JOINT DENSITY DEPENDENCE

B. Dennis,[1] W. P. Kemp,[2] and M. L. Taper[3]

[1] *Department of Fish and Wildlife Resources and Division of Statistics, University of Idaho, Moscow, Idaho 83844-1136 USA*
[2] *USDA/ARS Bee Biology and Systematics Laboratory, Utah State University, Logan, Utah 84322-5305 USA*
[3] *Department of Biology, Montana State University, Bozeman, Montana 59717-0346 USA*

Abstract. We present a new multivariate model for describing jointly fluctuating, density-dependent populations. The model is a stochastic, multivariate version of a discrete-time logistic model that explicitly accounts for spatial variation of growth rate parameters and covariances of fluctuations between populations. Statistical methods for applying the model to time series data on population abundances are described in detail. We derive formulas for maximum-likelihood estimates of model parameters, and we develop hypothesis tests for various reduced model structures, such as density independence or zero covariance. The applicability of the model to any given data set can be thoroughly evaluated with diagnostic procedures. As examples, we apply the model to two population systems: rangeland grasshoppers (Orthoptera: Acrididae) in three physiographic regions of Montana, USA, and bull trout (*Salvelinus confluentus*) in four tributaries of the Flathead River in Montana. The model and associated statistical methods have potentially important applications in conservation biology for describing metapopulations and for assessing joint jeopardy of multiple populations.

Key words: bull trout; density dependence; grasshoppers; locusts; logistic model; metapopulation; multivariate time series; nonlinear autoregressive model; population viability analysis; stochastic population model.

INTRODUCTION

The reader may have anticipated the general conclusion that we have been leading up to. The numbers of animals in a natural population may be limited in three ways: (a) by shortages of material resources, such as food, places in which to make nests, etc.; (b) by inaccessibility of these material resources relative to the animals' capacities for dispersal and searching; and (c) by shortage of time when the rate of increase is positive. Of these three ways, the first is probably the least, and the last is probably the most, important in nature. Concerning c, the fluctuations in the value of r may be caused by weather, predators, or any other component of the environment which influences the rate of increase.

—Andrewartha and Birch (1954: Section 14.2)

Density dependence has occupied the interests of ecologists for generations. Andrewartha and Birch synthesized this thinking in their massive and now classic 1954 tome. As illustrated in the above quotation, Andrewartha and Birch had a broad vision of the problem of abundances, in which the mechanisms of density regulation were embedded in a spatial matrix and impinged upon by stochastic influences.

A generation of ecologists has struggled with the quantitative treatment of the smallest part of this vision, that is, the question of whether natural populations are

regulated by density-dependent mechanisms. This has proven to be a much stickier statistical problem than its simple statement implies (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, Wolda 1989, den Boer 1990, 1991, Solow 1990, 1991, Turchin 1990, Berryman 1991, Turchin et al. 1991, Crowley 1992, Turchin and Taylor 1992, Hanski et al. 1993, Holyoak and Lawton 1993, Wolda and Dennis 1993, Wolda et al. 1994, Fox and Risdell-Smith 1995). Dennis and Taper (1994) reviewed many of the statistical and conceptual problems associated with the detection of density dependence in single populations, and proposed statistical methods for analyzing univariate time series data with a stochastic model of density-dependent population growth.

Frequently, though, population time series are not univariate. When related populations at different spatial locations are monitored, the data take the form of simultaneous time series of population abundances. It is tempting to pool such data and estimate common parameters for a univariate model (for instance, Perry et al. 1993) in order to gain power and shorten confidence intervals afforded by the usually short time series. However, pooling time series data from multiple sites has two potential drawbacks.

First, model growth rate parameters (e.g., r and k, Kemp and Dennis 1993) as well as noise variability

parameters (e.g., σ^2, Kemp and Dennis 1993) will typically have spatial variation. Ecological conditions affecting parameter values, in a heterogeneous world, are expected to be different from site to site.

Second, multiple time series can display coherence or spatial synchrony (Hanski 1991, Hanski and Woiwod 1993), that is, the fluctuations of population abundances at the different locations can appear correlated. Stochastic environmental factors that have region-wide influences on populations, termed regional stochasticity by Hanski (1991), are an important potential cause of spatial synchrony. Analyzing populations separately with univariate models in such situations will yield results that are interdependent in unknown ways.

Clearly, the next generation of ecological time series models should account for all three of Andrewartha and Birch's factors of density dependence, space, and stochasticity.

In an effort to begin the development of a quantitative treatment of Andrewartha's and Birch's vision, we extend the approach of Dennis and Taper (1994) to include populations at multiple sites. In this paper, we introduce a multivariate model of jointly fluctuating, density-dependent populations. The model is a stochastic, multivariate version of the discrete-time Ricker/logistic model. It explicitly accounts for spatial variation in growth-rate parameters and covariances of fluctuations between populations. We develop methods for using information from multiple time series in model identification, parameter estimation, and hypothesis tests concerning density dependence. We are able to measure density dependence in species that are buffeted by stochastic fluctuations in their growth rates and distributed in a spatial framework. With this model, many of the questions raised by Andrewartha and Birch concerning the relative importance of stochastic fluctuations vs. density-dependent forces ultimately can be addressed.

The model we present in this paper does not explicitly incorporate dispersal. In "metapopulations," subpopulations distributed in space are connected by migration. Migration can be an additional cause of spatial synchrony when it is a sizable proportion of population growth rates (see Hanski and Woiwod 1993). Thus, a full analysis of metapopulations would require a quantification of migration rates (Gilpin and Hanski 1991, Stacey and Taper 1992, Stacey et al. 1997). We will return in the *Discussion* to the potential advantages and limitations of using our methods in the analysis of metapopulations.

Model Description

We first describe briefly the simple univariate case of the model and then expand to the full multivariate version. Let N_t be population abundance at time t, as censused, estimated, or indexed. We assume that the population is observed at about the same time every year, producing a time series of length $q + 1$: N_0, N_1,

..., N_q. Let the log-transformed abundances be denoted $X_0 = \ln N_0$, $X_1 = \ln N_1$, ..., $X_q = \ln N_q$. Also, let $Y_t = X_t - X_{t-1} = \ln(N_t/N_{t-1})$; Y_t is the change in logarithmic population abundance at time t and is a discrete-time analogue to the per-unit-abundance growth rate in continuous-time models ($[1/n]dn/dt = d \ln n/dt$). The univariate model is

$$N_t = N_{t-1}\exp(a + bN_{t-1} + E_t) \qquad (1)$$

where a and b are constants, and E_t has a normal distribution with a mean $=0$ and variance $=\sigma^2$ [we write $E_t \sim$ normal$(0, \sigma^2)$]. The quantity E_t represents a random shock to the population growth rate caused by unspecified stochastic forces. The shocks are assumed uncorrelated through time. The population abundances, though, are correlated through time under the model.

On a logarithmic scale, the model becomes

$$X_t = X_{t-1} + a + b \exp(X_{t-1}) + E_t. \qquad (2)$$

Written in this form, the model is a first-order nonlinear autoregression (NLAR) model (Tong 1990). In terms of the growth increments, the model can be written as

$$Y_t = a + bN_{t-1} + E_t. \qquad (3)$$

The increment Y_t is seen to be a linear function of population abundance N_{t-1} at the beginning of the time interval, plus noise. In most applications, ecologists will be concerned with values of b such that $b \leq 0$ ($b > 0$ might be used to model an Allee effect at extremely low densities). The model is a stochastic version of Ricker's (1954) discrete-time logistic model. Dennis and Taper (1994) discuss the versatility and ecological ramifications of this model.

A special case of the univariate model occurs when $b = 0$. The model in that circumstance becomes a stochastic model of exponential increase ($a > 0$), random walk ($a = 0$), or exponential decline ($a < 0$). This stochastic exponential growth case was used by Dennis et al. (1991) for estimating extinction risks of endangered species. A statistical test of the hypothesis $b = 0$ (density-independent growth) vs. $b < 0$ (density-dependent growth) was developed by Dennis and Taper (1994).

One of the main limitations of the model defined by Eq. 1 is its univariate nature. Consider the population fluctuations depicted in Fig. 1. Portrayed are the estimated mean abundances of rangeland grasshoppers in three broad physiographic regions of Montana (see Kemp and Dennis 1993 for a map of the regions). From the standpoint of forage utilization in Montana rangelands, "grasshoppers" represent a functional group that contains a number of serious pest species (Kemp 1992a, b), and we shall refer to the group as a "population" in this sense. Ecologically, the grasshopper complex across Montana comprises ≥100 species, as many as 20 or more of which can be found coexisting in any given field (Kemp 1992a, b). The data in Fig. 1 are the total mean abundances for all species com-
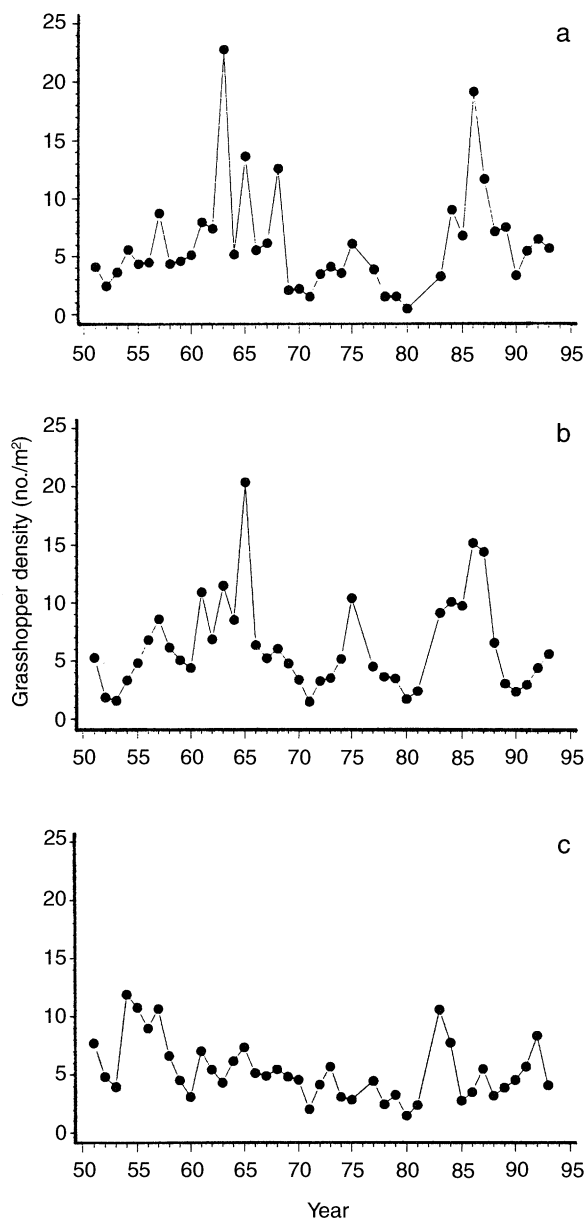
FIG. 1. Mean adult rangeland grasshopper densities for three physiographic regions of Montana, 1951–1993: (a) northern glaciated plains, (b) southern unglaciated plains, (c) western mountains. Data were described by Kemp and Dennis (1993).

bined. The overall populations in Fig. 1 appear to be fluctuating, albeit wildly, around some intermediate abundance level. Kemp and Dennis (1993) have shown that the univariate model (Eq. 1) separately describes the three populations quite well, and can serve as a valuable forecasting tool for rangeland managers. However, it is obvious from Fig. 1 that the three populations do not fluctuate independently of each other. A good/bad year for grasshoppers on the southern unglaciated plains of Montana is likely to be a good/bad

year on the northern glaciated plains, and vice versa. Conclusions about density dependence or predictions about outbreaks derived from statistical analyses of each population separately using Eq. 1 are not likely to be independent of each other.

