International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report IR-06-062

Stochastic analogues of deterministic single-species population models

Åke Brännström (brnstrom@iiasa.ac.at)
David J.T. Sumpter (david.sumpter@zoology.oxford.ac.uk)

Approved by

Ulf Dieckmann
Program Leader, Evolution and Ecology Program

December 2006

Contents

Stochastic analogues of deterministic single-species population models

Short title: Stochastic population models.

Å. Brännström (1,2,*) and D. J. T. Sumpter (1)

1: Mathematics Department, Umeå University, SE-901 87 Umeå, Sweden.
2: Evolution and Ecology Program, International Institute for Applied Systems Analysis, Schlossplatz 1, Laxenburg A-2361, Austria.

*: Author for correspondence (`brnstrom@iiasa.ac.at`).

Abstract

Although single-species deterministic difference equations have long been used in modeling the dynamics of animal populations, little attention has been paid to how stochasticity should be incorporated into these models. By deriving stochastic analogues to difference equations from first principles, we show that the form of these models depends on whether noise in the population process is demographic or environmental. When noise is demographic, we argue that variance around the expectation is proportional to the expectation. When noise is environmental the variance depends in a non-trivial way on how variation enters into model parameters, but we argue that if the environment affects individual fecundity then variance is proportional to the square of the expectation. We compare various stochastic analogues of the Ricker map model by fitting them, using maximum likelihood estimation, to data generated from an individual-based model and the weevil data of Utida. Our demographic models are significantly better than our environmental models at fitting noise generated by population processes where noise is mainly demographic. However, the traditionally chosen stochastic analogues to deterministic models—additive normally distributed noise and multiplicative lognormally distributed noise—generally fit all data sets well. Thus the form of the variance does play a role in the fitting of models to ecological time series, but may not be important in practice as first supposed.

Keywords: population models, stochastic population models, Ricker model, first principles.

1 Introduction

The goal of modeling the dynamics of animal populations is to understand how population change arises from the interplay of environmental forcing, density dependent regulation and inherent stochasticity (Bjørnstad & Grenfell, 2001; Hilborn & Mangel, 1997). Population fluctuations can arise from variation in the environment over time (Walther et al., 2002; Saether, 1997), from intrinsic 'demographic' stochasticity arising from variation in the number of offspring produced per individual (Bartlett, 1960; Royama, 1992), from errors in observations (Valpine & Hastings, 2002) and from deterministic non-linear dynamics, such as cycles and chaos (May, 1976; Turchin, 2003; Berryman, 1999). The combination of these effects, coupled with the fact that ecological time-series are often short, makes the construction of predictive mathematical models notoriously difficult (Turchin & Taylor, 1992). Indeed, the central requirement in developing mathematical models of population ecology is that stochastic and deterministic factors can be weighted appropriately.

Recent theoretical work on the construction of deterministic models has concentrated on the basis of 'laws' or 'first principles' of population growth (Berryman, 1999; Turchin, 2003). For example, that populations grow exponentially in the absence of environmental constraints is a simple consequence of reproduction. Such theory is most thoroughly developed for single-species, discrete and unstructured population models with first order feedback, i.e. of a form

$$a_{t+1} = f(a_t) = a_t g(a_t) \tag{1}$$

where a_t is the population on generation t and g has a natural interpretation as the net reproduction rate per individual for the population. Reasoning from a few simple, biologically reasonable postulates we can establish properties of g. For example, density independent reproduction is equal to $g(0)$ while, since populations cannot grow indefinitely, $g(a) < 1$ for sufficiently large a. Such reasoning results in the definition of two qualitatively different forms for f, compensatory models—where f is increasing—and overcompensatory models—where f obtains a maximum and then decreases (Bellows, 1981). A typical example of a compensatory model is the Beverton-Holt model, $g(a) = k_1/(k_2 + a)$ and overcompensatory the Ricker map, $g(a) = k_1 \exp(-a/k_2)$.

In deriving models from first principles of population dynamics, stochasticity in the population is not usually considered. However, when it comes to fitting these theoretical models to data, random variation—which is an undoubtable characteristic of most time series of natural populations—can no longer be ignored. The simplest manner in which noise can be incorporated is through additive normally distributed noise, i.e.

$$a_{t+1} = f(a_t) + \sigma\epsilon_t \tag{2}$$

where σ is the constant of standard deviation and the ϵ_t will be assumed throughout this paper to be independent random variables, normally distributed with mean 0 and variance 1 (see for example Hilborn & Mangel, 1997; Solow, 1998). For the Ricker map it is natural to use a multiplicative lognormally distributed noise term, such that

$$a_{t+1} = ba_t \exp(1 - a_t/K + \sigma\epsilon_t - \sigma^2/2) \tag{3}$$

This formulation has the convenient property that linear regression on the $\{a_t\}$ plotted against $\{\ln(a_{t+1}) - \ln(a_t)\}$ gives not only fitted values for b and K, but also the standard deviation, σ, which is equal to the error sum of squares from the regression (see for example Berryman, 1999).

Noise in population time series can be classified as demographic (arising from variation in the number of offspring produced per individual), environmental (arising from variation in the environment over time) and observational (arising from errors in measuring the population size). Despite these numerous ways in which stochasticity can arise in field data, it appears to be somewhat of an after-thought in the derivation of models such as equations 2 and 3. Indeed, the main influence on the choice of stochastic population model seems to be the ease with which it can be fitted to data. Since maximum likelihood techniques can be used to fit generalized models to data, the ease of using linear regression should not be a restriction on the type stochastic population models developed in ecology. Stochastic analogous of deterministic population models can be developed from first principles on the basis of the types of noise present in time series.

In order to address the need for models which more accurately capture demographic and environmental noise, Engen et al. (1998) introduced general definitions of demographic and environmental variance, as well as demographic covariance. A central point arising from their study is that environmental and demographic variance, just like the expectation, are functions of population size. Since both the deterministic and the stochastic components of a time series model must arise from the same set of reproductive actions and interactions between individuals in the population, we expect some relationship between the two. However, the exact nature of this relationship depends on the type of interactions as well as the relative importance of demographic and environmental noise. The aim of this paper is to develop practical stochastic population models of discrete, single-species population dynamics with first-order feedback, which combine appropriate deterministic and stochastic components.

