

# Reducing bias and improving precision in species extinction forecasts

KEVIN E. SEE<sup>1,3</sup> AND ELIZABETH E. HOLMES<sup>2</sup>

<sup>1</sup>*Quantitative Ecology and Resource Management, University of Washington, Box 352182, Seattle, Washington 98195 USA*

<sup>2</sup>*Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Boulevard East, Seattle, Washington 98112 USA*

**Abstract.** Forecasting the risk of population decline is crucial in the realm of biological conservation and figures prominently in population viability analyses (PVA). A common form of available data for a PVA is population counts through time. Previous research has suggested that improving estimates of population trends and risk from count data depends on longer observation periods, but that is often impractical or undesirable. Making multiple observations within a single time step is an alternative way to gather more data without extending the observation period. In this paper, we examine the trade-off between the length of the time period over which observations of the population have been taken and the total number of observations or samples that have been recorded through an analysis of simulated data. We found that when the ratio of process error to measurement error variance is high, more precise estimates of quasi-extinction risks can be obtained if replicated observations are taken at each time step, but when the ratio is low, replicated observations add little benefit in improving precision. These results can be used to efficiently design effective monitoring schemes for species of conservation concern.

**Key words:** *measurement error; monitoring design; population viability analysis; process error; PVA; quasi-extinction risk; repeated measures; sampling design; state-space model.*

## INTRODUCTION

Estimates of a species' risk of decline are crucial in identifying species of highest concern and to prioritize how quickly managers must act to prevent severe declines. For example, the IUCN (International Union for Conservation of Nature) classifies species as critically endangered when they are projected to decline by 80% over 10 years or three generations (IUCN Standards and Petitions Subcommittee 2014). As part of the classification process, the IUCN also requires a description of the uncertainty around these projections. Such projections and the surrounding uncertainty depend on estimates of the trend of population abundance and the variation around that trend. The accuracy and precision of these estimates rests in part on the quality and structure of data that have been gathered, which for many species of conservation concern take the form of population counts through time, developed from population surveys. Clearly more data improve the quality of an assessment, but given that a limited number of population surveys may be possible, for financial or logistical reasons, how those are structured in time may impact the assessment as well. The surveys could be spread through time, perhaps with one survey every few years, or could all occur within a small time frame with

multiple surveys each year. Understanding the effect of monitoring design on population assessments can help managers and conservationists to design more effective monitoring programs.

In recent years, ecologists have turned toward state-space models as a way to incorporate both environmental stochasticity (or process errors) and observation (or measurement) errors, while also accommodating missing data (Knape and de Valpine 2011, Pasinelli et al. 2011, Wilson et al. 2011, De Valpine 2012, Nadeem and Lele 2012, Russell et al. 2012, Hefley et al. 2013a, Larson et al. 2013). A state-space model consists of one component that describes the population process, including stochasticity, and a second component that models the observation process, including observation error. State-space models have outperformed process-error-only and observation-error-only models in a variety of non-observation situations (De Valpine 2002, De Valpine and Hastings 2002, Lele et al. 2007). They have been applied in a variety of ecological contexts, including fisheries (Meyer and Millar 1999, Millar and Meyer 2000, De Valpine and Hilborn 2005, Hinrichsen 2009, Russell et al. 2012), bird populations (Williams et al. 2003, J. W. Connelly et al. 2004 *unpublished report*; Jamieson and Brooks 2004, Dennis et al. 2006, Hefley et al. 2013b), animal movement (Newman 2000, Buckland et al. 2004, Newman et al. 2006, Buckland et al. 2007), and plankton communities (Ives et al. 2003).

A variety of state-space models have been described and utilized recently, including diffusion approximation models (Dennis et al. 2006, Holmes et al. 2007), *N*-

Manuscript received 17 October 2014; accepted 4 November 2014. Corresponding Editor: K. Gross.

