

ESTIMATION OF POPULATION GROWTH AND EXTINCTION PARAMETERS FROM NOISY DATA

STEVEN T. LINDLEY¹

National Oceanic and Atmospheric Administration, Santa Cruz Laboratory, 110 Shaffer Road,
Santa Cruz, California 95060 USA

Abstract. The random-walk-with-drift model of population dynamics is an important tool in conservation biology, partly because its parameters are easily estimated from periodic observations of population size. Estimating the model with noisy data is problematic, however, because the commonly used estimators of process variation are biased if population abundance measurements are imprecise, and a recently developed method that attempts to remove this bias is not robust. In this paper, I show how the random-walk-with-drift model can be applied to noisy time series of population estimates by converting the random-walk-with-drift model to state-space form and applying the Kalman filter to yield the likelihood of the data. The likelihood function allows the variances of the process error and measurement error and the growth rate of the population to be estimated in a way that is robust and fully supported by statistical theory. Comparative analysis using simulated data indicates that the Kalman-filter method reduces the bias in estimates of process variance without yielding negative variance estimates. I apply the method to California sea otter and Yellowstone grizzly bear data to illustrate how the method (and simple extensions) can be used to assess the status of real populations. California sea otters appear to have little risk of extinction over the next 100 years although the population may not be secure over the long term if a recent apparent cessation of population growth persists. The grizzly bear population appears to have responded positively to the 1988 Yellowstone fires, and if the population continues to grow at the average rate observed over the study period, it is extremely unlikely to go extinct.

Key words: diffusion approximation; extinction model; grizzly bear, Yellowstone population; Kalman filter; measurement errors bias parameter estimates; parameter estimation; random walk; sea otter, California population; state-space model.

INTRODUCTION

Ecologists are often faced with the problem of predicting the future of populations from periodic counts of abundance. Simple population models, like the random-walk-with-drift (Dennis et al. 1991) or ceiling (Foley 1994) models, can provide limited but important inferences from count data such as estimates of population growth rate and the probability of extinction. The utility of these models is limited in part by imprecise estimates of population size, which reduce the accuracy and precision of parameter estimates (Ludwig 1999).

When the parameters of a population model must be estimated from time-series data, it is usual to assume that stochasticity arises from measurement error or process error, but not both (Hilborn and Mangel 1997). A process-error model of population dynamics assumes that variation comes from the population growth process (e.g., random birth–death processes, environmentally driven variation in survival) and that observations are made without error. A measurement-error model assumes that changes in population size are determin-

istic and that measurements of population size are made with error. Measurement errors can be viewed as noise that obscures the signal of the true population dynamics. Some models, such as the random-walk-with-drift model of population growth (e.g., Dennis et al. 1991), are by definition process-error models; however, these models must be estimated from noisy data. If measurement error is significant but ignored, predictions about the future behavior of a population can be grossly in error because process variation is overestimated (Dennis et al. 1991, Meir and Fagan 2000, Holmes 2001).

Holmes (2001) proposes a method for estimating the process-error variance (σ_p^2) and rate of population growth (μ) for the random-walk-with-drift model from noisy estimates of population size. The method exploits the fact that, for a population following a random walk, $\text{Var}(X_t - X_{t-\tau})$ increases linearly with τ , where X_t is the log-transformed population size estimate and τ is the time gap between observations. Measurement error, on the other hand, does not cause $\text{Var}(X_t - X_{t-\tau})$ to increase as τ increases. Holmes's (2001) "slope" estimator of σ_p^2 is simply the slope of $\text{Var}(X_t - X_{t-\tau})$ vs. τ , which can be estimated by linear regression. While this method is a significant advance over the estimators of Dennis et al. (1991), it has a few problems. It is ad hoc,

Manuscript received 3 June 2002; revised 16 October 2002; accepted 25 October 2002. Corresponding Editor: B. A. Maurer.

¹ E-mail: Steve.Lindley@noaa.gov

and the properties of the estimator are not well understood. One must choose some maximum value of τ , and the results can be sensitive to that choice. The method is not based on likelihood, which complicates estimation of parameter uncertainty and precludes taking advantage of Bayesian analyses. Additionally, the slope estimator is not robust: it can generate negative variance estimates when applied to real data or to simulated data generated by a random-walk process.

State-space modeling, a method of time-series analysis, offers an alternative to the Holmes method. State-space population models treat the population process and the observation process separately and provide a way to estimate the variances of both the observation error and the process error. In this paper, I recast the random-walk-with-drift model as a state-space model and show how the Kalman filter can be used to calculate the likelihood of the data, providing a robust way to estimate σ_p^2 from noisy data. The performance of the Kalman-filter-based method is compared to that of the slope method and the maximum-likelihood method of Dennis et al. (1991) by application to simulated data. To further illustrate the state-space method, I apply it to population data for the threatened California sea otter (*Enhydra lutris* L.) and the Yellowstone grizzly bear (*Ursus arctos* L.).

