A state–space mixture approach for estimating catastrophic events in time series data

Eric J. Ward, Ray Hilborn, Rod G. Towell, and Leah Gerber

Abstract: Catastrophic events are considered a major contributor to extinction threats, yet are rarely explicitly estimated in population models. We extend the basic state–space population dynamics model to include a mixture distribution for the process error. The mixture distribution consists of a "normal" component, representing regular process error variability, and a "catastrophic" component, representing rare events that negatively affect the population. Direct estimation of parameters is rarely possible using a single time series; however, estimation is possible when time series are combined in hierarchical models. We apply the catastrophic state–space model to simulated time series of abundance from simple, nonlinear population dynamics models. Applications of the model to these simulated time series indicate that population parameters (such as the carrying capacity or growth rate) and observation and process errors are estimated robustly when appropriate time series are available. Our simulations indicate that the power to detect a catastrophe is also a function of the strength of catastrophes and the magnitude of observation and process errors. To illustrate one potential application of this model, we apply the state–space catastrophic model to four west coast populations of northern fur seals (*Callorhinus ursinus*).

Résumé : Bien qu'on considère que les événements catastrophiques constituent des menaces d'extinction importantes, ceux-ci sont rarement estimés de façon explicite dans les modèles démographiques. Nous étendons un modèle démographique de base de type état–espace pour inclure une distribution de mélange de l'erreur de processus. La distribution de mélange comprend une composante « normale » qui représente la variabilité de l'erreur de processus ordinaire et une composante « catastrophique » qui représente les événements rares qui affectent la population de façon négative. Il est rarement possible d'estimer directement les variables à partir d'une seule série chronologique, mais cette estimation est possible lorsque plusieurs séries chronologiques sont combinées dans des modèles hiérarchiques. Nous utilisons le modèle état–espace catastrophique avec des séries chronologiques simulées d'abondance provenant de modèles de dynamique de population simples et non linéaires. L'application de ces modèles aux séries chronologiques simulées indique que les variables de la population (telles que le stock limite et le taux de croissance), ainsi que les erreurs d'observation et de processus, sont estimées avec robustesse lorsque des séries chronologiques appropriées sont disponibles. Nos simulations indiquent que la capacité de prédire une catastrophe est aussi fonction de la force des catastrophes et de l'importance des erreurs d'observation et de processus. Afin d'illustrer une application possible du modèle, nous utilisons le modèle état-espace catastrophique avec quatre populations d'otaries à fourrure du Nord (*Callorhinus ursinus)* de la Côte ouest.

[Traduit par la Rédaction]

Introduction

State–space population dynamics models have become increasingly popular in ecology with the advancement of new statistical methods and software capable of computing high-dimensional integrals. The main advantage of using state–space models is that both observation error and process error can be modeled simultaneously. Before the advent of these models, ecologists were forced to model only one of the two processes, assuming the other to be negligible. While this assumption may be valid for some populations, it is likely to introduce bias (Walters and Ludwig 1981). The most common application of state–space models has been to time series of population abundance data, because understanding natural fluctuations in population size is critical for management (de Valpine and Hastings 2002; Calder et al. 2003; Clark and Bjørnstad 2004). While a variety of processes may be responsible for observed fluctuations in abundance, the most important factors determining a species' persistence may be the frequency and magnitude of catastrophic events (Mangel and Tier 1994). Not including these events in stock assessments and population viability analyses (PVA) may create an overly optimistic view of a population's status (Young 1994).

Received 8 June 2006. Accepted 10 March 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 12 July 2007. J19355

E.J. Ward[1] **and R. Hilborn.** School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA.
R.G. Towell. National Marine Mammal Laboratory, 7600 Sandpoint Way NE, Seattle, WA 98115, USA.
L. Gerber. School of Life Sciences, Arizona State University, Box 874501, Tempe, AZ 85287-4501, USA.

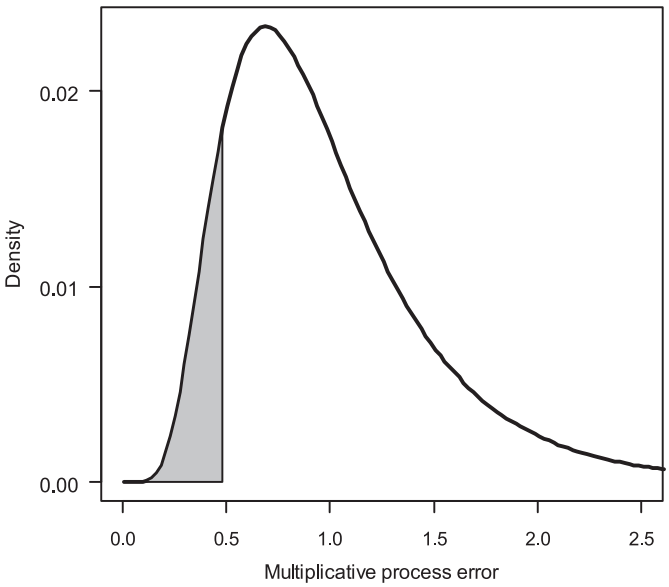
[1]Corresponding author (e-mail: eric.ward@noaa.gov).

 doi:10.1139/F07-060

All natural populations are subject to a variety of stochastic processes, including both demographic and environmental variability (Lande et al. 2003). The key difference between these two types of stochasticity is that demographic variability is simply the random process of survival or reproduction, with each individual having the same probability of surviving or reproducing. Environmental variability changes the probability of survival or reproduction for every member of a population, but all individuals still experience the stochastic process of survival and reproduction. In this paper, we assume environmental stochasticity to be the dominant process. Most PVAs and stock assessments assume that the extreme events may be represented by the tail of the process error distribution (Hilborn and Mangel 1997; Fig. 1). The trade-off of using single distribution is that the frequency of extreme negative events can only be increased by increasing the variance, thereby also increasing the frequency of extreme positive events. For many *K*-selected species that experience slow growth, incorporating extreme positive growth events into population models is not biologically reasonable (e.g., even in the presence of unlimited resources, very few large-mammal populations would be able to experience an annual doubling in population size). As a more flexible approach, we adopt a mixture distribution to model environmental variation. This mixture consists of a distribution representing the normal process error variability of the population and a distribution representing the variability that the population experiences during catastrophes. Catastrophes have been given multiple definitions in the literature (Mangel and Tier 1994; Young 1994; Erb and Boyce 1999). We choose to adopt the definition that has been generally accepted in the ecological literature (Gerber and Hilborn 2001): a catastrophe represents a single, unpredictable, rare, negative event that affects the probability of survival or reproduction in a single time step and reduces a population by more than 50%. Some populations that have experienced steady declines over multiple years, such as the Western Aleutian stock of Steller sea lion (*Eumetopias jubatus*) (Angliss and Lodge 2002), would not be considered catastrophic under our definition.

A variety of approaches have been developed to address parameter estimation in a state–space framework. Numerical search algorithms such as the Kalman filter (Kalman 1960) are limited to linear problems with normal error distributions. More recently, restricted maximum-likelihood methods have been suggested (Staples et al. 2004), but these approaches also have some limitations. An extension of the Kalman approach, the numerically integrated state–space method, has been developed as a more general solution to both linear and nonlinear systems (de Valpine 2002; de Valpine and Hastings 2002). A third approach to state–space models are Bayesian statistics, which have seen increased use in fisheries science (Meyer and Millar 1999*b*; Rivot et al. 2004), and will be the focus of this paper. One advantage of dealing with state–space models in a Bayesian or numerically integrated state–space framework is that non-normal error distributions may be applied to nonlinear population models (Punt 2003). In a hierarchical modeling framework, Bayesian methods allow for multiple data sets to be combined in an analysis and for prior distributions to be placed on shared population parameters. While Bayesian methods may give similar results to traditional maximum likelihood approaches, the key difference is that Bayesian sampling integrates over the parameter space rather than maximizing (Hobbs and Hilborn 2006).