2 Stochastic models from first principles

Consider a population of A_t individuals with discrete non-overlapping generations indexed by t. Let R_1, ..., R_{A_t} be the individual contributions to the population at the next generation and Z_t be a random vector describing various environmental factors (Athreya & Karlin, 1971). We assume that the R_i are identically distributed random variables, with expectation and variance conditioned on A_t and Z_t given by $\mu(A_t, Z_t)$ and $\sigma^2(A_t, Z_t)$ respectively.

The population at generation $A_{t+1} = R_1 + \ldots + R_{A_t}$ depends on the population density and the population abundance on the previous generation as well as the environmental conditions during the year, represented by A_t and Z_t respectively. If $Z_t = z_t$ and $A_t = a_t$ the expected population on the next generation is

$$\mathrm{E}[A_{t+1}|A_t = a_t, Z_t = z_t] = \mathrm{E}[R_1 + R_2 + \ldots + R_{a_t}] = a_t\mu(a_t, z_t) \tag{4}$$

From this point on, we will assume that $A_t = a_t$ and not state this explicitly. Unlike the expectation, the variance is only additive if the random variables are uncorrelated,

$$\mathrm{Var}[A_{t+1}] = \sum_{i=1}^{a_t} \mathrm{Var}[R_i] + \sum_{i \neq j} \mathrm{Cov}[R_i, R_j] \tag{5}$$

We will now give plausible forms of the variance as a function of a_t, first in the presence of demographic stochasticity only, and later for a combination of demographic and environmental stochasticity.

2.1 Demographic variance

In a constant environment, we have no environmental variance, and $Z_t = z_t = z$. If the correlation between the number of offspring produced by two individuals is $\rho(a_t, z)$ then from Equation 5

$$\mathrm{Var}[A_{t+1}] = \sigma^2(a_t, z)(a_t + \rho(a_t, z)a_t(a_t - 1)) \tag{6}$$

where $\rho(a_t, z) = \mathrm{Cov}[R_i, R_j]/\mathrm{Var}[R_i]$. Since σ is bounded, it follows that if ρ is at most of order a_t^{-1} the variance grows at most linearly with total population size. Correlation of order a_t^{-1} is not an unreasonable assumption since environmental and temporal constraints force most individuals to interact with only a fraction of the total population during a lifetime.

If we assume that each individual produces a binomially distributed number of offspring where each of b potential offspring is born and survives the year with probability p, then $\mu(a_t, z) = pb$ and $\sigma^2(a_t, z) = p(1-p)b$. Provided a_t is sufficiently large, we may assume that the correlation

$\rho(a_t, z) = (C_\rho - 1)/(a_t - 1)$. Then from equations 4 and 6, we get the expectation and the variance of A_t

$$\mathrm{E}[A_{t+1}] = pba_t \text{ and } \mathrm{Var}[A_{t+1}] = C_\rho p(1-p)ba_t \tag{7}$$

For our purposes, the important point to note here is how the expectation, pba_t, and the variance, $C_\rho p(1-p)ba_t$, are related to each other: both are expressed in terms of the same parameters, b and p, and both the expectation and the variance scale linearly with the population on the previous generation A_t. Specifically, $\mathrm{Var}[A_{t+1}]$ is proportional to $\mathrm{E}[A_{t+1}]$. This relationship between mean and variance contrasts with the model given by Equation 2 in the introduction, where the variance is constant for all a_t.

The model we have constructed here is a linear birth-death process with no density dependence. In order to construct general, non-linear stochastic models we now introduce density dependence by assuming that the number of offspring produced, b, depends on a_t. Specifically, given any deterministic population model, $f(a_t) = a_t g(a_t)$, we let $b(a_t) = g(a_t)/p$. In this formulation, $g(a_t)$, is interpreted as the expected net reproduction rate per individual, while $b(a_t)$, although not necessarily an integer, is interpreted as the gross reproductive gain per individual. Substituting into Equation 7 we get

$$\mathrm{E}[A_{t+1}] = f(a_t) \text{ and } \mathrm{Var}[A_{t+1}] = Cf(a_t) \tag{8}$$

where $C = C_\rho(1-p)$. Thus $\mathrm{Var}[A_{t+1}]$ is proportional to $f(A_t)$, the expected population in the next generation. Thus, if a population is expected to decrease, the variance in the outcome is smaller than were it to increase. Even though Equation 8 was derived under rather specific assumptions we hope that it captures the essential features of demographic stochasticity in a broad range of single-species populations.

2.2 Environmental variance

In the previous section we assumed the environment to be constant. We remove this assumption and again represent the environment with a random vector Z_t. With R_i as before, we follow Engen et al. (1998) and write R_i as a sum of three parts, $R_i = \mathrm{E}[R_i] + W_e + W_d^i$ where the first term is the expectation if we have no information about the environment, i.e. the deterministic part, the second is the deviation from this expectation due to environmental factors, i.e. Z_t and the third being the difference between the actual outcome R_i and the expectation in the realized environment this year:

$$R_i = \mathrm{E}[R_i] + (\mathrm{E}[R_i|Z_t] - \mathrm{E}[R_i]) + (R_i - \mathrm{E}[R_i|Z_t])$$

Here, $\mathrm{E}[R_i|Z_t] = \mu(a_t, Z_t)$ is the conditional expectation given Z_t. It can be shown that the two stochastic terms are uncorrelated (Engen et al., 1998). Then

$$\mathrm{Var}[A_{t+1}] = \mathrm{Var}\left[\sum_{i=1}^{a_t} R_i\right] = a_t(\sigma_d^2(a_t) - \tau(a_t)) + a_t^2(\sigma_e^2(a_t) + \tau(a_t)) \qquad (9)$$

where $\sigma_d^2(a_t) = \mathrm{Var}[W_d^i]$, $\sigma_e^2(a_t) = \mathrm{Var}[W_e]$ and $\tau(a_t) = \mathrm{Cov}[W_d^i, W_d^j]$. When populations are sufficiently large, there is usually little interaction between two randomly selected individuals and the demographic co-variance τ is small (see Engen et al., 1998, for some examples). We thus assume for now that $\tau = 0$ and Equation 9 takes the form

$$\mathrm{Var}[A_{t+1}] = a_t\sigma_d^2(a_t) + a_t^2\sigma_e^2(a_t) \qquad (10)$$

giving independent terms for demographic and environmental stochasticity.