A multivariate version of Eq. 1 is needed to account for such cofluctuations. Many different multivariate extensions could be proposed; the following one strikes a balance between realism and general usefulness. Let the abundances of m different populations at time t be the elements of a column vector: $\mathbf{N}_t = [N_{1t}, N_{2t}, \ldots, N_{mt}]'$. These elements could represent distinct, cofluctuating species, or populations of the same species (metapopulation). Let the log-transformed abundances be arranged in a column vector defined by $\mathbf{X}_t = [X_{1t}, X_{2t}, \ldots, X_{mt}]'$, where $X_{it} = \ln N_{it}$. The column vector $\mathbf{Y}_t$ will hold the corresponding increments: $Y_{it} = X_{it} - X_{i(t-1)}$. In the multivariate version, each population grows according to its own stochastic logistic equation:

$$N_{it} = N_{i(t-1)}\exp(a_i + b_i N_{i(t-1)} + E_{it}). \tag{4}$$

Here $E_{it} \sim$ normal$(0, \sigma_i^2)$ as in the univariate case; the multivariate version, however, allows for correlation between the fluctuations of the populations within a given time interval. Specifically, let $\mathbf{E}_t = [E_{1t}, E_{2t}, \ldots, E_{mt}]'$ denote the column vector of random shocks to the logarithmic growth rates of the m populations in year t. We assume that $\mathbf{E}_t$ has a multivariate normal (MVN) distribution with a mean vector of zero and a variance-covariance matrix of $\mathbf{\Sigma}$, where

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{12} & \sigma_2^2 & & \\ \vdots & & \ddots & \\ \sigma_{1m} & & & \sigma_m^2 \end{bmatrix} \tag{5}$$

with σ_i^2 denoting the variance of E_{it}, and σ_{ij} denoting the covariance of E_{it} and E_{jt}. We write $\mathbf{E}_t \sim$ MVN$(\mathbf{0}, \mathbf{\Sigma})$. We assume that the elements in $\mathbf{E}_t$ are not correlated through time. We expect such autocorrelations to be small compared to the within-time correlations, provided the underlying population trends are reasonably well described by the logistic/Ricker model. These assumptions can be evaluated for any given data set (see *Diagnostics and evaluation* section).

The model can be written in matrix form on the logarithmic scale. Letting $\mathbf{D}(\mathbf{N}_t)$ represent a matrix with the elements of $\mathbf{N}_t$ on the main diagonal and zeros elsewhere, the m model equations (Eq. 4) can be expressed as

$$\mathbf{X}_t = \mathbf{X}_{t-1} + \mathbf{a} + \mathbf{D}(\mathbf{N}_{t-1})\mathbf{b} + \mathbf{E}_t. \tag{6}$$

Here $\mathbf{a} = [a_1, a_2, \ldots, a_m]'$ and $\mathbf{b} = [b_1, b_2, \ldots, b_m]'$. The model is a system of m stochastic difference equations, in which the populations are basically growing separately, but the fluctuations in their growth rates are coupled through the covariances of the elements in $\mathbf{E}_t$.

Written in the form of Eq. 6, it is a type of multivariate, nonlinear autoregression (MNLAR) model.

Population trajectories under this model are easy to simulate. We describe this feature first, as it is a key to convenient parameter estimation, hypothesis testing, and joint-risk analysis (*Parameter estimation, Diagnostics and evaluation,* and *Discussion* sections). A matrix-based programming language, such as GAUSS (Aptech Systems 1993) or MATLAB (Math Works 1993) renders the programming effort almost trivial. The steps for such simulation are as follows: (0) In this initialization step, provide values for the elements of $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{\Sigma}$. (Methods for estimating these parameters from data are presented in the *Parameter Estimation* section.) Provide initial population sizes and arrange them into the vector $\mathbf{N}_0$. Calculate $\mathbf{X}_0$ by log-transforming the elements of $\mathbf{N}_0$. Factor the matrix $\mathbf{\Sigma}$ into a product of an upper ($\mathbf{T}$) and a lower ($\mathbf{T}'$) triangular matrix:

$$\mathbf{\Sigma} = \mathbf{T}'\mathbf{T}. \tag{7}$$

This factorization is called the Cholesky decomposition and is a library subroutine in various matrix programming languages. If the calculations are being performed with a primitive language, one may use a simple algorithm for the factorization (Graybill 1976:231, Press et al. 1992:89). Once the factorization is completed, the population time series are then generated as follows: (1) Generate m independent normal(0, 1) random variables: $\mathbf{Z} = [Z_1, Z_2, \ldots, Z_m]'$. (2) Calculate the vector of fluctuations, $\mathbf{E}_1$, as

$$\mathbf{E}_1 = \mathbf{T}'\mathbf{Z}. \tag{8}$$

According to an elementary property of the multivariate normal distribution, $\mathbf{E}_1 \sim \mathrm{MVN}(\mathbf{0}, \mathbf{\Sigma})$ (Graybill 1976:98). (3) Use Eq. 6 to calculate $\mathbf{X}_1$. Exponentially transform the elements of $\mathbf{X}_1$ to obtain $\mathbf{N}_1$. (4) Repeat steps 1–3 to obtain $\mathbf{X}_2$ and $\mathbf{N}_2$ from $\mathbf{X}_1$ and $\mathbf{N}_1$, etc.

The model (Eq. 6) represents a multivariate, stochastic logistic model. It is a model of what we term "joint density dependence." The populations cofluctuate in a long-term statistical equilibrium. Each population has a return tendency centered at a point, $-a_i/b_i$; the conditional expected value of the logarithmic change $Y_{i(t-1)}$, given $N_{i(t-1)} = n_{i(t-1)}$, is positive if $n_{i(t-1)} < -a_i/b_i$ and negative if $n_{i(t-1)} > -a_i/b_i$. However, $-a_i/b_i$ is not a point equilibrium. A point equilibrium is not a well-defined concept in a stochastic model (Dennis and Patil 1984, Dennis and Costantino 1988, Wolda 1989, Kemp and Dennis 1993, Dennis and Taper 1994). Each population also has its own scale of fluctuations under the model. The diagonal elements, $\{\sigma_i^2\}$, in $\mathbf{\Sigma}$ measure the intensity of the random growth rate shocks for the different populations. The off-diagonal elements, $\{\sigma_{ij}\}$, measure the covariance of the growth rate shocks between pairs of populations and can be positive or negative.

A special case of the model is a multivariate, sto-

chastic exponential growth model. By setting $\mathbf{b} = \mathbf{0}$ in Eq. 6, one obtains

$$\mathbf{X}_t = \mathbf{X}_{t-1} + \mathbf{a} + \mathbf{E}_t. \tag{9}$$

Under this model, each population on average experiences exponential increase or decrease, depending on whether its rate constant, a_i, is positive or negative. The fluctuations represented by $\mathbf{E}_t$ retain the covariance structure of the multivariate logistic. The special case given by Eq. 9 can serve as a null hypothesis that all the populations have density-independent growth rates, to be tested against the alternative that some populations are density dependent ($\mathbf{b} \neq \mathbf{0}$).

PARAMETER ESTIMATION

Data for the multivariate stochastic logistic model would consist of a time series of abundances for each of the m populations. At each time t, the abundances $n_{1t}, n_{2t}, \ldots, n_{mt}$ (as censused, estimated, or indexed) would be recorded. Let $\mathbf{n}_t = [n_{1t}, n_{2t}, \ldots, n_{mt}]'$ be the vector of recorded abundances at time t, and let $\mathbf{x}_t = [\ln n_{1t}, \ln n_{2t}, \ldots, \ln n_{mt}]'$. Also, let $\mathbf{y}_t = [\ln(n_{1t}/n_{1(t-1)}), \ln(n_{2t}/n_{2(t-1)}), \ldots, \ln(n_{qt}/n_{q(t-1)})]'$. The series of vectors $\mathbf{n}_0, \mathbf{n}_1, \ldots, \mathbf{n}_q$ constitute a multivariate time series of length $q + 1$.

Data and model are connected through a likelihood function. The likelihood function gives the relative chance that a realization of the multivariate stochastic process $\mathbf{N}_t$, with starting point fixed at $\mathbf{n}_0$, would result in the outcome $\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_q$ that was actually observed. The likelihood function is built from the transition probability density function (pdf), that is, the pdf for $\mathbf{N}_t$ conditional on $\mathbf{n}_{t-1}$. Because of the autoregressive nature of the model on a logarithmic scale (Eq. 6), it is easiest to work with the transition pdf for $\mathbf{X}_t$. From Eq. 6, the distribution of $\mathbf{X}_t$ given $\mathbf{X}_{t-1} = \mathbf{x}_{t-1}$ is $\mathrm{MVN}(\mathbf{x}_{t-1} + \mathbf{a} + \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}, \mathbf{\Sigma})$. The multivariate normal transition pdf can be written as

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = |\mathbf{\Sigma}|^{-1/2}(2\pi)^{-m/2}$$
$$\times \exp[-(\mathbf{x}_t - \mathbf{x}_{t-1} - \mathbf{a} - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b})'$$
$$\times \mathbf{\Sigma}^{-1}(\mathbf{x}_t - \mathbf{x}_{t-1} - \mathbf{a} - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b})/2].$$
$$\tag{10}$$

The likelihood function is the product of transition pdf's, that is, the probability of a transition from $\mathbf{x}_0$ to $\mathbf{x}_1$, times the probability of a transition from $\mathbf{x}_1$ to $\mathbf{x}_2$, and so on:

$$L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma}) = p(\mathbf{x}_1|\mathbf{x}_0)p(\mathbf{x}_2|\mathbf{x}_1) \cdots p(\mathbf{x}_q|\mathbf{x}_{q-1}). \tag{11}$$

The likelihood function is denoted $L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma})$ to emphasize its dependence on the unknown parameters in $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{\Sigma}$. In calculations, the log-likelihood function, $\ln L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma})$ is typically used:

$$\ln L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma}) = \sum_{t=1}^{q} \ln p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$$

$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\mathbf{\Sigma}|$$

$$- (1/2) \sum_{t=1}^{q} [\mathbf{y}_t - \mathbf{a} - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}]'\mathbf{\Sigma}^{-1}$$

$$\times [\mathbf{y}_t - \mathbf{a} - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}]. \qquad (12)$$