Fig. 1. Lognormal density with an expected value of 1.0 and coefficient of variation (CV) = 0.5. The region of the curve <0.5 has been shaded to illustrate those process errors that might be considered catastrophes (i.e., in this region, the population is reduced by more than 50%).



The objectives of this paper are to (*i*) develop a model for incorporating catastrophic events into time series of abundance data, (*ii*) evaluate the ability of Bayesian methods to estimate the underlying processes, (*iii*) evaluate the power of this method to detect catastrophic events, and (*iv*) apply the catastrophic model to west coast populations of northern fur seals (*Callorhinus ursinus*).

Materials and methods

The catastrophic state–space model

State–space models require the formulation of two models: an observation model and a process model describing the transition between unobserved population states. For time series of population abundance, the observation model relates the set of population abundance estimates based on observed data to the true unobserved states of the population. Typically, the observation model takes the form $\hat{N}_t = S_t\epsilon_t$, where $\hat{N}_t$ is the estimated abundance (e.g., from a survey), S_t represents the true unobserved abundance in year t, and ϵ_t is a multiplicative error term. The set of observation error terms, $\boldsymbol{\epsilon}$, are assumed independent between years and may be assigned any valid probability distribution (lognormal, gamma, etc.):

$$(1) \qquad P(\epsilon_t \mid S_t, \sigma_o) \sim \text{lognormal}\left(-\frac{\sigma_o^2}{2}, \sigma_o\right)$$

where S_t is the true population size.

The general form of the state equation for the process component of the model is $S_{t+1} = g(S_t)\,\delta_t$, where the function $g(\cdot)$ represents a transitional relationship or growth function between states S_t and S_{t+1}, and δ_t is a multiplicative process error term. If a discrete theta-logistic growth model is used, for example (Gilpin and Ayala 1973), the process model can be written as

$$S_{t+1} = S_t\left[1 + r\left(1 - \frac{S_t}{K}\right)^{\theta}\right]\delta_t$$

where r is the logistic growth parameter, K is the logistic carrying capacity, θ controls the shape of density dependence, and the error term may again be assigned any distribution, e.g.

$$\delta_t \sim \text{lognormal}\left(-\frac{\sigma_p^2}{2}, \sigma_p\right)$$

Our catastrophic state–space model extends the basic state–space model by also assuming that each year of a particular time series is an unknown categorical state; it may be a normal year in which the population experiences regular process error variability, or it may be a catastrophic year in which the population experiences the variability associated with catastrophic events. There is no constraint placed on the magnitude of process errors in catastrophic years; the variability in those years may be less than or greater than the expected variability in normal years. Assuming independence between years, the chance of a catastrophe occurring in any particular year can be modeled as a Bernoulli trial with parameter ϕ, where ϕ represents the probability of a catastrophe (catastrophes representing successful trials). Using the laws of conditional probability, we can write this process as

$$P(S_{t+1} | Y_t, \boldsymbol{\beta}) = P(S_{t+1} | Y_t, \boldsymbol{\beta})\,P(Y_t | \phi\,P(\phi)$$

where $\boldsymbol{\beta}$ represents a vector of all model parameters (population parameters, variances, etc.), and Y_t represents a categorical variable (normal or catastrophic year). The vector of categorical states may be treated as nuisance parameters, because they are integrated out of the model. The magnitude of a catastrophic event (θ) is modeled as a multiplicative term, so if the process errors are lognormally distributed, we have

$$\delta_t \sim \begin{cases} \text{lognormal}\left(-\frac{\tau_{\text{norm}}^2}{2}, \tau_{\text{norm}}\right) & \text{with } P(1-\phi) \\ \text{lognormal}\left(\theta - \frac{\tau_{\text{cat}}^2}{2}, \tau_{\text{cat}}\right) & \text{with } P(\phi) \end{cases}$$

The parameter controlling θ represents the average percent reduction in the population size; values close to 1.0 indicate that the population is nearly driven extinct. The probability density function of this mixture is described by the equation

$$(2)\qquad P(\delta_t | S_{t-1}, \phi, \theta, \tau_{\text{cat}}, \tau_{\text{norm}}) = \frac{\phi}{\sqrt{2\pi}\,\tau_{\text{cat}}\,\delta_t}\exp\left\{-\frac{\left[\ln(\delta_t) - \theta + \frac{\tau_{\text{cat}}^2}{2}\right]^2}{2\tau_{\text{cat}}^2}\right\} + \frac{1.0-\phi}{\sqrt{2\pi}\,\tau_{\text{norm}}\,\delta_t}\exp\left\{-\frac{\left[\ln(\delta_t) + \frac{\tau_{\text{norm}}^2}{2}\right]^2}{2\tau_{\text{norm}}^2}\right\}$$

Priors and likelihood

A joint prior distribution must be assigned to all unobservable parameters and states in the model. There are two approaches to dealing with the prior distribution on the first state of the system: a prior may be placed on the state S_0, one time period before the first observation (yielding N conditional elements in the prior), or a prior may be placed on S_1, the state in the same time period as the first observation (yielding $N - 1$ conditional elements in the prior). In the catastrophic state–space model, we choose the latter approach, placing a prior on the first state, S_1. Assuming independence between parameters and conditional states, the joint prior for a simple logistic model can be expressed as follows:

$$(3)\qquad P(r, K, \phi, \theta, \tau_{\text{cat}}, \tau_{\text{norm}}, S_1, S_2, \ldots, S_N) = P(r)\,P(K)\,P(\phi)\,P(\theta)P(\sigma_o)\,P(\tau_{\text{cat}})\,P(\tau_{\text{norm}})\,P(S_1)\prod_{t=2}^{N} P(\delta_t | S_{t-1}, r, K, \phi, \theta, \sigma_o, \tau_{\text{cat}}, \tau_{\text{norm}})$$

If data on a particular species consisted of 10 observations over a 10-year period, there would be a total of 17 parameters (10 population states, 2 logistic parameters, 3 variance parameters, 2 catastrophe parameters). Choosing appropriate prior distributions must be done with caution, because arbitrary priors that are not informative with respect to one parameter may be informative with respect to another (Walters and Ludwig 1994; Punt and Hilborn 1997; McAllister and Kirkwood 1998). For the catastrophic parameters, we assigned a uniform (0, 0.5) prior to the probability of a catastrophe (ϕ) so that catastrophes would be constrained to be rare events. A uniform (0, 1) prior was placed on the magnitude of a catastrophe (θ) to express our vague knowledge about how much catastrophes reduce populations. Although this prior allows for events that are small in magnitude (e.g., a negative event might reduce a population by 3%), we do not consider values of $\theta < 0.5$ to represent catastrophes. For the purpose of this analysis, we assigned an uninformative uniform (0, 1) prior to the growth rate (r) and an informative lognormal prior to carrying capacity (K). Because we are assuming errors in the model to be lognormally distributed, the standard deviation of the error in log space can be interpreted as the coefficient of variation (CV) in normal space. We assumed that all three CVs were assigned inverse gamma prior distributions, because these priors are a common

choice in Bayesian linear models (Gelman et al. 1995). The likelihood function of the data given the parameters is derived from eq. 1:

$$(4)\qquad L(r, K, \phi, \theta, \tau_{cat}, \tau_{norm}, S_1, S_2, \ldots | \hat{N}_2, \hat{N}_1, \ldots) = \prod_{t=1}^{N} P(\epsilon_t | S_t) = \left(\frac{1}{\sqrt{2\pi}\sigma_o}\right)^N \exp\left[-\frac{1}{2\sigma_o^2}\sum_{t=1}^{N}\left(\epsilon_t - 1.0\right)^2\right]$$

or distribution (eq. 3) and the likelihood (eq. 4).