METHODS

In this section I briefly review the random-walk-with-drift model and show how it can be placed into state-space form. I then present the likelihood function and describe two ways it can be used to estimate the unknown parameters. Finally, I discuss some diagnostics useful for testing whether the assumptions of the random-walk-with-drift model are met by the data, and a way to compare formally the simple random-walk-with-drift model to extensions of the random-walk-with-drift model.

The random-walk-with-drift model

In the random-walk-with-drift model, a population grows or declines exponentially with time according to a randomly varying growth rate:

$$N_t = N_{t-1} \exp(\mu + \varepsilon_t) \quad (1)$$

where N_t is the population size at time t , μ is the mean rate of population growth, and ε_t is a random deviation in growth rate (a "process error") with mean 0 and variance σ_p^2 . When ε_t is a normal random variable, N_t is a lognormal random variable. If we define $X_t = \log(N_t)$, the process can be represented as a linear function:

$$X_t = X_{t-1} + \mu + \varepsilon_t. \quad (2)$$

The behavior of the discrete-time random-walk-with-drift process can be approximated by a continuous-time diffusion process. Based on the diffusion approximation, analytical expressions for various extinction

parameters (e.g., probability of crossing certain thresholds within certain times and the mean time to cross these thresholds) are available that are functions of X_t , μ , and σ_p^2 ; see Dennis et al. (1991). The random-walk-with-drift model is attractive because it can be applied in situations where only population abundance data are available and because its theoretical properties are well understood.

The random-walk-with-drift model of population growth has several implicit assumptions. One assumption is that the population dynamics can be represented adequately by a first-order autoregressive process, i.e., the age structure of the population is unimportant. Other assumptions are that the effects of density dependence can be ignored and that serially uncorrelated Gaussian noise adequately represents the process variation. Finally, it is assumed that X_t is measured without error. Dennis et al. (1991) discuss why the first three assumptions are often reasonable. The last assumption can be relaxed by relying on a probability structure for the observations associated with the estimation of X_t . State-space theory and the Kalman filter provide a statistically rigorous framework for doing exactly that.

State-space models

In this subsection, I show how the random-walk-with-drift model can be recast as a state-space model. The term "state" refers to the true state of nature, such as the real size of a population. Typically, we cannot know the true state of nature, but must make inferences about it from observations that may be somewhat uncertain. State-space models describe two time series evolving in parallel. The true state of nature produces the first time series. The state transition equations describe how the true state of nature evolves stochastically with time as a first-order Markov process. Observations of the state form the second time series. The observation equations describe how observations of the state are produced. These observations might be imprecise and inaccurate.

Eq. 2 describes how the true population size changes with time, and is therefore a state equation. The Kalman filter operates on matrices, however, which requires that Eq. 2 be put into matrix notation. This is done by constructing the vector $\mathbf{x}_t = [X_t \ \mu]^T$. The vector \mathbf{x}_t is known as the state vector. The state evolves stochastically over time according to

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \boldsymbol{\eta}_t. \quad (3)$$

The transition matrix \mathbf{A} is $[1 \ 1; 0 \ 1]$ and $\boldsymbol{\eta}_t$ is a vector with mean = 0 and variance-covariance $\mathbf{Q} = [\sigma_p^2 \ 0; 0 \ 0]$ that represents process-error variation. X_t is a normal random variable with mean $\mu + X_{t-1}$ and variance σ_p^2 . Note that, as in Eq. 2, μ is a fixed parameter.

Typically, we cannot know X_t directly, but we can estimate it with some sort of population survey. A univariate observation at time t (y_t) is related to \mathbf{x}_t by the measurement equation:

$$y_t = \mathbf{c}\mathbf{x}_t + \varepsilon_t^* \quad (4)$$

where \mathbf{c} is a row vector $[1 \ 0]$ relating the observation to the state, and ε_t^* is a measurement error with mean $= 0$ and variance $= r = \sigma_m^2$. The observation y_t is therefore a random variable with mean X_t and variance σ_m^2 . Note that $\mathbf{c} = [1 \ 0]$ implies that μ is not observed.

To complete the definition of the probabilistic model, we must specify an initial distribution for the state, defined by a mean $\mathbf{x}_0 = [X_0 \ \mu]^\top$ and variance–covariance $\mathbf{P}_0 = [\text{Var}(X_0) \ 0; 0 \ 0]$. Given the initial state distribution and \mathbf{A} , \mathbf{c} , \mathbf{Q} , and r , the Kalman filter provides optimal estimates of \mathbf{x}_t , as well as the likelihood of the data (Harvey 1989).

The Kalman filter

The Kalman filter is a recursive algorithm that estimates the mean and covariance of the state at time t from the observation at time t and the mean and covariance of the state at time $t - 1$. In this subsection, I provide the form of the Kalman filter as applicable to the random-walk-with-drift model; Harvey (1989) provides an accessible explanation of the general form and a complete derivation of the Kalman filter.