Maximum likelihood (ML) estimates are the values of the unknown parameters that jointly maximize $L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma})$. In the Appendix, we derive the equations listed below which yield the ML estimates as solutions. The solutions must be calculated iteratively; we describe below two easy methods for calculating the ML estimates. In the formulas, we use the following notation:

$$\bar{\mathbf{y}} = [\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_m]' \qquad (13)$$

$$\bar{\mathbf{n}} = [\bar{n}_1, \bar{n}_2, \ldots, \bar{n}_m]' \qquad (14)$$

where $\bar{y}_i$ is the sample mean of $y_{i1}, y_{i2}, \ldots, y_{iq}$, and $\bar{n}_i$ is the sample mean of $n_{i0}, n_{i1}, \ldots, n_{i(q-1)}$ (note starting point at $t = 0$ and stopping point at $t = q - 1$). The ML estimates of the parameters in $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{\Sigma}$ are solutions to the following equations:

$$\hat{\mathbf{b}} = \left[\sum_{t=1}^{q} \mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\mathbf{\Sigma}}^{-1}\mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}}) \right]^{-1} \qquad (15)$$

$$\times \sum_{t=1}^{q} \mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\mathbf{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}}),$$

$$\hat{\mathbf{a}} = \bar{\mathbf{y}} - \mathbf{D}(\bar{\mathbf{n}})\hat{\mathbf{b}} \qquad (16)$$

$$\hat{\mathbf{\Sigma}} = \hat{\mathbf{R}}\hat{\mathbf{R}}'/q. \qquad (17)$$

In Eq. 17, the matrix $\hat{\mathbf{R}}$ is a matrix of "residuals," that is, a matrix of the deviations of the log-population sizes in $\mathbf{x}_t$ from their estimated conditional expected values. Specifically

$$\hat{\mathbf{R}} = \begin{Bmatrix} \hat{e}_{11} & \hat{e}_{12} & \cdots & \hat{e}_{1q} \\ \hat{e}_{21} & \hat{e}_{22} & \cdots & \hat{e}_{2q} \\ \vdots & \vdots & & \vdots \\ \hat{e}_{m1} & \hat{e}_{m2} & \cdots & \hat{e}_{mq} \end{Bmatrix} \qquad (18)$$

where $\hat{e}_{it} = x_{it} - x_{i(t-1)} - \hat{a}_i - \hat{b}_i n_{i(t-1)}$ is the residual of the ith population at time t (x_{it} minus its predicted value).

One easy way of calculating the solutions to Eqs. 15–17 is suggested by the structure of the formula for $\hat{\mathbf{b}}$. If $\hat{\mathbf{\Sigma}}$ were a known matrix, then Eq. 15 is in the form of a weighted least squares estimate. If an initial approximation for $\hat{\mathbf{\Sigma}}$ could be found, then approximations for $\hat{\mathbf{b}}$ and $\hat{\mathbf{a}}$ could be calculated from Eqs. 15 and 16. These approximate values for $\hat{\mathbf{b}}$ and $\hat{\mathbf{a}}$ could then be used to calculate a better approximation for $\hat{\mathbf{\Sigma}}$ through Eq. 17, and the whole process repeated iteratively until convergence. This "iteratively reweighted least squares" algorithm (Green 1984) has the following

specific steps: (1) Calculate initial values of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ using Eqs. 15 and 16, with $\hat{\mathbf{\Sigma}}$ taken as an identity matrix (ones on main diagonal, zeros elsewhere). These initial values of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are known as "conditional least squares" (CLS) estimates (Klimko and Nelson 1978, Dennis et al. 1995) and are usually close to the final ML estimates. Calculate the initial value of $\hat{\mathbf{\Sigma}}$ through Eq. 17, using the CLS estimates for $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. (2) Calculate new values of $\hat{\mathbf{b}}$ and $\hat{\mathbf{a}}$ using Eqs. 15 and 16 and the latest value of $\hat{\mathbf{\Sigma}}$. Calculate the new value of $\hat{\mathbf{\Sigma}}$ using Eq. 17 and the latest values of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$. Repeat step 2 until convergence. At each iteration, calculate the log-likelihood function (Eq. 12). Convergence is attained when the relative increase in the log-likelihood is small enough (say, $[\ln L_j - \ln L_{j-1}]/|\ln L_{j-1}| < 10^{-10}$, where $\ln L_j$ is the value of the log-likelihood at the jth iteration) and when the relative change in the parameter values is small.

A second easy way of obtaining the ML estimates is to maximize the log-likelihood directly using the Nelder-Mead simplex algorithm (Press et al. 1992:402). The algorithm only requires a routine to calculate the function being maximized (i.e., Eq. 12); full computer code and explanations for this surprisingly simple routine are provided by Press et al. (1992). Because Eq. 17 provides the ML estimate of the matrix $\hat{\mathbf{\Sigma}}$, one need only use $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ as the unknown quantities in the function being maximized. Use Eq. 17 in place of $\hat{\mathbf{\Sigma}}$ in the log-likelihood during the iterations, with the residuals (Eq. 18) calculated using the current values of $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$.

For the special case of the multivariate exponential growth model (Eq. 9), the transition pdf is given by Eq. 10, with $\mathbf{b} = \mathbf{0}$. The likelihood function, $L_0(\mathbf{a}, \mathbf{\Sigma})$, for the parameters in $\mathbf{a}$ and $\mathbf{\Sigma}$ is then given by Eq. 11, evaluated at $\mathbf{b} = \mathbf{0}$:

$$L_0(\mathbf{a}, \mathbf{\Sigma}) = L(\mathbf{a}, \mathbf{0}, \mathbf{\Sigma}). \qquad (19)$$

The ML estimates of $\mathbf{a}$ and $\mathbf{\Sigma}$ under the multivariate exponential growth model jointly maximize $L_0(\mathbf{a}, \mathbf{\Sigma})$. The estimates are easy to calculate; the following formulas are derived in the Appendix:

$$\hat{\mathbf{a}} = \bar{\mathbf{y}}, \qquad (20)$$

$$\hat{\mathbf{\Sigma}} = \hat{\mathbf{R}}\hat{\mathbf{R}}'/q. \qquad (21)$$

Here, the matrix $\hat{\mathbf{R}}$ is given by Eq. 18, but with $\hat{e}_{it} = x_{it} - x_{i(t-1)} - \hat{a}_i = y_{it} - \bar{y}_i$.

HYPOTHESIS TESTING AND MODEL IDENTIFICATION

Various statistical analyses can be used to study the structure of the model and to evaluate its performance for a given data set. We describe in this section techniques for hypothesis testing and model selection.

The hypothesis of density dependence becomes a multivariate concept when two or more populations are fluctuating in a statistically dependent fashion. Conclusions about one population will be related to con-

clusions about another population. Various different model structures have different degrees of density dependence: the model could have zero populations growing logistically, one population growing logistically, and so on.

One decision in population data analysis is whether the joint density dependence model is a significant improvement over the joint density independence model. The decision can be formally stated as a statistical hypothesis test of the null hypothesis H_0: $b_1 = 0$, $b_2 = 0$, . . . , $b_m = 0$ vs. the alternative hypothesis H_1: $b_1 \neq 0$, $b_2 \neq 0$, . . . , $b_m \neq 0$. The null hypothesis is an assertion that the time series of all populations in the group are adequately described by the multivariate exponential growth model, with some populations perhaps increasing, others decreasing. The alternative hypothesis states that the populations are better described by a multivariate stochastic logistic model.

The parametric bootstrap likelihood ratio (PBLR) procedure is a numerically intensive method of testing H_0 vs. H_1. The advantages and limitations of PBLR tests have been discussed by Dennis and Taper (1994). The following PBLR test is a multivariate extension of Dennis and Taper's univariate PBLR test for density dependence.

The multivariate PBLR test revolves around a statistic, denoted Λ, known as the likelihood ratio (LR) statistic. The LR statistic is the ratio of the likelihood function $L_0(\mathbf{a}, \mathbf{\Sigma})$ (Eq. 19) maximized under the null hypothesis, and the likelihood function $L(\mathbf{a}, \mathbf{b}, \mathbf{\Sigma})$ (Eq. 11) maximized under the alternative hypothesis:

$$\Lambda = L_0(\hat{\mathbf{a}}_0, \hat{\mathbf{\Sigma}}_0)/L(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\mathbf{\Sigma}}). \qquad (22)$$

Here $\hat{\mathbf{a}}_0$ and $\hat{\mathbf{\Sigma}}_0$ are the ML estimates (Eqs. 20 and 21) under the null hypothesis model, and $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, $\hat{\mathbf{\Sigma}}$ are the ML estimates (Eqs. 15, 16, and 17) under the alternative hypothesis model. Any monotone function of Λ can be used for the LR test; a frequently used function, $-2 \ln \Lambda$, reduces algebraically to a fairly simple form involving matrix determinants (the Appendix):

$$G^2 = -2 \ln \Lambda = q \ln(|\hat{\mathbf{\Sigma}}_0|/|\hat{\mathbf{\Sigma}}|). \qquad (23)$$

If G^2 is large (i.e., Λ is small), the alternative hypothesis is favored; if G^2 is small (Λ is large), the null hypothesis is favored.

The statistical distribution of G^2 is unknown. For many time series models, G^2 has an approximate chi-square distribution under the null hypothesis (see Tong 1990). However, the theorems giving the chi-square distribution do not hold for testing the particular null hypothesis $\mathbf{b} = \mathbf{0}$, because the model under the null hypothesis does not have a stationary distribution (Tong 1990).

Instead, the distribution of G^2 under the null hypothesis can be estimated by simulation (parametric bootstrap). The steps for conducting the PBLR test of the joint density independence model (H_0) vs. the joint density dependence model (H_1) are as follows: (0) Ob-

tain ML parameter estimates for both the null and alternative models ($\hat{\mathbf{a}}_0$, $\hat{\mathbf{\Sigma}}_0$; $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$, $\hat{\mathbf{\Sigma}}$). Calculate the Cholesky decomposition of $q\hat{\mathbf{\Sigma}}_0/(q - 1)$: $q\hat{\mathbf{\Sigma}}_0/(q - 1) = \hat{\mathbf{T}}'\hat{\mathbf{T}}$ (the $q/(q - 1)$ term is a bias correction). Calculate G^2 (Eq. 23). (1) Starting at the observed initial population sizes ($\mathbf{n}_0$), generate a multivariate time series from the null hypothesis model (Eq. 9) using $\hat{\mathbf{a}}_0$ for $\mathbf{a}$ and $\hat{\mathbf{T}}'$ for $\mathbf{T}'$. The simulated time series should have the same number of observations as the data: $\mathbf{n}_0$, $\mathbf{n}_1^*$, $\mathbf{n}_2^*$, . . . , $\mathbf{n}_q^*$. (2) Calculate ML parameter estimates for both the null and alternative models with the simulated data ($\hat{\mathbf{a}}_0^*$, $\hat{\mathbf{\Sigma}}_0^*$; $\hat{\mathbf{a}}^*$, $\hat{\mathbf{b}}^*$, $\hat{\mathbf{\Sigma}}^*$). This step requires iterative ML calculations for the alternative hypothesis model. Calculate by substituting $\hat{\mathbf{\Sigma}}_0^*$ and $\hat{\mathbf{\Sigma}}^*$ in Eq. 23. (3) Repeat steps 1–2 at least 2000 times, thereby obtaining thousands of G^{2*} values. (4) Calculate the estimated P value of the test, denoted, $\hat{P}$, as the proportion of the G^{2*} values that exceed G^2. (5) Reject H_0 if $\hat{P}$ is less than a selected critical value (such as 0.05).