To construct a stochastic analogue to a given deterministic model,

$$a_{t+1} = f(a_t) = a_t g(a_t)$$

under the assumption that noise is environmental, we need to find how $\sigma_e^2(a_t)$ is related to $g(a_t)$. It seems reasonable to assume that environmental stochasticity affects one or more of the parameters in the deterministic model. For example, the non-dimensionalised Ricker model can be written as $g(a_t) = b\exp(-a_t)$, where b is the density-independent growth rate. Assume that this rate is determined by and actually equal to the environmental variable Z_t, so that $\mathrm{E}[R_i|Z_t] = Z_t \exp(-a_t)$ and, consequently, the expected net reproduction per individual is $g(a_t) = \mathrm{E}[Z_t]\exp(-a_t)$. Then

$$\sigma_e^2(a_t) = \mathrm{Var}\left[\mathrm{E}[R_i|Z_t]\right] = \exp(-2a_t)\,\mathrm{Var}[Z_t] = Cg(a_t)^2 \tag{11}$$

where C is a constant depending on $\mathrm{E}[Z_t]$ and $\mathrm{Var}[Z_t]$. Assuming that there is no demographic variance, that is $\sigma_d^2(a_t) = 0$, Equation 9 and 11 imply that

$$\mathrm{Var}[A_{t+1}] = Cf(a_t)^2 \tag{12}$$

Thus if environmental stochasticity affects the density-independent growth rate, the variance is proportional to the square of the expectation.

Many deterministic models of population dynamics incorporate a carrying capacity, a parameter related to equilibrium population density. Unlike the density-independent growth rate, there does not seem to be any general relationship between $\sigma_e^2(a_t)$ and $g(a_t)$ when environmental change affects the carrying capacity, and the relationship therefore has to be determined on a case-by-case basis. For example, the Ricker model can be written with $g(a_t) = b\exp(-\ln(b)a_t/K)$, where K is the carrying capacity. If we assume that the environmental variable Z_t equals the fluctuations in carrying capacity, so that $\mathrm{E}[Z_t] = 0$, we can write $\mathrm{E}[R_i|Z_t] = b\exp(-\ln(b)a_t/(K + Z_t))$. Using the linear approximations $(1+x)^{-1} \approx 1 - x$ and $\exp(x) \approx 1 + x$ we get

$$\sigma_e^2(a_t) = \mathrm{Var}\left[b\exp\left(-\frac{a_t \ln b}{K}\frac{1}{1 + Z_t/K}\right)\right] \approx Ca_t^2 g(a_t)^2 \tag{13}$$

where C is a constant depending on $\mathrm{Var}[Z_t]$, K and b. Equation 13 gives a good description of the variance provided the quotient Z_t/K is small. Assuming that there is no demographic variance, that is $\sigma_d^2(a_t) = 0$, Equation 9 and 13 implies that

$$\mathrm{Var}[A_{t+1}] = Ca_t^2 f(a_t)^2 \tag{14}$$

is the environmental variance.

The difference between equations 13 and 14 highlight a difficulty that environmental noise presents. The relationship between variance and expected population is dependent on how environmental change affects the population. However, this relationship can be determined provided that the important environmental driving forces are well understood. In the case that environmental stochasticity affects density-independent growth rate or maximum fecundity of individuals, the relationship is largely independent of $g(a_t)$ and we can assume that the variance is proportional to the square of the expectation.

2.3 Stochastic population models

The preceding discussion argued that for many biologically realistic deterministic models of population dynamics, demographic noise can be introduced by the addition of a noise term with variance proportional to $f(a_t)$, while environmental noise affecting the density-independent growth rate can be introduced by the addition of a noise term proportional to $f(a_t)^2$. As yet, we have not discussed the distribution underlying environmental and demographic noise. While a normal distribution is natural in view of the central limit theorem, a lognormal distribution may be favored on biological grounds, as it is always positive. We thus examine models with both distributions. Table 1 lists the demographic and environmental stochastic models we now investigate, along with other models taken from the literature.

The demographic models D1 and D2 have approximately the same variance, with the variance of D1 proportional to the expectation, and a normal and lognormal distribution respectively. The third demographic model in Table 1, Model DW, has its origins in diffusion approximations of continuous population dynamics (Engen et al., 1998). It was used in an extended form by Sæther et al. (2000) for fitting bird populations and Bjørnstad &

Grenfell (2001) suggested it could be used in fitting of discrete generation ecological time series provided the dynamics is not overcompensatory.

The models representing environmental stochasticity, models E1 and E2, both have variance proportional to $f(a_t)^2$. Model E1 is normally distributed and Model E2 is lognormally distributed. Model N1, is commonly used normally distributed additive noise where the variance is constant, independent of population size. It can be classified neither as environmental nor demographic, since it predicts that variance is independent of population size. It is however a plausible model of observation error, which can be independent of a_t (Solow, 1998).

3 Testing the models

We fit the models in Table 1 to time-series first from an individual-based model and then from experimental data. The reason for fitting to data from an individual-based model is that we can control the proportion of demographic and environmental noise in the time series, while retaining some degree of biological realism. In the absence of environmental stochasticity, we may also compare the results with those obtained when fitting a theoretically derived model. The model we consider is one of scramble competition for discrete resources, originally introduced by Sumpter & Broomhead (2001) for the parsitism of honey bee brood cells by *Varroa* mites.

In the model, A_t individuals are distributed randomly with uniform probability over n resource sites. Reproduction then takes place at each site independently, such that if the number of individuals at site i after distribution is determined by the random variable C_i, then

$$\phi(C_i) = \begin{cases} b_{\min} + Z_t & \text{if } C_i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

gives the number of individuals passing on to the next generation, i.e.

$A_{t+1} = \sum_{i=1}^{n} \phi(C_i)$. Environmental noise is represented by $Z_t \sim \text{Bin}(b_r, p)$, a random variable determining the number of individuals produced per site in the absence of competition. b_{min} is the minimum number of individuals produced, and we define $b = b_{\text{min}} + pb_r$ to be the expected number of individuals per site. Note that the demographic noise in this model is entirely due to the distribution process, i.e. the distribution of the C_i. Figure 1 shows typical time-series from the model for three different values of b, in the absence of environmental noise, i.e. $p = 1$. As b increases, the population dynamics changes from stable ($b < 8$) to periodic ($8 < b < 15$) and chaotic ($b > 15$).

