



Predicting the Time to Quasi-extinction for Populations far below their Carrying Capacity

STEINAR ENGEN*† AND BERNT-ERIK SÆTHER‡

†Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway and ‡Department of Zoology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

(Received on 8 December 1999, Accepted in revised form on 5 May 2000)

Populations threatened by extinction are often far below their carrying capacity. A population collapse or quasi-extinction is defined to occur when the population size reaches some given lower density. If this density is chosen to be large enough for the demographic stochasticity to be ignored compared to environmental stochasticity, then the logarithm of the population size may be modelled by a Brownian motion until quasi-extinction occurs. The normal-gamma mixture of inverse Gaussian distributions can then be applied to define prediction intervals for the time to quasi-extinction in such processes. A similar mixture is used to predict the population size at a finite time for the same process provided that quasi-extinction has not occurred before that time. Stochastic simulations indicate that the coverage of the prediction interval is very close to the probability calculated theoretically. As an illustration, the method is applied to predict the time to extinction of a declining population of white stork in southwestern Germany.

© 2000 Academic Press

1. Introduction

The current species extinction rate is, at least for some groups, among the highest ever experienced in the history of the earth (May *et al.*, 1995), mainly due to various kinds of human impact on the environment. An important tool in the conservation of endangered or threatened species has been population viability analysis, which uses data to model the risk of extinction [see recent reviews in Beissinger & Westphal (1998) and Groom & Pascual (1998)]. A natural consequence of population viability analysis was the introduction of quantitative criteria for risk assessment, provided by Mace & Lande (1991) where populations were assigned to different risk

categories according to their probability of extinction during a given time period.

In practice, several problems are encountered when estimating the risk of extinction of a population by a population viability analysis (Beissinger & Westphal, 1998). One major problem is to estimate the uncertainty in the population projections which may be caused by uncertainties in the parameter estimates (Ludwig, 1996) as well as stochasticity in the population dynamics (Ludwig, 1999) due to random effects in the demography and temporal environmental fluctuations (Lande, 1993). Stochastic effects on the population dynamics are commonly modelled by using discrete time steps. Let the relative expected change in the population size from one season to the next be $E(\Delta N/N|N) = r - g(N)$, where $g(N)$ is an

* Author to whom correspondence should be addressed.

increasing function describing the density regulation which is approximately zero for small values of N . Hence, r is the specific growth rate for small population sizes. The variance of ΔN conditioned on N may be partitioned into two components, the demographic and the environmental variance (Engen *et al.*, 1998). These components are often assumed to be constants, say σ_d^2 and σ_e^2 , corresponding to $\text{var}(\Delta N|N) = \sigma_d^2 N + \sigma_e^2 N^2$. The demographic variance is the component arising due to stochastic variation in reproduction and survival between individuals within a season, whereas the environmental variance is the component arising due to the environmental conditions acting simultaneously on all individuals. For more precise definitions, see Engen *et al.* (1998).

Using Ito calculus (Karlin & Taylor, 1981), the diffusion approximation to this process has infinitesimal mean and variance $N[r - g(N)]$ and $\sigma_d^2 N + \sigma_e^2 N^2$, respectively. The approximation is valid if the absolute value of the specific growth rate is small, say less than 0.1 (Turelli, 1977; Lande *et al.*, 1999). For most populations, the environmental term is the dominating one (Lande, 1998), at least when the populations are not too small. Ignoring the demographic variance, the diffusion $X(t) = \ln[N(t)]$ has infinitesimal mean and variance $\mu(x) = r - \sigma_e^2/2 - g(e^x)$ and $\sigma^2 = \sigma_e^2$, respectively. For populations that are threatened by extinction, the population size is often far below the carrying capacity so that $g(e^x) \approx 0$ (see Dennis *et al.*, 1991). Then, $X(t)$ is actually a Brownian motion with drift $\mu = r - \sigma_e^2/2$, and infinitesimal variance $\sigma^2 = \sigma_e^2$.

For a population with size $N(0) = e^{x_0}$ at time zero, one will often be interested in predicting the time to extinction, or the population size in the future. For sexually reproducing species, the extinction barrier is often chosen to be at $N = 1$, corresponding to $X = 0$. However, quite often one will also be interested in predicting the time it takes for the process to reach some other barrier larger than 1, say at N_q , e.g. to avoid unfortunate genetic consequences of small population sizes (Lande, 1995). Reaching this barrier, which is usually called quasi-extinction (Ginzburg *et al.*, 1982), would then be considered to have some serious consequences for the management of the species. Mathematically, quasi-extinctions may be analysed in the same way as extinctions at

$X = 0$ due to the invariance property of Brownian motions. We then define the process $X(t) = \ln N(t) - \ln N_q$, which still has the same drift and infinitesimal variance, and still define extinction to occur at $X = 0$. As in Dennis *et al.* (1991), the following analysis ignores demographic stochasticity. For the methods to be reliable, the quasi-extinction barrier N_q should therefore be chosen large enough for the environmental stochasticity to be the major stochastic component. The environmental term in the infinitesimal variance is larger than the demographic term for $N > \sigma_d^2/\sigma_e^2$. For small passerine birds the available estimates of demographic variance range from 0.18 to 0.66, whereas the corresponding figures for the environmental variance vary between 0.024 and 0.41 (Tufto *et al.*, 2000). If our method is to be applied, the barrier defining quasi-extinction should therefore be chosen at population sizes larger than 10, say in the region 10–100, depending on the magnitude of the different components of the stochastic fluctuations.