Several methods are available to generate samples from the posterior. Thomas et al. (2005) implemented a sequential-importance-sampling (SIS) approach, which appears to give similar results to Markov chain Monte Carlo (MCMC) based approaches. Metropolis sampling methods have also been applied to state–space models; however, the most popular methods have been Gibbs sampling (or adaptive rejection sampling) implemented in WinBUGS (Meyer and Millar 1999*a*; Spiegelhalter et al. 2003). We chose to implement the catastrophic state–space model in WinBUGS because the batch-mode feature of WinBUGS 1.4 enabled us to do large-scale simulations on many machines.

Simulations

Relative to population models that only consider observation error or process error, state–space models are generally data-hungry. Adding a mixture distribution to the basic state–space model introduces a further layer of complexity. When the catastrophic state–space model is fit to a single time series of data, one of the two process error variances generally converges to zero, because the model is unable to separate the components of the mixture. An alternative approach is to combine multiple time series from multiple populations. In a hierarchical setting, this would allow each population to have unique population parameters (e.g., growth rates, carrying capacities), with the only assumption being that process error variances are shared between time series. A further layer of complexity may be to combine time series from multiple species in multiple environments, each with different probabilities of catastrophes.

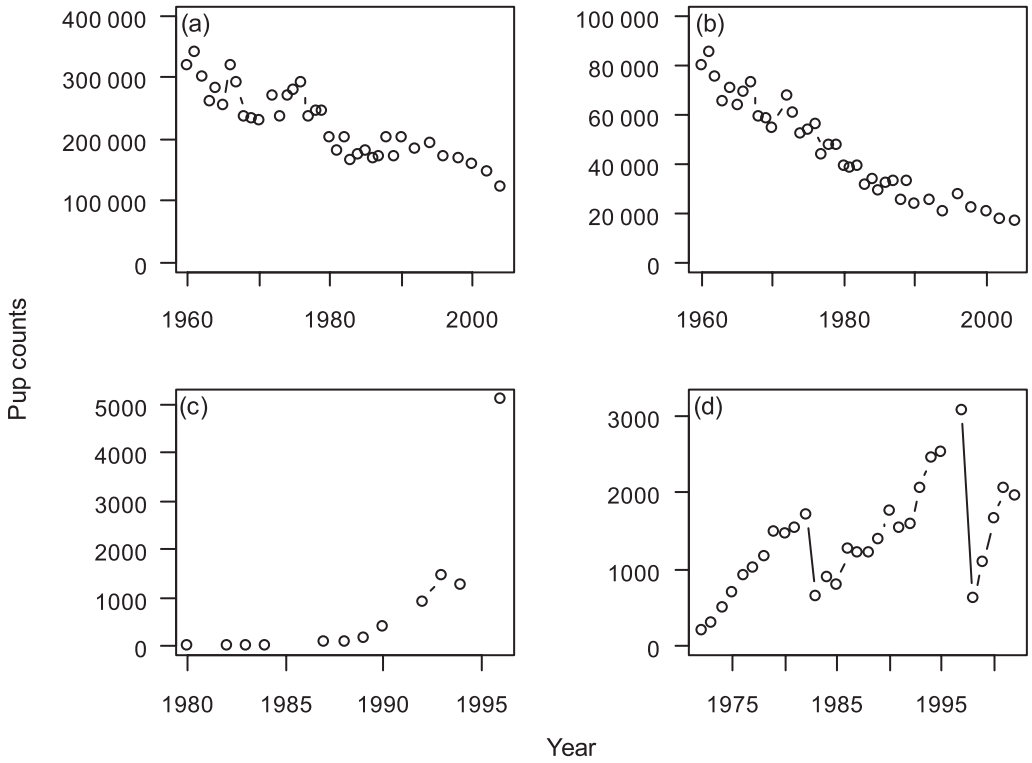
We generated 10 000 random data sets to test the catastrophic state–space model, where each data set represented a collection of two–five time series, representing data from similar species in the same environment. We assumed that these populations experienced the same observation and process errors and were affected by catastrophes similarly (e.g., the probability and magnitude of a catastrophe is assumed constant across time series in a data set). Each of the time series within a data set consisted of 5–25 simulated observations. We chose to initially consider only the exponential growth and theta-logistic models as simulation models, but each time series was allowed to have unique population parameters (initial population sizes, growth rates, carrying capacities). Observation error and process error were assumed to be lognormally distributed. The probability and magnitude of catastrophes was independent between years and time series (ϕ ranged from 0.01 to 0.2, θ ranged from 0.6 to 0.9). Instead of drawing random values of the observation and process error standard deviations for each data set, we considered each of the standard deviations to be random factors (each ranging from 0.0 to 0.6 in steps of 0.05). For some of the time series, we generated large negative process error events in the absence of catastrophes to evaluate the ability of the model to distinguish between these two types of decline. Although they are unlikely to occur, these large negative events represent the tail area of a single process error distribution (Fig. 1).

To monitor MCMC convergence, we implemented the CODA package in R (www.r-project.org). In addition to monitoring the trace plots, autocorrelation plots, and densities of each parameter, the Geweke and Heidelberger–Welch statistic were used to determine whether or not chains had converged (Gelman et al. 1995; Gilks et al. 1996; Carlin and Louis 2000). Following a burn-in of 20 000 elements, we conducted 30 000 sampling iterations, saving every 10th draw to create a posterior sample of 3000 elements of each parameter per data set. After parameters for each simulated data set were estimated, the bias was calculated for each parameter as $\hat{\theta} - \theta$ (the true value of the parameter subtracted from the parameter estimate from the posterior distribution).

Application: northern fur seals

To illustrate one potential application of our state–space catastrophe model, we considered the west coast population of the northern fur seal. Previous analyses have estimated the total probability of catastrophes in the otariid family to be in the range of 0.25–2.1%·year^{-1} (Gerber and Hilborn 2001). Historically, pup counts have been used as indices of abundance for many otariid species (e.g., Loughlin et al. 1994). The number of pups in a given year may be estimated with precision, because pups usually remain on land, concentrated in small spatial locations. Northern fur seal pup abundance data has been collected by the National Marine Fisheries Service (NMFS) at four rookeries: the Pribilof Islands (St. George, St. Paul), Bogoslof Island, and San Miguel Island. Surveys on the Pribilof Islands have been conducted since the 1950s, while counts on Bogoslof and San Miguel have been initiated more recently (1980 and 1968, respectively, when the rookeries were discovered). The *C. ursinus* populations that breed on the Pribilof Islands constitute approximately 90% of the US population and 60% of the global population (York et al. 2000). The remaining 40% of the global population is found in the western Pacific (Kuril Islands, Tyuleniy Island, Commander Islands); however, annual surveys do not exist for these populations. Data from NMFS surveys since 1960 (annual counts 1960–1990, biennial counts after 1990) indicate that pup production on the St. Paul and St. George rookeries has been declining (Towell et al. 2006; Fig. 2), and the cause for the decline remains unknown. Many of the Pribilof pup estimates have associated standard error estimates, which are computed based on the traditional Lincoln–Peterson mark–recapture estimator (Seber 1982). The CVs based on these standard er-