In the case where there is no environmental noise Johansson & Sumpter (2003) have shown that the population dynamics of the model are well-approximated by the stochastic dynamical system

$$A_{t+1} = n\Phi\left(\frac{A_t}{n}\right) + \sqrt{nv\left(\frac{A_t}{n}\right)}\epsilon_t \tag{16}$$

where $\Phi(x) = bx\exp(-x)$ is the well-known Ricker map, $\epsilon_t \sim \text{N}(0, 1)$ is a normally distributed random variable with mean 0 and expectation 1, and $v(x)$ is defined by

$$v(x) = b\Phi(x)(1 - e^{-x}) + \Phi(x)^2(1 - x) \tag{17}$$

We use this function $v(x)$ to define Model T1 in Table 1. In the absence of environmental noise, T1 serves as a ‘benchmark’ against which the performance of the other models can be measured.

3.1 Demographic noise

We fitted each of the models in Table 1 with $f(a) = ba\exp(a/n)$ to time-series of length 20 sampled after 100 generations from the individual-based model with $n = 500$ sites. By setting $p = 1$ we ensured that all stochasticity generated for the individual-based model was demographic. All three parameters—b, n and s—were estimated using the method of maximum-likelihood (see Appendix A), and the estimates are written as $\hat{b}$, $\hat{n}$ and $\hat{s}$ respectively. We repeated the simulations we now describe for $n = 100$ sites, thus introducing more noise. The results were similar though the relative difference between the log-likelihood of the various models were reduced.

Figure 2 shows the difference between the estimated value $\hat{b}$ and the actual value b used in the simulations, for models fitted to data from simulations with a range of b values. Model T1 gives, for nearly all values of b, the best fit to the simulation data, in the sense of being least biased in the estimate $\hat{b}$. For $b < 8$ the individual-based model has stable population dynamics (see Figure 1a). In this case, all models fit approximately equally well and exhibit the same consistent bias in parameter estimate. For $b \geq 8$, individual-based model produces periodic then chaotic population dynamics (see Figure 1b and c) and differences appear between the fit of the models. Contrary to our predictions about its applicability, the model N1 gives values of $b - \hat{b}$ closest to that of T1, for all b. It is however, only slightly better than the demographic models, D1 & D2, which produce almost identical estimates $\hat{b}$ to each other. The environmental models E1 and E2 are somewhat worse than D1 & D2. $\hat{b} - b$ for these models differs in sign, but not greatly in magnitude. The model DW is inaccurate for $b > 8$, consistently over-estimating b.

Since all our models are based on the same deterministic framework, we can compare their fitted variance to the theoretically derived variance $v(a_t/n)$ (see Equation 17). Figure 3 shows variance as a function of population size for models in Table 1 fitted to time-series from the individual-based model with $b = 10$ and $n = 500$, and the theoretically derived variance $nv(a_t/n)$ for these parameters. Models D1 and D2 are similar in variance and best approximate $nv(a_t/n)$. Despite its good fit to the simulation data, model N1 does not have a variance that lies particularly close to that of $nv(a_t/n)$. The variance of DW is $f(a_t)^2(\exp(s^2/a_t) - 1)$ which tends to infinity as the population goes to 0. Furthermore, Model DW does not approximate the theoretically derived variance well, which may explain the large bias in the estimate $\hat{b}$.

When fitting models with the same number of parameters to a time-series using the method of maximum likelihood, the likelihood can be used to select one of the models as the best fit. Thus, in order to distinguish the ability of the various models to fit simulation data we can compare the log-likelihood. This is done in Figure 4. For $b \geq 8$, model T1 is consistently the most likely model over a large number of trials. The likelihood of models D1 and D2 lie nearest to that of T1, followed by N1, with models E1 and E2 fitting less well. Model DW is consistently the least likely model. Figure 4 also shows the probability that model D1 is selected over model E1 (i.e. the proportion of fittings that D1 has a higher log-likelihood than E1). Despite the absence of environmental noise, model E1 has a slightly higher probability of being selected than D1 for $b < 8$. As b increases, however, the demographic models are more likely to be selected, such that for $b > 12$ there is more than 90% probability of selecting Model D1 over model E1.

3.2 Environmental noise

By setting $p < 1$ in the individual-based model we introduced environmental noise into the time series. For a fixed population a_t the individual-based model's demographic variance scales with n, while the environmental variance scales with n^2 (see equations 16 and 10).

Thus by increasing n we increase the relative proportion of environmental noise. We compared the models D1 and E1 by fitting them to time-series generated by the individual-based model with increasing values of n.

Figure 5 shows the probability that Model D1 has higher log-likelihood than E1 as a function of n, with $p = 0.5$ and $b_r = 8$. When $b = 5$ (i.e. $b_{\text{min}} = 1$) the individual-based model has stable population dynamics and there is little difference between the models. However, as in the case of demographic noise, model E1 is selected more often than D1, even for small n. When the population dynamics are periodic ($b = 12$), the probability of selecting D1 over E1 decreases with n. A similar effect is observed for chaotic population dynamics ($b = 16$), with E1 being chosen over D1 only for $n > 2000$. In this last case, a typical time-series will contain many transitions from very large to very small populations (see Figure 1c).

3.3 Experimental data

To test the techniques against experimental laboratory data, we model the population dynamics of the southern cowpea weevil, *Callosobruchus maculatus* (Utida, 1967). The population data consists of four carefully arranged experiments where weevils were kept in a constant environment. The population was censured and supplied with fresh food every 25 days, approximately the start of each new generation. The advantage of fitting our models to this data set is that it should be free of environmental noise and observation error, thus allowing us to test the applicability of our demographic noise models.

By calculating the autocorrelation function and the partial rate correlation function we could confirm that the weevils' dynamics are driven by first-order feedback (Berryman, 1999). We thus modelled the expectation with the theta-Ricker map, $f(a_t) = nba_t \exp(a_t^\theta/n^\theta)$ and fitted the stochastic population models in Table 1 using the method of maximum likelihood. Parameter estimates for the model with highest likelihood in each of the four replicates are given in Table 6. In three out of four cases an environmental model gave the best fit to the data, while in the other case the observation error model, N1, gave the best fit. Although the demographic models had only slightly lower log-likelihoods, they failed to provide the best fit to any of the laboratory based data.