Dennis *et al.* (1991) have previously used Brownian motions to estimate the risk of extinction of several endangered or threatened species. However, they only analysed the expected time to extinction and made no attempt to derive prediction intervals for the actual time it takes until the population reaches the extinction barrier, although forecasting the population size ignoring extinctions was dealt with. Here we extend the approach of Dennis *et al.* (1991) by deriving prediction intervals for the time of extinction based on an observed time series at discrete points of the Brownian motion. We then apply those to predict extinctions and population sizes. Prediction intervals are stochastic intervals covering the unknown stochastic variable (Sæther *et al.*, 2000), which in this case may be the population size or the extinction time, with probability $(1 - \alpha)$. We also derive predictions of the population size taking into account the possibility of extinction. We illustrate our approach by predicting the population development of a declining population of white stork *Ciconia ciconia* (Bairlein & Zink, 1979).

2. Some Properties of Brownian Motions

Let $X(t)$ be a Brownian motion with infinitesimal mean and variance μ and σ^2 , respectively,

and initial value $X(0) = x_0 > 0$. Then, by definition (Karlin & Taylor, 1981), the differences $X(t + h) - X(t)$, $h > 0$, are normally distributed with mean μh and variance $\sigma^2 h$, and the distribution is the same if conditioned on $X(t)$. Furthermore, the differences are independent for non-overlapping intervals.

If zero is an absorbing state, the distribution of the time to extinction at zero is the inverse Gaussian distribution

$$f(t) = \frac{x_0}{\sqrt{2\pi t^3} \sigma} e^{-(x_0 + \mu t)^2 / (2\sigma^2 t)} \quad (1)$$

(Cox & Miller, 1970, p. 221), which integrates to one for $\mu \leq 0$, whereas for $\mu > 0$ there is a positive probability $1 - \exp(-2\mu x_0 / \sigma^2)$ for the process to be absorbed at infinity (Karlin & Taylor, 1981). For a theoretical treatment and for a number of applications of this distribution, see Seshadri (1999).

We shall also be concerned with the distribution of $X(t)$ for a given $t > 0$, conditioned on extinction at $X = 0$ not occurring before time t ,

$$g(x) = \frac{1}{\sqrt{2\pi t} \sigma} [1 - e^{-2xx_0/\sigma^2 t}] e^{-(x - x_0 - \mu t)^2 / (2\sigma^2 t)}, \quad (2)$$

which is the proper reparameterization of the corresponding distribution given by Cox & Miller (1970, eqn (71), p. 221). If $X = 0$ represents quasi-extinction, so that the real process actually does not go extinct at this barrier, the above distribution may still be applicable when we make the distinction between those paths reaching, and those not reaching, the extinction barrier before time t . For the latter ones, $X(t)$ has the above distribution for $x > 0$. If the barrier for quasi-extinction is chosen at some small value and the estimate of the expected growth rate is negative, it is very unlikely that a population passing the barrier should not actually go extinct unless actions are taken to prevent this. Hence, the distribution of $X(t)$ used in this case is very close to the distribution found when conditioning on extinction not occurring before time t .

If we only consider the finite time interval $[0, t]$, the process is either absorbed at zero before time t when crossing the horizontal axis at a time with distribution given by eqn (1), or it takes a positive value $X(t)$ with distribution given by eqn (2) when crossing the vertical axis at time t as illustrated in Fig. 1.

