Guidelines when estimating temporal changes in density dependent populations

Erik Blystad Solbu [a,*], Steinar Engen [a], Ola Håvard Diserud [b]

[a] *Centre for Biodiversity Dynamics, Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway*
[b] *Norwegian Institute for Nature Research, N-7485 Trondheim, Norway*

ABSTRACT

Anthropogenic activity can cause changes in the population dynamics of species. The changes can be modelled by density dependent models with time varying parameters. The following study looks at the accuracy of model parameter estimates using primarily simulations, in addition to a real data set of Grey Heron. A key point is the amount of data required to detect deterministic changes, either step-wise or gradual, in parameters for species with different population dynamics. The theta-logistic model is used to simulate the data and fitted to realizations of step-wise change in growth rate, and a linear model is fitted to gradual or step-wise changes in carrying capacity. Bayesian analysis is applied to study the effect of different prior distributions on the strength of density regulation. The range of the data is especially important when trying to detect step-wise changes in growth rate. Detection of changes in carrying capacity depends on the dynamics of the population, e.g. it is difficult to observe change for species with long return time to equilibrium within short time frames. The estimates of change in carrying capacity can become more accurate using a strong prior on the strength of density regulation. However, the prior may give more conservative estimates, if the prior assumes a weak density regulation. The results provide ecologists and decision makers with a general idea of what to expect of analyses of time series data of populations in changing environments.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In population ecology, stochastic density dependent models are used to describe fluctuations in population size and ultimately predict how species are expected to persist in the future. Some species have a large growth rate and are able to increase their abundance rapidly after environmental perturbations, while other species produce fewer offspring, are less abundant and will have a slow return to carrying capacity. The range of possible species dynamics can be described by different parameterizations of the stochastic population models. Provided that the model used is reasonable, the inference about the species dynamics is based on some method of estimation. The estimation methods often assume that the environment of the species is stationary, which means that the parameters of the stochastic models are constant over the time period the species is observed. The stationary assumption may be reasonable over shorter time intervals, or when there is no reason to believe that any changes to the environment have occurred.

However, anthropogenic activity can create a non-stationary environment, causing temporal changes in population parameters, such as habitat reduction because of deforestation, change in number of offspring due to food availability, change in age of maturity following harvesting and greater variation in weather conditions as a consequence of climate change. When the cause of the potential change in population dynamics is assumed to be known, deterministic changes in the parameters can be considered, e.g. step-wise change at a known time point. How temporal changes in different model parameters can affect the dynamics of a density dependent population have been discussed in Solbu et al. (2013). Basically, the species' response to change in the parameters depends on the initial dynamics of the species, so that those with high reproduction/short lifetime respond quicker to change than those with low reproduction/long lifetime. Because species specific characteristics determine the response to change in the environment, it is important to assess how accurately these parameters, in addition to any temporal change, can be estimated.

Several methods are available for estimating parameters in density dependent population models, both when the relationship between growth rate and log population size is linear (Dennis et al., 2006; Dennis and Ponciano, 2014) and non-linear (Wang, 2007;

Pedersen et al., 2011). These approaches account for observation error and include both frequentist and Bayesian methods. Following Bolker et al. (2013), the software and packages applied in this work are straightforward to use for ecologists and can be easily checked and modified. The non-linear model uses a slightly modified JAGS procedure as described by Pedersen et al. (2011) and Bolker et al. (2013), while the linear model uses the Integrated Nested Laplace Approximation (INLA) (Rue et al., 2009). Both methods are within the Bayesian framework so that the effect prior knowledge have on the estimated temporal changes can be investigated. The difficulties with non-stationary time series, and the pitfalls of the common de-trending approach, is well known in ecology (Turchin, 2003). While a time series can indicate a trend in abundance, the observations could simply be caused by a natural perturbation of a species with long return time to equilibrium. Similarly, failing to account for a trend could result in over-estimating the return time to equilibrium of the species. A remedy for separating a trend in abundance from natural long perturbations away from equilibrium is obtaining a larger sample size. However, acquiring more observations is often impossible in ecology, e.g. when assessments of species viability must be made now, as a basis for management of natural resources. The goal of this work is to quantify the challenges of detecting change in a population parameter using simulations. Realistic sample sizes of biological data, and different population dynamics of species, are key aspects that will be investigated. In addition, the effect of sampling error and prior knowledge on the species' parameters will be studied. Finally, the range of the data will be considered, i.e. whether the population abundance is observed in a growth phase from small population size up to carrying capacity, or only observed fluctuating around the equilibrium.

2. Materials and methods

2.1. Theta-logistic model with time varying parameters

The theta-logistic population model (Gilpin and Ayala, 1973), can be approximated by a diffusion process, provided that the yearly fluctuations are small or moderate (20–30%) (Lande et al., 2003),

$$dN_t = \mu_N(n)dt + \sqrt{\nu_N(n)}dB_t \qquad (1)$$

where the infinitesimal mean is $\mu_N(n) = r_1 n[1 - (n^\theta - 1)/(K^\theta - 1)]$ for $\theta > 0$ and $\mu_N(n) = r_1[1 - \ln(n)/\ln K]$ for $\theta = 0$, which is the Gompertz model (Gompertz, 1825). Here r_1 is the growth rate at population size $N = 1$, θ the shape of density regulation and K the carrying capacity, defined as the population size where the deterministic growth rate is equal to zero. The fluctuations in population size are caused by environmental and demographic stochasticity, modeled by the infinitesimal variance as $\nu_N(n) = \sigma_e^2 n^2 + \sigma_d^2 n$, where σ_e^2 and σ_d^2 are the environmental and demographic variance, respectively. The environmental variance is defined as the variation in mean individual fitness across years, while the demographic variance is the mean variation in individual fitness within years (Engen et al., 1998). Equation (1) transformed to the log scale $X_t = \ln N_t$ becomes