Fig. 2. Estimates of northern fur seal (*Callorhinus ursinus*) pup abundance for four US populations. The Pribilof Islands — (*a*) St. Paul, (*b*) St. George — and (*c*) Bogoslof Island are located in the Bering Sea, while (*d*) San Miguel Island is located off the coast of southern California.



rors are generally quite small because of large sample sizes (CVs < 0.10) and may underestimate the true measurement error of the NMFS surveys. Surveys of *C. ursinus* pups on Bogoslof Island were initiated by NMFS when the rookery was established in 1980, and although the surveys have not been conducted annually, the observations appear to show an increasing exponential trend in production (Fig. 2). Production for the San Miguel stock also has shown an increasing trend, but has experienced several catastrophic events. These negative events may be partially caused by El Niño (York 1991), but the effects are not completely understood. While the two largest declines in San Miguel pup production occurred during strong El Niño years, the *C. ursinus* population on San Miguel has experienced several other El Niño years (such as the ones that occurred over the period 1990–1994), with no apparent decline in pup numbers. While interannual fluctuations in the number of births is expected, pup production of both stocks may also be impacted by occasional catastrophic events. Catastrophic reductions in the number of *C. ursinus* pups may be an indicator that pup survival is impacted by catastrophes, but may also be an indicator of changes in the population dynamics of mature adults. If a catastrophe such as a disease outbreak were to reduce, for instance, adult survival (fewer adults results in less production) or adult fecundity, the catastrophe would be reflected by a drop in the number of births.

In developing a group of candidate models for this analysis, we extended the models used for the simulated time series by considering both population-specific parameters and ecosystem-specific parameters. Each of the four rookeries was allowed to have an island-specific growth equation. Pup counts at Bogoslof and San Miguel appeared to be increasing exponentially and were represented by the two-parameter model: $S_{t,j} = S_{0j} \exp(tr_j)$, where r_j is the rate of change for population j. Counts at St. George and St. Paul appeared to show the opposite trend (Fig. 2). Two alternative models were considered for the St. George and St. Paul rookeries: a two-parameter exponential model (similar to the one used for Bogoslof and San Miguel) and a three-parameter exponential decay model that was used in previous population assessments (York et al. 2002). This latter equation can be represented as $S_{t,j} = a_j + b_j\, 2^{(1960-t)/c_j}$, where a_j is a parameter representing equilibrium pup level (asymptote) for population j, b_j is a parameter representing the total decrease in annual pup production between 1960 and the equilibrium level a_j, and c_j is a decay term representing the half-life for population j. Growth rates in the models (r, b) were assigned uniform prior distributions, while the equilibrium (a) and decay term (c) were assigned gamma prior distributions. The equilibrium parameter represents the asymptotic production for each population, and the decay parameter represents the time necessary for the number of pups to decrease to half the difference between the number of pups at time zero and the asymptotic number of pups.

We assumed that the two catastrophe parameters in the model were ecosystem-specific, so that ϕ_j and θ_j represented the probability and magnitude, respectively, of catastrophes affecting the jth population. The primary justification in not allowing the catastrophe parameters to be shared across all populations is that the San Miguel Island population experiences El Niño events, while the Bering Sea populations do not. We were forced to assume that the process

Table 1. Deviance information criterion (DIC) values for a set of 48 candidate models describing trends in northern fur seal (*Callorhinus ursinus*) pup production.

St. Paul parameters	St. George parameters	Catastrophe	Extra CV	1 CV	2 CVs	3 CVs
3	3	Y	Y	114.9	67.5	50.1
3	3	N	Y	164.6	102.6	107.0
3	**3**	**Y**	**N**	**48.8**	**15.6**	**0.0**
3	3	N	N	104.6	52.4	52.8
2	2	Y	Y	102.7	93.7	84.8
2	2	N	Y	171.3	134.4	135.0
2	2	Y	N	50.4	27.7	12.4
2	2	N	N	109.5	66.4	66.9
2	3	Y	Y	94.4	68.2	51.1
2	3	N	Y	143.2	108.4	111.4
2	3	Y	N	48.9	20.2	5.0
2	3	N	N	108.1	57.7	70.5
3	2	Y	Y	113.3	70.2	63.6
3	2	N	Y	165.1	122.7	123.2
3	2	Y	N	45.3	20.9	17.5
3	2	N	N	106.7	63.4	60.1

Note: Models may be represented as basic state–state models or as state–space mixture models (including catastrophes) and may contain additional observation error in the form of an additive coefficient of variation (CV) parameter. For each model, the state transition equations on Bogoslof Island and San Miguel Island are assumed to be two-parameter exponential models. Pribilof populations (St. Paul, St. George) may be modeled as two-parameter or three-parameter exponential decay models. Scenarios with 1 CV constrain the observation error CVs for all surveys in years without standard error estimates to be constant across the four populations; scenarios with 2 CVs allow the Pribilof surveys to have separate observation error CVs from the Bogoslof and San Miguel surveys; and scenarios with 3 CVs allow Bogoslof and San Miguel to have island-specific observation error CVs (the Pribilof populations are assumed to share the same CV). The best-fit model with the lowest DIC score is highlighted in bold.

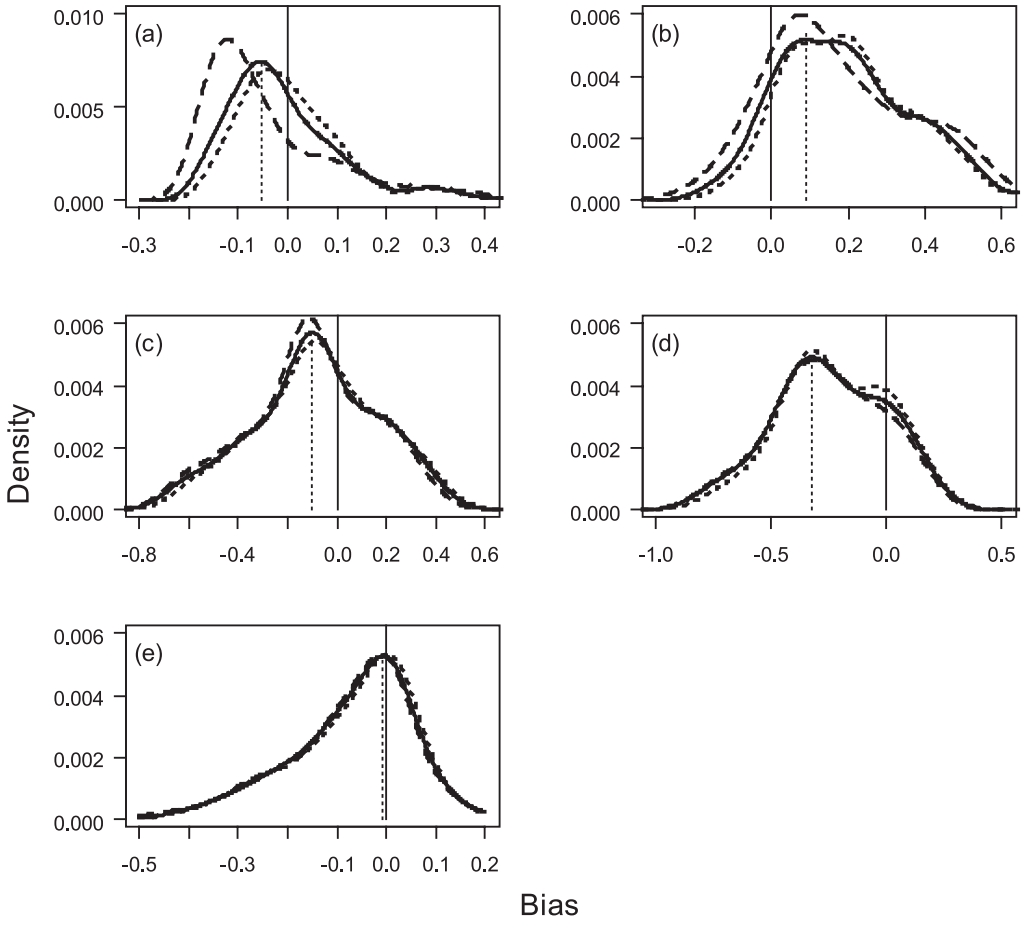
error variances (not the process errors themselves) were shared across all four populations, because the San Miguel time series alone would not be able to separate the components of the process error mixture for a model that allows ecosystem-specific variances. For each ecosystem, we assigned uniform (0, 0.5) priors to ϕ, and a uniform (0, 1) prior to θ. This latter prior was chosen to include all negative events; however, only those in the range 0.5–1.0 will be interpreted as catastrophes.