Let $\hat{\mathbf{x}}_t$ denote the estimate of \mathbf{x}_t based on observations up to and including y_t . Let \mathbf{P}_t denote the covariance matrix of the estimation error:

$$\mathbf{P}_t = E[(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)^\top]. \quad (5)$$

Application of the Kalman filter consists of predicting the new state value before a new observation is obtained, and then updating the prediction when the new observation is available. The state is predicted by propagating the previous state estimate through the transition equation:

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{A}\hat{\mathbf{x}}_{t-1}. \quad (6)$$

The covariance of the prediction error is the sum of the covariance of the estimation error associated with $\hat{\mathbf{x}}_{t-1}$ and the covariance of the process error:

$$\mathbf{P}_{t|t-1} = \mathbf{P}_{t-1} + \mathbf{Q}. \quad (7)$$

Eqs. 6 and 7 are called the “prediction equations.” Note that in general, Eq. 7 would be $\mathbf{P}_{t|t-1} = \mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^\top + \mathbf{Q}$, but for the random-walk-with-drift model, $\mathbf{A}\mathbf{P}_{t-1}\mathbf{A}^\top = \mathbf{P}_{t-1}$.

When a new observation y_t is obtained, the prediction error v_t is evaluated as

$$v_t = y_t - \mathbf{c}\hat{\mathbf{x}}_{t|t-1} \quad (8)$$

and $\hat{\mathbf{x}}_t$ and \mathbf{P}_t are updated. The state estimate is moved towards the observation according to the observation accuracy:

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{c}^\top f_t^{-1}v_t \quad (9)$$

where

$$f_t = \mathbf{c}\mathbf{P}_{t|t-1}\mathbf{c}^\top + r. \quad (10)$$

Similarly, the uncertainty about $\hat{\mathbf{x}}_t$ is reduced by information provided by the new observation:

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{c}^\top f_t^{-1}\mathbf{c}\mathbf{P}_{t|t-1}. \quad (11)$$

Eqs. 9 and 11 are termed the “updating equations.”

The recursion is initiated with the initial distribution of the state defined by a mean \mathbf{x}_0 and covariance \mathbf{P}_0 . For population data, usually nothing is known about \mathbf{x}_0 . In such situations, the filter can be initiated with a diffuse prior distribution.

Parameter estimation

The key to estimating the unknown parameters μ , σ_p^2 , and σ_m^2 is the likelihood (L) of the data conditional on these parameters, which is a function of f_t and v_t (Harvey 1989):

$$\log L = -\frac{T}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^T \log f_t - \frac{1}{2}\sum_{t=1}^T \frac{v_t^2}{f_t} \quad (12)$$

where T is the length of the observation series. This likelihood could be used directly for parameter estimation, although the dimension of the problem can be reduced by reparameterizing the model in terms of a signal-to-noise ratio (Harvey 1989). This is accomplished by defining \mathbf{Q} as $[s \ 0; 0 \ 0]$ and letting $\text{Var}(\boldsymbol{\eta}_t) = \sigma_m^2\mathbf{Q}$. Likewise, set $r = 1$ and let $\text{Var}(\varepsilon_t^*) = \sigma_m^2r$. The parameter s is interpreted as the signal-to-noise ratio. The unknown parameters are then μ and s , denoted as the vector $\boldsymbol{\psi}$; σ_p^2 and σ_m^2 are functions of $\boldsymbol{\psi}$ and the data. The concentrated likelihood of y_t for $t = 1, \dots, T$, conditional on $\boldsymbol{\psi}$, is given by

$$\log L(\boldsymbol{\psi}) = -\frac{T}{2}\log(2\pi + 1) - \frac{1}{2}\sum_{t=1}^T \log f_t - \frac{T}{2}\log \sigma_m^2(\boldsymbol{\psi}) \quad (13)$$

where

$$\sigma_m^2(\boldsymbol{\psi}) = \frac{\sigma_p^2(\boldsymbol{\psi})}{s} = \frac{1}{T}\sum_{t=1}^T \frac{v_t^2}{f_t}. \quad (14)$$

To initiate the Kalman filter for the purpose of calculating the likelihood, the first observation is used to form a prior distribution for \mathbf{x}_0 . This is accomplished by setting $\mathbf{x}_1 = [y_1 \ \mu]^\top$ and $\mathbf{P}_1 = [\sigma_p^2 + \sigma_m^2 \ 0; 0 \ 0]$ and running the sums in Eq. 13 from $t = 2, \dots, T$ (Durbin and Koopman 2001). Numerical stability can be improved by log-transforming the signal-to-noise ratio s in order to avoid bounding s to nonnegative values.

Point estimates of $\boldsymbol{\psi}$ can be obtained by numerically maximizing Eq. 13 with respect to $\boldsymbol{\psi}$, and subsequently these point estimates could be used to calculate estimates of various extinction-risk metrics. Point estimates can be misleading, however, because without large amounts of data, σ_p^2 estimates are often not precise, regardless of the estimation algorithm used. Moderate uncertainty in σ_p^2 can translate into large uncertainty in extinction parameters (Ludwig 1996), making it desirable to account for parameter uncertainty when assessing real populations of conservation concern.

This entails estimating the joint probability density of Ψ conditional on the observations ($p(\Psi|y_T)$), and integrating risk metric equations over this joint probability distribution. Markov chain Monte Carlo methods (MCMC), such as the Metropolis-Hastings algorithm (Metropolis et al. 1953, Hastings 1970), offer a convenient way to perform this integration. The Metropolis-Hastings algorithm generates a sample from $p(\Psi|y_T)$ that can be used to estimate the moments of $p(\Psi|y_T)$ as well as the moments of functions of Ψ (Gills et al. 1996). While this approach is computationally intensive, it is easily handled by modern personal computers and is not difficult to implement.