<sup>3</sup> Present address: Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, Washington 98112 USA. E-mail: kevin@qcinc.org

mixture models, (Kery et al. 2009, Dail and Madsen 2011, Sólymos et al. 2012, Hefley et al. 2013a), as well as other formulations (Nadeem and Lele 2012, Hefley et al. 2013b). The state models all include a few demographic parameters and some process error variance, whereas the observation model usually takes one of three forms for the observation error structure: a log-normal, binomial, or Poisson model. In practice, some combination of biological knowledge and model selection can determine the most appropriate model to use. We chose to focus on a common diffusion approximation model that has been modified into a state-space model (Dennis et al. 2006), sometimes referred to as a corrupted stochastic exponential growth model, CSEG (Holmes et al. 2007). The purpose of this work was to investigate the impact of data structure on the precision of trend and quasi-extinction risk using the CSEG model, which has been validated and shown to correctly estimate quasi-extinction risk for large data sets of population count data from species of conservation concern. This model provided unbiased and fairly precise estimates of the risk of quasi-extinction for vertebrate species of conservation concern with declining populations when forecasting large declines, e.g., 50% and 80% (Holmes et al. 2005, 2007). Simulation studies have also shown that the quasi-extinction risk can be reliably approximated with a CSEG model in the presence of age-structured, density-dependent, spatially structured metapopulation or predator-prey population dynamics, as well as imperfect detection (Holmes and Semmens 2004, Sabo et al. 2004, Holmes et al. 2007, Sabo and Gerber 2007). If one is working with data that strongly violate the assumption of log-normal observation error or have strong cycles, a CSEG model is not appropriate; our results should not be assumed to extend to non-CSEG models applied to non-CSEG-like data.

Previous research using CSEG models has suggested that improving estimates of population trends and process variance depends on longer time series (Fieberg and Ellner 2000, Holmes et al. 2007, Ellner and Holmes 2008, Humbert et al. 2009, Connors et al. 2014), but there are often limits to the number of surveys that can be conducted and collecting a longer time series may be undesirable if it means delaying an assessment. Making multiple observations within a single time step is an alternative way to gather more data without extending the observation period. Several recent studies have demonstrated how repeated measurements can improve the uncertainty and identifiability of CSEG parameter estimates. Dennis et al. (2010) used a form of the CSEG model that includes density dependence to demonstrate how one additional observation per time step leads to steeper likelihood profiles for each parameter, meaning that parameters are more identifiable, even if the time series were half as long. Hinrichsen (2009) examined the effect of incorporating data from nearby populations and assuming correlated process errors to improve the estimates of mean trend and process error variance.

Knape et al. (2013) explored the effect of a range of spatial replicates (2–10) on parameter estimates from a model similar to the model employed by Dennis et al. (2010), using several different maximum likelihood techniques. They demonstrated improved estimation performance with an increasing number of spatial replicates for three different time series lengths.

Although the benefits of repeated measurements and longer observation periods are clear, researchers and managers must often design a monitoring program under financial constraints: replicated observations are not free, and the best allocation of surveys across time and space is a key question for monitoring design. Our paper examines the trade-offs between the total length of the observation period and the number of total observations on each parameter estimate, as well as the impact of these trade-offs on estimates of the probability of quasi-extinction, under a variety of relative sizes of process error and measurement error variances. We consider a range of scenarios, from short to long time series and from multiple observations within years to years with missing data. Our analysis allows researchers to weigh the effects of different monitoring design choices on the expected uncertainty of various viability metrics when limited resources must be allocated across time and space. For managers tasked with making decisions about which species are at the greatest risk and how quickly to take action, a shorter monitoring period before an assessment could make the difference between acting in time to save a species from decline or not.

This paper proposes to examine the effects of two different structures in the data—the length of the observation period (e.g., number of years) and the total number of observations—on the estimates of the probability of quasi-extinction ( $P_e$ ). Process error and measurement error are difficult to disentangle (Dennis et al. 2006, Knape 2008), but doing so is crucial in estimating the risk of quasi-extinction, because only process error impacts that risk. Longer observation periods can provide less biased and more precise estimates of variance in process error (Holmes et al. 2007, Ellner and Holmes 2008), but more observations per time step, rather than a longer observation period, may provide more precise estimates of variance in measurement error, leading in turn to more precise estimates of process error variance.