$$dX_t = \mu_X(x)dt + \sqrt{\nu_X(x)}dB_t \qquad (2)$$

where $\mu_X(x) = r_1\{1 - [(e^x)^\theta - 1]/[K^\theta - 1]\} - (\sigma_e^2 + \sigma_d^2/e^x)/2$ for $\theta > 0$, $\mu_X(x) = r_1\{1 - x/\ln K\} - (\sigma_e^2 + \sigma_d^2/e^x)/2$ for $\theta = 0$ and $\nu_X(x) = \sigma_e^2 + \sigma_d^2/e^x$. The example of deterministic temporal change in the parameters considered for the theta-logistic model, substituting $\mu_X(x)$ by $\mu_X(x, t)$ in Eq. (2), is a step-wise change in growth rate, replacing r_1 by $r_1(t) = r_1 \rho^{I(t \geq s)}$. Here, I is an indicator

function equal to one if true and zero otherwise, and s, which is assumed to be known, is the first time point after the parameter value is altered. The proportional change is determined by $\rho > 0$, where $\rho < 1$ and $\rho > 1$ are decrease and increase in growth rate, respectively.

2.2. Linearized model with time varying parameters

The linearization of $\mu_X(x)$ around $\ln K$ is $\mu_X^*(x) = \beta(\alpha_0 - x)$ where $\alpha_0 = \ln K - (\sigma_e^2 + \sigma_d^2/K)/[2\gamma + \sigma_d^2/K]$ and $\beta = \gamma + \sigma_d^2/(2K)$ for $\theta > 0$ and where $\gamma = r_1\theta/(1 - K^{-\theta})$ is the strength of density regulation (May, 1981). For the Gompertz model, $\alpha_0 = \ln K - (\sigma_e^2 + \sigma_d^2/K)/(2r_1/\ln K + \sigma_d^2/K)$ and $\beta = r_1/\ln K + \sigma_d^2/(2K)$, thus $\mu_X(x) = \mu_X^*(x)$. When demographic variance is assumed to be zero, the linearized model is known as an Ornstein–Uhlenbeck process (Karlin and Taylor, 1981). Since r_1 is confounded with θ by the return time $\approx r_1\theta$, the analysis for this model is limited to step-wise and gradual changes in carrying capacity, where the changes are parametrized by replacing K with $K(t) = K\kappa^{I(t \geq s)}$ and $K(t) = K\kappa^t$, respectively. The magnitude of change is determined by $\kappa > 0$, where $\kappa < 1$ and $\kappa > 1$ expresses decrease and increase in carrying capacity, respectively. When demographic variance is accounted for, K occurs in the denominator of both α_0 and β. To simplify the calculations, the fraction σ_d^2/K is approximated by $\sigma_d^2/\bar{K}$, where $\bar{K}$ is the average carrying capacity: $\bar{K} = K[s + (n - s)\kappa]/n$ or $\bar{K} = [K\sum_{t=0}^{n} \kappa^t]/n$ for step-wise and gradual change, respectively. The resulting time varying infinitesimal mean is, $\mu_X^*(x, t) = \beta(\alpha_0 + I_{(t \geq s)}\alpha_1 - x)$ for the step-wise and $\mu_X^*(x, t) = \beta(\alpha_0 + t\alpha_1 - x)$ for gradual change in carrying capacity, where $\alpha_1 = \ln \kappa$. A detailed discussion of the dynamics of temporal changes in carrying capacity and other parameters for the linear model is found in Solbu et al. (2013).

2.3. Estimation

Although demographic variance is defined in the model, time series of species abundance contains no information about individuals' number of offspring or survival from one census to the next, which is needed to estimate demographic variance with some accuracy (Engen et al., 1998). Therefore, it is assumed that demographic variance can be estimated from individual data from the population or similar populations, or from expert information, and used as a known parameter in the model. Ignoring the effect of demographic variance could result in over-estimation of the environmental variance in the population (Sæther et al., 2000). However, if the population is above $10\sigma_d^2/\sigma_e^2$, the effect of demographic variance is usually negligible (Lande et al., 2003).

To account for measurement uncertainty of population size, the observed log abundance Y_t is assumed normally distributed with expectation X_t and variance τ^2, corresponding to the coefficient of variation being approximately constant (Seber, 1982). There are several methods available, both frequentistic and Bayesian, for estimating the parameters in non-linear state-space, or hierarchical, models such as the theta-logistic model, giving similar results for large sample sizes ($n = 200$) (Wang, 2007; Pedersen et al., 2011). Following the guidelines by Bolker et al. (2013), extensively used estimation methods are chosen, which are simple to examine and modify by other researchers. Since one of the objectives is to study the effect of incorporating prior knowledge from biologists and ecologists, Bayesian inference must be applied. For the theta-logistic model, a modification of the code provided by the National Center for Ecological Analysis and Synthesis web site (Bolker et al., 2013) is applied using JAGS (Plummer, 2003) and run in R (R Core Team, 2014).

The parameters of the linear model are estimated using the Integrated Nested Laplace Approximation (INLA) method (Rue et al.,