

The theta-logistic is unreliable for modelling most census data

Francis Clark¹, Barry W. Brook¹, Steven Delean¹, H. Reşit Akçakaya² and Corey J. A. Bradshaw^{1,3*}

¹The Environment Institute and School of Earth & Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia; ²Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794, USA; and ³South Australian Research and Development Institute, P.O. Box 120, Henley Beach, SA 5022, Australia

Summary

1. The theta-logistic is a simple and flexible model for describing how the growth rate of a population slows as abundance increases. Starting at r_m (taken as the maximum population growth rate), the growth response decreases in a convex or concave way (according to the shape parameter θ) to zero when the population reaches carrying capacity.
2. We demonstrate that fitting this model to census data is not robust and explain why. The parameters θ and r_m are able to play-off against each other (providing a constant product), thus allowing both to adopt extreme and ecologically implausible values.
3. We use simulated data to examine: (i) a population fluctuating around a constant carrying capacity (K); (ii) recovery of a population from 10% of carrying capacity; and (iii) a population subject to variation in K . We show that estimates of extinction risk depending on this or similar models are therefore prone to imprecision. We refute the claim that concave growth responses are shown to dominate in nature.
4. As the model can also be sensitive to temporal variation in carrying capacity, we argue that the assumption of a constant carrying capacity is both problematic and presents a fruitful direction for the development of phenomenological density-feedback models.

Key-words: abundance, density dependence, feedback, measurement error, population growth rate, Ricker, theta-logistic, time series

Introduction

Historically, the need for quantitative fisheries management led Ricker (1954, 1958) to develop his highly influential phenomenological model of population dynamics (Geritz & Kisdi 2004), encapsulating the key idea of density dependence as the mechanism preventing unbounded growth. We set out the Ricker model first, and then introduce the shape parameter θ to arrive at the theta-logistic model. Following Lotka (1925) and Volterra (1926) who had formalized and explored simple and specific models of predator–prey interactions, Ricker also built on the work of Verhulst (1838) and Pearl (1828) in generalizing the idea of a population tending to fluctuate around an equilibrium abundance.

What has come to be known as Ricker's model can be written formally (Cook 1965; Royama 1992) as a difference equation:

$$N_{t+1} = N_t e^{r_m \left(1 - \frac{N_t}{K}\right) + \varepsilon(0, \sigma^2)}, \quad \text{eqn 1}$$

that is, the abundance at the next time point N_{t+1} is modelled as being equal to the current abundance N_t multiplied by a value e^r , where $r = r_m \left(1 - \frac{N_t}{K}\right) + \text{process error}$ is the growth rate over the time step. The deterministic component of the growth rate declines linearly from a maximum (intrinsic) growth rate r_m at low abundance, to zero when $N_t = K$ (i.e. K quantifies the carrying capacity), and thence negative for $N_t > K$ (identical to the $\theta = 1$ trace in Fig. 1a). The *process error* (also called the *environmental variation*) is used in model fitting to capture the real-world differences between population census data and the idealization of the model (see Methods; also Appendix S1, Sections 2 and 6, Supporting information).

*Correspondence author. E-mail: corey.bradshaw@adelaide.edu.au
Correspondence site: <http://www.respond2articles.com/MEE/>

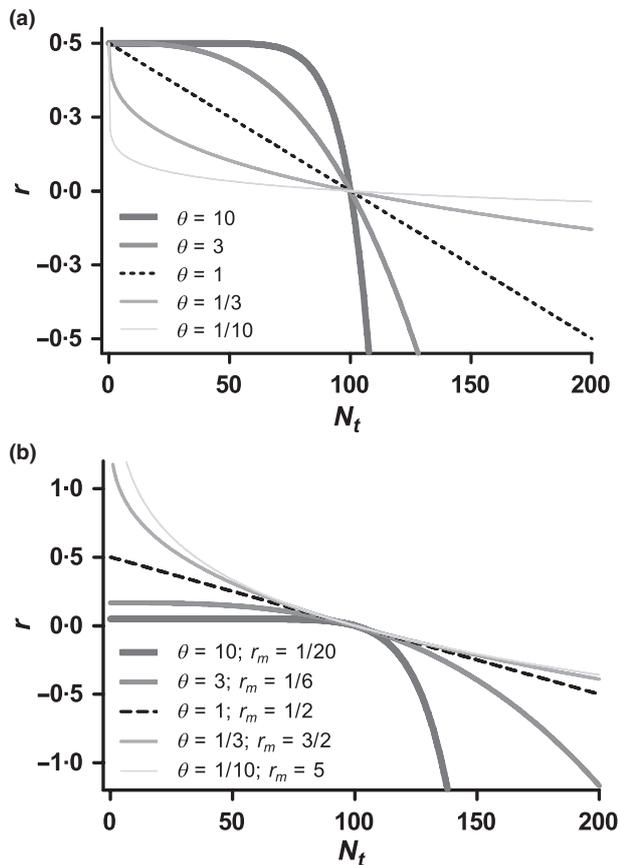


Fig. 1. Different forms of the r - N (growth-response) curve. (a) A typical demonstration of how various values of θ alter the shape of the curve (generating model: $r_m = 0.5$, $K = 100$); (b) using the same values of θ , but with r_m chosen to preserve the slope of the linear region around $N = K$ where most of the data tend to cluster (i.e. $\theta r_m = 0.5$).

Although it can be mathematically ‘derived’ from first principles in various ways (Royama 1992), the Ricker model and its variants are most usefully understood as a mathematical formulation of the ideas they represent (Fig. 1). It is important to appreciate that the Ricker family of models are phenomenological, as opposed to physical or mechanistic. One particularly desirable feature of Eqn 1 is the maintenance of non-negative abundance values (i.e. $e^r > 0$, for all r). Another is that simple linear regression of $r_{\text{obs}} = \log_c \left(\frac{N_{t+1}}{N_t} \right)$ (sometimes called pgr for per-capita growth rate – e.g. Sibly *et al.* 2005) against N_t allows estimates of K and r_m to be extracted from abundance time series: the intercept of the regression (i.e. the $N = 0$ intercept) equates to r_m ; the slope equates to $-r_m/K$. Note that the fitting approach we demonstrate here can allow K to take any form that the modeller chooses to apply.

Ricker’s phenomenological model is one of a family that includes the Gompertz (Pollard, Lakhani & Rothery 1987), Beverton–Holt (Beverton & Holt 1957) and theta-logistic models, which collectively are used often in applied ecology to estimate maximum sustainable yield targets (Cameron & Benton 2004), temporal abundance patterns (Sæther, Engen & Matthysen 2002), the most effective wildlife management

interventions (Caughley & Sinclair 1994), extinction risk (Philippi *et al.* 1987) and epidemiological patterns (Anderson & May 1991). This family of models encapsulates three principal abstractions: that (i) carrying capacity K is constant, (ii) factors driving population dynamics can be split into intrinsic (density-dependent) and extrinsic (environmental variation) categories and (iii) the population growth rate r can be expressed as a simple function of abundance N . It is this last abstraction, being the form of the r - N relationship (what we call the ‘growth response’ in ecological context) that differentiates the various models; in the Ricker model, this is simply a linear relationship.