Diagnostics for state-space models

The assumptions of the random-walk-with-drift model must be met for the model to provide reasonable predictions of the distribution of future population size. For this reason Dennis et al. (1991) suggest that model diagnostics be used to test assumptions when interpreting the results of analyses based on the random-walk-with-drift model. The algorithm proposed by Dennis et al. (1991) is based on linear regression, making available the full suite of well-developed regression model diagnostics, many of which are available in common statistics software packages. For state-space models, diagnostics are based on the prediction errors, which by definition should be $\text{Normal}(0, \sigma^2)$ independent random variables (Harvey 1989). Brown et al. (1975) describe graphical procedures for evaluating the cumulative prediction errors and cumulative squared prediction errors that can detect temporal instability of the model parameters. The prediction errors can also be checked for nonnormality, heteroscedasticity, and serial autocorrelation with the usual tests. Together, these procedures test for violation of the assumptions underlying the random-walk-with-drift model.

Another approach to testing some of the assumptions of the random-walk-with-drift model is to compare its fit to that of models that relax these assumptions. For example, one might want to consider models where μ is a function of time or some other covariate. Alternative models can be compared rigorously using Akaike's Information Criterion (AIC), which is a simple function of the maximized log likelihood and the number of estimated parameters (K):

$$\text{AIC} = -2 \log(L(\Psi|y)) + 2K. \quad (15)$$

An AIC for an alternative model that is substantially lower than that of the simple random-walk-with-drift model (by, say, more than 2) indicates substantial support for the alternative model (Burnham and Anderson 1998). Because the Kalman filter supplies $\log(L(\Psi|y))$ and the state-space form can accommodate many models, the state-space framework facilitates such model comparisons.

Simulation study of estimator performance

While the application of the Kalman filter to models cast in state-space form offers a theoretically sound basis for estimating σ_p^2 and μ from noisy observations, intuition suggests that short data series might not contain enough information to allow separation of σ_p^2 and σ_m^2 . I conducted simulations to test the performance of the Kalman filter under different conditions of σ_p^2 , σ_m^2 , and T , and to compare this performance to that of the Dennis et al. (1991) and Holmes (2001) estimators. In all cases, simulated data were created using Eqs. 3 and 4. Several values of σ_p^2 , σ_m^2 , and time-series length were explored, and 1000 time series were simulated for each combination. The Kalman-filter point estimates were produced by minimizing the negative log likelihood using the Nelder-Mead simplex algorithm for unconstrained minimization of a function with several unknowns (Nelder and Mead 1965).

Examples

The method outlined above was applied to California sea otter and Yellowstone grizzly bear population data. In both cases, the population might have experienced a change in μ during the period of observation, so a step-change model that admits such a change is compared to the basic constant- μ version of the random-walk-with-drift model. In the step-change model, μ is assumed to be constant within time periods separated by a change point t_c , but may vary between time periods:

$$\mu_t = \begin{cases} \mu & \text{for } t < t_c \\ \mu + \delta & \text{for } t \geq t_c. \end{cases} \quad (16)$$

For both models, I used the Metropolis algorithm to generate a sample from the posterior distribution of $p(\Psi|y_T)$ (note that Ψ now contains δ for the step-change model). I then used the sample to estimate central probability intervals and modes of $p(\Psi|y_T)$ and functions of $p(\Psi|y_T)$, such as σ_p^2 , σ_m^2 , and various extinction parameters. Weak prior distributions ($\text{Normal}(0, 5)$) were placed on the log-transformed s to create a proper posterior. This can be important for models containing the parameter s because the log likelihood of some data is flat as $s \rightarrow \infty$ and as $s \rightarrow 0$.

Sea otters.—Sea otters (*Enhydra lutris*) were once abundant in near-shore waters from Baja California to northern Japan, but were hunted to the brink of extinction by the late 1800s. Hunting was banned in 1911, and sea otter populations subsequently increased. The California sea otter population comprises a distinct subspecies (*E. lutris nereis*) and has grown at a slower rate than other populations (Estes 1990). The population now numbers somewhat more than 2000 individuals. The California sea otter was listed as a threatened species under the U.S. Endangered Species Act (ESA) in 1973. It has been hypothesized that the California sea otter population ceased growing about 1995, per-

haps due to increased mortality as indicated by observations of beach-cast carcasses (USGS, *public communication*).² The constant- μ and step-change model, with $t_c = 1995$, were fit to the data.

The entire California sea otter population has been counted twice annually since 1983. Observations are made from shore and fixed-wing aircraft by two-person teams. Surveys extend from Santa Barbara to Half Moon Bay. The data used in this analysis are total springtime otter abundance, obtained from the USGS Western Ecological Research Center (Sacramento, California, USA). Estes and Jameson (1988) estimated that these surveys have a sighting probability of 0.945, which implies that, given binomial sampling variation, estimates of a sea otter population of 2000 individuals have a coefficient of variation of about 0.005. This is a lower bound for measurement error because other sources of variation are presumably present in the survey estimates.