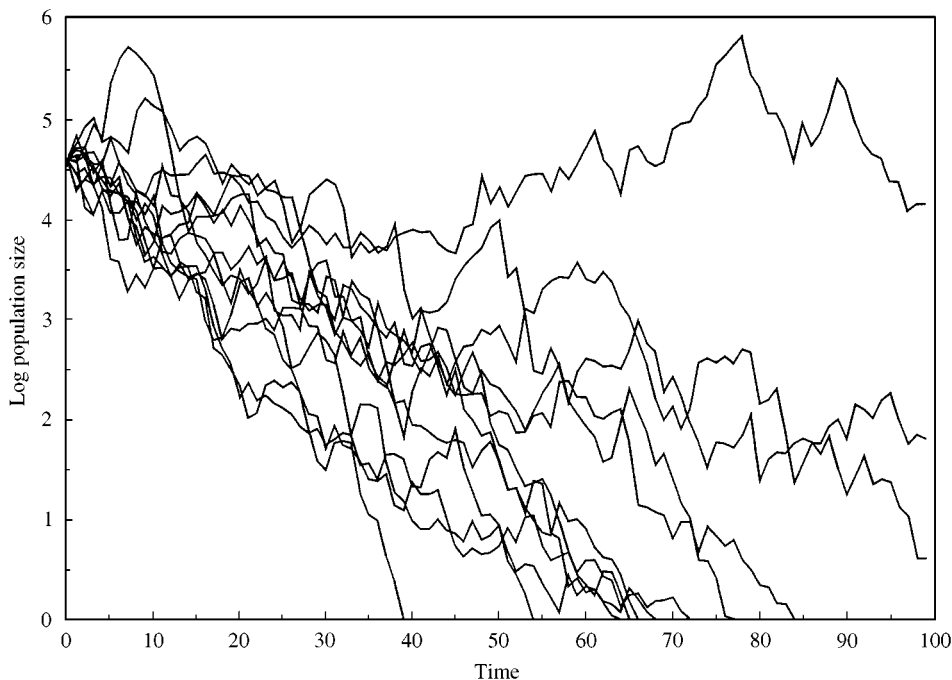


FIG. 1. The paths of 12 simulations of a Brownian motion with $x_0 = \ln 100 \approx 4.61$, $\mu = -0.06$ and $\sigma^2 = 0.04$. Absorbing barriers are at $X = 0$ and $t = 100$.

3. Predictions in Processes with no Extinction Barriers

Following Dennis *et al.* (1991) we first consider the problem of estimating the parameters μ and σ^2 from an observed time series X_1, X_2, \dots, X_n at times $t_1 < t_2 < \dots < t_n$. Ignoring the possibility of extinction during the time interval the observations are taken, the maximum likelihood estimator for μ is $\hat{\mu} = (X_n - X_1)\theta$, where $\theta = 1/(t_n - t_1)$. Hence, for estimating the drift only the first and the last observation is required, the other $(n - 2)$ observations being redundant. Writing $\delta_i = \sqrt{t_{i+1} - t_i}$ for $i = 1, 2, \dots, n - 1$, the standard regression estimator for σ^2 is $s^2 = (1/(n - 2))\sum_{i=1}^{n-1} (D_i - \delta_i\hat{\mu})^2$, where $D_i = (X_{i+1} - X_i)/\delta_i$. Furthermore, $\hat{\mu}$ and s^2 are independent, $\hat{\mu}$ is normal with mean μ and variance $\sigma^2\theta$, and $s^2(n - 2)/\sigma^2$ is χ^2 -distributed with $v = (n - 2)$ degrees of freedom.

It is now straightforward to find prediction intervals for the population size $N(t) = e^{X(t)}$ at time t when the last known population size is $N_0 = e^{x_0}$ at time zero, provided that we ignore the possibility of extinction before time t . Hence, these predictions are only valid if extinction is unlikely to occur before time t . This is the case either if t is small, or if the population size at the time we make the predictions is still fairly large. Further, extinction before time t is unlikely if the specific growth rate is positive.

For most practical applications, the last known population size will actually be X_n , in which case t_n by definition is zero. Since $X(t)$ conditioned on $X(0) = x_0$ is a normal variate with mean $x_0 + \mu t$ and variance $\sigma^2 t$, this problem is practically equivalent to the problem of predicting the value of a normal variate on the basis of independent observations from the same normal distribution. The best prediction of $X(t)$ is $x_0 + \hat{\mu} t$ and the symmetric prediction interval with coverage $(1 - \alpha)$ for $X(t)$ is

$$(L_t, U_t) = (x_0 + \hat{\mu} t - s\sqrt{t(\theta t + 1)}\gamma_{\alpha/2, n-2}, x_0 + \hat{\mu} t + s\sqrt{t(\theta t + 1)}\gamma_{\alpha/2, n-2}),$$

where $\gamma_{\alpha, v}$ denotes the α -quantile of the T -distribution with v degrees of freedom. The interpretation of this prediction interval is the same as for

a confidence interval, that is, L_t and U_t are two statistics with the property that

$$P(L_t < X(t) < U_t) = 1 - \alpha.$$

In viability analysis, we may be more interested in an upper one-sided interval

$$(L_t, \infty) = (x_0 + \hat{\mu} t - s\sqrt{t(\theta t + 1)}\gamma_{\alpha, n-2}, \infty).$$

The corresponding intervals for $N(t)$ are (e^{L_t}, e^{U_t}) and (e^{L_t}, ∞) in the two- and one-sided case, respectively.

4. Predicting the Time to Extinction

If extinction is a possibility, these prediction intervals are no longer correct because sample paths with $X(t)$ taking values in the prediction interval may have passed the extinction barrier at an earlier point of time. This has the effect that our confidence in (L_t, U_t) should actually be smaller than $(1 - \alpha)$. In such cases, we should rather try to construct a prediction interval for the time to extinction. More precisely, if the process goes extinct at time W , we should find a statistic, say $\tilde{W}_{1-\alpha}$, with the property that $P(W > \tilde{W}_{1-\alpha}) = (1 - \alpha)$ for any value of the unknown nuisance parameters μ and σ^2 , in which case $(\tilde{W}_{1-\alpha}, \infty)$ is a prediction interval for W with coverage $1 - \alpha$.