The hypothesis of correlation among populations can also be studied with a PBLR test. Investigators could use such a test to determine whether a multivariate approach to modeling the populations was necessary, or whether the populations were essentially fluctuating independently. The null hypothesis of the test is that $\mathbf{\Sigma}$ is a diagonal matrix. That is, H_0: $\sigma_{ij} = 0$, $i, j = 1, 2, . . . , m$, $i \neq j$, where σ_{ij} is the covariance of the noise terms E_i and E_j in the model (Eq. 4). The alternative hypothesis is that all population pairs have growth rate fluctuations that covary. That is, H_1: $\sigma_{ij} \neq 0$, $i, j = 1, 2, . . . , m$, $i \neq j$. Generally, the base model for this test should be the multivariate logistic (Eq. 6) rather than the multivariate exponential (Eq. 9), because the exponential is included in the logistic as a possible case.

Under the null hypothesis of no correlations, the ML estimates of $\mathbf{a}$ and $\mathbf{b}$ do not require iterative calculations. The ML estimates are identical to CLS estimates and are found from Eqs. 15 and 16 with an identity matrix substituted for $\hat{\mathbf{\Sigma}}$. Under the null hypothesis, the ML estimates of the parameters in $\mathbf{\Sigma}$ are the diagonal elements of $\hat{\mathbf{\Sigma}}$ (Eq. 17): $\hat{\sigma}_i^2 = (\hat{e}_{i1}^2 + \hat{e}_{i2}^2 + . . . + \hat{e}_{iq}^2)/q$. The ML estimates of $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{\Sigma}$ under the alternative hypothesis are calculated in the usual iterative fashion (Eqs. 15–17). The LR statistic G^2 is then calculated with Eq. 23.

To conduct the PBLR test of H_0: $\sigma_{ij} = 0$, $i \neq j$, vs. H_1: $\sigma_{ij} \neq 0$, $i \neq j$, one carries out the steps zero to five, described earlier for the density dependence test, with the new null hypothesis model and parameter estimates substituted throughout. In step zero, the Cholesky decomposition is simplified, because the estimate of the matrix $\mathbf{\Sigma}$ under the null hypothesis is diagonal. Simply calculate $\hat{\mathbf{T}}' = \mathbf{D}([\hat{\sigma}_1, \hat{\sigma}_2, . . . , \hat{\sigma}_m])$ and then calculate the noise vectors in step one with the expression $\mathbf{E} = \hat{\mathbf{T}}'\mathbf{Z}$ (Eq. 8). In step one, the bootstrap time series are generated from the logistic (Eq. 6), using parameters estimated under the null model.

Frequently, a mix of different model structures will

be of interest. Some populations might be growing or declining, while others might be stationary. Or, the fluctuations of some populations might be correlated, while those of other populations are not. The previously described hypothesis tests do not indicate which parameters in **b** or **Σ** are nonzero.

Classical hypothesis testing methods of statistics are not well suited to the problem of identifying which parameters among many should be included in a model. For instance, suppose there are four populations. If one or more parameters in **b** are nonzero, there are 15 possible combinations of zero and nonzero parameter values, meaning 15 possible models from which to choose. A statistical hypothesis test only chooses between a pair of models (or model classes) at a time. Sequential pairwise testing schemes do not necessarily converge to the best model, and are impractical if the tests must be bootstrapped.

Instead, various model selection indexes have been devised for choosing among many candidate models. We find that the Akaike Information Criterion (AIC) has intuitive appeal for sorting out different substructures in the multivariate stochastic logistic model. The AIC is a model selection index consisting of the maximized likelihood penalized for the number of parameters included in the model (Akaike 1973, Sakamoto et al. 1986):

$$\text{AIC} = -2 \ln \hat{L} + 2p. \quad (24)$$

Here $\hat{L}$ is the maximized likelihood function, and p is the number of parameters in the model estimated from the data. A bias adjustment suggested by Bozdogan (1987) resulted in the consistent AIC (CAIC) defined by

$$\text{CAIC} = -2 \ln \hat{L} + p[(\ln q) + 1] \quad (25)$$

where q is the number of observations (or time steps in a time series model). The procedure is to select the model with the lowest CAIC out of a parametric family of models. If the "true" model is contained in the family (or is well approximated by one of the family members), the CAIC procedure provides an (asymptotically) unbiased and consistent procedure for estimating the true model.

For the multivariate stochastic logistic model, the CAIC can be written as (see the Appendix)

$$\text{CAIC} = mq[1 + \ln(2\pi)] + q \ln|\hat{\mathbf{Σ}}| + p[(\ln q) + 1]. \quad (26)$$

The CAIC for the lth sub-model would be

$$\text{CAIC}_l = mq[1 + \ln(2\pi)] + q \ln|\hat{\mathbf{Σ}}_l| + p[(\ln q) + 1] \quad (27)$$

where $\hat{\mathbf{Σ}}_l$ is the estimated variance–covariance matrix of the fluctuations in submodel l. To choose among, say, 15 models (all models of four populations for which one or more parameters in **b** is nonzero), ML estimates are calculated for all 15 models. Those pop-

ulations in a given model growing exponentially have values of b_i set to zero in the log-likelihood function (Eq. 12); maximizations can be accomplished with the Nelder-Mead algorithm. The model with the smallest value of CAIC$_l$, according to this information criterion, represents the best balance of parsimony and data description out of the 15 candidates. An informal rule of thumb is that one can be indifferent about models for which the AIC or CAIC values differ by <2 (Akaike 1973).

Analysis of covariance structure in the noise vector proceeds along similar lines. CAIC values would be calculated for models with different combinations of covariances in **Σ** fixed at zero. The estimate $\hat{\mathbf{Σ}}_l$ in Eq. 27 is calculated using residuals (Eq. 18), except with the appropriate entries in the matrix set equal to zero. The value of p must reflect only parameters estimated, and so should be reduced by the number of parameters (covariances) in **Σ** set to zero.

DIAGNOSTICS AND EVALUATION

The multivariate stochastic logistic model must be used cautiously, like other more familiar types of multivariate analyses. In particular, as in other multivariate models, parameters proliferate rapidly as variables are added. With m populations being described, the multivariate stochastic logistic model has $m(m + 5)/2$ parameters. For instance, to model just four populations takes 18 parameters. With small sample sizes, the estimates of those parameters will have large standard errors, and tests concerning those parameters will have low power. Ecological circumstances might sometimes permit the use of a model substructure with fewer parameters, such as a constant covariance parameter for all population pairs, or a covariance that declines geometrically as a function of distance between populations. A reduced-parameter model might be the only practical solution when the populations are too numerous or the time series are too short.

Attention to diagnostic analyses is essential. The model is a parametric model, and its strengths and failings in describing any given data set should be carefully evaluated. Diagnostic techniques center on the model residuals (Eq. 18). Let $\hat{\mathbf{e}}_t = [\hat{e}_{1t}, \hat{e}_{2t}, \ldots, \hat{e}_{mt}]'$ represent the residuals of the m populations at time t (tth column of the matrix $\hat{\mathbf{R}}$, Eq. 18). If the model fits well, then $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \ldots, \hat{\mathbf{e}}_q$ should be, approximately, vector observations from a multivariate normal distribution. Though the residual vectors are not independent, they can be treated as approximately independent for purposes of diagnostic screening (Tong 1990:323).

Each set of residuals for each population $\hat{e}_{i1}, \hat{e}_{i2}, \ldots, \hat{e}_{iq}$ for the ith population) should be plotted separately and analyzed for autocorrelation and univariate normality. Tong (1990:323) reviews useful white noise tests. A plot of residuals from each pair of populations should reveal an elliptical cloud if the residuals from the pair are bivariate normal. Multivariate normality of
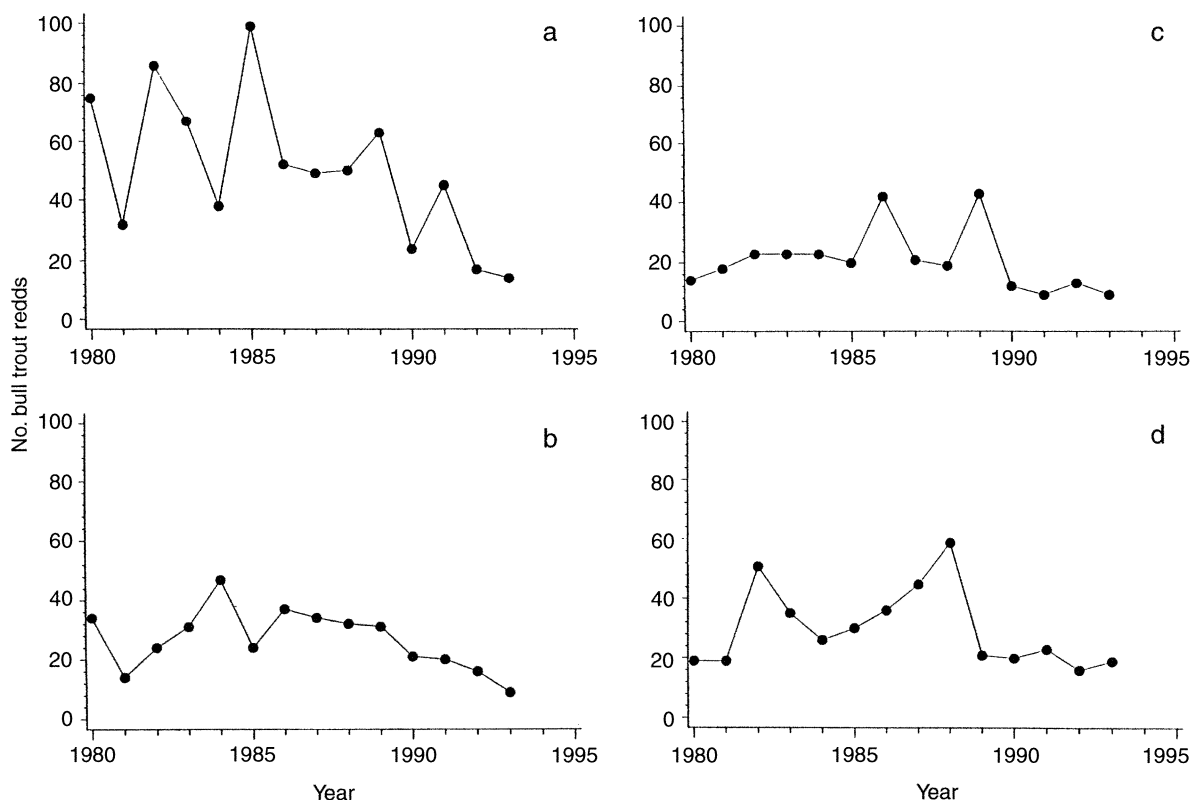
FIG. 2. Bull trout redd counts for four tributaries of the Middle Fork of the Flathead River, Montana, 1980–1993: (a) Morrison, (b) Granite, (c) Lodgepole, (d) Ole. Data were described by Rieman and McIntyre (1993).

the residual vectors can be examined with methods discussed by Cox and Small (1978) and Seber (1984: 148). A quick scan for multivariate outliers can be performed by calculating the quadratic forms: $s_t = \hat{e}_t'\hat{\Sigma}^{-1}\hat{e}_t$, $t = 1, 2, \ldots, q$. The values of s_t should be observations from an approximate chi-square distribution (m df), and those values exceeding, say, the 95th chi-square percentile should be studied carefully.