Grizzly bears.—Grizzly bears (*Ursus arctos horribilis*) once ranged over a third of the continental United States, but have been reduced to a few isolated populations in wilderness areas of Idaho, Montana, Washington, and Wyoming. Grizzly bears in the continental United States are currently listed as threatened under the ESA. Yellowstone National Park contains the largest, but most isolated, of the five remaining populations of grizzly bears in the continental United States. The Yellowstone grizzly bear population data were analyzed by Dennis et al. (1991) using the data available up to 1987, and the random-walk-with-drift model indicated that Yellowstone's grizzly bear population was doomed to extinction, although extinction was not expected within 500 years. The data consist of visual counts of unique females with cubs made by observers on the ground and in the air. The field methods are described by Knight et al. (1995), and more recent data are reported in Haroldson (2000). A three-year running sum is applied to the annual counts to reflect the fact that females reproduce no more frequently than every three years (Knight and Eberhardt 1985).

Dennis et al. (1991) speculated that effects of the large fires in 1988 might be detectable in a future analysis of the grizzly bear population data. I therefore analyzed the grizzly bear data with the constant- μ and step-change models (with $t_c = 1988$).

RESULTS AND DISCUSSION

Analysis of simulated data

Estimation accuracy and precision for the mean rate of population growth, μ , was similar for all three algorithms (results not shown), but different for the process-error variance, σ_p^2 . Fig. 1 shows the distribution of $\hat{\sigma}_p^2$ for different combinations of true σ_p^2 and σ_m^2 with different amounts of data (T). Several patterns are apparent. As expected, the Dennis algorithm (Dennis et

al. 1991) overestimates σ_p^2 when series contain substantial measurement error (Fig. 1A and C). The mean and variance of σ_p^2 estimates are similar for both the Kalman-filter algorithm and Holmes (2001) algorithm. The Holmes algorithm, however, is prone to negative σ_p^2 estimates when $\sigma_m^2 > \sigma_p^2$ (Fig. 1B) and T is short. Both the Kalman-filter algorithm and Holmes algorithm tend to underestimate σ_p^2 when T is short, but this bias declines with increasing T . When $\sigma_p^2 \leq \sigma_m^2$ (Fig. 1A and C), both the Kalman-filter algorithm and Holmes algorithm do a much better job of estimating σ_p^2 than the Dennis algorithm.

Negative bias in σ_p^2 estimates when time series are short is worrisome because extinction risk will tend to be underestimated in these cases. If any independent information is available on measurement precision, bias could be reduced and precision increased by specifying an informative prior distribution for σ_m^2 . Taking a Bayesian approach would also address the problem of imprecision in σ_p^2 estimates. None of the three algorithms provide very precise estimates of σ_p^2 , even when T is fairly large (≥ 60 observations). Because extinction-risk metrics are highly nonlinear functions of σ_p^2 , the uncertainty in σ_p^2 estimates should be accounted for in any population assessment used for management. A Bayesian assessment is therefore appropriate (Ludwig 1996).

Analysis of empirical data

California sea otter.—Fig. 2 shows the time series of sea otter observations and estimates of the true population size produced by the Kalman filter using the maximum-likelihood estimates (MLEs) for μ , σ_p^2 , and σ_m^2 for the constant- μ model. Table 1 gives MLEs and 90% central probability intervals of μ , σ_p^2 , σ_m^2 , and associated extinction parameters for the step-change and constant- μ models. The AIC for the step-change model is 3.22 lower than for the constant- μ model, which indicates the step-change model is a substantially better approximation to the data than the constant- μ model. For both models, about half the variance is attributed to process variation, although the estimate of the signal-to-noise ratio (s) is not very precise (cv [coefficient of variation] = 0.43–0.47). This indicates substantial uncertainty about the relative importance of process and measurement variation in generating variation in abundance estimates.

The estimate of σ_m^2 can be compared to the cv expected from binomial sampling, noting that $\sqrt{\sigma_m^2}$ is the standard deviation of the (multiplicative) measurement error assumed by the random-walk-with-drift model and therefore comparable to a cv . The cv expected from sampling 94.5% of a population of roughly 2000 individuals falls towards the lower end of the 90% central probability interval of $\sqrt{\sigma_m^2}$, implying that other sources of sampling variation are present (e.g., variation in observation conditions) or that the proportion of the population observed is <0.945 .