The extrinsic factors (e.g. environmental variation in food abundance, predation, disease, breeding conditions) overlay, or can even dominate, the deterministic density dependence. Early debate centred on whether (Andrewartha & Birch 1954) or not (Nicholson 1957) extrinsic forcings overwhelmed intrinsic ones, with modern work establishing that density dependence often plays an important role in species’ dynamics (Turchin & Taylor 1992; Woiwod & Hanski 1992; Zeng *et al.* 1998; Lande *et al.* 2002; Brook & Bradshaw 2006).

While the Ricker model assumes a linear decline in the growth rate with abundance, both theoretical and empirical work (Fowler 1981; Johst, Berryman & Lima 2008) have shown that some populations have concave (as viewed from above) r - N curves. Concave r - N curves are typical of so-called ‘ r -selected’ organisms where density dependence acts strongly at lower densities, whereas convex curves arise from ‘ K -selected’ species where density dependence acts to reduce growth only at higher densities (although these terms have fallen out of favour). Studies of laboratory fruit fly populations by Gilpin & Ayala (1973) led them heuristically to add a shape parameter θ to the growth term in Ricker’s model, such that;

$$r = r_m \left(1 - \left(\frac{N_t}{K} \right)^\theta \right), \quad \text{eqn 2}$$

which resulted in the Theta-Ricker or theta-logistic model (Gilpin & Ayala 1973; Thomas, Pomerantz & Gilpin 1980). Thus, the growth response is parameterized jointly by r_m and θ (Fig. 1). For a population growing from low abundance, the parameter θ reflects those aspects of a species’ evolved life history (demographic rates) that determine how abruptly growth slows as abundance interacts with resource availability (Freckleton *et al.* 2003; Sibly *et al.* 2005) and type of competition (Johst, Berryman & Lima 2008). When concave ($\theta < 1$), the growth response (i.e. the ‘ r - N curve’) ideally characterizes a population unable to recover quickly from extrinsic perturbations, whereas a convex growth response ($\theta > 1$) implies that density feedback occurs mainly above some (relatively large) threshold abundance (Fowler 1981; Owen-Smith 2006), typical, for instance, of the population dynamics in large mammals (McCullough 1999; Owen-Smith 2006). For $\theta = 0.1$, the growth response at $N = K/2$ is 6.7% of r_m compared with 50% for the Ricker ($\theta = 1$), and 99.9% when $\theta = 10$; in the latter case, growth drops to 90% of r_m at $N \sim 0.8 K$ and to 50% at

$\sim 0.93 K$. We henceforth use this range ($0.1 \leq \theta \leq 10$) as sufficiently inclusive for most real populations (Fig. 1).

Census data for real populations usually show a fluctuating population rather than one growing steadily from low abundance. Thus, when fitting models to such data, what is actually being measured is only that part of the growth response for $N \sim K$. Also, as these phenomenological models describe only a population's tendency to return to carrying capacity, the actual dynamics of a real population over each time step will often differ greatly from the tendency suggested by the model.

Although mechanistic models of population dynamics are typically more desirable to investigate the role of density feedback, for most populations census data are more readily available than detailed life-history data, especially those of sufficient duration to estimate process variability in vital rates (Brook & Bradshaw 2006). As such, a primary applied use of the theta-logistic is to describe growth dynamics at population sizes distant from K : for instance, to determine viability of small populations or sustainable yields of harvested populations (Philippi *et al.* 1987; Cameron & Benton 2004). Because maximum sustainable yield occurs where the product of abundance and growth rate are maximal, the use of an incorrect growth response can lead to a calculated yield that is sub-optimal or unsustainable (i.e. if curvature is less or more concave than supposed).

Sibly *et al.* (2005) analysed abundance time series for mammals, birds, fish and insects and controversially (Doncaster 2006; Getz & Lloyd-Smith 2006; Ross 2006; Sibly *et al.* 2006) found that concave r - N curves dominated (i.e. concave growth responses if taken as ecologically real). This substantive work has been important in bringing attention to the problems we examine. Despite differences in methodology between Sibly *et al.* (2005) and that used here (examined and compared explicitly in the Appendix S1, Section 4, Supporting information), we also observe a dominance of concave r - N curves when fitting real abundance data; however, we show that these cannot be taken as ecologically realistic and thus reject the conclusion that concave growth responses are shown to dominate in nature.

In this work, we demonstrate and clearly explain our observation that estimating the curvature of the growth response from abundance data, in particular using the theta-logistic model, is highly problematic. The specific technical point at the core of this explanation has been independently shown by Polansky *et al.* (2009), and we thus focus on providing a complementary and less technical explanation of the specific problem and its consequences, while also exploring some general problems in obtaining meaningful results when applying phenomenological models to census data. While we include some analysis and discussion of real data, the foundational demonstrations are based on simulated data; the logic being that simulated data have known dynamics and thus allow the results of model fitting to be interpreted against a known reality. We use simulated data to examine: (i) a population fluctuating around a constant carrying capacity (i.e. stationary time series), (ii) recovery of a population from 10% of carrying capacity, and (iii) a population subject to sinusoidal variation in K . Further, we: (i) demonstrate the usefulness of confidence intervals from

resampling, (ii) contrast the simulation results with those from real abundance data, and (iii) provide an example that demonstrates the implications of poorly defined θ for estimating extinction risk. Throughout we stress the relationship between the parameters θ and r_m in jointly parameterizing the r - N curve for the theta-logistic model. We also (iv) examine what can happen when the assumption of a constant K is not met.

The overall approach – back-fitting simulated data to examine how readily generating parameters can be recaptured – is a standard technique that has broad applicability in ecological modelling. The contrast between the analysis of simulated and real data illustrates this point, and also the need for careful consideration of model assumptions, measurement error and confidence intervals when using models to estimate real ecological parameters. An extensive Appendix S1 (Supporting information) contains (i) further details and explanations in relation to the practical aspects of model fitting that will be useful for those not well-versed in fitting mathematical models to ecological data, and (ii) some more technical aspects, in particular, the geometry of the objective landscape (likelihood profile) which is characterized by $r_m - \theta$ gullies containing an arc of parameter fits of similar likelihood (equivalent to the ‘ridges’ described in Polansky *et al.* 2009).

Materials and methods

SCRIPTS AND DATA

Aside from the use of RAMAS (Akçakaya 2007) in the population viability example, all of the simulation and fitting work was implemented with carefully written and debugged Matlab[®] (Mathworks, Natick, Massachusetts, USA) scripts and validated with similar scripts developed semi-independently in R (R Development Core Team 2009). All scripts and data files are available in the Appendix S1 (Supporting information).