Statistical methods based on the normal model for the noise vector are not necessarily robust to model misspecification. If the true model is not normal, the parameter estimates and hypothesis tests have unknown statistical properties. The degree to which conclusions based on the normal likelihood remain valid varies from case to case. The consequences of fitting this model to data arising from other chance mechanisms have not been studied. Investigators should therefore use the analyses we have presented as they would use other multivariate techniques: for exploratory purposes, for data summarization, and for discerning structure and patterns in complex, variable systems. Should the residuals not seem sufficiently normal, hypothesis testing can be attempted by using CLS parameter estimates (known to be statistically consistent for a wider class of models) combined with nonparametric bootstrapping. This involves estimating the multivariate noise distribution directly from the residuals using a non-

parametric technique. One such technique is simply to sample residual vectors with replacement. The sampled vectors are used as noise vectors for generating the time series needed to estimate the distribution of G^2 (or other test statistic) under a null hypothesis. Because of the nonparametric nature of this suggested test, longer time series would be desirable.

EXAMPLES

The two systems analyzed in this section illustrate the range of uses of the model. One system involves a functional group containing species that are serious agricultural pests; the other is a metapopulation facing a risk of extinction. The grasshopper abundances in Fig. 1 are used by rangeland managers in three physiographic regions of Montana as indexes of the regional potential for local and large-scale grasshopper outbreaks. The sampling program has been conducted annually since 1951, with the exception of the years 1976, 1981, and 1982. The bull trout (*Salvelinus confluentus*) data (Fig. 2) represent counts of redds (breeding sites) recorded during 1980–1993 on four tributaries of the Middle Fork of the Flathead River, Montana (see Rieman and McIntyre 1993). The rates of migration in these two systems are unknown. The problems arising from missing data, the use of abundance indexes, and the lack of information about migration rates in the

TABLE 1. Maximum-likelihood parameter estimates and diagnostic statistics for the multivariate stochastic logistic model fitted to observations of grasshopper abundances in three regions of Montana: northern glaciated plains (NGP), southern unglaciated plains (SUP), and western mountains (WM). Data were described by Kemp and Dennis (1993).

Region	$\hat{a}$†	$\hat{b}$‡	$\hat{\Sigma}$§			Normality‖	First order¶	Second order#
NGP	0.519	−0.0886	0.370	0.214	0.127	0.85 (0.40)	0.03	0.37
SUP	0.385	−0.0640	(0.726)	0.234	0.111	0.52 (0.60)	0.16	−0.06
WM	0.360	−0.0792	(0.483)	(0.534)	0.186	0.23 (0.82)	−0.08	−0.10

† 3×1 vector of estimates given by Eq. 16.
‡ 3×1 vector of estimates given by Eq. 15.
§ 3×3 symmetric matrix of estimates given by Eq. 17; correlations are given in parentheses below the main diagonal.
‖ Lin-Mudholkar test statistic (and P value) for null hypothesis that residuals have a normal distribution (Lin and Mudholkar 1980, Tong 1990:324).
¶ First-order autocorrelations of residuals. Critical value at 0.05 significance level is 0.33.
# Second-order autocorrelations of residuals. Critical value at 0.05 significance level is 0.33.

grasshopper and trout data sets are common in other systems.

The multivariate stochastic logistic model fits the grasshopper data well, according to diagnostic statistics (Table 1). ML estimates of the parameters were calculated by omitting transitions, that is, pairs of adjacent observations (x_{t-1}, x_t), with missing observations from the likelihood function [Eq. 11]. Thus, the 1975–1976, 1976–1977, 1980–1981, 1981–1982, and 1982–1983 transitions were omitted from the ML formulas [Eqs. 15, 16, and 17]. The remaining 37 transitions nonetheless constitute a long time series by ecological standards. The residuals from each region (Eq. 18) do not depart significantly from univariate normal distributions, according to the Lin and Mudholkar (1980) normality test (Table 1). Normal quantile–quantile plots of the univariate residuals are highly satisfactory (Fig. 3). None of the multivariate residuals (quadratic forms: s_t) exceeds the 95th percentile of a chi-square distribution with 3 df. First- and second-order autocorrelations (Table 1) were estimated using those residuals that were separated by one and two years, respectively. No significant first- or second-order autocorrelation was detected in the residuals for the southern unglaciated plains or the western mountains regions (Table 1). Some detectable but small amount of second-order autocorrelation was noted in residuals from the northern glaciated plains region (Table 1).

The model parameter estimates indicate that the fluctuations of grasshopper abundances in the three regions are noticeably correlated ($\hat{\Sigma}$, Table 1). A PBLR test rejects the null hypothesis that Σ is diagonal (Eq. 23, $G^2 = 40.5$, $\hat{P} < 0.0001$). Likewise, the CAIC values for the null and alternative hypotheses indicate that the off-diagonal parameters in Σ improve the model substantially (CAIC[null] = 202.9, CAIC[alternative] = 176.3). Thus, outbreaks and crashes tend to occur jointly in the three regions. The strong synchrony between the three populations suggests that grasshopper abundances are influenced by environmental factors that operate on a superregional scale. Precipitation and temperature are obvious candidates for such factors and have been implicated in various studies (Lockwood and

Lockwood 1991, Kemp and Cigliano 1994, Cigliano and Kemp 1995).

The multivariate stochastic logistic model fits better than the multivariate stochastic exponential growth model, according to the PBLR test of joint density dependence (Eq. 23; $G^2 = 32.9$, $\hat{P} < 0.0001$). The CAIC (Eq. 25) for the exponential model is much larger than the CAIC for the logistic model (CAIC[0, 0, 0] = 195.4, CAIC[1, 1, 1] = 176.3, where "0" indicates exponential growth model and "1" indicates logistic growth model, for northern glaciated plains, southern unglaciated plains, and western mountains, respectively). The completely density-dependent combination (1, 1, 1) produces a sharp minimum among all possible combinations of logistic and exponential models; the second lowest is CAIC(1, 1, 0) = 181.3. The results suggest that grasshopper abundances are in a state of statistical equilibrium, in which the vector of abundances have long-run behavior described by a trivariate stationary probability distribution. The univariate (marginal) stationary distributions for the three regions resemble right-skewed gamma distributions (Kemp and Dennis 1993).

The bull trout data are also well described by the multivariate stochastic logistic model (Table 2, Fig. 4). The residuals for all four subpopulations do not depart significantly from normal distributions, and do not show any significant first- or second-order autocorrelation (Table 2). None of the multivariate residuals (s_t) exceeds the 95th percentile of a chi-square distribution with 4 df. The fluctuations of subpopulation abundances appear correlated; the estimated correlations range 0.265–0.656 (Table 2). However, a PBLR test of whether or not Σ is diagonal is only marginally significant (Eq. 23; $G^2 = 20.7$; $\hat{P} = 0.067$). The CAIC values for the null (Σ diagonal) and alternative hypotheses are very close (CAIC[null] = 99.5, CAIC[alternative] = 100.2), suggesting that the separate and joint models are approximately equivalent in overall quality.

The multivariate stochastic logistic model fits the bull trout data better than the multivariate stochastic exponential growth model, according to the PBLR test
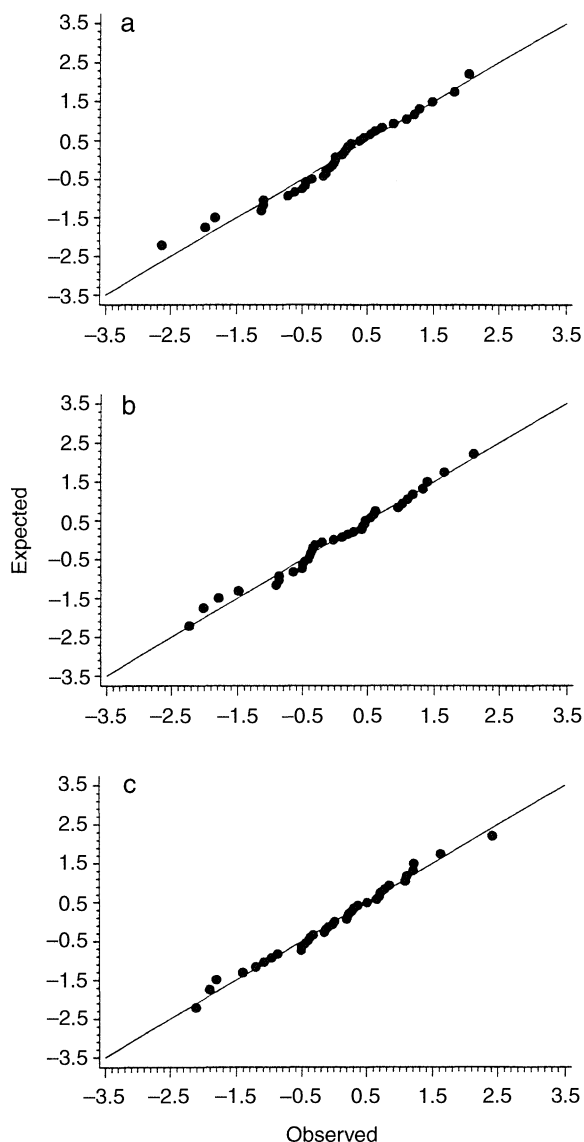
FIG. 3. Normal distribution quantile–quantile plots of residuals resulting from fitting multivariate stochastic logistic model to time series grasshopper abundances in three physiographic regions of Montana: (a) northern glaciated plains, (b) southern unglaciated plains, (c) western mountains. Axis scales are in standard deviation units.

of joint density dependence (Eq. 23; $G^2 = 36.1$; $\hat{P} = 0.006$). Indeed, the CAIC (Eq. 25) for the exponential model is much larger than the CAIC for the logistic model (CAIC[0, 0, 0, 0] = 122.0, CAIC[1, 1, 1, 1] = 100.2, where "0" indicates exponential growth model and "1" indicates logistic growth model, for Morrison, Granite, Lodgepole, and Ole subpopulations, respectively). Among all possible combinations of logistic and exponential models, the completely density-dependent combination (1, 1, 1, 1) produces a sharp minimum CAIC value, beating the next lowest combination by >2 (CAIC[1, 0, 1, 1] = 104.0).

DISCUSSION

Model extensions

The model can be expanded in various ways to make it more applicable in specific situations.