As an extension of the catastrophe model previously applied to simulated data, we assumed that observation error and process errors in all candidate models were multiplicative gamma distributions. There are several possible parameterizations of the gamma distribution; however, we chose to adopt a two-parameter version dependent on a shape parameter and CV parameter. To evaluate whether there was evidence for the catastrophic state–space model over the standard state–space model, we considered some models with a single process error distribution. All process error CVs were assigned gamma prior distributions with an expected value of 0.2.

To assess whether a model that incorporated additional observation uncertainty was appropriate, we developed several models that included additional observation error parameters. Because not all years had standard error estimates present, we were forced to assume that the CV for years with missing estimates was constant (CV_c). In the first set of models (1 CV, Table 1), we assumed that the CV for the *i*th year and *j*th population could be modeled as $CV_{i,j} = \sqrt{CV_{est(i,j)} + CV_{extra}^2}$ if the standard error estimate did exist ($CV_{est(i,j)}$ being the CV calculated from the abundance and standard error estimates) and $CV_{i,j} = \sqrt{CV_c^2 + CV_{extra}^2}$ otherwise. One problem with assuming that the CV is constant when standard errors don't exist is that surveys conducted on the Pribilof Islands may have a different CV than surveys conducted on Bogoslof Island or San Miguel Island. The justification for this difference is that pup estimates from the Pribilofs represent mark–recapture estimates, while Bogoslof and San Miguel counts are direct counts based on single observations. To account for this additional variation, we developed a second set of models (2 CVs, Table 1) that allowed the observation error CV for Bogoslof and San Miguel to differ from that of the Pribilof surveys (but which was assumed constant between years and between the two islands). Finally, we developed a third set of models (3 CVs, Table 1) that allowed the Bogoslof and San Miguel populations to have unique, island-specific observation error CVs. All observation error CVs were assigned gamma prior distributions with a mean of 0.2, with the exception of the additive parameter (CV_{extra}), which was assigned a uniform prior distribution.

All combinations of population models, process errors, and observation error scenarios yielded a total of 48 candidate models. MCMC parameter estimation for each model was done via MCMC in WinBUGS 1.4. After a burn-in of 2.0×10^5 iterations, we thinned the next 5.0×10^5 elements of the MCMC chain, retaining every 100th sample. The deviance information criterion (DIC) was also computed as a model selection tool (Spiegelhalter et al. 2003). In some instances, DIC may behave similarly to the widely used

Fig. 3. Density plots of the bias in posterior estimates of the probability of a catastrophe (ϕ, panel *a*), the magnitude of a catastrophe (θ, panel *b*), the observation error coefficient of variation (CV) (σ_o, panel *c*), process error CV in normal years (τ_{norm}, panel *d*), and process error CV in catastrophic years (τ_{cat}, panel *e*). The solid line represents the distribution of the median bias (posterior median minus true value), the short broken line represents the bias of the posterior mode, and the longer broken line represents the bias of the posterior mean.



Akaike's information criterion (AIC, Akaike 1973), with the major difference being that DIC integrates over the parameter space (instead of maximizing). The DIC computation can be expressed as DIC = $D(\hat{\theta}) + 2p_D$, where $D(\hat{\theta})$ represents the deviance evaluated at the posterior mean, and p_D represents the effective number of model parameters. The term p_D can also be expressed as the deviance function evaluated at the expected posterior values of the parameters subtracted from the mean deviance across all parameter vectors, $p_D = \overline{D(\theta)} - D(\hat{\theta})$.

Results

Of all parameters in the catastrophic state–space model, we found the population parameters to be estimated robustly, but parameter estimates were sensitive to the initial state in the simulation model. This result is expected, as the same result holds under simpler error models (e.g., models that consider observation or process error only). For the case where the simulation model is logistic, a time series with low relative population size provides information about the initial state and population growth rate, but little or no information about carrying capacity. The opposite is also true — when the initial population state is close to carrying capacity, estimates of the initial state and growth rate are biased; however, estimates of K are relatively unbiased. Bias in estimating ϕ was slightly negative, indicating that the parameter was underestimated more frequently than it was overestimated (Fig. 3). The opposite result was true for estimating θ — the bias tended to be positive, indicating that the parameter tended to be overestimated. Bias in the observation error is centered close to zero, indicating that this parameter and process error in normal years is centered close to zero; however, bias in estimating the process error in catastrophic years appears to be positive (Fig. 3).

One necessary feature of the catastrophic state–space model is that it must be able to distinguish catastrophes from anomalies caused by negative process errors. Because we know each catastrophe in the simulated time series, it is possible to quantify both the ability of the model to detect catastrophes and the error rate with which the model classifies normal variation as catastrophes. Statistical power has been applied to many situations in biology (Taylor and Gerrodette 1993), and for the purpose of this analysis we define power as the ability to detect a catastrophe when it actually occurs. The parameters that affect the power to detect catastrophes

Fig. 4. The relationship between the power to detect a trend (solid circles) and percent reduction in population size (i.e., the magnitude of a catastrophe). Also included is the probability of falsely detecting a catastrophe (open circles). Power is calculated across all levels of observation error and process errors.