THE THETA-LOGISTIC MODEL

Following the basic form of the Ricker model, the theta-logistic model can be written as:

$$N_{t+1} = N_t e^{r_m \left(1 - \left(\frac{N_t}{K}\right)^\theta\right) + \varepsilon(0, \sigma^2)}, \quad \text{eqn 3}$$

or rearranged for fitting as:

$$r_{\text{obs}} = \log_e \left(\frac{N_{t+1}}{N_t} \right) = r_m \left(1 - \left(\frac{N_t}{K} \right)^\theta \right) + \varepsilon(0, \sigma^2), \quad \text{eqn 4}$$

where N_t is the population abundance at time t ; r_m is the intrinsic maximum growth rate; r_{obs} is observed growth rate; K is carrying capacity; θ is the shape parameter; and with the stochastic variable $\varepsilon(0, \sigma^2)$ quantifying the process error (also called *environmental variation*) as Gaussian with mean zero and variance σ^2 . While it is possible to apply the error to N rather than r , it is conventional to regress r_{obs} against N_t , and in any case exogenous factors that contribute to process error in r_{obs} affect ‘birth’ and ‘death’ rates additively (Turchin 2003). We show in the Appendix S1 (Section 8, Supporting information) that our results are robust to this choice. It might also be appropriate in some circumstances to consider the process error as autocorrelated

or deriving from some other distribution; however, we do not consider these issues here. Note that there are two reasons for including the process error in the expression of the model: first, this details explicitly how the differences between the model and the data are being quantified, and second, to manage these differences implicitly as residuals in model fitting can give a misleading impression that they are small (they are not).

SIMULATED ABUNDANCE TIME SERIES

Abundance time series are simulated directly from the theta-logistic model. For stationary series, an initial population size is established based on 10 steps of pre-simulation starting with $N = K$. For a recovering population, we start at $0.1K$ and set r_m (θ and r_m are tightly correlated – see Results) to give an expected recovery to $0.9K$ over 15 time steps (within 20-step time series). This gives r_m ranging from 1.9 at $\theta = 0.1$ to 0.15 at $\theta = 10$. When applying measurement error to simulated time series, an error for each abundance value is drawn from a Gaussian distribution with mean zero and standard deviation equal to the percent measurement error applied to the abundance value.

REAL ABUNDANCE TIME SERIES

We refined a data set of 1198 abundance time series originally used in Brook & Bradshaw (2006) derived principally from the Global Population Dynamics Database (Imperial College London; <http://www3.imperial.ac.uk/cpb/research/patternsandprocesses/gpdd/>). We restricted these data to include only series with at least 19 transitions (i.e. consecutive non-null abundance measurements). The resulting 570 time series were examined manually and 184 were rejected principally due to runs of extremely low counts (values of 1 and 2). The remaining 386 abundance series were then fitted against the theta-logistic model, with 22 failing to fit (i.e. giving negative parameter values or an estimate of K that did not pass through the data range). This left 364 real abundance time series to compare with the simulation results.

These remaining series were further examined manually to obtain a smaller number of series from which to draw example cases. Many of the 364 census time series contained questionable extreme values, including cases of large increases from low values in a single (yearly) time step (e.g. $3 \rightarrow 10$). In 53% of the time series, the range of abundance values was greater than an order of magnitude. A reduced set of 99 series was developed using the following heuristic: (i) set to null any abundance values $N_t < 5$, (ii) restrict the range of abundance values to no more than 10-fold by iteratively setting to null any values $N_t < K/5$ or $N_t > 2K$, (iii) reject any series with a median of < 20 , and (iv) consider blocks of consecutive time series of the same length to be potentially redundant and to select from these only a small sample of cases. The manual reduction proceeded, as above, on the basis of the series themselves (i.e. no consideration of species), and with time series too severely disrupted by the imposition of null values rejected. Although we expect this procedure both excluded biologically meaningful data and maintained inclusion of some problematic data, this reduced set is used only for selecting manageable numbers of examples for illustration.

FITTING PROCEDURE

To fit the theta-logistic model to abundance data, we used standard methods (see also Appendix S1, particularly Sections 2 and 6, Supporting information). We use maximum likelihood via numerical

optimization of a log-likelihood (\mathcal{L}) function subject to the parameter vector $[r_m, \theta, K, \sigma]$. That is, for each transition $N_t \rightarrow N_{t+1}$, calculate:

$$\varepsilon_t = \log_e \left(\frac{N_{t+1}}{N_t} \right) - r_m \left(1 - \left(\frac{N_t}{K} \right)^\theta \right), \quad \text{eqn 5}$$

and evaluate the Gaussian probability density function at ε_t (for the given σ) to obtain P_t , the likelihood of that transition. The objective function (negative log-likelihood) is constructed by summing these likelihoods as: $\mathcal{L} = -\sum \log_e(P_t)$, which we minimize against the parameter vector by numerical optimization using *fminsearch* in Matlab. This requires specification of initial parameter values, and in all cases this is taken as $\theta = 1$, with r_m and K taken from linear regression to the Ricker model, and with σ estimated directly from the residuals of this regression. Further commentary on, and explanation of, the model fitting method can be found in the Appendix S1 (Sections 2–4, Supporting information), including a comparison with the approach used by Sibly *et al.* (2005); while the Sibly *et al.* approach differs from ours in major ways, including their use of negative values of θ , these issues are peripheral to the conclusions drawn here because our results remain unchanged using either method (see Appendix S1, Section 4, Supporting information).

Of particular note are the ' $r_m - \theta$ gullies' within the objective landscape (i.e. the surface defined by the value of the objective function at each point in the parameter space, and thus the surface upon which the fitting takes place; this is also sometimes called a 'likelihood profile'). The fitting minima existed within these $r_m - \theta$ gullies, containing either shallow pools or a gradual slope to more extreme play-off between r_m and θ . Understanding these features was central to developing the robust analysis we present (see Appendix S1, Section 6 for details).

CONFIDENCE INTERVALS

Confidence intervals (CI) were calculated by bootstrapping against each value N_t and the associated $r_{\text{obs}} = \log_e \left(\frac{N_{t+1}}{N_t} \right)$. One thousand resamples were used to estimate the 90% CIs given in Table 1. An analysis of the reliability of the CI calculations is given in the Appendix S1 (Section 7, Supporting information), and shows that these are reasonably reliable overall. Confidence intervals can also be calculated directly from the likelihood profile (i.e. from likelihood ratios; see, e.g. Polansky *et al.* 2009); however, we prefer bootstrapping because it is a simple and general-purpose technique that explicitly addresses the possibility of individual data dominating the model fitting. That one calculates confidence intervals by some reasonable method is much more important than the specific choice of method.