4 Discussion

A recent trend in ecology is that general ‘principles’ or ‘laws’ of population dynamics (Berryman, 1999; Turchin, 2003) are starting to gain acceptance. The ‘principles’ or ‘laws’ that have been considered so far are all concerned with the expected population change. In this paper we search for corresponding principles for stochasticity arising from demographic and environmental sources. If firm principles of population change can be established, it would significantly facilitate the reconstruction of underlying ecological processes from experimental data (Jonzén et al., 2002).

The link between deterministic and stochastic models is not as strong as one may initially think. From a stochastic model we can recover a deterministic skeleton in a number of ways; usually by taking the expectation or the mode. The opposite task, to construct a stochastic population model incorporating a fair amount of biological realism from a deterministic skeleton, is considerably more difficult. The ‘first-principles’ approach taken here is to state assumptions about individual interactions and derive the stochastic population dynamics as a consequence. Such reasoning allowed us to derive a generic model of demographic stochasticity, where the variance is proportional to the expectation (Model D1 in Table 1). Furthermore, building on previous foundational work by Engen et al. (1998), we constructed a generic model for populations with intrinsic growth-rate affected by environmental stochasticity, such that the variance is proportional to the square of the expectation (Model E1 in Table 1).

In the absence of environmental stochasticity, our generic model of demographic noise fitted data from the individual-based model nearly as well as the theoretically derived ‘benchmark’. The demographic models were significantly better than the environmental models, E1 and E2, when fitting to periodic or chaotic time series. This is not surprising since the variations in population size brought about by the underlying periodic or chaotic dynamics cause profound changes in the variance, resulting in greater difference in likelihood. Somewhat surprising however was the minor influence of the distribution when compared to the importance of the relationship between expectation and variance. Both normally and lognormal distributed stochasticity fitted the data equally well.

Most ecological time-series are likely to contain significant amounts of both demographic and environmental stochasticity. As the size of a population increases, the dominant noise component changes from being demographic to environmental. Our models captured this effect: as we increased the number of resource sites the best fitting model changed from being demographic to environmental. Even when there are a relatively large number of resource sites, up to 2000, the demographic model still proved the best fit when population dynamics were chaotic. This is because the uncertainty in the outcome of transitions from large populations to small populations is mainly demographic. A time-series can thus be viewed as a series of transitions which differ in the amount of demographic and environmental stochasticity. In an attempt to capture all transitions, we have made preliminary attempts (not shown here) to fit models that include both demographic and environmental components. The maximum likelihood of these models did not converge reliably, and the only known robust approach to separating environmental and demographic noise is estimation of demographic stochasticity from individual reproductive data (Sæther et al., 2000).

The difficulty of applying these ideas in practice was evident when we constructed stochastic population models from Utida's cowpea weevil data. Although the weevils were kept in a constant environment, the environmental models, E1 and E2, had the highest likelihood in all but one case where Model N1 fitted the data best. A possible explanation for this could be demographic covariance. With relatively few individuals in a small space, it is likely that reproductive success is highly correlated between individuals. Given the way demographic covariance would manifest itself (as $\rho > 1/a_t$ in Equation 6 or $\tau > 0$ in Equation 9) this could conceivably lead to a variance-abundance scaling relationship resembling that of the environmental models. Were the experiment to be conducted with a large group in a large area this effect would likely disappear as two given individuals from a large group are not likely to interact much, thereby reducing the covariance.

Model N1, additive normally distributed noise, fitted the data best in one of the four time-series and was the best model at estimating b from the individual-based model, although not with the highest log likelihood. Despite its versatility and its common application in fitting ecological time series, we could not find a first-principles argument, or any other justification on biological grounds, for using model N1 to fit to data free from measurement error. Similarly, the other commonly used stochastic model, model E2, also proved a good choice when fitted to the laboratory data. In light of the widespread use of these two models in the literature, these conclusions should be comforting.

Less comforting is the implications of model N1's good performance for our 'first principles' approach. It may be argued that if model N1 fitted the data well, there is no need to consider more complicated demographic models. If the sole aim of an ecological investigation is to fit parameters of the deterministic skeleton, then this view may hold some truth. However, models with noise independent of population size overestimate stochastic effects for small populations and underestimate them for large populations. In one highly important ecological endeavor—the estimation of extinction time—such errors would have serious consequences for predictions. Our current results have demonstrated the robust performance of both demographic and environmental stochastic models. The further development of these techniques—to include age-structured populations, population dynamics with higher order feedback, and environmental change correlated between years—will ultimately increase the certainty with which extinction events can be predicted.

Acknowledgements

We thank Anders Johansson and Kjell Leonardsson for useful discussion and Paul Glenndinning for comments on an earlier draft. We thank the Swedish Foundation for International Cooperation in Research and Higher Education and Umeå University for providing funding for this research.

References

Athreya, K. B. & Karlin, S. (1971). On branching processes with random environments: Extinction probabilities. *The Annals of Mathematical Statistics*, **42**, 1499–1520.

Bartlett, M. S., editor (1960). *Stochastic Population Models in Ecology and Epidemiology*. Metheuen, London.

Bellows, T. S. (1981). The descriptive properties of some models for density dependence. *Journal of Animal Ecology*, **50**, 139–156.

Berryman, A. A. (1999). *Principles of Population Dynamics and Their Applications*. Stanley Thornes Ltd.

Bjørnstad, O. N. & Grenfell, B. T. (2001). Noisy clockwork: Time series analysis of population fluctuations in animals. *Science*, **293**, 638–643.

Engen, S., Bakke, O., & Islam, A. (1998). Demographic and environmental stochasticity–concepts and definitions. *Biometrics*, **54**, 840–846.

Hilborn, R. & Mangel, M. (1997). *The Ecological Detective*. Princeton University Press.

Johansson, A. & Sumpter, D. J. T. (2003). From local interactions to population dynamics in site-based models of ecology. *Theoretical Population Biology*, **64**, 497–517.