² URL: (<http://www.werc.usgs.gov/news/1999-06-21.html>)

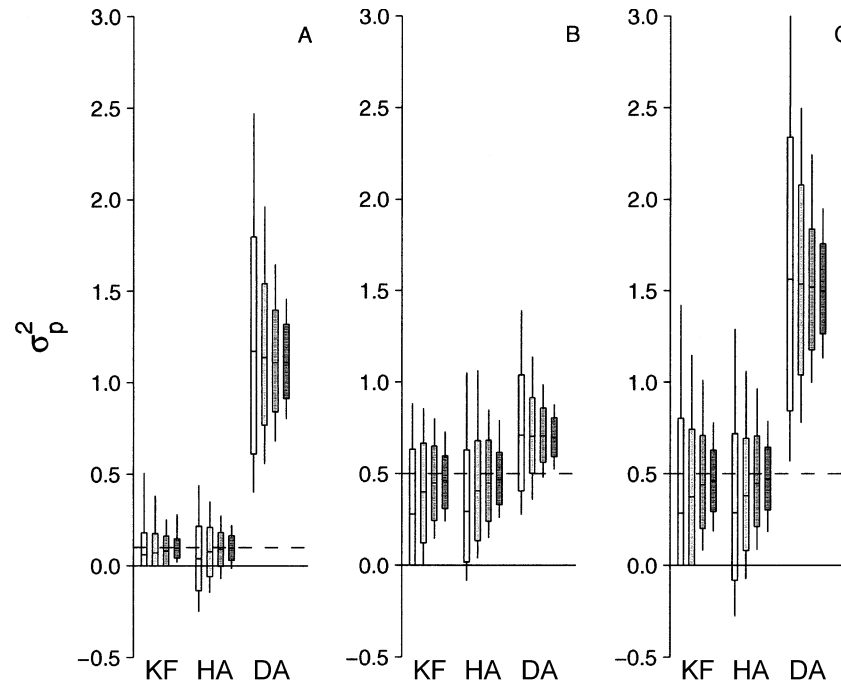


FIG. 1. Distribution of process-error variance (σ_p^2) estimates for the Kalman filter (KF), Holmes algorithm (HA; Holmes 2001), and Dennis algorithm (DA; Dennis et al. 1991) when applied to simulated random-walk-with-drift data of different length created with various levels of process variation and measurement variation. Boxes span the central 75th percentile interval, and whiskers span the 95% interval; the horizontal line inside the box marks the mean; the shade of gray denotes the length of the data series: white = 15 yr, light gray = 30 yr, medium gray = 60 yr, and dark gray = 120 yr. The dashed horizontal line denotes the actual value of σ_p^2 used to generate the simulated data. (A) $\sigma_p^2 = 0.10$, $\sigma_m^2 = 0.50$; (B) $\sigma_p^2 = 0.50$, $\sigma_m^2 = 0.10$; (C) $\sigma_p^2 = \sigma_m^2 = 0.50$.

While the constant- μ model indicates essentially no risk of extinction, the better-fitting step-change model suggests that the sea otter population is doomed to extinction because the current growth rate is negative ($\mu + \delta = -0.015$). The step-change model predicts a mean time to extinction of 518 years. While it is likely that the growth rate of this population will change again

in the future, the post-1995 data should motivate investigations of the causes of the population decline.

Although the step-change model appears to fit the data better, model diagnostics are satisfactory for both models. The plot of the cumulative sum of normalized residuals does not indicate any change in μ or $\text{Var}(X_t)$ during the time period of observation, and the plot of the cumulative sum of squared normalized residuals shows no sign of heteroscedasticity. Prediction errors appear to be normally distributed according to the Shapiro-Wilks test ($P = 0.37$). Neither the Box-Ljung Q test ($P = 0.74$) or the prediction-error autocorrelation function indicate significant serial correlation in the prediction errors.

Grizzly bear.—Fig. 3 shows the observed running-summed counts of female grizzly bears with cubs and the number estimated according to the constant- μ model. Parameter estimates for the two models are shown in Table 2. As in the sea otter example, the step-change model is a better approximation of the data, with an AIC 2.45 less than the constant- μ model. For the constant- μ model, the MLE of s (the signal-to-noise ratio) suggests that roughly 67% of the variation in grizzly bear counts comes from process variation, while the step-change model assigns 58% of the variation to process variation. For both models, the upper bound of the

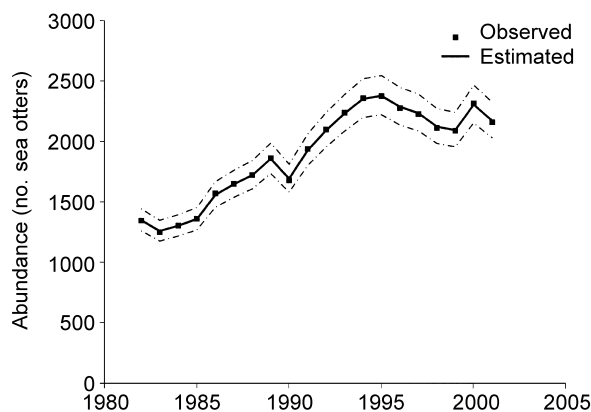


FIG. 2. Observed (solid squares) and model-estimated (thick line) California sea otter abundance. Dashed lines represent the standard deviation (± 1 SD) of the model estimation error.

TABLE 1. California sea otter population growth and extinction parameter estimates.