POPULATION VIABILITY EXAMPLE

What are the ecological implications of imprecise and biased estimates of θ derived from the theta-logistic? A simple example demonstrates these. Our example is based on a well-studied population of magpie geese (*Anseranas semipalmata*) in the Northern Territory of Australia where associated demographic data were available (see Appendix S1, Section 10, Supporting information), but any equally well-studied population would suffice to demonstrate our points. That is, this example is not intended to be broadly representative of all demographic models; it is merely illustrative of the biases that can arise (in this case, for extinction risk estimates) when trying to extract the θ 'signal' by analysing the phenomenological output of a more

Table 1. Example cases of time series [from the Global Population Dynamics Database (GPDD) and Brook, Traill & Bradshaw 2006 – see Methods] concluded as strongly concave (first 7) or strongly convex (latter 2) based on the estimate of θ from the theta-logistic model

Species	Common name	Source	GPDD ID	θ (90% CI)	r_m (90% CI)
Concave					
<i>Myodes glareolus</i>	Bank vole	GPDD	47	0.036 (0.014–0.86)	39 (2.3–97)
<i>Oncorhynchus keta</i>	Chum salmon	GPDD	1667	0.009 (0.006–0.052)	64 (8.8–92)
<i>Panulirus interruptus</i>	California spiny lobster	GPDD	7026	0.007 (0.004–0.15)	35 (11–60)
<i>Parus major</i>	Great tit	GPDD	9200	0.013 (0.008–0.40)	55 (2.3–79)
<i>Martes pennanti</i>	Fisher	GPDD	9724	0.007 (0.005–0.014)	62 (21–90)
<i>Mustela sibirica</i>	Siberian polecat	Reading et al. (1998)	10299	0.019 (0.009–0.31)	34 (2.1–75)
<i>Odocoileus virginianus</i>	White-tailed deer	McCullough (1979)	20125	0.12 (0.010–0.56)	3.8 (1.0–46)
Convex					
<i>Connochaetes taurinus</i>	Blue wildebeest	GPDD	7060	8.7 (6.0–24)	0.11 (0.092–0.13)
<i>Rostrhamus sociabilis</i>	Snail kite	Beissinger (1995)	20063	5.9 (1.0–53)	0.21 (0.001–0.32)

Columns 5 and 6 show the parameter estimates for θ and r_m with 90% confidence intervals based on 1000 resamples.

complex underlying process (with real-world populations being far more complex again). We used a demographic model to project the population through time and create artificial census time series of 20 years' duration into the future. The model, incorporating spatial and temporal variation and related sources of uncertainty, built on previous analyses of the plausible rates of increase for this species (Brook & Whitehead 2005a), and was implemented using the RAMAS Metapop v5 population viability analysis software system (Akçakaya 2007). We fit the theta-logistic (using our Matlab code) to 1000 replicate time series that had been generated by the RAMAS model with a theta-logistic function imposed on the demographic survival and fertility matrix elements for each of a selection of θ values. We then used the estimate of θ obtained from the time series fit as a 'plug-in' value to re-parameterize the original RAMAS model (similar in principle to a parametric bootstrap). We then compared expected minimum abundance (EMA), which is a robust index of extinction risk (McCarthy & Thompson 2001), between the two RAMAS runs: those based on the initial θ value and those using the plug-in estimate.

Results

To demonstrate the problems using the theta-logistic model to estimate a growth response, we start with a simulation run of 1000 time series with $\theta = 1$ and fit the theta-logistic to these; all results, unless otherwise stated, are based on simulated series of length 20, $r_m = 0.5$, $\sigma = 0.1$ and $K = 100$. The distribution of fitted K is unremarkable with a median of 100.6 (70% CI of 97.7–104.1, similar to ± 1 SD); for σ , there is a small downward bias (median of 0.088, 70% CI of 0.068–0.108). The distributions of r_m and θ , however, are bimodal and mirror images of each other (Fig. 2), each with a high mode and a low mode (Fig. 2a). Critically, while the fitted values of r_m and θ individually yield estimates spanning four orders of magnitude, their product remains ~ 0.5 (Fig. 2b). Clearly, these are not independent parameters; rather, the value of each depends on the other. Although the problem of measurement error is more complex than presented here (see Discussion), it is straightforward to illustrate the known effect of measurement error (Freckleton *et al.* 2006) introducing a bias towards concave fits, with the application of 10% mea-

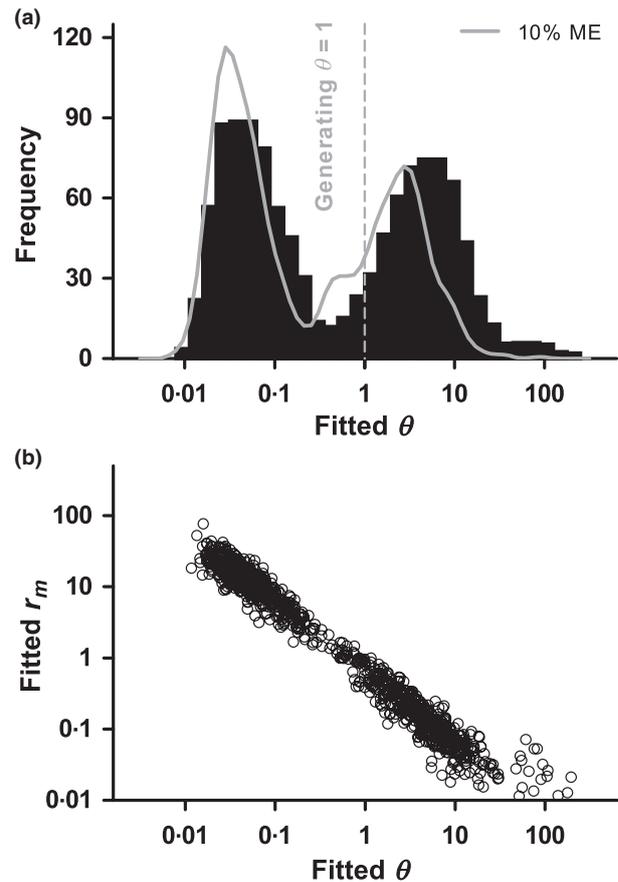


Fig. 2. Bimodality of fitted θ and relationship with r_m . (a) Fitted θ for 1000 simulations of 20 time-steps length, with generating model parameters: $r_m = 0.5$, $\theta = 1.0$, $\sigma = 0.1$ and $K = 100$ (vertical line at generating $\theta = 1.0$). Trace shows the effect of 10% measurement error (ME) applied to the abundance values; (b) \log_{10} - \log_{10} scatter plot of fitted $[r_m, \theta]$ pairs.

surement error to the abundance values increasing the number of concave fits from 53% to 64% in this case (Fig. 2a).

