McLain, D. H. 1974. Drawing contours from arbitrary data points. Computer Journal 17:318–324.

McLaren, I. A., editor. 1971. Natural regulation of animal populations. Atherton, New York, New York, USA.

Mountford, M. D. 1988. Population regulation, density dependence, and heterogeneity. Journal of Animal Ecology 57:845–858.

Neyman, J., and E. S. Pearson. 1933. On the problem of the most efficient tests of statistical hypotheses. Philosoph-

ical Transactions of the Royal Society Series A 231:289–337.

Nicholson, A. J. 1933. The balance of animal populations. Journal of Animal Ecology 2:132–178.

Pianka, E. R. 1974. Evolutionary ecology. Harper and Row, New York, New York, USA.

Pollard, E., K. H. Lakhani, and P. Rothery. 1987. The detection of density-dependence from a series of annual censuses. Ecology 68:2046–2055.

Reddingius, J. 1971. Gambling for existence. A discussion of some theoretical problems in animal population ecology. Acta Biotheoretica 20 (Supplement):1–208.

———. 1990. Models for testing: a secondary note. Oecologia 83:50–52.

Reddingius J., and P. J. den Boer. 1989. On the stabilization of animal numbers. Problems of testing. 1. Power estimates and estimation errors. Oecologia 78:1–8.

Rice, J. A. 1988. Mathematical statistics and data analysis. Wadsworth, Belmont, California, USA.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada 11:559–623.

Roughgarden, J. 1979. Theory of population genetics and evolutionary ecology: an introduction. Macmillan, New York, New York, USA.

Royama, T. 1977. Population persistence and density dependence. Ecological Monographs 47:1–35.

———. 1981. Fundamental concepts and methodology for the analysis of population dynamics, with particular reference to univoltine species. Ecological Monographs 51:473–493.

Schork, N. 1992. Bootstrapping likelihood ratios in quantitative genetics. Pages 389–396 in R. LePage and L. Billard, editors. Exploring the limits of bootstrap. John Wiley & Sons, New York, New York, USA.

Serfling, R. J. 1980. Approximation theorems of mathematical statistics. John Wiley & Sons, New York, New York, USA.

Shaffer, M. L. 1981. Minimum population sizes for species conservation. BioScience 31:131–134.

Simberloff, D. 1988. The contribution of population and community ecology to conservation science. Annual Review of Ecology and Systematics 19:473–511.

Slade, N. A. 1977. Statistical detection of density dependence from a series of sequential censuses. Ecology 58:1094–1102.

Solow, A. R. 1990. Testing for density dependence: a cautionary note. Oecologia 83:47–49.

———. 1991. Response. Oecologia 86:146.

Solow, A. R., and J. H. Steele. 1990. On sample size, statistical power, and the detection of density dependence. Journal of Animal Ecology 59:1073–1076.

Stacey, P. B., and M. Taper. 1992. Environmental variation and the persistence of small populations. Ecological Applications 2:18–29.

Strong, D. R. 1986a. Density vagueness: abiding the variance in the demography of real populations. Pages 257–268 in J. Diamond and T. J. Case, editors. Community ecology. Harper and Row, New York, New York, USA.

———. 1986b. Density-vague population change. Trends in Ecology and Evolution 1:39–42.

Stuart, A., and J. K. Ord. 1991. Kendall's advanced theory of statistics. Volume 2: classical inference and relationship. Fifth edition. Oxford University Press, New York, New York, USA.

Tong, H. 1990. Non-linear time series: a dynamical system approach. Oxford University Press, Oxford, England.

Turchin, P. 1990. Rarity of density dependence or population regulation with lags? Nature 344:660–663.

Turchin, P., P. L. Lorio Jr., A. D. Taylor, and R. F. Billings. 1991. Why do populations of southern pine beetles (Co-

leoptera: Scolytidae) fluctuate? Environmental Entomology **20**:401–409.

Turchin, P., and A. D. Taylor. 1992. Complex dynamics in ecological time series. Ecology **73**:289–305.

Vickery, W. L., and T. D. Nudds. 1984. Detection of density-dependent effects in annual duck censuses. Ecology **65**:96–104.

Vickery, W. L., and T. D. Nudds. 1991. Testing for density-dependent effects in sequential censuses. Oecologia **85**:419–423.

Wissel, C., and S. Stöcker. 1991. Extinction of populations by random influences. Theoretical Population Biology **39**:315–328.

Woiwod, I. P., and I. Hanski. 1992. Patterns of density dependence in moths and aphids. Journal of Animal Ecology **61**:619–629.

Wolda, H. 1989. The equilibrium concept and density dependence tests. What does it all mean? Oecologia **81**:430–432.

———. 1991. The usefulness of the equilibrium concept in population dynamics. A reply to Berryman. Oecologia **86**:144–145.

Wolda, H., and B. Dennis. 1993. Density dependence tests, are they? Oecologia **95**:581–591.