Parameter†	Model	
	Constant μ	Step-change μ
μ	0.0253 (0.00677, 0.0467)	0.0442 (0.0234, 0.0655)
δ	NA NA	-0.0596 (-0.0985, -0.0189)
s	1.23 (0.638, 2.50)	0.978 (0.542, 2.29)
σ_p^2	0.00231 (0.00179, 0.00329)	0.00149 (0.00113, 0.00269)
σ_m^2	0.00188 (0.00000289, 0.00448)	0.00152 (0.00106, 0.00233)
$P_{100}(\text{ext})$	0.0 ...	0.0 ...

Notes: Point estimates of μ are maximum-likelihood estimates; 90% central probability intervals (PI) are in parentheses. Values $< 1 \times 10^{-6}$ are indicated as 0.0. NA = not applicable.

† Parameters: μ = mean rate of population growth; δ = change in rate of population growth; s = signal-to-noise ratio; σ_p^2 = process-error variance; σ_m^2 = measurement-error variance; and $P_{100}(\text{ext})$ = probability of extinction within 100 years.

95% central probability interval (PI) for σ_p^2 's is quite large, which means that almost all of the variation in the data is potentially process variation; the lower bounds indicate that it is unlikely that measurement error contributes more than 52% or 35% to the total variation for the constant- μ and step-change models, respectively.

The population, in light of the post-1998 data, appears to be growing, although for the constant- μ model, the 90% PI includes negative values. It is possible to compare these model-estimated growth rates to independent, demographically based estimates of annual rate of population growth on the arithmetic scale (λ ; $\lambda = \exp(\mu + \sigma_p^2/2)$). Estimates of μ and σ_p^2 for the constant- μ model yield a $\lambda = 1.02$, in close agreement with a recent estimate of λ based on mortality and fecundity observations (Pease and Mattson 1999).

The most likely probability of ultimate extinction, according to the constant- μ model, is $< 1.0 \times 10^{-9}$, but the 90% PI for this parameter includes 1.0, and there is a 0.1% chance of extinction within 100 years. The step-change model, which fits the data better, paints a somewhat different picture. If the post-1988 growth rate continues indefinitely, the population would be exceedingly unlikely ever to go extinct. While the step-change model may be better in the sense of having a lower AIC, the constant- μ model might be better at capturing long-term processes and should be considered when making extinction predictions. It is likely that the grizzly bear population will not continue to grow at recent high rates because (1) the beneficial effects of the 1988 fires are probably transitory and (2) grizzly bear mortality is expected to increase as the

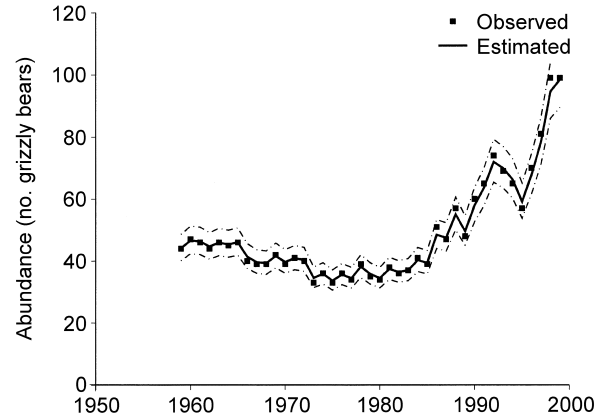


FIG. 3. Observed (solid squares) and model-estimated (thick line) grizzly bear abundance in Yellowstone Park. Dashed lines represent the standard deviation (± 1 SD) of the model estimation error.

size of the human population increases and the abundance of whitebark pine declines in the Yellowstone area (Pease and Mattson 1999).

Model diagnostics are not entirely satisfactory for the grizzly bear data: the plot of cumulative squared prediction errors for both models indicates that the variance of X_t is not constant throughout the time series. Visual inspection of the squared prediction errors shows that the variance is higher in the last third of the series, with a change perhaps occurring in 1988. This may be due to a change in the population or a change in the observations related to the fire effects, or to other changes in the Yellowstone ecosystem. While there are not enough data or auxiliary information available to identify the cause of the change in variance, the effects of different sources of variance on population projections can be examined. $\text{Var}(X_t - X_{t-1})$

TABLE 2. Yellowstone grizzly bear population growth and extinction parameter estimates.

Parameter	Model	
	Constant μ	Step-change μ
μ	0.0200 (-0.00536, 0.0465)	0.0221 (-0.0252, 0.0313)
δ	NA NA	0.0586 (0.00583, 0.110)
s	2.71 (1.12, 4400)	1.37 (0.532, 1450)
σ_p^2	0.00719 (0.00504, 0.0130)	0.00413 (0.00292, 0.0125)
σ_m^2	0.00251 (0.00000289, 0.00448)	0.00387 (0.00000853, 0.00551)
$P_{100}(\text{ext})$	0.0 (0.0, 0.000145)	0.0 ...

Notes: Point estimates of μ are maximum-likelihood estimates; 90% central probability intervals (PI) are in parentheses. Values $< 1 \times 10^{-6}$ are indicated as 0.0. NA = not applicable. Parameter definitions are as in Table 1.

was about 0.010 for $t \leq 1987$ and 0.020 for $t > 1987$. Taking a worst-case view, if all of the increased variance was due to an increase in process variation, and we assume that the additional variance is applicable in the future, then the most likely probability of ultimate extinction predicted by the constant- μ model increases from $<1.0 \times 10^{-9}$ to 6.6×10^{-4} . Other diagnostics are acceptable for both models.