The same play-off between r_m and θ occurs for the real abundance data (Fig. 3), albeit with the low- θ mode dominating

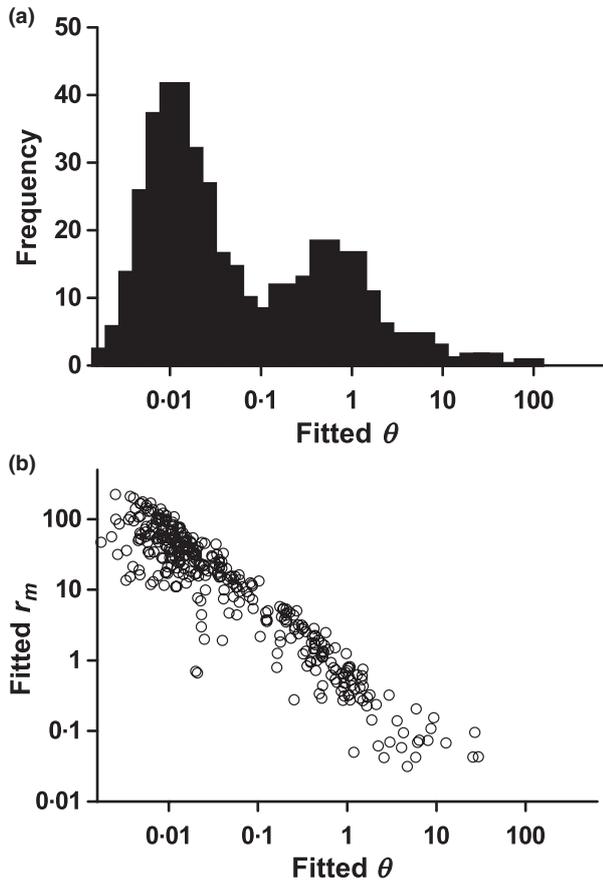


Fig. 3. Fitted θ and r_m for the 364 real abundance (census) time series and the effect of measurement error. (a) Histogram of fitted θ for the real abundance data. (b) \log_{10} - \log_{10} scatter plot of fitted $[r_m, \theta]$ pairs.

and an overall shift towards lower values of θ . The scatter plot of estimated θ vs. r_m for the real data (Fig. 3b) shows that these have a product at around 0.5, thus justifying the choice of parameters (that maintain the product) for simulations. The dominance of low θ fits observed (87% concave) is much the same as that which lead Sibly *et al.* (2005) to conclude that concave growth responses dominate in nature (see Discussion).

To provide some specific examples of strongly concave and convex r - N curves, we used the reduced set of 99 census time series (see Methods), and following Sibly *et al.* (2005), we identified those cases where the θ estimate differs from 1 at a nominal Type I error threshold of $P < 0.05$ (corresponding to the 90% CI for θ being wholly on one side or the other of $\theta = 1$); there were seven clear cases (and another three marginal cases) of concavity (Table 1), and two convex cases at around the threshold. For the seven populations estimated to have concave growth curves using the above threshold, the corresponding estimates of r_m are all > 3.8 (most > 10), which is ecologically implausible (Table 1). Because the concept of concavity or convexity can only exist in relation to a value of r_m , the estimates of θ thus also lack ecological meaning.

A key issue is that time series, both simulated and real, tend to form a cloud of data centred on $r = 0$, $N = K$,

and it is through this cloud that we seek to establish a unique curve when in fact different $[r_m, \theta]$ pairs can produce similar curves across the data domain (Fig. 1). The curves in Fig. 1a, showing the growth response for different values of θ , are close to linear around $N = K$ and these linear regions will approximately overlap given an appropriate choice of r_m (Fig. 1b). The slope at $N = K$ is preserved by maintaining the product of r_m and θ as a constant, and this relationship defines an arc of parameter pairs that produce r - N curves differing predominantly in their behaviour at extremes.

To explore whether fitting the theta-logistic to data might enable the principal form of the growth response to be distinguished correctly – that is, as concave ($\theta < 1$) or convex ($\theta > 1$) – we simulated ensembles of stationary abundance data with different generating θ (with θ from 0.1 to 10, and with the product of r_m and θ maintained at 0.5). The results are severely problematic because there is only modest departure from equal chance attribution (Fig. 4a). Increasing the product of r_m and θ to 1 only marginally improved this capacity (Fig. 4a). We also simulated the recovery of populations growing from 10% of K (see Methods) and find these can fit reasonably at low process error ($< 5\%$) (Fig. 4b; i.e. good categorization into concave or convex curvature across the range of θ). As process error increases, however, the adequacy of these fits deteriorates rapidly (Fig. 4b). While ‘carrying capacity’ (K) represents the abstraction of an average over time, real populations will experience variation in K . An example demonstrating that the model can be sensitive to variation in K is shown in Fig. 4b. With K varying as a sine wave, with one full cycle over the length of the time series and amplitude equal to $K/3$ (with the product of r_m and θ maintained at 0.5, and reduced $\sigma = 0.05$), the split into concave and convex model fits was around 75% concave across the range of θ (Fig. 4b). Finally, increasing process error can provide modest improvement (Fig. 4c). It thus remains in almost all cases that model fitting does not provide robust estimates of θ , managing at best only moderate success in distinguishing between concave and convex.

Although the central issues of imprecision, bimodality, sensitivity to departure from model assumptions, and the interdependence of r_m and θ stand on their own, the potential demographic consequences of naively using θ values derived from fitting abundance data need to be appreciated. For a mechanistic RAMAS population model for magpie geese (see Methods), the distribution of fitted θ was bimodal in all cases, with close to one half of the data associated with each of the high- θ (> 1) and low- θ (< 1) modes. When the two theta-logistic-derived modal estimates were used as the plug-in values for a re-parameterized RAMAS model (Fig. 5), the most extreme (upward) deviation in EMA was 42% for the high- θ mode at large generating θ ($\theta = 4$). That is, extinction risk was underestimated when the true curvature of the growth response was convex, and overestimated when it was concave, although percent deviation in EMA was highly variable depending on the essentially random outcome of fitting into the low- θ or high- θ mode (Fig. 5).