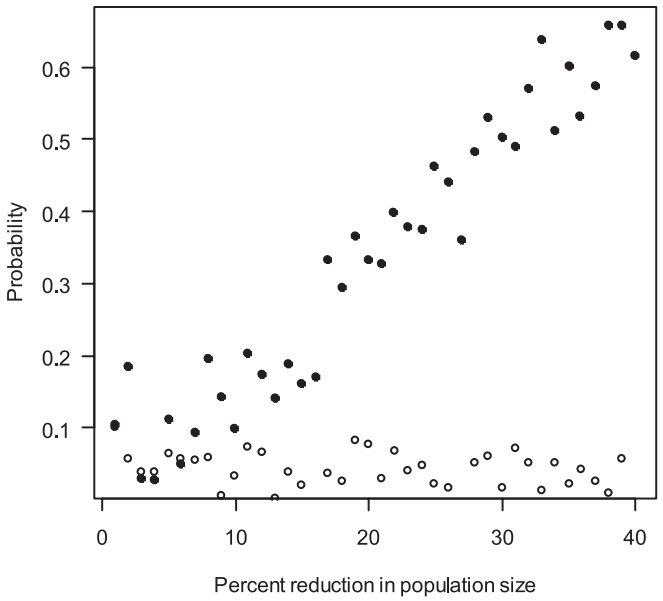
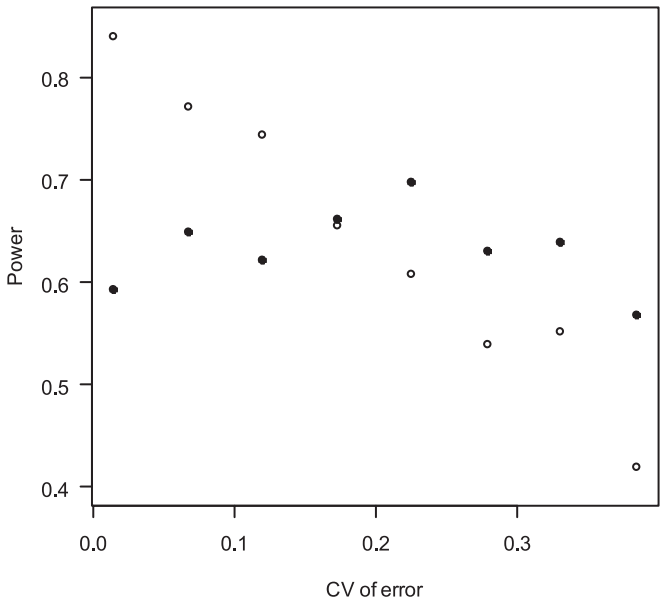
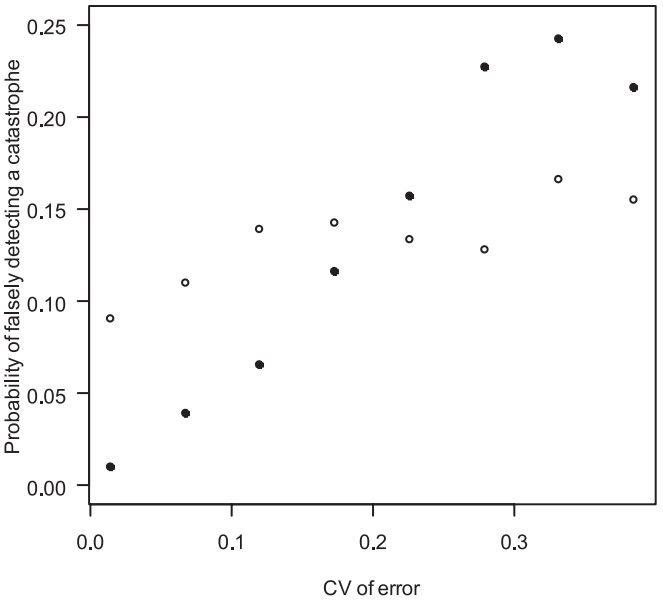


Fig. 5. The relationship between the coefficient of variation (CV) of the errors (when the errors are lognormally distributed) and the power to detect a catastrophe for 1000 data sets. The magnitude of a catastrophe is held constant at 0.6 across all data sets. Open circles represent varying levels of observation error (across all levels of process error), and solid circles represent varying levels of process error (across all levels of observation error).



in our state–space model are θ and the magnitude of the observation and process errors. As catastrophes become more severe, the probability of detecting them increases (Fig. 4). Observation error appears to have a larger effect on power than process error (Fig. 5); however, both influence the rate

Fig. 6. The relationship between the coefficient of variation (CV) of the errors (when the errors are lognormally distributed) and the probability of falsely detecting a catastrophe for 1000 data sets. The magnitude of a catastrophe is held constant at 0.6 across all data sets. Open circles represent varying levels of observation error (across all levels of process error), and solid circles represent varying levels of process error (across all levels of observation error).

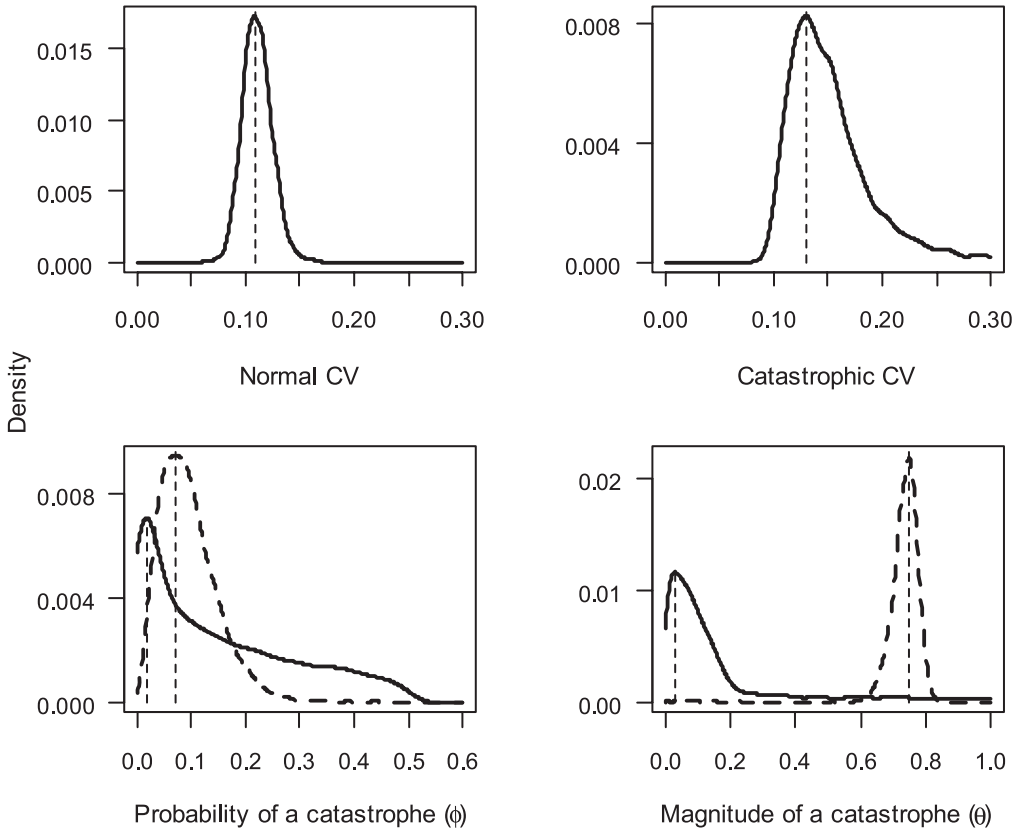


of falsely detecting a catastrophe (Fig. 6). At intermediate values of the observation error (CV > 0.35), the probability of detecting a catastrophe that reduces the population by 60% becomes less than would be expected by chance alone. As expected, the power to detect catastrophes is highest under low levels of observation and process error and when catastrophes are large in magnitude.

Of the remaining 48 candidate models, the model with the lowest DIC score appeared to be a model that included catastrophes, allowed both Pribilof populations to follow three-parameter exponential curves, and allowed both Bogoslof Island and San Miguel Island to have unique island-specific observation errors (Table 1). This model was strongly favored over the alternative models, receiving 97.5% of the normalized DIC weight. One of the most important results from the model selection procedure was that the state–space model that incorporated catastrophes was always favored over the simpler model with a single process error distribution. This result is important because it indicates that catastrophes should be included in our analysis, regardless of which population models are chosen. A second important result is that for nearly all scenarios, there appears to be no evidence for the inclusion of an additional observation error parameter, suggesting that the estimated CVs for the Pribilof surveys are appropriate.