CONCLUSIONS AND FURTHER DIRECTIONS

The state-space approach and Kalman filter provide a practical, robust, and rigorous way to estimate the parameters of the random-walk-with-drift model from noisy data. Arbitrarily accurate estimates of parameter uncertainty and the uncertainty of functions of these estimated parameters can be obtained conveniently using MCMC (Markov chain Monte Carlo) methods. This is possible because the estimates of uncertainty are based on the likelihood function for the data rather than on large-sample approximations. An additional advantage of state-space models, not exploited in the models and data presented here, is the ease with which missing data can be handled (Harvey 1989, Durbin and Koopman 2001).

The random-walk-with-drift model can be extended within the state-space framework, allowing models to be compared quantitatively. Obvious extensions include time-varying parameters, more elaborate state representations (such as stage- or age-structured populations), nonlinear state transitions (e.g., density-dependent reproduction), non-normal error structures, and inclusion of covariates influencing growth rate. In particular, an age-structured state-space model could be used to overcome the problem that arises when only semelparous breeding adults are observed, such as observations of spawning salmon.

One problem with more complex models is that they tend to be nonlinear, precluding the use of the Kalman filter. Other filtering methods, such as the extended Kalman filter (Kitagawa 1981), grid-based numerical approaches (Rein 1993, de Valpine and Hastings 2002), and sequential Monte Carlo filtering methods (Doucet et al. 2001), while more computationally intensive, may be useful for estimating more complex models of population abundance.

ACKNOWLEDGMENTS

I thank E. Bjorkstedt, D. Boughton, P. Crone, and R. C. Tuckfield for thoughtful reviews of the manuscript.

LITERATURE CITED

- Brown, R. L., J. Durbin, and J. M. Evans. 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B (statistical methods)* **37**:149–192.
- Burnham, K. P., and D. R. Anderson. 1998. *Model selection and inference: a practical information-theoretic approach*. Springer-Verlag, New York, New York, USA.
- Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. *Ecological Monographs* **61**:115–143.
- de Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. *Ecological Monographs* **72**:57–76.
- Doucet, A., N. de Freitas, and N. Gordon, editors. 2001. *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York, New York, USA.
- Durbin, J., and S. J. Koopman. 2001. *Time series analysis by state space methods*. Oxford University Press, Oxford, UK.
- Estes, J. A. 1990. Growth and equilibrium in sea otter populations. *Journal of Animal Ecology* **59**:385–401.
- Estes, J. A., and R. J. Jameson. 1988. A double-survey estimate for sighting probability of sea otters in California. *Journal of Wildlife Management* **52**:70–76.
- Foley, P. 1994. Predicting extinction times from environmental stochasticity and carrying capacity. *Conservation Biology* **8**:124–137.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. Introducing Markov chain Monte Carlo. Pages 1–19 in W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov chain Monte Carlo in practice*. Chapman & Hall, London, UK.
- Haroldson, M. A. 2000. Unduplicated females. Pages 9–13 in C. C. Schwarz and M. A. Haroldson, editors. *Yellowstone grizzly bear investigations: annual report of the Interagency Grizzly Bear Study Team, 1999*. U. S. Geological Survey, Bozeman, Montana, USA.
- Harvey, A. C. 1989. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge, UK.
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**:97–109.
- Hilborn, R., and M. Mangel. 1997. *The ecological detective*. Princeton University Press, Princeton, New Jersey, USA.
- Holmes, E. E. 2001. Estimating risks in declining populations with poor data. *Proceedings of the National Academy of Sciences (USA)* **98**:5072–5077.
- Kitagawa, G. 1981. A nonstationary time series model and its fitting by a recursive filter. *Journal of Time Series Analysis* **2**:103–116.
- Knight, R. R., B. M. Blanchard, and L. L. Eberhardt. 1995. Appraising status of the Yellowstone grizzly bear population by counting females with cubs-of-the-year. *Wildlife Society Bulletin* **23**:245–248.
- Knight, R. R., and L. L. Eberhardt. 1985. Population dynamics of Yellowstone grizzly bears. *Ecology* **66**:323–334.
- Ludwig, D. 1996. Uncertainty and the assessment of extinction probabilities. *Ecological Applications* **6**:1067–1076.
- Ludwig, D. 1999. Is it meaningful to estimate a probability of extinction? *Ecology* **80**:298–310.
- Meir, E., and W. F. Fagan. 2000. Will observation error and biases ruin the use of simple extinction models? *Conservation Biology* **14**:148–154.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equations of state calculations by fast computing machine. *Journal of Chemical Physics* **21**:1087–1091.
- Nelder, J. A., and R. Mead. 1965. A simplex method for function minimization. *Computer Journal* **7**:308–313.
- Pease, C. M., and D. J. Mattson. 1999. Demography of the Yellowstone grizzly bears. *Ecology* **80**:957–975.
- Rein, S. R. 1993. *A model of chinook population dynamics*. Dissertation. University of California, Berkeley, California, USA.