Jonzn, N., Lundberg, P., Ranta, E., & Kaitala, V. (2002). The irreducible uncertainty of the demography-environment interaction in ecology. *Proc. R. Soc. Lond. B.*, **269**, 221–225.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, **261**, 459–467.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., William, Y., & Flannery, B. P. (1992). *Numerical recipes in C.* Cambridge University Press, Cambridge, second edition. The art of scientific computing.

Royama, T. (1992). *Analytical Population Dynamics.* Chapman & Hall.

Saether, B. E. (1997). Environmental stochasticity and population dynamics of large herbivores: A search for mechanisms. *Trends in Ecology & Evolution*, **12**, 143–149.

Sæther, B.-E., J. Tufto, K. J., Røstad, O. W., & Skåtan, J. E. (2000). Population dynamical consequences of climate change for a small temperate songbird. *Science*, **287**, 854–856.

Solow, A. R. (1998). On fitting a population model in the presence of observation error. *Ecology*, **79**, 1463–1466.

Sumpter, D. J. T. & Broomhead, D. S. (2001). Relating individual behaviour to population dynamics. *Proc. R. Soc. Lond. B*, **268**, 925–932.

Tong, H. (1990). *Non-Linear Time Series: A Dynamical Systems Approach.* Oxford University Press.

Turchin, P. (2003). *Complex Population Dynamics: a Theoretical/Empirical Synthesis.* Princeton University Press.

Turchin, P. & Taylor, A. D. (1992). Complex dynamics in ecological time-series. *Ecology*, **73**, 289–305.

Utida, S. (1967). Damped oscillation of population density at equilibrium. *Researches on Population Ecology*, **9**, 1–9.

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

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J. M., Hoegh-Guldberg, O., & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, **416**, 389–395.

A Model fitting

Let $Y^N = (Y_1, \ldots, Y_N)$ denote a sample of N consecutive observations from a stochastic process $\{Y_t\}$. We assume that the probability density function f_{Y^N} depends on a vector of parameters Θ, and that $\{Y_t\}$ is Markovian. Thus

$$f_{Y_t|Y^{t-1}} = f_{Y_t|Y_{t-1}} \tag{18}$$

where $Y^{t-1} = (Y_1, \ldots, Y_{t-1})$. Intuitively, this means that the process lacks memory so that once we know Y_{t-1} our a priori knowledge of Y_t cannot be improved even if we are given all the samples up to this time.

By definition of conditional distributions and using 18 we have

$$f_{Y^N} = f_{Y_1} \prod_{t=2}^{N} \frac{f_{Y^t}}{f_{Y^{t-1}}} = f_{Y_1} \prod_{t=2}^{N} f_{Y_t|Y_{t-1}}$$

Thus, if $y^N \in \mathbb{R}^N$ is a given time-series we consider it to be a realisation of Y^N and interpret $L(\Theta|y^N) := f_{Y^N}(y^N)$ as the likelihood of the parameters Θ given the data. If $L(\cdot, y^N)$ is maximised by some parameters Θ^* we call Θ^* a maximum likelihood estimate for y^N. For stationary and ergodic time-series, the theoretical properties of the likelihood estimator is well-understood, see Tong (1990) for an outline of the theory.

We point out that f_{Y_1} should be interpreted as the stationary distribution of the time-series. This distribution can be determined numerically by iterating the model as many time as needed to reach some termination criteria for stability. Valpine & Hastings (2002) takes this approach. We consider only the likelihood of $(y_2, \ldots, y_N)$ given y_1.

To maximise the likelihood, we minimised the negative log likelihood using Matlab's implementation of the Nelder-Mead simplex algorithm (Press et al., 1992). We initialised the algorithm with a random set of parameters drawn uniformly from a bounded set containing the true parameters. To discourage the algorithm from finding local minimum outside the set of valid parameters we defined the log likelihood to be negative infinity for values outside this set. The algorithm was then restarted until no improvement could be found in 20 consecutive runs.

To determine the set of valid parameters we examined several first-principles derivations of the Ricker map and included parameters that were ecologically realistic. The carrying capacity was bounded at 10 times the average of the time-series.

Label	Stochastic population model	Applicability	References	See
D1	$f(a_t) + s\sqrt{f(a_t)}\epsilon_t$	Demographic noise	This paper	Eq. 8
D2	$f(a_t)\exp\left(\sqrt{\frac{s^2}{f(a_t)}}\epsilon_t - \frac{1}{2}\frac{s^2}{f(a_t)}\right)$	Demographic noise	This paper	
DW	$f(a_t)\exp\left(\sqrt{\frac{s^2}{a_t}}\epsilon_t - \frac{1}{2}\frac{s^2}{a_t}\right)$	Demographic noise	3	
E1	$f(a_t) + sf(a_t)\epsilon_t$	Environmental noise	This paper	Eq. 11
E2	$f(a_t)\exp(s\epsilon_t - s^2/2)$	Environmental noise	1,2	Eq. 3
N1	$f(a_t) + s\epsilon_t$	Observation error	1	Eq. 2
T1	$bna_t\exp(-a_t/n) + s\sqrt{nv(a_t/n)}\epsilon_t$			Eq. 16

References: 1 – Hilborn & Mangel (1997), 2 – Royama (1992), 3 – Sæther et al. (2000).

Table 1: Stochastic population models studied in this paper. Here $f(a_t)$ is the expected population in generation $t+1$ given a_t individuals in generation t, and s is a constant which, together with the parameters of f is estimated from data, while ϵ_t is a normally distributed variable with expectation 0 and variance 1. The classification of the models after applicability follows the discussion in Section 2.