First, the population growth functions can be made more flexible. One candidate function to use as the base growth model for each population is the "θ-Ricker" model (see Turchin 1991). Each term $b_i N_{i(t-1)}$ in Eq. 4 is changed to $b_i[N_{i(t-1)}]^{\theta_i}$, where θ_i is a nonnegative parameter ($i = 1, 2, \ldots, m$). The linear density-dependence expression in the exponential function of Eq. 4 becomes a nonlinear, flexible (concave up or down) function of $N_{i(t-1)}$.

Second, environmental covariates can be incorporated. In one approach, the intercept parameters a_i in Eq. 4 can be written as linear functions of extrinsic environmental factors prevailing in the time interval. Such factors might include precipitation, temperature, or even other species' abundances. Elkinton et al. (1996), for example, construct spatially explicit models of gypsy moth populations in which a covariate was the abundance of white-footed mice. Incorporating environmental covariates is one possible means of attempting to identify the sources of covariation among population growth rates.

Third, sampling variability can be accomodated through a "state-space" model (see Harvey 1989). In one such approach, the observed log-abundance of the ith population would not be X_{it} but rather X_{it} + sampling error. A state-space model might offer improved description when the populations are estimated or indexed, and when the variability resulting from sampling is a substantial proportion of the overall variability.

Finally, migration can be included. We consider the effects of migration in greater detail later in this *Discussion*.

The drawback of all these model extensions is the added parameters. Parameters are costly in statistical inference. The types of parameters described above will typically be estimable only in special circumstances when the data carry explicit information about the processes involved. For instance, the θ parameter in the θ-Ricker model is hard to estimate unless the population displays transient dynamic behaviors far from equilibrium (θ determines the location of the inflection point in a deterministic sigmoid population trajectory). In addition, state-space models often require long time series to estimate properly, because the process variability and the sampling variability are difficult to disentangle. Also, migration parameters are easier to estimate if most locations in the system are initially empty, as in models of the spread of an introduced species (Matis et al. 1996, Lele et al. 1998). The joint density-dependence model we have presented herein has numerous parameters already, and for many situations it

TABLE 2. Maximum-likelihood parameter estimates and diagnostic statistics for the multivariate stochastic logistic model fitted to bull trout redd counts in four tributaries of the Middle Fork of the Flathead River. Data were described by Rieman and McIntyre (1993).

Tributary	$\hat{a}$†	$\hat{b}$‡	$\hat{\Sigma}$§				Normality‖	First order¶	Second order#
Morrison	1.11	−0.0230	0.310	0.142	0.124	0.132	1.08 (0.28)	0.28	0.20
Granite	0.674	−0.0276	(0.644)	0.156	0.107	0.0823	1.15 (0.25)	0.26	0.24
Lodgepole	0.893	−0.0430	(0.542)	(0.656)	0.169	0.0408	−0.20 (0.84)	0.10	0.10
Ole	0.874	−0.0284	(0.635)	(0.557)	(0.265)	0.140	−1.06 (0.29)	0.22	−0.07

† 4 × 1 vector of estimates given by Eq. 16.
‡ 4 × 1 vector of estimates given by Eq. 15.
§ 4 × 4 symmetric matrix of estimates given by Eq. 17; correlations are given in parentheses below the main diagonal.
‖ Lin-Mudholkar test statistic (and P value) for null hypothesis that residuals have a normal distribution (Lin and Mudholkar 1980, Tong 1990:324).
¶ First-order autocorrelations of residuals. Critical value at 0.05 significance level is 0.55.
# Second-order autocorrelations of residuals. Critical value at 0.05 significance level is 0.55.

is likely to represent a practical upper limit on complexity.

Assessing joint risks

The joint density-dependence model proposed herein can be used for assessing risks of attaining different population levels, such as endangered levels or outbreak levels. Associated with every stochastic population model is a set of first-passage distributions. A first-passage distribution is the distribution of time it takes for a population to reach some abundance level v_1 for the first time, starting from some abundance level v_0. The level v_1 can be greater than or less than v_0, depending on the problem under study; different values of v_0 and v_1 lead to different first-passage distributions. The basic procedure for estimating a first-passage distribution is first to estimate parameters in the stochastic population model, and then to obtain by simulation the first-passage distribution for the selected values of v_0 and v_1.

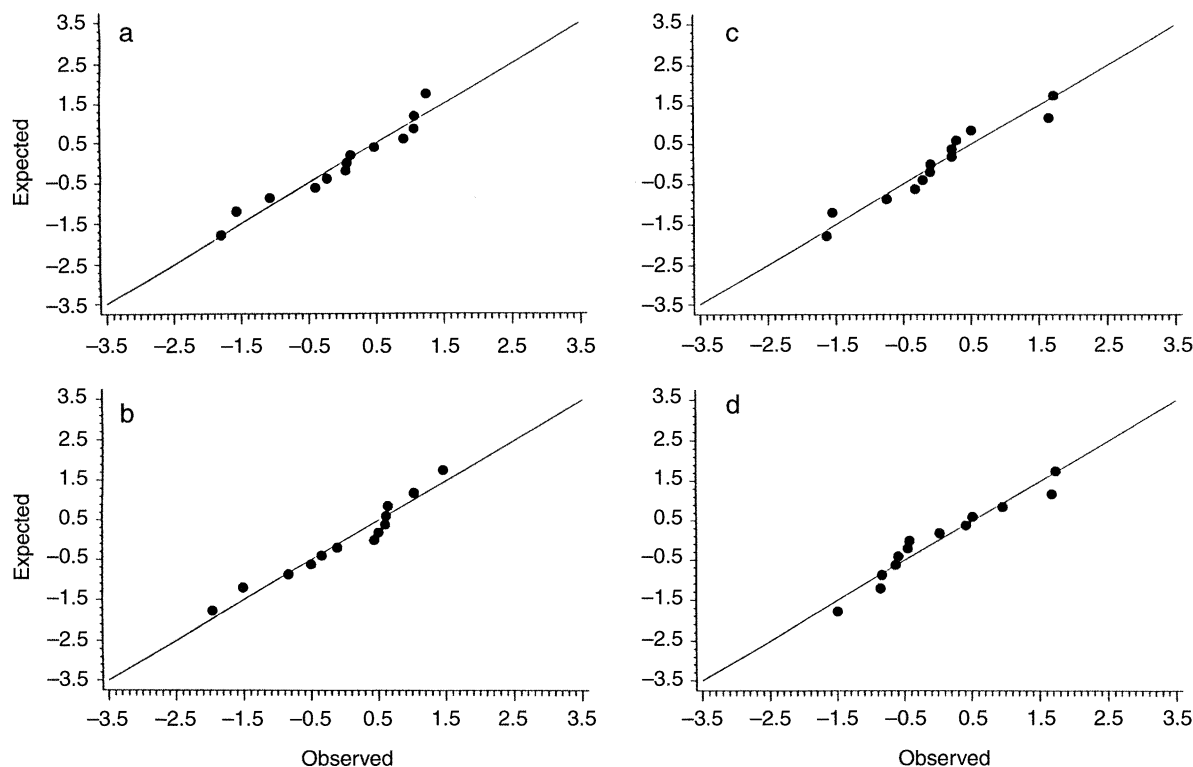Many different first-passage distributions can be defined for multiple populations. We list four ex-

FIG 4. Normal distribution quantile–quantile plots of residuals resulting from fitting multivariate stochastic logistic model to time series bull trout abundances in four tributaries of the Middle Fork of the Flathead River, Montana: (a) Morrison, (b) Granite, (c) Lodgepole, (d) Ole. Axis scales are in standard deviation units.

amples. Suppose v_0 is a vector ($m \times 1$) of starting abundances (one for each population), and v_1 is a vector of target abundances. First, one can study the univariate (marginal) first-passage distributions separately for each population. Second, one can study the joint distribution of the m first-passage times. Third, one can study the distribution of the smallest first-passage time (time for the first population to attain its target) or of the largest first-passage time. Finally, one can study first-passage distributions for the sum of the population abundances. Once parameters in the multivariate stochastic logistic model are estimated, each of these first-passage distributions can be estimated by computer simulation.

Ignoring covariances among subpopulations can cause substantial errors in assessing the jeopardy of metapopulations. Any kind of joint first-passage property is affected by covariances (Goodman 1987a). In particular, under the multivariate stochastic logistic model, positive covariances in Σ enhance the variability of the total (summed) population size. The risk that the total size attains some lower (or higher) target size within a fixed time is therefore increased. Use of a diagonal matrix for Σ (that is, no covariances) would then underestimate the risk of attaining the target.

P. J. den Boer (1968, 1981) suggested that, due to the phenomenon of "spreading of risk," population persistence could increase with fractionation into multiple populations. The spreading-of-risk hypothesis asserts that the variability of the total population is diminished because the fluctuations of the constituent subpopulations tend to cancel each other out in the sum. However, such cancellation occurs only if increases in some subpopulations tend to be compensated for by decreases in other populations, that is, only if there are negative covariances in Σ. Positive covariances (corresponding to increasing spatial synchrony) decrease the expected lifetime of metapopulations (Harrison and Quinn 1989, Gilpin 1990). Because subpopulations are often affected similarly by region-wide environmental conditions (droughts, extreme winters, and so on), it is likely that positive covariances are more prevalent than negative covariances in metapopulations or multiple population systems. However, negative covariances are possible if local habitats respond differently to regional perturbations, or if migration between subpopulations forms a substantial portion of the changes in subpopulation sizes. Extensive data analysis with the multivariate stochastic logistic would help determine what kinds of covariance patterns, if any, are widespread.

Effects of migration

The multivariate stochastic logistic model documented here does not explicitly account for migration between subpopulations. The effects of ignoring migration can be negligible provided the uses of the model are carefully delimited. First, the subpopulation equilibrium abundances should be moderate or large, not small. Births and deaths, as opposed to migration, should be the dominant contribution to the population growth rate. Some pooling of subpopulations might be necessary to use the model in some circumstances. At moderate to large abundance levels, migration will typically represent only a minor portion of the growth rate (but see Stacey and Taper 1992). Migration then becomes important mainly for keeping the gene pool mixed and for allowing subpopulations to rescue each other from extinction. Second, risk assessments ideally should not extend to extinction levels of abundance. The lower abundance level v_1 in the first-passage distribution for any subpopulation should be set above levels at which migration dominates the growth rate. The value v_1 might be set, for instance, at a level of critical concern for policymakers. The practice of using higher targets in viability analysis should be adopted in any case even when using other models; at very low abundance levels, Allee effects (Dennis and Patil 1984, Dennis 1989), extinction vortices (Gilpin and Soulé 1986), and demographic accidents are likely to rival migration in impacts (Goodman 1987b) and complicate viability analysis. When investigators are forced by policy or biological considerations to model metapopulations at low abundances, the likelihood-based methods outlined in this paper might still provide useful parameter estimates to other joint population models that account for such impacts (e.g., Stacey et al. 1997).

LITERATURE CITED

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petrov and F. Csaki, editors. Second international symposium on information theory. Akademia Kiado, Budapest, Hungary.

Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, Illinois, USA.

Aptech Systems. 1993. GAUSS. Aptech Systems, Maple Valley, Washington, USA.

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

Bozdogan, H. 1987. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika **52**:345–370.