In Appendix A we give an explanation why $\tilde{W}_{1-\alpha}$ should be chosen as the upper α -quantile of the distribution

$$\hat{f}(t) = \frac{x_0\Gamma((v + 1)/2)}{\sqrt{\pi v s^2 t^3(\theta t + 1)}\Gamma(v/2)} \times \left\{ \frac{(x_0 + \hat{\mu} t)^2}{s^2 v t(\theta t + 1)} + 1 \right\}^{-(v+1)/2}. \quad (3)$$

More precisely, provided that $\int_0^\infty \hat{f}(t) dt > \alpha$ the variable $\tilde{W}_{1-\alpha}$ is now defined by the relation

$$\int_0^{\tilde{W}_{1-\alpha}} \hat{f}(t) dt = \alpha, \quad (4)$$

otherwise $\tilde{W}_{1-\alpha} = \infty$, which means that the entire interval is in infinity. The distribution $\hat{f}(t)$ was named the normal-gamma mixture of inverse

Gaussian distributions by Whitmore (1986). A further discussion and some applications of the distribution are given by Seshadri (1999). This distribution does not integrate to unity. For an explanation see Appendix A.

Note that $\tilde{W}_{1-\alpha}$ is a function of the known parameters v and θ as well as the statistics $x_0, \hat{\mu}$ and s^2 . Hence, $\tilde{W}_{1-\alpha}$ is a random variable, and the observed value of this variable can be calculated by one simple numerical integration using eqn (4).

5. Joint Predictions of Extinction and Population Size

As indicated in Section 2 we may, for a fixed value of t , order the sample paths by their value of the time to extinction, say W , if extinction occurs before time t , and otherwise by $X(t)$. When our main interest is to make predictions in the finite time interval $(0, t)$, a natural generalization of predicting extinction under an infinite time horizon is to make stochastic prediction of the crossing point for the process of the horizontal axis at zero or vertical axis at time t . In population viability analysis, it is most appropriate to find an upper interval of the type (\tilde{W}, \tilde{X}) , where either $\tilde{W} < t$ or $\tilde{X} > 0$. Our aim is to define the functions \tilde{W} and \tilde{X} so that the probability that the real unknown process has value larger than (\tilde{W}, \tilde{X}) by our ordering of sample paths, is $1 - \alpha$.

The ordering of the vectors (W, X) may be further illustrated by considering Fig. 1. Here we have given ten sample paths and $t = 100$. We see that in this case seven paths reach extinction before time $t = 100$, so that their time to extinction which is their realized value of W , is less than 100, while $X(t) = 0$. These seven paths are ordered according to their time to extinction, that is, the value of W . The other three paths has not reached extinction at time $t = 100$, so we say that these take larger values than the other seven paths. These paths have positive values of $X(t)$, and they are ordered according to these values.

Our general problem is then to construct a statistic $(\tilde{W}, \tilde{X})_{1-\alpha}$ with the property that $P[(W, X) > (\tilde{W}, \tilde{X})_{1-\alpha}] = \alpha$.

By the argument we have given in Appendix A justifying the result in Section 4, we now define

the component \tilde{W} by the distribution \hat{f} as in Section 4, while the component \tilde{X} can be defined by the same mixture performed on the distribution g defined by eqn (2), giving

$$\hat{g}(x) = \frac{\Gamma((v + 1)/2)}{\Gamma(v/2)\sqrt{\pi v s^2 t(\theta t + 1)}} \times \left[\left(1 + \frac{(x - x_0 - \hat{\mu}t)^2}{s^2 t v(\theta t + 1)} \right)^{-(v+1)/2} - \left(1 + \frac{(x - x_0 - \hat{\mu}t)^2}{s^2 t v(\theta t + 1)} + \frac{4xx_0}{s^2 vt} \right)^{-(v+1)/2} \right]. \tag{5}$$

Then, if $\int_0^t \hat{f}(t) dt > \alpha$, we define \tilde{W} by $\int_0^{\tilde{W}} \hat{f}(t) dt = \alpha$. Otherwise, the prediction of the crossing of the vertical axis at time t is defined by $\int_0^t \hat{f}(t) dt + \int_0^{\tilde{X}} \hat{g}(x) dx = \alpha$. In the next section, we perform stochastic simulations indicating strongly that these prediction intervals along the two axis has coverage $1 - \alpha$ for any value of the nuisance parameters μ and σ^2 .

6. Verifications by Stochastic Simulations

To check the coverage probability for the prediction interval, we simulate the data as well as the process. We choose the parameters μ and σ^2 , as well as x_0 , which is the initial value of the logarithm of the population size at time zero. We also fix the parameter t , which is the length of the finite time interval to be analysed. Finally, we choose the parameters related to the sampling process, that is, the number of observations, $n = v + 2$, and $\theta = 1/(t_n - t_1)$.

The lower bound of the prediction interval defined by (\tilde{W}, \tilde{X}) is simulated by first simulating one data set, giving $\hat{\mu}$ and s^2 , and then performing the numerical integrations defining the bound. Second, we simulate the process with its true known parameters until it is absorbed at the chosen horizontal or vertical axis, and check if it ends up in the prediction interval or not. The whole process is repeated a large number of times with the same parameters to give the frequency of coverage to be compared to $1 - \alpha$.

We have performed these simulations for a large number of different parameters and different values of α , two of which are shown in Fig. 2.