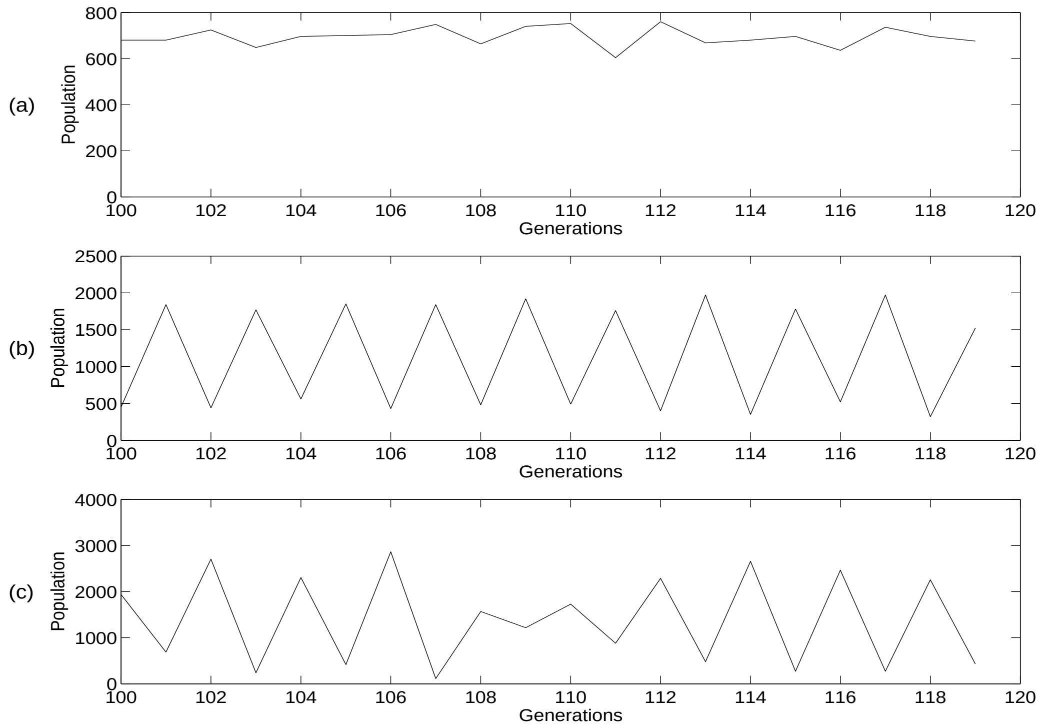


Figure 1: Time-series from the individual-based model in the absence of environmental noise, with $n = 500$ sites and a) $b = 4$, b) $b = 10$ and c) $b = 18$ offspring, corresponding to stable, periodic and chaotic dynamics respectively.

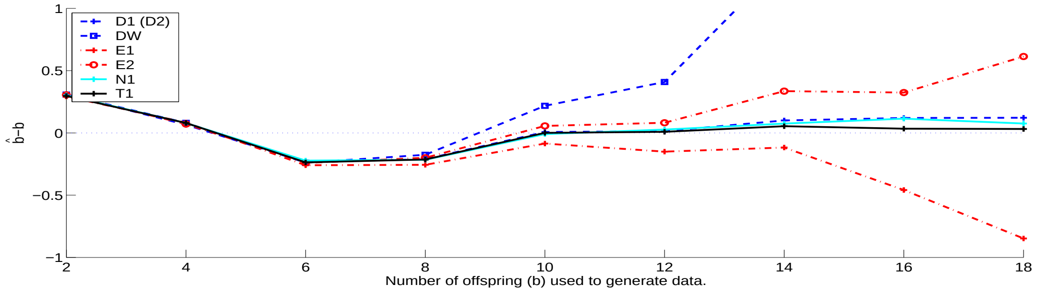


Figure 2: Bias in the maximum likelihood estimate of the number of offspring for the models in Table 1 when fitting to time-series of length 20, sampled after 100 generations from the individual based model in the absence of environmental stochasticity and with $n = 500$ sites. 100 time-series from populations that did not go extinct were used to estimate the bias. Models for which the parameter estimate differed only slightly have been grouped together.

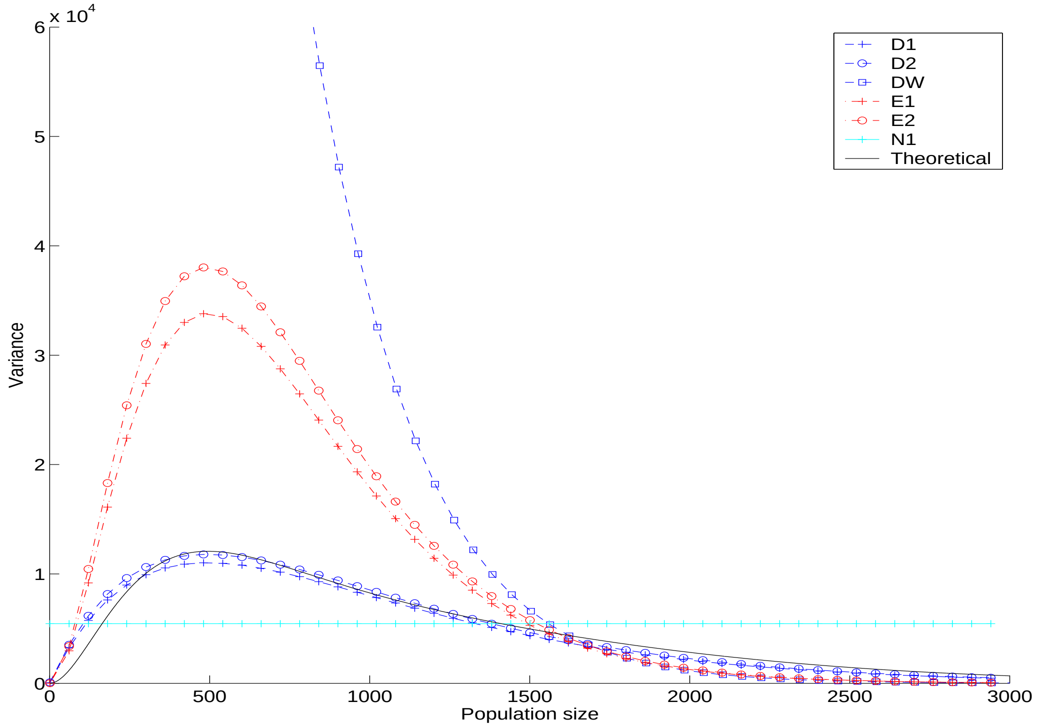


Figure 3: Variance as a function of population size for the generic models in Table 1 fitted to a non-extinct time-series of length 20 sampled after 100 generations from the individual based model in the absence of environmental stochasticity and parameters $n = 500$ and $b = 10$. Also shown for comparison is the theoretically derived variance for these parameter values given by $nv(a_t/n)$ with v defined by Equation 17.