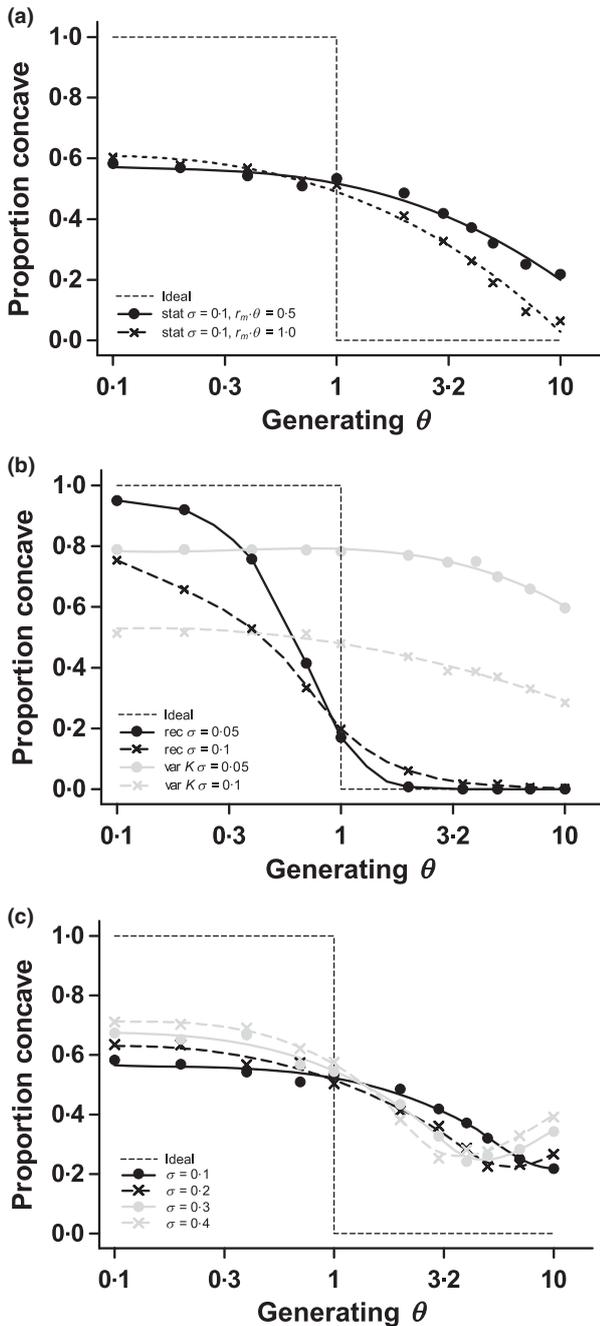


Fig. 4. Imprecision in estimated θ . Results based on 1000 simulations for each generating θ with standard error less than $\pm 2\%$ for all data. Shown is the proportion of theta-logistic fits estimating concave curvature ($\theta < 1$) in the r - N (growth–response) relationship for various scenarios: (a) stationary (stat) population trajectory with $\sigma = 0.1$ and the product of r_m and θ at 0.5 or 1.0. (b) Recovering (rec) population trajectories [from ($N_1 = K/10$) at $\sigma = 0.05$ or 0.1] or where carrying capacity is sinusoidal (var K) (see Methods). (c) Proportion concave for increasing process error ($\sigma = 0.1, 0.2, 0.3$ or 0.4).

Discussion

We have shown that abundance time series do not allow for reliable estimation of the growth response using the theta-logistic model, even categorically [i.e. distinguishing concave

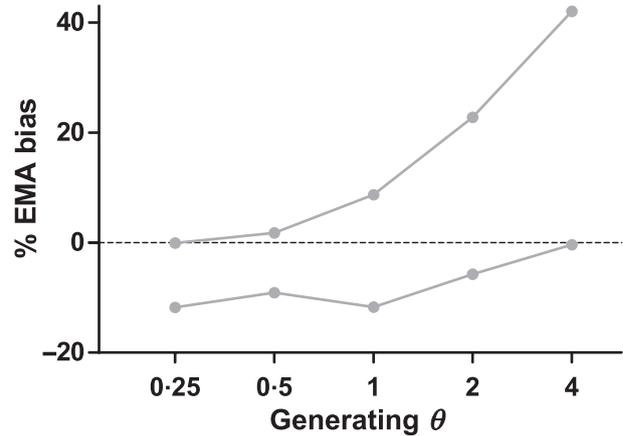


Fig. 5. Percent deviation in expected minimum abundance (EMA) as an index of extinction risk for demographic model projections of magpie goose (*Anseranas semipalmata*) populations in northern Australia (Brook & Whitehead 2005b). RAMAS-Metapop generated stochastic simulation data over 20 years with $r_m = 0.09$, $K = 42000$ for a range of generating θ values (see Methods and Appendix S1, Supporting information for details). We then fit the theta-logistic in Matlab to 1000 replicate time series derived from the RAMAS model, estimating θ , r_m and K to provide new estimates of θ with which to re-parameterize the original RAMAS model. Fitted θ is bimodal; the two modal values of θ are incorporated separately as ‘plug-in’ values for a re-parameterized RAMAS model to give the % deviation in EMA (upper and lower lines indicate predictions based on the two θ modes).

($\theta < 1$) from convex ($\theta > 1$)], except perhaps in the few cases where the population recovers from a small size and environmental variability is low. As simulation data, which are a precise realization of the model, cannot be back-fitted to give robust estimates of the generating parameters (θ and r_m in particular), we refute the notion that real-world data can be reliably interrogated using the theta-logistic model. The basic problem is that θ and r_m do not act as independent parameters; rather, it is their product that describes the slope of the growth response around carrying capacity (with census data concentrated around this region in most instances). The theta-logistic model does not fail as such – the confidence intervals do reflect the precision achieved, at least for simulated data – rather, the precision achieved is low due principally to the play-off between r_m and θ . These difficulties in obtaining meaningful parameter values from the theta-logistic model have been examined recently by others (e.g. Eberhardt, Breiwick & Demaster 2008; Sæther *et al.* 2008), but not explained prior to the work of Polansky *et al.* (2009) and, independently, this paper.

The analysis of real abundance time series shows a dominant low- θ mode (Fig. 3), as also reported by others (Sibly *et al.* 2005; Brook & Bradshaw 2006), raising an obvious question: Does this work support the claim (Sibly *et al.* 2005) that concavity in the growth response (i.e. r -selection) dominates among taxa? The lack of precision in the individual estimates of θ combined with the real data fits clustering around values of $r_m > 10$ and $\theta < 0.1$ (Fig. 3b) means that such a conclusion cannot be sustained. The r_m estimates are ecologically fanciful (Table 1). The claim that concavity in the growth

response is dominant in nature is based on faulty logic: it supposes that the biological interpretation of θ can be sustained while abandoning it for r_m .

As many different $[r_m, \theta]$ pairs can produce similar curves across the data domain (Fig. 1), it can become essentially a matter of chance as to which particular curve passes through given data to achieve a minimum in fitting. Our principal observation about the theta-logistic model is straightforward: fitting a line or curve through noisy data is inherently prone to large uncertainties, and parameterizing the growth response using the ‘flexible’ theta-logistic model allows for wild extrapolations to r_m , and thus also of θ (Fig. 1b). In sum, abundance data do not generally provide sufficient information for a robust definition of the growth response curve, it being innocuous variation in r_{obs} at the extremes of N_{obs} in particular that dominates the fitting. For real data, there might be a single extreme datum that acts to anchor the curve, giving a misleading impression that the data set as a whole supports particular values of r_m and θ (although such a misleading impression can be caught by calculating bootstrapped confidence intervals).

Given these limitations, a simpler model such as the Ricker (equivalent to setting $\theta = 1$) is preferred as a starting point; however, a nuanced interpretation of the model is needed. In fitting work, the parameter r_m arises as the zero intercept and generally represents a substantial extrapolation from the actual data. What is often of practical interest, and what is really being measured in any case, is the slope of the growth response around $N = K$. With r_m from model fitting as no more than a tentative estimate of the maximum intrinsic growth rate, it can be compared with biological expectations (e.g. Brook & Whitehead 2005a): if it is untenably high (or low), then a convex (concave) growth response might be inferred (but note following discussion on the constant- K assumption). Similarly, if a reasonable value of r_m can otherwise be established (e.g. via allometric predictions or demographic models), this can be held constant in fitting the theta-logistic model (e.g. Sæther, Engen & Matthysen 2002). Alternatively, the use of prior distributions in a Bayesian estimation framework might provide a general approach to applying constraints, especially if this incorporates variation in K as part of a prior distribution (e.g. based perhaps on territory size constraints and seasonal variation in food sources).