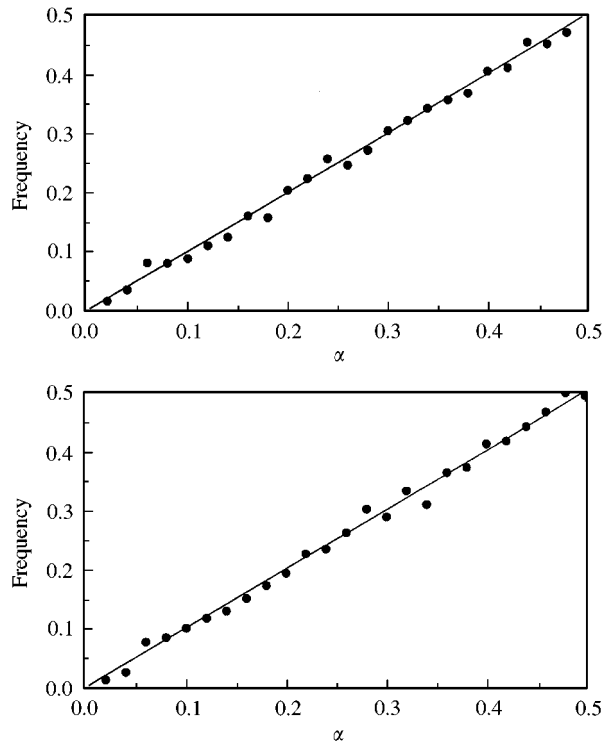


FIG. 2. Simulations of the frequency of coverage for different values of α . The frequencies are independent and each based on 1000 simulations. Both panels have common parameters $N(0) = 100$, $\sigma^2 = 0.04$, $n = 10$, and $t = 100$. For the upper panel, $\mu = -0.06$ and $\chi^2_{25} = 26.32$, and for the lower panel $\mu = 0$ and $\chi^2_{25} = 24.28$.

There are no indications that the method is not exact even if each graph is based on 25 000 simulations. The χ^2 -values for each graph (with 25 degrees of freedom) are all insignificant. This is a fairly strong indication of the method being exact, taking into account that even rather small errors in the numerical approximations will generally lead to an increase in χ^2 when the number of simulations is large.

7. Applications to Population Viability Analysis

When applying the method in practice, we may plot the lower bounds of the prediction intervals for $X(t)$, that is \tilde{X} , against t . Then, for a given value of α this line crosses the time axis at the lower bound \tilde{W} of the corresponding prediction interval for the time to extinction. As an illustration, we consider the white stork which has declined over larger part of its range of distribution in Europe during the last decades, probably

due to food shortage in modern agricultural landscapes (Rheinwald *et al.*, 1989). In Baden-Württemberg in Germany, long-term data are available that document this decline (Bairlein & Zink, 1979). We use the data for the period 1950–1965 to predict the future population fluctuations of this populations, introducing a quasi-extinction barrier at $N_q = 20$ individuals. The bias in the estimates of the population size is probably negligible because of the species conspicuous nesting habits and the high proportion of ringed individuals in the population. The estimates are $\hat{\mu} = -0.0537$ and $s^2 = 0.0172$. The prediction intervals are shown in Fig. 3. In this case, it is interesting to compare the results with the true development of the population which passed the extinction barrier in 1974, at $t = 9$. This is at the lower bound of the upper 90% prediction interval for the time to quasi-extinction. One reason why extinction occurs so quickly may be that we have not taken the demographic stochasticity into account.

8. Discussion

We have not been able to present a mathematical proof for the exactness of the method of calculating prediction intervals. However, the simulations presented here (Fig. 2), and a number of other simulations, show that the method, if it should turn out not to be exact, gives the probability of coverage with sufficient accuracy for practical applications. A different approach based on conditional parametric bootstrapping was checked by Lillegård & Engen (1999). Although the method seemed to be a rather good approximation, their simulations also demonstrated that their method was not exact.

It ought to be mentioned that we have assumed that there is no extinction barrier during the sampling period. On the other hand, the populations one want to analyse are unlikely to go extinct during the sampling period, so in practice this approximation will only have a negligible effect on the sampling properties of the estimators.

If one wants to predict extinctions at $N = 1$, one should generally try to take the demographic stochasticity into account (Lande, 1998). For some species, such as birds, one may obtain data