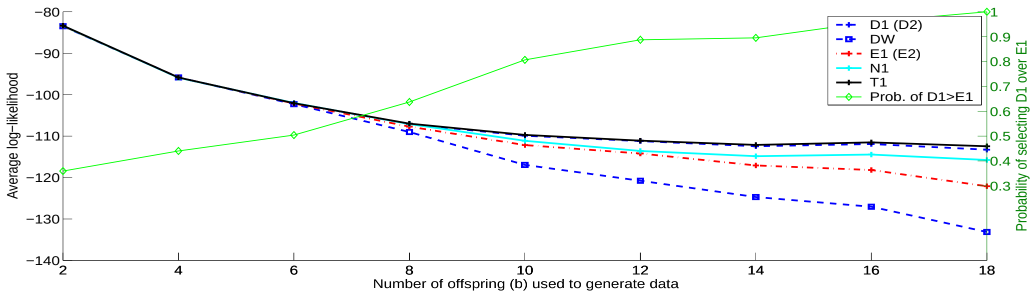


Figure 4: Average log-likelihood for the models in Table 1, and probability that Model D1 is given a higher likelihood than Model E1, when fitting to time-series of length 20 sampled after 100 generations from the individual based model in the absence of environmental stochasticity and with $n = 500$ sites. 100 times-series from populations that did not go extinct were used to calculate the average log-likelihood. Models for which the parameter estimate differed only slightly have been grouped together.

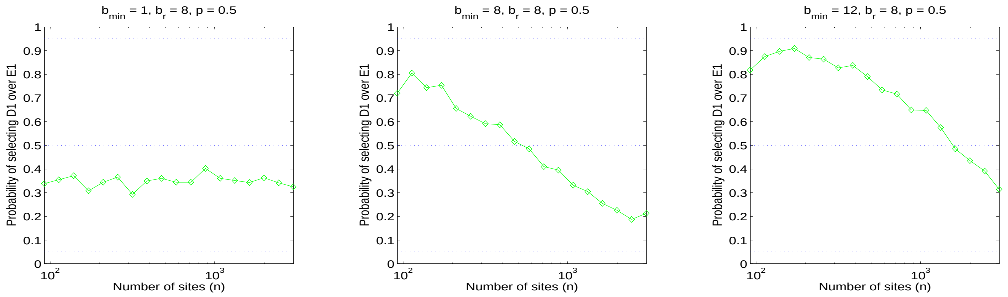
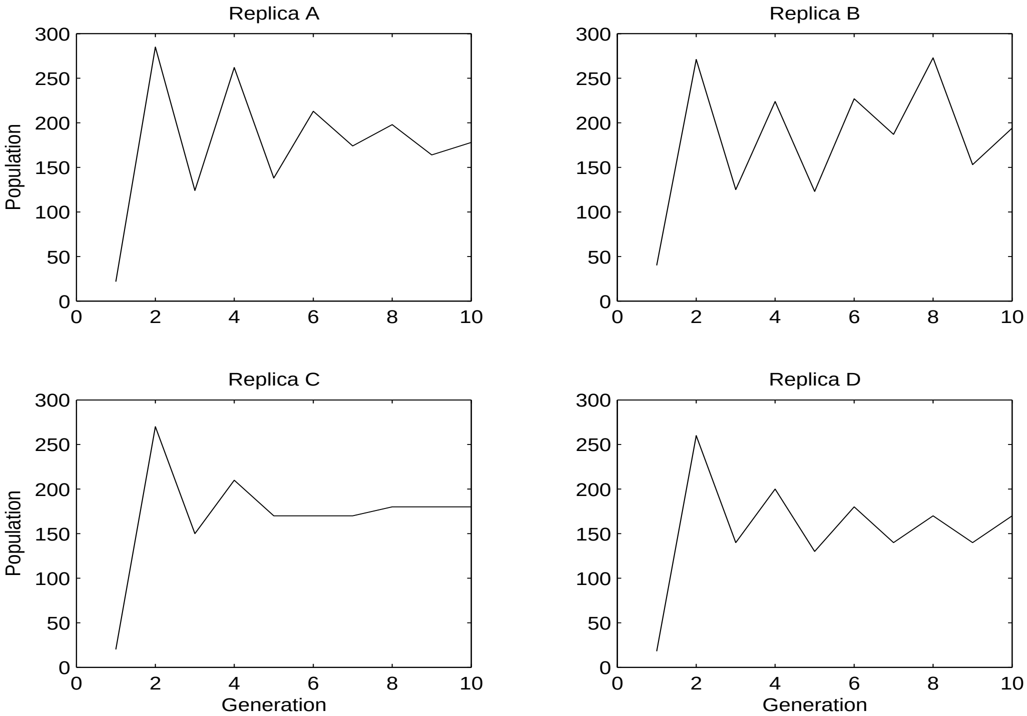


Figure 5: Probability that the demographic model D1 is given a higher likelihood than the environmental model E1 when fitting to time-series of length 20 sampled after 100 generations from the individual based model in the presence of environmental stochasticity with parameters $b_{\min} = 1$, $b_r = 8$, $p = 0.5$ (left figure), $b_{\min} = 8$, $b_r = 8$, $p = 0.5$ (mid figure) and $b_{\min} = 12$, $b_r = 8$, $p = 0.5$ (right figure). With the exception of the first four points in the figure on the right-hand side, 100 times-series from populations that did not go extinct were used to estimate the probability. Note that the number of sites, n is plotted on a log-scale. Dotted lines represent a probability of 0.05, 0.5 and 0.95 respectively.



Replica	Best-fit	Log-likelihood	σ	b	n	θ
A	E1	-34.8145	0.06196	102.6	3.263	0.3780
B	E2	-44.3005	0.1750	48.32	8.478	0.4356
C	E2	-32.0774	0.04639	10180	7.603×10^{-05}	0.1515
D	N1	-36.3543	13.74	1789	7.329×10^{-3}	0.2014

Figure 6: Population time-series for *Callosobruchus maculatus*, the four laboratory experiments under identical conditions performed by Utida (1967) and corresponding parameter estimates. Each time-series were fitted to each of the generic models in Table 1 with the theta-Ricker map as the expectation. For each of the four time-series, the parameter estimates for the model with the highest likelihood are given above.