Biological reality aside, we see no *a priori* expectation for the observed dominance of concavity in r - N curves estimated from the real data. However, two potentially substantial problems with the application of these phenomenological models to census data are likely to be the keys to this observation: (i) the effect of measurement error in the abundance estimates and (ii) equating K to carrying capacity and assuming it remains constant. These are discussed in what follows.

Including measurement error (sometimes also called observation error) into the model fitting is not necessarily straightforward (Dennis *et al.* 2006); nonetheless, as a matter of principle and practice census data should have associated estimates of measurement error; much as derived parameters must be accompanied by a measure of spread (e.g. a standard deviation or confidence interval) to be meaningful. At the very least,

this allows for basic checks and engenders some appreciation for what is (and is not) a reasonable use of the data. In particular, most census time series represent a sample of the population, and in the absence of specifically associated errors we can reasonably default to Poissonian sampling error, for which measurement error in N_t is $\sqrt{N_t}$. In other words, the amount of measurement error depends on the magnitude of the counts, dropping below 10% only for counts greater than 100. What then is to be made of census time series laden with values less than 20, often including 1s and 2s? The problem is that transitions between low values (e.g. $2 \rightarrow 5$) produce observed growth rates that are the same as for a population growing from, in this example, 2000 \rightarrow 5000 in a single time step. Absolute scale matters for measurement error, and for transitions based on small values the error in r_{obs} can easily be larger than r_{obs} itself. It is thus essential to recognize that there are issues with real census data that make it dangerous to assume these are accurate representations of population trajectories.

Indeed, observed concavity can be partially explained by the existence of only modest measurement error (Fig. 2a). Other potential issues with the real abundance data are beyond the scope of this article (see filtering procedure described in Methods), but it is important to recall the three assumptions of the model: (i) extrinsic factors are neatly handled as process error, (ii) the form of the modelled growth response is appropriate, and (iii) density dependence is captured by the return tendency to a constant carrying capacity. It is this last assumption that we consider most tractable and fruitful for future exploration.

The concept of a constant carrying capacity is an abstraction born of necessity, and represents an average over time; this is conceptually different to ‘ K ’ in the model (e.g. Chamaillé-Jammes *et al.* 2008), which fundamentally describes the temporally varying ‘capacity’ of the environment to support the population. We contend that in most ecological contexts K will vary in time, except perhaps in controlled laboratory experiments (e.g. Gilpin & Ayala 1973). We have shown with a simple example that the theta-logistic model can be sensitive to varying K (Fig. 4b) although much investigation into the effects of such variation remains to be done. Conceptually, when K varies, it essentially becomes smeared along a range of abundance values (what we call ‘ K vagueness’), and in this case a concave r - N curve that obliquely intersects the $r = 0$ axis can provide a better fit to the data than a convex r - N curve that drops steeply through K .

Finally, as process error rises (in simulated data), estimated θ becomes less imprecise (Fig. 4c). This occurs because the model only draws the population towards K ; it is the magnitude of the process error that defines how far from K the population moves. Note that in simulation the process error is ‘pure’, as distinct from the effects of varying K and measurement error being captured as process error in model fitting. When σ is small, the range of abundance values is mostly restricted to the essentially linear part of the r - N curve (Fig. 1b), thus providing little information as to the shape of the growth response at more distal abundances. As process error increases, so too does the range of observed N , and this provides more information about the underlying growth

response. Of course, this breaks down for sufficiently high process error (Fig. 4c).

In conclusion, of the three principal abstractions of phenomenological density-feedback models, the functional form of the growth response is arguably the least tractable and most prone to misestimation. This may best be treated as linear in the region $N = K$, but with the understanding that linearity might not hold at more distal densities. If management decisions are based on the characteristic behaviour of the growth response at low densities (e.g. maximum sustainable yield at $0.5 K$ for $\theta = 1$), and if observed data do not cover this region adequately, then it is especially important to consider the uncertainties involved in the extrapolation. While the second assumption (that the 'extrinsic' or density-independent factors can be handled as process error) is amenable to careful treatment in specific instances, it is the third assumption (that density dependence acts in relation to a constant carrying capacity) that we consider most tractable and fruitful for future exploration. Major progress requires a new generation of models, within which the more or less fluctuating K experienced by real populations is explicitly acknowledged and treated. This will require ecological clarification of the concept of *carrying capacity*, experimental quantification in actual systems and mathematical work integrating the use and analysis of varying K into model fitting.

Acknowledgements

We thank S. Gregory, S. Herrando-Pérez, L. Eberhardt and R. Sibly for helpful comments. Funded by an Australian Research Council Discovery Project grant (DP0878582) to C.J.A.B.