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

Cigliano, M. M., and W. P. Kemp. 1995. Spatiotemporal characteristics of rangeland grasshopper (Orthoptera: Acrididae) regional outbreaks in Montana. Journal of Orthoptera Research **4**:111–126.

Cox, D. R., and N. J. H. Small. 1978. Testing multivariate normality. Biometrika **65**:263–282.

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

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

———. 1981. On the survival of populations in a heterogeneous and variable environment. Oecologia **50**:39–53.

———. 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. 1989. Allee effects: population growth, critical density, and the chance of extinction. Natural Resource Modeling **3**:481–538.

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., R. A. Desharnais, J. M. Cushing, and R. F. Costantino. 1995. Nonlinear demographic dynamics: mathematical models, statistical methods, and biological experiments. Ecological Monographs **65**:261–281.

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.

Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. Ecological Monographs **64**:205–224.

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

Elkinton, J. S., W. M. Healy, J. P. Buonaccorsi, G. H. Boettner, A. M. Hazzard, H. R. Smith, and A. M. Liebhold. 1996. Interactions among gypsy moths, white-footed mice, and acorns. Ecology **77**:2332–2342.

Fox, D. R., and J. Ridsdell-Smith. 1995. Tests for density dependence revisited. Oecologia **103**:435–443.

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.

Gilpin, M. E. 1990. Extinction of finite populations in correlated environments. Pages 177–186 *in* B. Shorrocks and R. I. Swingland, editors. Living in a patchy environment. Oxford University Press, Oxford, UK.

Gilpin, M. E., and I. Hanski, editors. 1991. Metapopulation dynamics. Academic Press, New York, New York, USA.

Gilpin, M. E., and M. E. Soulé. 1986. Minimum viable populations: the process of species extinctions. Pages 13–34, *in* M. E. Soulé, editor. Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, Massachusetts, USA.

Goodman, D. 1987*a*. Consideration of stochastic demogra-

phy in the design and management of biological reserves. Natural Resource Modeling **1**:205–234.

———. 1987*b*. The demography of chance extinction. Pages 11–34 *in* M. E. Soulé, editor. Variable populations for conservation. Cambridge University Press, Cambridge, Massachusetts, USA.

Graybill, F. A. 1976. Theory and application of the linear model. Wadsworth, Belmont, California, USA.

Green, P. J. 1984. Iteratively reweighted least squares for maximum likelihood estimation, and some robust and resistant alternatives. Journal of the Royal Statistical Society B **46**:149–192.

Hanski, I. 1991. Single-species metapopulation dynamics: concepts, models, and observations. Pages 17–38 *in* M. E. Gilpin and I. Hanski, editors. Metapopulation dynamics. Academic Press, New York, New York, USA.

Hanski, I., and I. P. Woiwod. 1993. Spatial synchrony in the dynamics of moth and aphid populations. Journal of Animal Ecology **62**:656–668.

Hanski, I., I. Woiwod, and J. Perry. 1993. Density dependence, population persistence, and largely futile arguments. Oecologia **95**:595–598.

Harrison, S., and J. F. Quinn. 1989. Correlated environments and the persistence of metapopulations. Oikos **56**: 293–298.

Harvey, A. C. 1989. Forecasting, structural time series models, and Kalman filter. Cambridge University Press, Cambridge, UK.

Holyoak, M., and J. H. Lawton. 1993. Comments arising from a paper by Wolda and Dennis: using and interpreting the results of tests for density dependence. Oecologia **95**: 592–594.

Kemp, W. P. 1992*a*. Rangeland grasshopper (Orthoptera: Acrididae) community structure: a working hypothesis. Environmental Entomology **21**:461–470.

———. 1992*b*. Temporal variation in rangeland grasshopper (Orthoptera: Acrididae) communities in the steppe region of Montana, USA. Canadian Entomologist **124**:437–450.

Kemp, W. P., and M. M. Cigliano. 1994. Drought and rangeland grasshopper species diversity. Canadian Entomologist **126**:1075–1092.

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

Klimko, L. A., and P. I. Nelson. 1978. On conditional least squares estimation for stochastic processes. Annals of Statistics **6**:629–642.

Lele, S., M. L. Taper, and S. H. Gage. 1998. Statistical analysis of population dynamics in space and time. Ecology **79**, *in press*.

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

Lockwood, J. A., and D. R. Lockwood. 1991. Rangeland grasshopper population dynamics: insights from catastrophe theory. Environmental Entomology **20**:970–980.

Math Works. 1993. MATLAB, Natick, Massachusetts, USA.

Matis, J. H., T. R. Kiffe, and R. Hengeveld. 1996. Estimating parameters for birth–death–migration processes from spatio-temporal abundance data: the case of muskrat spread in the Netherlands. Journal of Agricultural, Biological, and Environmental Statistics **1**:1–20.

Perry, J. N., I. P. Woiwod, and I. Hanski. 1993. Using response-surface methodology to detect chaos in ecological time series. Oikos **68**:329–339.

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

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in FORTRAN: the art

of scientific computing. Second edition. Cambridge University Press, Cambridge, UK.

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. I. Power estimates and estimating errors. Oecologia **78**:1–8.

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

Rieman, B. E., and J. D. McIntyre. 1993. Demographic and habitat requirements for conservation of bull trout. U.S. Forest Service Intermountain Research Station General Technical Report **INT-302.**

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.

Sakamoto, Y., M. Ishiguro, and G. Kitagawa. 1986. Akaike information criteria statistics. D. Reidel, New York, New York, USA.

Seber, G. A. F. 1984. Multivariate observations. John Wiley & Sons, New York, New York, USA.

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.

Stacey, P. B., V. A. Johnson, and M. L. Taper. 1997. Migration within metapopulations: the impact upon local population dynamics. Pages 267–291 *in* I. Hanski, and M. Gilpin, editors. Metapopulation biology: ecology, genetics, and evolution. Academic Press, San Diego, California, USA.

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

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

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

———. 1991. Reconstructing endogenous dynamics of a laboratory population of *Drosophila.* Journal of Animal Ecology **60**:1091–1098.

Turchin, P., P. L. Lorio Jr., A. D. Taylor, and R. F. Billings. 1991. Why do populations of southern pine beetles (Coleoptera: Scolytidae) fluctuate? Environmental Entomology **20**:401–409.

Turchin, P., and A. Taylor. 1991. 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.

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

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

Wolda, H., B. Dennis, and M. L. Taper. 1994. Density dependence tests, and largely futile comments: answers to Holyoak and Lawton (1993) and Hanski, Woiwod and Perry (1993). Oecologia **98**:229–234.

APPENDIX

Here, we derive ML parameter estimates for the multivariate stochastic logistic and the multivariate stochastic exponential growth models (Eqs. 6 and 9). The formulas for the likelihood ratio statistic (Eq. 23) and the consistent Akaike information criterion (Eq. 26) follow directly from the derivation. The derivation is greatly simplified by using the vector derivative notation widely used in statistical theory (e.g., Seber 1984).

We first briefly state three standard results about vector derivatives that are used in our derivation. Let $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_m]'$ be an $m \times 1$ column vector, and let $h(\boldsymbol{\theta})$ be a scalar-valued function of $\boldsymbol{\theta}$. The vector derivative, $\partial h(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$, is defined as a column vector of partial derivatives:

$$\partial h(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = [\partial h/\partial \theta_1, \partial h/\partial \theta_2, \ldots, \partial h/\partial \theta_m]'. \quad (A.1)$$

The vector derivative is essentially the gradient of the function $h(\boldsymbol{\theta})$, except with row and column distinctions of vectors maintained. The first two results stem directly from Eq. A.1 (Seber 1984:530).

Result 1: For $h(\boldsymbol{\theta}) = \mathbf{c}'\boldsymbol{\theta} = \boldsymbol{\theta}'\mathbf{c}$, where $\mathbf{c} = [c_1, c_2, \ldots, c_m]'$ is a column vector of constants:

$$\partial h(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = \mathbf{c}. \quad (A.2)$$

Result 2: For $h(\boldsymbol{\theta}) = \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}$, where $\mathbf{A}$ is a symmetric matrix of constants:

$$\partial h(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = 2\mathbf{A}\boldsymbol{\theta}. \quad (A.3)$$

The third result used in our derivation is proved in most multivariate statistics theory texts (e.g., Seber 1984:523). In the result, "positive definite" refers to a symmetric matrix,

say $\mathbf{A}$, for which $\boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta} > 0$ for all nonzero vectors $\boldsymbol{\theta}$. Also, tr[$\mathbf{A}$] denotes the trace of the matrix $\mathbf{A}$, that is, the sum of the elements on the main diagonal of $\mathbf{A}$.

Result 3: Let $\mathbf{C}$ and $\boldsymbol{\Sigma}$ be positive definite matrices. The scalar function

$$g(\boldsymbol{\Sigma}) = \ln|\boldsymbol{\Sigma}| + \text{tr}[\boldsymbol{\Sigma}^{-1}\mathbf{C}] \quad (A.4)$$

is minimized uniquely at $\boldsymbol{\Sigma} = \mathbf{C}$.

The ML estimates of $\mathbf{a}$, $\mathbf{b}$, and $\boldsymbol{\Sigma}$ jointly maximize the likelihood function $L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ (Eq. 11), or equivalently, the log-likelihood function $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ (Eq. 12). To maximize, we first find the derivatives of $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ with respect to $\mathbf{a}$ and $\mathbf{b}$ and set the derivatives jointly equal to zero. With respect to the vector $\mathbf{a}$, terms in the summation portion of Eq. 12 are in the form

$$(\mathbf{c} - \mathbf{a})'\boldsymbol{\Sigma}^{-1}(\mathbf{c} - \mathbf{a})$$
$$= \mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{c} - \mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{a} - \mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{c} + \mathbf{a}'\boldsymbol{\Sigma}^{-1}\mathbf{a} \quad (A.5)$$

where $\mathbf{c} = \mathbf{y}_t - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}$ is a column vector that does not depend on $\mathbf{a}$. The derivative with respect to $\mathbf{a}$ of a term in the form of Eq. A.5 becomes, from Eqs. A.2 and A.3

$$(\partial/\partial\mathbf{a})(\mathbf{c} - \mathbf{a})'\boldsymbol{\Sigma}^{-1}(\mathbf{c} - \mathbf{a}) = \mathbf{0} - \boldsymbol{\Sigma}^{-1}\mathbf{c} - \boldsymbol{\Sigma}^{-1}\mathbf{c} + 2\boldsymbol{\Sigma}^{-1}\mathbf{a}$$
$$= -2\boldsymbol{\Sigma}^{-1}\mathbf{c} + 2\boldsymbol{\Sigma}^{-1}\mathbf{a}. \quad (A.6)$$

Thus, the derivative of $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ with respect to $\mathbf{a}$ is

$\partial \ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})/\partial \mathbf{a}$

$$= -(1/2) \sum_{t=1}^{q} \{-2\boldsymbol{\Sigma}^{-1}[\mathbf{y}_t - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}] + 2\boldsymbol{\Sigma}^{-1}\mathbf{a}\}$$

$$= \boldsymbol{\Sigma}^{-1}\left\{\sum_{t=1}^{q} [\mathbf{y}_t - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}]\right\} - \boldsymbol{\Sigma}^{-1}q\mathbf{a}. \qquad (A.7)$$

Setting equal to zero, multiplying by $\boldsymbol{\Sigma}$, and solving for the ML estimate $\hat{\mathbf{a}}$ in terms of $\hat{\mathbf{b}}$ gives

$$\hat{\mathbf{a}} = (1/q) \sum_{t=1}^{q} \mathbf{y}_t - \left[(1/q) \sum_{t=1}^{q} \mathbf{D}(\mathbf{n}_{t-1})\right]\hat{\mathbf{b}} \qquad (A.8)$$

or identically, Eq. 16.