For the model with the lowest DIC score, the posterior mode of ϕ was found to be 0.019 for Bering Sea populations (2.5% = 0.001, 97.5% = 0.462) and 0.072 for the San Miguel population (2.5% = 0.021, 97.5% = 0.228; Fig. 7). The posterior mode of θ for the San Miguel population was estimated to be 0.751 (2.5% = 0.591, 97.5% = 0.801). The

Fig. 7. Posterior probability distributions for the process error parameters and catastrophe parameters for the model with the lowest deviance information criterion (DIC) score. Posteriors of the process error coefficients of variation (CVs) indicate that the variation in catastrophic years (catastrophic CV) is more skewed than the normal variability (normal CV). Because we allow the probability of a catastrophe and magnitude of a catastrophe to vary by ecosystem, two posterior distributions are presented (solid line, Bering Sea; broken line, San Miguel Island). The broken vertical line represents the posterior mode of each distribution.



value of θ in the Bering Sea was estimated to be close to zero, indicating that while the Bering Sea populations might experience chance reductions in the number of pups, the populations do not experience true catastrophes. There are several important differences between the estimated process error CVs in normal and catastrophic years (Fig. 8). First, the posterior mode of the CV in catastrophic years (0.129) is larger than the mode of the CV in normal years (0.109), implying that there is slightly more process uncertainty in catastrophic years. The variance of the estimated process error CV in catastrophic years is greater than that in normal years. For example, approximately 17% of the estimated process error CV in catastrophic years is greater than the upper bound of the process error CV in normal years (~0.16). This result is somewhat intuitive. Because the number of catastrophic years is relatively small compared with the number of normal years, the process error variation in catastrophic years is expected to be estimated with less precision.

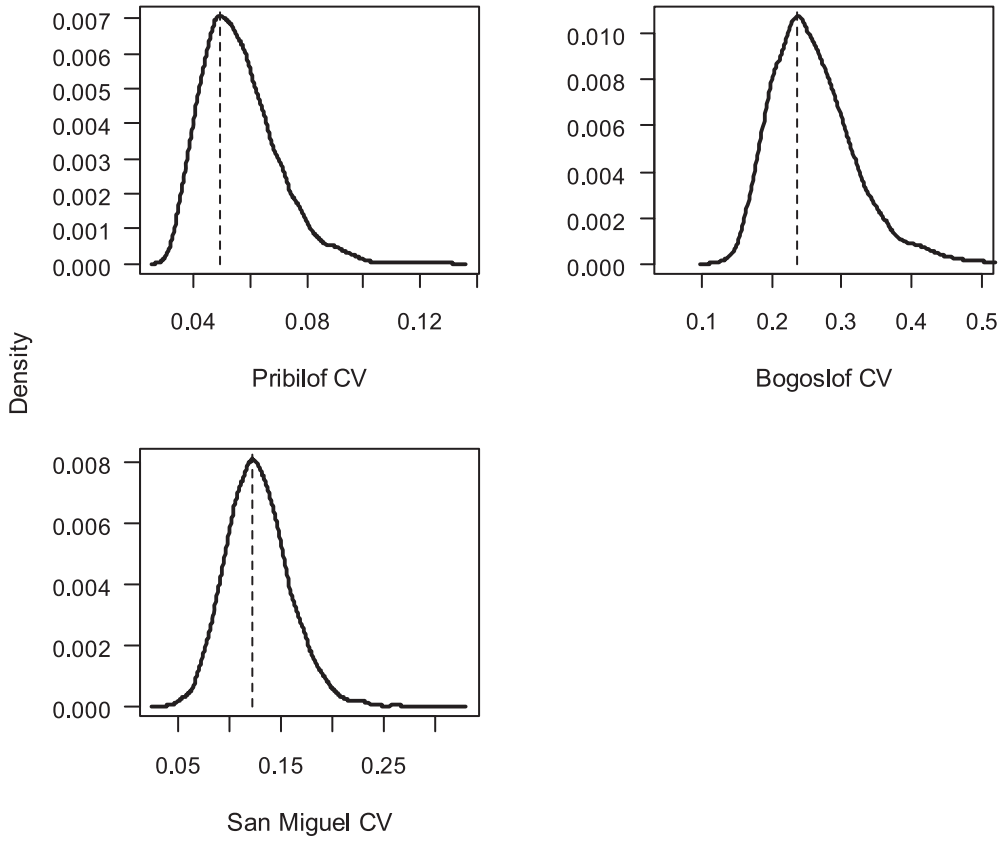
Discussion

The catastrophic state–space model could be applied to a wide range of species, including both terrestrial and marine organisms. One potential source for combining continuous time series in a meta-analysis is the Global Population Dynamics Database (NERC 1999), which has collected over 5000 time series from the ecological and fisheries literature. Applications of the catastrophic state–space model should not be limited to parameter estimation or hypothesis testing — incorporating a more realistic form of catastrophic events will have dramatic implications for population projections and management decisions. The most useful case of the model is the scenario where lengthy time series from similar populations (or even metapopulations) are combined in a hierarchical model.

The data sets in our simulations included a range of realistic observation and process errors, with CVs ranging from 0.01 to 0.5. Across all simulations, the population parameters (e.g., logistic parameters), observation error, and process error in normal years were estimated robustly. The ability of the model to estimate the catastrophic parameters (ϕ, θ, τ_{cat}) was highly dependent on the frequency and magnitude of catastrophes involved. The relationship between the magnitude of the catastrophic parameters and the bias in estimating them is somewhat intuitive — as catastrophes become more frequent and stronger, the model is better able to discriminate them from normal years with extreme process error, thereby decreasing the bias.

One potential criticism of this model is that for some data sets (particularly those in which no real catastrophes exist), the magnitude of a catastrophe may be estimated as being approximately 1.0 and the probability of a catastrophe may

Fig. 8. Posterior probability distributions of estimated observation error coefficient of variation (CV) parameters for the model with the lowest deviance information criterion (DIC) score, which allows Bogoslof Island and San Miguel Island to each have unique observation error CVs. The Pribilof Islands have the lowest observation error, while Bogoslof Island appears to have the largest observation error. The broken vertical line represents the posterior mode of each distribution.



be non-zero. Our mixture model could still be applied to such data, but results would have to be interpreted differently. Instead of interpreting the process error mixture as the error for normal and catastrophic years, the mixture might be interpreted as a measure of data quality. In this setting, some small fraction of the data would be treated by the model as outliers, and those data points would be assigned to the component of the mixture associated with higher, larger variance (Chen and Fournier 1999). For each data set, researchers should consider the relationship of catastrophic parameters, magnitudes of observation, and process errors on power. A second potential criticism of the model is that the exponential or logistic population dynamics models in this analysis are too simple to represent reality. For many species, data that would allow the use of age-structured models have only been collected recently (Holmes and Fagan 2002). A potential future application of the catastrophic state–space model would be to apply it to age-structured data, with catastrophes only affecting one particular age class (e.g., juvenile survival).

In this paper, we demonstrate the application of a novel technique to estimate the probability of catastrophes. Estimation of catastrophes is important in both basic analyses of population dynamics and in applied analyses of extinction risk. Our analyses demonstrate the potential application of our approach in a meta-analysis, and similar studies should be conducted for taxa for which data are available for multiple populations. Our catastrophe model is able to incorporate catastrophes that affect total population size or specific components of the population (as in the fur seal example presented here). Although we cannot identify the cause for catastrophic events in the three time series of northern fur seal pups, the DIC weights provide strong support for the state–space model that includes catastrophes. It is not surprising that in the best model, there are large differences between the estimated CVs for each island. Because the Pribilof surveys involve multiple counts, they are more precise than direct counts from other islands (essentially a minimum estimate). The modes of the posterior distributions indicated that catastrophes occur on San Miguel with a 7.2% probability, and the average magnitude of those catastrophes is a 75% reduction in the number of pups. This estimated probability of a catastrophe is greater than the estimates previously suggested by Gerber and Hilborn (2001) for all otariid species.