References

- Akçakaya, H.R. (2007) *RAMAS Metapop: Viability Analysis for Stage-Structured Metapopulations (Version 5.0)*. Applied Biomathematics, Setauket, NY, USA.
- Anderson, R.M. & May, R.M. (1991) *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford, UK.
- Andrewartha, H.G. & Birch, L.C. (1954) *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago, IL, USA.
- Beissinger, S.R. (1995) Modeling extinction in periodic environments: everglades water levels and snail kite population viability. *Ecological Applications*, **5**, 618–631.
- Beverton, R.J.H. & Holt, S.J. (1957) On the dynamics of exploited fish populations. *MAFF Fisheries Investigations Series 2*, **19**, 1–533.
- Brook, B.W. & Bradshaw, C.J.A. (2006) Strength of evidence for density dependence in abundance time series of 1198 species. *Ecology*, **87**, 1445–1451.
- Brook, B.W., Traill, L.W. & Bradshaw, C.J.A. (2006) Minimum viable population size and global extinction risk are unrelated. *Ecology Letters*, **9**, 375–382.
- Brook, B.W. & Whitehead, P.J. (2005a) Plausible bounds for maximum rate of increase in magpie geese (*Anseranas semipalmata*): implications for harvest. *Wildlife Research*, **32**, 465–471.
- Brook, B.W. & Whitehead, P.J. (2005b) Sustainable harvest regimes for magpie geese (*Anseranas semipalmata*) under spatial and temporal heterogeneity. *Wildlife Research*, **32**, 459–464.
- Cameron, T.C. & Benton, T.G. (2004) Stage-structured harvesting and its effects: an empirical investigation using soil mites. *Journal of Animal Ecology*, **73**, 996–1006.
- Caughley, G. & Sinclair, A.R.E. (1994) *Wildlife Ecology and Management*. Blackwell Scientific, Boston, MA, USA.
- Chamaillé-Jammes, S., Fritz, H., Valeix, M., Murindagomo, F. & Clobert, J. (2008) Resource variability, aggregation and direct density dependence in an open context: the local regulation of an African elephant population. *Journal of Animal Ecology*, **77**, 135–144.
- Cook, L.M. (1965) Oscillation in the simple logistic growth model. *Nature*, **207**, 316.
- Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L. & Staples, D.F. (2006) Estimating density dependence, process noise, and observation error. *Ecological Monographs*, **76**, 323–341.
- Doncaster, C.P. (2006) Comment on "On the regulation of populations of mammals, birds, fish, and insects" III. *Science*, **311**. URL <http://www.sciencemag.org/cgi/content/full/311/5764/1100c>, accessed 1 April 2010.
- Eberhardt, L.L., Breiwick, J.M. & Demaster, D.P. (2008) Analyzing population growth curves. *Oikos*, **117**, 1240–1246.
- Fowler, C.W. (1981) Density dependence as related to life history strategy. *Ecology*, **62**, 602–610.
- Freckleton, R.P., Matos, D.M.S., Bovi, M.L.A. & Watkinson, A.R. (2003) Predicting the impacts of harvesting using structured population models: the importance of density-dependence and timing of harvest for a tropical palm tree. *Journal of Applied Ecology*, **40**, 846–858.
- Freckleton, R.P., Watkinson, A.R., Green, R.E. & Sutherland, W.J. (2006) Census error and the detection of density dependence. *Journal of Animal Ecology*, **75**, 837–851.
- Geritz, S.A. & Kisdi, E. (2004) On the mechanistic underpinning of discrete-time population models with complex dynamics. *Journal of Theoretical Biology*, **228**, 261–269.
- Getz, W.M. & Lloyd-Smith, J.O. (2006) Comment on "On the regulation of populations of mammals, birds, fish, and insects" I. *Science*, **311**. URL <http://www.sciencemag.org/cgi/content/full/311/5764/1100a>, accessed 1 April 2010.
- Gilpin, M.E. & Ayala, F.J. (1973) Global models of growth and competition. *Proceedings of the National Academy of Sciences of the USA*, **70**, 3590–3593.
- Johst, K., Berryman, A. & Lima, M. (2008) From individual interactions to population dynamics: individual resource partitioning simulation exposes the causes of nonlinear intra-specific competition. *Population Ecology*, **50**, 79–90.
- Lande, R., Engen, S., Sæther, B.-E., Filli, F., Matthysen, E. & Weimerskirch, H. (2002) Estimating density dependence from population time series using demographic theory and life-history data. *American Naturalist*, **159**, 321–337.
- Lotka, A. (1925) *Elements of Physical Biology*. Williams and Wilkins, Baltimore, MD, USA.
- McCarthy, M.A. & Thompson, C. (2001) Expected minimum population size as a measure of threat. *Animal Conservation*, **4**, 351–355.
- McCullough, D.R. (1979) *The George Reserve Deer Herd: Populations Ecology of a K-Selected Species*. University of Michigan Press, Ann Arbor, MI, USA.
- McCullough, D.R. (1999) Density dependence and life-history strategies of ungulates. *Journal of Mammalogy*, **80**, 1130–1146.
- Nicholson, A.J. (1957) The self-adjustment of populations to change. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 153–172.
- Owen-Smith, N. (2006) Demographic determination of the shape of density dependence for three African ungulate populations. *Ecological Monographs*, **76**, 93–109.
- Pearl, R. (1828) *The Rate of Living, Being an Account of Some Experimental Studies on the Biology of Life Duration*. Alfred A. Knopf, New York, NY, USA.
- Philippi, T.E., Carpenter, M.P., Case, T.J. & Gilpin, M.E. (1987) Drosophila population dynamics: chaos and extinction. *Ecology*, **68**, 154–159.
- Polansky, L., de Valpine, P., Lloyd-Smith, J.O. & Getz, W.M. (2009) Likelihood ridges and multimodality in population growth rate models. *Ecology*, **90**, 2313–2320.
- Pollard, E., Lakhani, K.H. & Rothery, P. (1987) The detection of density-dependence from a series of annual censuses. *Ecology*, **68**, 2046–2055.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>, accessed 1 November 2009.
- Reading, R.P., Mix, H., Lhagvasuren, B. & Tseveenmyadag, N. (1998) The commercial harvest of wildlife in Dornod Aimag, Mongolia. *Journal of Wildlife Management*, **62**, 59–71.
- Ricker, W.E. (1954) Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, **11**, 559–623.
- Ricker, W.E. (1958) *Handbook of Computations for Biological Statistics of Fish Populations*. Fisheries Research Board of Canada, Ottawa, Canada.
- Ross, J.V. (2006) Comment on "On the regulation of populations of mammals, birds, fish, and insects" II. *Science*, **311**. URL <http://www.sciencemag.org/cgi/content/full/311/5764/1100b>, accessed 1 April 2010.

- Royama, T. (1992) *Analytical Population Dynamics*. Springer-Verlag, New York, NY, USA.
- Sæther, B.-E., Engen, S. & Matthysen, E. (2002) Demographic characteristics and population dynamical patterns of solitary birds. *Science*, **295**, 2070–2073.
- Sæther, B.-E., Engen, S., Grøtan, V., Bregnballe, T., Both, C., Tryjanowski, P., Leivits, A., Wright, J., Møller, A.P., Visser, M.E. & Winkel, W. (2008) Forms of density regulation and (quasi-) stationary distributions of population sizes in birds. *Oikos*, **117**, 1197–1208.
- Sibly, R.M., Barker, D., Denham, M.C., Hone, J. & Pagel, M. (2005) On the regulation of populations of mammals, birds, fish, and insects. *Science*, **309**, 607–610.
- Sibly, R.M., Barker, D., Denham, M.C., Hone, J. & Pagel, M. (2006) Response to comments on “On the regulation of populations of mammals, birds, fish, and insects”. *Science*, **311**. URL <http://www.sciencemag.org/cgi/content/full/311/5764/1100d>, accessed 1 April 2010.
- Thomas, W.R., Pomerantz, M.J. & Gilpin, M.E. (1980) Chaos, asymmetric growth and group selection for dynamical stability. *Ecology*, **61**(6), 1313–1320. *Ecology*, **61**, 131–1320.
- Turchin, P. (2003) *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton University Press, Princeton, NJ, USA.
- Turchin, P. & Taylor, A.D. (1992) Complex dynamics in ecological time-series. *Ecology*, **73**, 289–305.
- Verhulst, P.F. (1838) Notice sur la loi que la population poursuit dans son accroissement. *Correspondance mathématique et physique*, **10**, 113–121.
- Volterra, V. (1926) Fluctuations in the abundance of a species considered mathematically. *Nature*, **118**, 558–600.
- Woiwod, I.P. & Hanski, I. (1992) Patterns of density dependence in moths and aphids. *Journal of Animal Ecology*, **61**, 619–629.
- Zeng, Z., Nowierski, R.M., Taper, M.L., Dennis, B. & Kemp, W.P. (1998) Complex population dynamics in the real world: modeling the influence of time-varying parameters and time lags. *Ecology*, **79**, 2193–2209.

Received 9 March 2010; accepted 1 April 2010
 Handling Editor: Robert P. Freckleton

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Supplementary methods and results.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.