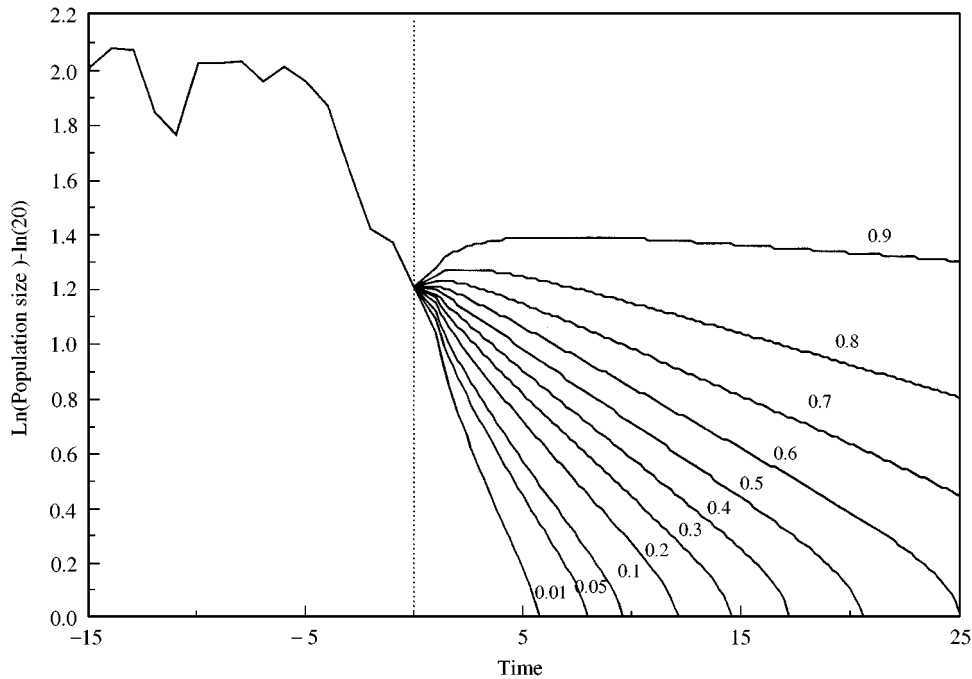


FIG. 3. The observed process on log scale minus $\ln 20$ of the population of white storks from 1950 to 1965, the last year corresponding to zero on the time axis. For $t > 0$ the lower bounds of the prediction intervals \hat{X} are plotted against t for $\alpha = 0.01, 0.05, 0.10, 0.20, \dots, 0.90$. The corresponding lower bound, \tilde{T} , for the prediction intervals for the time to quasi-extinction at 20 are where the lines cross the horizontal axis.

on survival and reproduction on the individual level, from which σ_a^2 may be estimated (Sæther *et al.*, 1998, 2000; Tufto *et al.*, 2000). Provided that the population during sampling is large enough for σ_a^2 to be ignored, the effect of σ_a^2 close to extinction may be taken into account by performing stochastic simulations. One should then redefine the process $\hat{X}(t)$ by adding the demographic term, using the stochastic differential equation

$$d\hat{X}(t) = (\hat{\mu} + \hat{U}s\sqrt{\theta v/\hat{Z}})t + s\sqrt{v/\hat{Z}} d\hat{B}(t) + \sigma_a/\sqrt{\hat{X}(t)} dB_d(t)$$

that can easily be simulated. The \hat{U} and \hat{Z} are chosen independently for each simulated path, and $\hat{B}(t)$ and $B_d(t)$ are independent standard Brownian motions. From a large number of simulations until the process reaches the chosen horizontal or vertical axis, the upper quantiles of the distribution may be approximated by the corresponding simulated frequencies (Sæther *et al.*, 2000). It acquires time-consuming simulations to check the probability of the interval

covering the extinction time, and usually it may also be necessary to perform a correction of the method to give an approximately correct coverage. Hence, the threshold of knowledge in statistics, numerical computations and programming for performing such an analysis is much higher than for the method given here. Actually, the present approach is only based on one numerical integration for each point defining the curves in Fig. 3. Although the functions to be integrated requires a procedure for the gamma function, it is a fairly simple task to write a procedure for these expressions.

One may claim that the number of observations (number of years) are generally too small to perform a population viability analysis (e.g. Ludwig, 1999), since prediction intervals often become very wide (Figs 2 and 3; see also Sæther *et al.*, 2000). On the other hand, we see no alternative to making such a statistical analysis. According to the precautionary principle (IUCN, 1994), all types of uncertainties must be taken into account to give more cautious decisions for the management of threatened or endangered species.

Several authors (see references in Ludwig, 1999; Tufto *et al.*, 1999) have pointed out the importance of taking into account the uncertainty in parameter estimation. Dennis *et al.* (1991) and Ludwig (1999) incorporated this uncertainty by finding confidence intervals with rather rough approximations for the coverage, for the probability of extinction before some given time. Dennis *et al.* (1991) also derived approximate confidence intervals for the expected time to extinction. Since the basic problem is not to find the probabilities or expectations, but rather to draw inference about the actual time to extinction, we believe that the concept of prediction intervals is much more appropriate. This approach is also conceptually simpler for research workers applying the theory. In practice, one will have problems figuring out how to apply a confidence interval for the probability of extinction or for the expected time to extinction, in population viability analysis. Ludwig (1999) only claims that the confidence region is too large to be applicable, and Dennis *et al.* (1991) do not really deal with this problem.

For density-regulated populations the presented method is no longer applicable. Sæther *et al.* (2000) did a similar type of analysis for a density-regulated population based on parametric bootstrapping. An alternative way of dealing with uncertain estimates is of course to rely on Bayesian methods, which from some points of view are conceptually simpler, and may also be easier to perform in practice (Taylor *et al.*, 1996). However, regardless of the choice of approximate method, the properties of the method, in the frequentistic sense, should always be checked by extensive stochastic simulations.

This work has been supported by the Research Council of Norway and EU (Project Metabird).