In applying model selection tools such as DIC, it is important to stress the fact that even though the catastrophic model received the majority of the model weights, none of the models in this analysis should be considered the true model. As our simulations illustrated, the power to detect catastrophes is strongly affected by observation error and both the frequency and magnitude of the catastrophic events. Process error also appears to have an effect on power, but increasing the magnitude of the process error also increases

the chance of falsely detecting catastrophes. Under the range of errors and catastrophic parameters that affect *C. ursinus* populations and surveys, the power to detect catastrophes is greater than chance alone, but probably in the range of 80%–90%. These results underscore the importance of quantitative analyses to determine data needed to detect catastrophes. Data requirements are expected to differ both between species and between population dynamics models — if the model were to be extended to include catastrophes in age-structured population dynamics, a single, annual abundance estimate for each population would likely not be sufficient.

Acknowledgements

Funding for this project was provided by the NMFS and Washington State Sea Grant. All simulations were made possible by computer labs in the School of Aquatic and Fishery Sciences at the University of Washington. Kristin Marshall provided editorial comments that greatly improved the quality of this manuscript.

References

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. *In* 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, 2–8 September 1971. *Edited by* B.N. Petrov and F. Csaksi. Akademiai Kiado, Budapest, Hungary. pp. 267–281.

Angliss, R.P., and Lodge, K.L. 2002. Alaska marine mammal stock assessments, 2002. US Department of Commerce, Seattle, Wash. NOAA Tech. Memo. NMFS-AFSC-133.

Calder, C., Lavine, M., Muller, P., and Clark, J.S. 2003. Incorporating multiple sources of stochasticity into population dynamics models. Ecology, **84**: 1395–1402.

Carlin, B.P., and Louis, T.A. 2000. Bayes and empirical Bayes methods for data analysis. Chapman and Hall, New York.

Chen, Y., and Fournier, D. 1999. Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries. Can. J. Fish. Aquat. Sci. **56**: 1525–1533.

Clark, J.S., and Bjørnstad, O.N. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. Ecology, **85**: 3140–3150.

de Valpine, P. 2002. Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-gaussian state–space models. Bull. Mar. Sci. **70**: 455–471.

de Valpine, P., and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. Ecol. Monogr. **72**: 57–76.

Erb, J.D., and Boyce, M. 1999. Distribution of population declines in large mammals. Conserv. Biol. **13**: 199–201.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. 1995. Bayesian data analysis. Chapman and Hall, New York.

Gerber, L.R., and Hilborn, R. 2001. Catastrophic events and recovery from low densities in populations of otariids: implications for risk of extinction. Mammal Rev. **31**: 131–150.

Gilks, W.R., Richardson, S., and Spiegelhalter, D. 1996. Markov Chain Monte Carlo in practice. Chapman and Hall, New York.

Gilpin, M.E., and Ayala, F.J. 1973. Global models of growth and competition. PNAS, **70**: 3950–3593.

Hilborn, R., and Mangel, M. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, N.J.

Hobbs, N.T., and Hilborn, R. 2006. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. Ecol. Appl. **16**: 5–19.

Holmes, E.E., and Fagan, W.E. 2002. Validating population viability analysis for corrupted data sets. Ecology, **83**: 2379–2386.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. Journal of Basic Engineering (Transactions of the ASME), **82**: 35–45.

Lande, R., Steinar, E., and Saether, B.E. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, New York.

Loughlin, T.R., Antonelis, G.A., Baker, J.D., York, A.E., Fowler, C.W., DeLong, R.L., and Braham, H.W. 1994. Status of the northern fur seal population in the United States during 1992. *In* Fur seal investigations, 1992. *Edited by* E.H. Sinclair. US Department of Commerce, Seattle, Wash. NOAA Tech. Memo. NMFS-AFSC-45.

Mangel, M., and Tier, C. 1994. Four facts every conservation biologist should know about persistence. Ecology, **75**: 607–614.

McAllister, M.K., and Kirkwood, G.P. 1998. Bayesian stock assessment: a review and example application using the logistic model. ICES J. Mar. Sci. **55**: 1031–1060.

Meyer, R., and Millar, R.B. 1999*a*. BUGS in Bayesian stock assessments. Can. J. Fish. Aquat. Sci. **56**: 1078–1087.

Meyer, R., and Millar, R.B. 1999*b*. Bayesian stock assessment using a state–space implementation of the delay difference model. Can. J. Fish. Aquat. Sci. **56**: 37–52.

NERC. 1999. The global population dynamics database [online]. NERC Center for Population Biology, Imperial College London, Ascot, Berkshire, UK. Available from www.sw.ic.ac.uk/cpb/gpdd.html [accessed 1 June 2006; updated September 2005].

Punt, A.E. 2003. Extending production models to deal with process error in the population dynamics. Can. J. Fish. Aquat. Sci. **60**: 1217–1228.

Punt, A.E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. **7**(1): 35–63.

Rivot, E., Prevost, E., Parent, E., and Bagliniere, J.L. 2004. A Bayesian state–space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. Ecol. Model. **179**: 463–485.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. Griffin, London.

Spiegelhalter, D., Thomas, A., and Best, N. 2003. WinBUGS version 1.4 user manual [online]. Available from www.mrc-bsu.cam.ac.uk/bugs [accessed 01 June 2006].

Staples, D.F., Taper, M.L., and Dennis, B. 2004. Estimating population trend and process variation for PVA in the presence of sampling error. Ecology, **85**: 923–929.

Taylor, B.L., and Gerrodette, T. 1993. The uses of statistical power in conservation biology: the Vaquita and Northern Spotted Owl. Conserv. Biol. **7**: 489–500.

Thomas, L., Buckland, S.T., Newman, K.B., and Harwood, J. 2005. A unified framework for modelling wildlife population dynamics. Aust. N.Z. J. Stat. **47**: 19–34.

Towell, R.G., Ream, R.R., and York, A.E. 2006. Decline in northern fur seal (*Callorhinus ursinus*) pup production on the Pribilof Islands. Mar. Mamm. Sci. **22**: 486–491.

Walters, C.J., and Ludwig, D. 1981. Effects of measurement errors on the assessment of stock–recruitment relationship. Can. J. Fish. Aquat. Sci. **38**: 704–710.

Walters, C.J., and Ludwig, D. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. **51**: 713–722.

York, A.E. 1991. Sea surface temperatures and their relationship to the survival of juvenile male northern fur seals from the Pribilof Islands. *In* Pinnipeds and El Niño. *Edited by* F. Trillmich and K.A. Ono. Springer-Verlag, Berlin, Germany. pp. 94–106.

York, A.E., Towell, R.G., Ream, R.R., Baker, J.D., and Robson, B.W. 2000. Population assessment, Pribilof Islands, Alaska. *In* Fur seal investigations, 1998. *Edited by* B.W. Robson. US Department of Commerce, Seattle, Wash. NOAA Tech. Memo. NMFS-AFSC-113. pp. 7–26.

York, A.E., Towell, R.G., Ream, R.R., and Fowler, C.W. 2002. Population assessment, Pribilof Islands, Alaska. *In* Fur seal investigations, 2000–2001. *Edited by* B.W. Robson. US Department of Commerce, Seattle, Wash. NOAA Tech. Memo. NMFS-AFSC-134. pp. 7–32.

Young, T.P. 1994. Natural die-offs of large mammals: implications for conservation. Conserv. Biol. **8**: 410–418.