With respect to the vector $\mathbf{b}$, terms in the summation portion of Eq. 12 are in the form

$(\mathbf{c} - \mathbf{D}\mathbf{b})'\boldsymbol{\Sigma}^{-1}(\mathbf{c} - \mathbf{D}\mathbf{b})$

$$= \mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{c} - \mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{b} - \mathbf{b}'\mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{c} + \mathbf{b}'\mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{b}. \qquad (A.9)$$

Here $\mathbf{c} = \mathbf{y}_t - \mathbf{a}$ is a column vector that does not depend on $\mathbf{b}$, and $\mathbf{D} = \mathbf{D}(\mathbf{n}_{t-1})$ is a diagonal matrix. The derivative with respect to $\mathbf{b}$ of such an expression is, from Eqs. A.2 and A.3

$(\partial/\partial\mathbf{b})(\mathbf{c} - \mathbf{D}\mathbf{b})'\boldsymbol{\Sigma}^{-1}(\mathbf{c} - \mathbf{D}\mathbf{b})$

$$= \mathbf{0} - (\mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{D})' - \mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{c} + 2\mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{b}$$

$$= -2\mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{c} + 2\mathbf{D}\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{b}. \qquad (A.10)$$

The derivative of $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ with respect to $\mathbf{b}$ then becomes

$$\partial \ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})/\partial \mathbf{b} = \sum_{t=1}^{q} \{\mathbf{D}(\mathbf{n}_{t-1})\boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \mathbf{a})$$

$$- 2\mathbf{D}(\mathbf{n}_{t-1})\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}\}. \qquad (A.11)$$

Substituting $\hat{\mathbf{a}}$ (Eqs. A.9, 15) for $\mathbf{a}$, $\hat{\mathbf{b}}$ for $\mathbf{b}$, $\hat{\boldsymbol{\Sigma}}$ for $\boldsymbol{\Sigma}$, and setting $\partial \ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})/\partial \mathbf{b}$ equal to zero produces

$$\sum_{t=1}^{q} \{\mathbf{D}(\mathbf{n}_{t-1})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}}) + \mathbf{D}(\mathbf{n}_{t-1})\hat{\boldsymbol{\Sigma}}^{-1}$$

$$\times [\mathbf{D}(\bar{\mathbf{n}}) - \mathbf{D}(\mathbf{n}_{t-1})]\hat{\mathbf{b}}\} = \mathbf{0}. \qquad (A.12)$$

Here, $\bar{\mathbf{y}}$ and $\bar{\mathbf{n}}$ are the sample mean vectors defined by Eqs. 13 and 14. By adding $-\mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}}) + \mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}})$ and $-\mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{D}(\bar{\mathbf{n}}) - \mathbf{D}(\mathbf{n}_{t-1})]\hat{\mathbf{b}} + \mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{D}(\bar{\mathbf{n}}) - \mathbf{D}(\mathbf{n}_{t-1})]\hat{\mathbf{b}}$ to each term in the sum in Eq. A.12, we obtain

$$\sum_{t=1}^{q} \{[\mathbf{D}(\mathbf{n}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}}) + \mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}})$$

$$+ \mathbf{D}(\bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{D}(\bar{\mathbf{n}}) - \mathbf{D}(\mathbf{n}_{t-1})]\hat{\mathbf{b}}$$

$$- [\mathbf{D}(\bar{\mathbf{n}}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{D}(\mathbf{n}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\hat{\mathbf{b}}\} = \mathbf{0}. \qquad (A.13)$$

Note that $(\mathbf{y}_t - \bar{\mathbf{y}})$ and $[\mathbf{D}(\bar{\mathbf{n}}) - \mathbf{D}(\mathbf{n}_{t-1})]$, when summed from $t = 1$ to $t = q$, become zero (deviations from average sum to zero). Thus, Eq. A.13 reduces to

$$\sum_{t=1}^{q} [\mathbf{D}(\mathbf{n}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}})$$

$$- \left\{\sum_{t=1}^{q} [\mathbf{D}(\mathbf{n}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{D}(\mathbf{n}_{t-1}) - \mathbf{D}(\bar{\mathbf{n}})]\right\}\hat{\mathbf{b}}$$

$$= \sum_{t=1}^{q} \mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}})$$

$$- \left\{\sum_{t=1}^{q} [\mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{D}(\bar{\mathbf{n}}_{t-1} - \bar{\mathbf{n}})]\right\}\hat{\mathbf{b}} = \mathbf{0}. \qquad (A.14)$$

Solving for $\hat{\mathbf{b}}$ yields

$$\hat{\mathbf{b}} = \left\{\sum_{t=1}^{q} [\mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})]\right\}^{-1}$$

$$\times \sum_{t=1}^{q} \mathbf{D}(\mathbf{n}_{t-1} - \bar{\mathbf{n}})\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{y}_t - \bar{\mathbf{y}}) \qquad (A.15)$$

or just Eq. 15.

The ML equation for $\boldsymbol{\Sigma}$ is found by rearranging the log-likelihood function (Eq. 12), and by noting that a 1×1 matrix equals its trace, that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ when $\mathbf{A}$ and $\mathbf{B}$ are conformable to the multiplications, and that $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$ (Seber 1984:517).

In the following, we write $\mathbf{e}_t = \mathbf{y}_t - \mathbf{a} - \mathbf{D}(\mathbf{n}_{t-1})\mathbf{b}$:

$$\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma}) = -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2) \sum_{t=1}^{q} \mathbf{e}_t'\boldsymbol{\Sigma}^{-1}\mathbf{e}_t$$

$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2)$$

$$\times \sum_{t=1}^{q} \text{tr}(\mathbf{e}_t'\boldsymbol{\Sigma}^{-1}\mathbf{e}_t)$$

$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2)$$

$$\times \sum_{t=1}^{q} \text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{e}_t\mathbf{e}_t')$$

$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2)$$

$$\times \text{tr}\left[\sum_{t=1}^{q} (\boldsymbol{\Sigma}^{-1}\mathbf{e}_t\mathbf{e}_t')\right]$$

$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2)$$

$$\times \text{tr}\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{t=1}^{q} \mathbf{e}_t\mathbf{e}_t'\right)\right]. \qquad (A.16)$$

We see that $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$ is maximized with respect to $\boldsymbol{\Sigma}$ when

$$\ln|\boldsymbol{\Sigma}| + (1/q)\text{tr}\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{t=1}^{q} \mathbf{e}_t\mathbf{e}_t'\right)\right] = \ln|\boldsymbol{\Sigma}| + \text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) \qquad (A.17)$$

is minimized, where $\mathbf{C} = (1/q) \sum_{t=1}^{q} \mathbf{e}_t\mathbf{e}_t'$. By applying the previously quoted minimization result (see Eq. A.4), and by substituting the ML estimates $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, we find that the log-likelihood is maximized at

$$\hat{\boldsymbol{\Sigma}} = (1/q) \sum_{t=1}^{q} [\mathbf{y}_t - \hat{\mathbf{a}} - \mathbf{D}(\mathbf{n}_{t-1})\hat{\mathbf{b}}][\mathbf{y}_t - \hat{\mathbf{a}} - \mathbf{D}(\mathbf{n}_{t-1})\hat{\mathbf{b}}]'. \qquad (A.18)$$

An identical expression for $\hat{\boldsymbol{\Sigma}}$ is given by Eq. 17.

For the multivariate stochastic exponential growth model (Eq. 9), the log-likelihood function is $\ln L(\mathbf{a}, \mathbf{0}, \boldsymbol{\Sigma})$:

$$\ln L(\mathbf{a}, \mathbf{0}, \boldsymbol{\Sigma}) = -(mq/2)\ln(2\pi) - (q/2)\ln|\boldsymbol{\Sigma}| - (1/2)$$

$$\times \sum_{t=1}^{q} (\mathbf{y}_t - \mathbf{a})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \mathbf{a}). \qquad (A.19)$$

This is identical to the log-likelihood of a series of independent observations $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_q$ arising from a multivariate normal distribution with a mean vector of $\mathbf{a}$ and a variance–covariance matrix of $\boldsymbol{\Sigma}$. Indeed, under the multivariate exponential model, the $\mathbf{y}_t$'s are increments of a multivariate Brownian motion process with drift (see Dennis et al. 1991, Dennis and Taper 1994) and are therefore independent observations from a multivariate normal distribution. Note that the $\mathbf{y}_t$'s are not independent under the multivariate logistic (Eq. 6).

ML estimates for the multivariate exponential model are derived in a fashion similar to that of the estimates for the multivariate logistic (set $\mathbf{b} = \mathbf{0}$ in Eqs. A.5, A.6, A.7, A.8, A.16, A.17, and A.18). Because of the multivariate normal

distribution of the $\mathbf{y}_t$'s, the derivation is well-known in statistics (e.g., Seber 1984:60) and yields Eqs. 20 and 21.

The formulas for the likelihood ratio statistic (Eq. 23) and the CAIC (Eq. 26) can be obtained from the rearranged expression for the log-likelihood, $\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma})$, of the stochastic logistic model (Eq. A.16). The maximized likelihoods in G^2 and the CAIC are found by substituting the ML estimates of parameters for $\mathbf{a}$, $\mathbf{b}$, and $\boldsymbol{\Sigma}$ in Eq. A.16. If the ML estimates were obtained under a hypothesized constraint (such as $\mathbf{b} = \mathbf{0}$, or off-diagonal elements of $\boldsymbol{\Sigma}$ are zero), then the constrained parameter values are substituted in Eq. A.16 as well. After substitution, the last expression in Eq. A.16 becomes

$$\ln L(\mathbf{a}, \mathbf{b}, \boldsymbol{\Sigma}) = -(mq/2)\ln(2\pi) - (q/2)\ln|\hat{\boldsymbol{\Sigma}}|$$
$$- (1/2)\mathrm{tr}\left[\hat{\boldsymbol{\Sigma}}^{-1} \sum_{t=1}^{q} \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t'\right]$$
$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\hat{\boldsymbol{\Sigma}}|$$
$$- (1/2)\mathrm{tr}[\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}q]$$
$$= -(mq/2)\ln(2\pi) - (q/2)\ln|\hat{\boldsymbol{\Sigma}}| - (mq/2). \quad (A.20)$$

Here $\hat{\boldsymbol{\Sigma}}$ and $\hat{\mathbf{e}}_t$ represent $\boldsymbol{\Sigma}$ and $\mathbf{e}_t$ with ML estimates plus constrained parameter values substituted. The formulas for G^2 (Eq. 23) and CAIC (Eq. 26) result from Eq. A.20.