REFERENCES

- BAIRLEIM, F. & ZINK, G. (1979). Der Bestand des Weiss-torchs *Ciconia ciconia* in Südwestdeutschland: eine Analyse der Bestandsentwicklung. *J. Ornithol.* **120**, 1–11.
- BEISSINGER, S. R. & WESTPHAL, M. I. (1998). On the use of demographic models in endangered species management. *J. Wildlife Manage.* **62**, 821–841.
- COX, D. R. & MILLER, H. D. (1970). *The Theory of Stochastic Processes*. London: Methuen.
- DENNIS, B., MUNHOLLAND, P. L. & SCOTT, J. M. (1991). Estimation of growth and extinction parameters for endangered species. *Ecol. Monographs* **61**, 115–123.
- ENGEN, S., BAKKE, Ø. & ISLAM, A. (1998). Demographic and environmental stochasticity — concepts and definitions. *Biometrics* **54**, 840–846.
- GINZBURG, L. R., SLOBODKIN, L. B., JOHNSON, K. & BINDMAN, A. G. (1982). Quasiextinction probabilities as a measure of impact on population growth. *Risk Anal.* **21**, 181–191.
- GROOM, M. J. & PASCUAL, M. (1998). The analysis of population persistence: an outlook on the practice of population viability analysis. In: *Conservation Biology for the Coming Decade*. FIEDLER, P. & KAREIVA, P. (eds), 2nd Edn. New York: Chapman & Hall.
- IUCN/SSC (1994). *IUCN Red List Categories*. Switzerland: Gland, IUCN/SSC. Available on the internet, URL: http://www.iucn.org/themes/ssc/redlist/ssc-rl_c.htm.
- KARLIN, S. & TAYLOR, H. M. (1981). *A Second Course in Stochastic Processes*. London: Academic Press.
- LANDE, R. (1993). Risk of population extinction from demographic and environmental stochasticity and random catastrophes. *Am. Nat.* **142**, 911–927.
- LANDE, R. (1995). Mutation and conservation. *Conservation Biol.* **9**, 782–791.
- LANDE, R. (1998). Demographic stochasticity and Allee effect on a scale with isotropic noise. *Oikos* **83**, 353–358.
- LANDE, R., ENGEN, S. & SÆTHER, B.-E. (1999). Spatial scale of population synchrony: environmental correlation versus dispersal and density regulation. *Am. Nat.* **154**, 271–281.
- LILLEGÅRD, M. & ENGEN, S. (1999). Exact confidence intervals based on conditional parametric bootstrapping. *J. Appl. Stat.* **26**, 447–459.
- LUDWIG, D. (1996). The distribution of population survival times. *Am. Nat.* **147**, 506–526.
- LUDWIG, D. (1999). Is it meaningful to estimate probability of extinction? *Ecology* **80**, 298–310.
- MACE, G. M. & LANDE, R. (1991). Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. *Conservation Biol.* **5**, 148–157.
- MAY, R. M., LAWTON, J. H. & STORK, N. E. (1995). Assessing extinction rates. In: *Extinction Rates* (J. H. Lawton & R. M. Mays eds), pp. 1–24. Oxford: Oxford University Press.
- RHEINWALD, G., OGDEN, J. & SCHULZ, H. (1989). *White Stork, Status and Conservation*. Bonn: ICBP.
- SÆTHER, B.-E., ENGEN, S., ISLAM, A., MCCLEERY, R. & PER-RINS, C. (1998). Environmental stochasticity and extinction risk in a population of a small songbird, the great tit. *Am. Nat.* **151**, 441–450.
- SÆTHER, B.-E., ENGEN, S., LANDE, R., ARCESE, P. & SMITH, J. N. M. (2000). Estimating time to extinction in an island population of song sparrows. *Proc. Roy. Soc. London Ser. B* **267**, 621–626.
- SESHADRI, V. (1999). *The Inverse Gaussian Distribution. Statistical Theory and Applications*. New York: Springer.
- TAYLOR, B. L., WADE, P., R., STEHN, R. A. & COCHRANE, J. F. (1996). A Bayesian approach to classification criteria for spectacled eiders. *Ecol. Appl.* **6**, 1077–1089.
- TUFTO, J., SÆTHER, B.-E., ENGEN, S., SWENSON, J. E. & SANDEGREN, F. (1999). Harvesting strategies for conserving minimum viable populations based on World Conservation Union criteria: brown bears in Norway. *Proc. Roy. Soc. London B* **266**, 961–967.

TUFTO, J., SÆTHER, B.-E., ENGEN, S., ARCESE, P., JERSTAD, K., RØSTAD, O. W. & SMITH, J. N. M. (2000). Bayesian meta-analysis of demographic parameters in three small temperate passerines. *Oikos* **88**, 273–281.
 TURELLI, M. (1977). Random environments and stochastic calculus. *Theor. Popul. Biol.* **12**, 140–178.
 WHITMORE, G. A. (1986). Normal-gamma mixtures of inverse Gaussian distributions. *Scand. J. Stat.* **13**, 211–220.

APPENDIX

Theoretical Derivations

In order to give an intuitive explanation of the main results, we first rewrite the process as $X(t) = x_0 + \mu t + \sigma B(t)$, where $B(t)$ is a standard Brownian motion with infinitesimal mean and variance zero and one, respectively (Karlin & Taylor, 1981). Then the variable

$$T(t) = \frac{(\mu - \hat{\mu}) + \sigma B(t)}{s\sqrt{t(\theta t + 1)}}$$

is T -distributed with ν degrees of freedom. Introducing the standard normal variate $U = (\mu - \hat{\mu})/(\sigma\sqrt{\theta})$ and writing $Z = s^2\nu/\sigma^2$ for the χ^2 -variable with ν degrees of freedom, we find

$$T(t) = \frac{t\sqrt{\theta}U + B(t)}{\sqrt{(Z/\nu)t(\theta t + 1)}}$$

Now, let \hat{U} and \hat{Z} be independent of U and Z but with the same distributions, and let $\hat{B}(t)$ be another independent standard Brownian motion. Then

$$\hat{T}(t) = \frac{t\sqrt{\theta}\hat{U} + \hat{B}(t)}{\sqrt{(\hat{Z}/\nu)t(\theta t + 1)}}$$

is another independent variable with the same distribution as $T(t)$. Then we compare the process

$$X(t) = x_0 + \hat{\mu}t + s\sqrt{t(\theta t + 1)}T(t)$$

with

$$\hat{X}(t) = x_0 + \hat{\mu}t + s\sqrt{t(\theta t + 1)}\hat{T}(t).$$

Imagine a large number of independent realizations of $\hat{X}(t)$ and one realization of $X(t)$. Then, for any positive value of t , the rank of $X(t)$ among the $\hat{X}(t)$ is uniformly distributed since the rank of $X(t)$ among the $\hat{X}(t)$ is the same as the rank of $T(t)$ among the $\hat{T}(t)$, and the $\hat{T}(t)$ and $T(t)$ are all independent with the same distribution. The question is now whether this uniformity of ranks also transforms to the horizontal axis when we introduce $X(t) = 0$ and $\hat{X}(t) = 0$ as absorbing barriers. That is, will the rank of the extinction time W for $X(t)$ be uniformly distributed among the extinction times \hat{W} for the extinction of the processes $\hat{X}(t)$? This may at first seem obvious, but we are currently not able to present a formal proof for the conjecture. Actually, the problem is that the process $\hat{T}(t)$ is independent of $\hat{\mu}$ and s^2 , while, for the given values of μ and σ^2 , $T(t)$ depends on these two statistics.

If the result is exact, we simply have to derive the conditional distribution (conditioning on $\hat{\mu}$ and s^2) of the time to extinction for the process $\hat{X}(t)$. Then the corresponding conditional probability that W is larger than the α -quantile $\hat{W}_{1-\alpha}$ of this distribution is exactly $(1 - \alpha)$.

To find this distribution, we first express $\hat{T}(t)$ in the expression for $\hat{X}(t)$ by \hat{U} , \hat{Z} and $\hat{B}(t)$ giving

$$\hat{X}(t) = s\sqrt{\nu/\hat{Z}}\hat{B}(t) + x_0 + (\hat{\mu} + \hat{U}s\sqrt{\theta\nu/\hat{Z}})t.$$

Hence, if we first condition on \hat{U} and \hat{Z} , this is a Brownian motion starting at x_0 at time zero and with drift $(\hat{\mu} + \hat{U}s\sqrt{\theta\nu/\hat{Z}})$ and infinitesimal variance $s^2\nu/\hat{Z}$, which inserted into eqn (1) gives that the conditional time to extinction has the inverse Gaussian distribution

$$f(t|\hat{U}, \hat{Z}) = \frac{x_0\hat{Z}^{1/2}}{\sqrt{2\pi t^3\nu s^2}} \times \exp\left(-\frac{\hat{Z}(x_0 + \hat{\mu}t + \hat{U}st\sqrt{\theta\nu/\hat{Z}})^2}{2s^2t\nu}\right).$$

Finally, we find the unconditional distribution by integrating over the distribution of \hat{U} , which is a standard normal, and \hat{Z} , which is χ^2 with ν

degrees of freedom, giving

$$\hat{f}(t) = \frac{x_0 \Gamma((v+1)/2)}{\sqrt{\pi v s^2 t^3 (\theta t + 1)} \Gamma(v/2)} \times \left\{ \frac{(x_0 + \hat{\mu}t)^2}{s^2 v t (\theta t + 1)} + 1 \right\}^{-(v+1)/2}. \quad (\text{A.1})$$

Provided that $\int_0^\infty \hat{f}(t) dt > \alpha$, the variable $\tilde{W}_{1-\alpha}$ is now defined by the relation

$$\int_0^{\tilde{W}_{1-\alpha}} \hat{f}(t) dt = \alpha, \quad (\text{A.2})$$

otherwise $\tilde{W}_{1-\alpha} = \infty$, which means that the entire interval is in infinity. The distribution $\hat{f}(t)$ was named the normal-gamma mixture of the inverse Gaussian distributions by Whitmore (1986). A further discussion and some applications of the distribution are given by Seshadri (1999). Notice that the distribution does not integrate to unity, since the underlying inverse Gaussian does not integrate to unity when $(\hat{\mu} + \hat{U}s\sqrt{\theta v/\hat{Z}})$ is positive, corresponding to the fact that Brownian motions with nonnegative drift may never go extinct but rather be absorbed at infinity.