In this paper, we will address three questions related to the design of monitoring programs designed to generate population counts.

- 1) How do changes in the data structure (longer time series vs. shorter with repeated observations) impact the bias and precision of parameter estimates such as the rate of population decline ( $\mu$ ), process error variance ( $\sigma^2$ ), and measurement error variance ( $\tau^2$ )?
- 2) Under what conditions (i.e., relative size of process error and measurement error variance) do changes in

the data structure alter the bias and precision of estimates of a population’s risk of decline?

- 3) How quickly can one gather enough data to make a reasonably precise estimate of the probability of quasi-extinction?

It should be noted that, in this paper, precision refers to how reproducible an estimate is across a series of simulations (coefficient of variation across simulations), rather than the size of the standard errors for any particular simulation.

METHODS

*CSEG model*

The CSEG model is

$$\log X_{t+1} = \log X_t + \mu + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2) \quad (1)$$

$$\log Y_t = \log X_t + \eta_t, \quad \eta_t \sim N(0, \tau^2). \quad (2)$$

This state-space model describes the exponential growth of a population’s abundance,  $X_t$ , at time  $t$ , with the mean long-term annual growth rate,  $\mu$ , as well as process error,  $\varepsilon_t$ , due to environmental stochasticity. Observed counts,  $Y_t$ , are based on the state,  $X_t$ , overlaid with another form of random noise,  $\eta_t$ , which derives from measurement or observation error. Both  $\varepsilon_t$  and  $\eta_t$  are assumed to be drawn from a normal distribution with mean 0 and variance  $\sigma^2$  and  $\tau^2$ , respectively. To include  $m$  observations of a state at each time step, Eq. 2 can be extended to describe a multivariate state-space model:

$$\log \mathbf{Y}_t = \mathbf{Z} \log X_t + \mathbf{F}_t, \quad (3)$$

where  $\mathbf{Y}_t$  is an  $m \times 1$  vector of the multiple observations taken at time  $t$ .  $\mathbf{Z}$  is an  $m \times 1$  vector of 1’s, describing that all the observations are of the same underlying (population) state.  $\mathbf{F}_t$  is also an  $m \times 1$  vector of measurement errors, assumed to come from a multivariate normal distribution with mean  $\mathbf{0}$  and covariance matrix  $\mathbf{R}$ . These measurement errors could be independent, identical, or correlated within each time step. Further details of this model can be found in Hinrichsen and Holmes (2010).

Diffusion approximation methods have been used in conservation biology for over 20 years to estimate several PVA metrics, including mean time to extinction, distribution of extinction times and the probability of dropping below a quasi-extinction threshold (Dennis et al. 1991, Lande et al. 2003, Schultz and Hammond 2003, Holmes 2004, Snover and Heppell 2009). The probability of quasi-extinction over a given time-horizon,  $T$ , is estimated by using the inverse Gaussian distribution of first passage times for Brownian motion with drift (Dennis et al. 1991, Fieberg and Ellner 2000). This probability is given by:

$$P_e = \Phi(U - V) + \exp(2UV)\Phi(-(U - V)) \quad (4)$$

where  $\Phi$  is the standard normal cumulative distribution function,  $U = -\mu\sqrt{T}/\sigma$ ,  $V = a/\sigma\sqrt{T}$ , and  $a$  is  $\log(\text{initial population abundance}/\text{quasi-extinction threshold})$ . Note that because the measurement errors do not impact the true state’s trajectory through time, their variance does not impact the probability of quasi-extinction. Diffusion approximation techniques make the assumption, as do most forecasting methods, that the past is representative of the future, or that the population will continue to experience the same mean growth rate,  $\mu$ , and the same demographic stochasticity described by the process error variance,  $\sigma^2$ . This type of estimate is used as a “baseline” in a PVA, namely “if conditions as reflected in the time period used for estimation continue, what is the population risk of quasi-extinction?” This is another reason that a long time series can be undesirable; it means that one is more likely to include data (the far past) that are not reflective of the conditions that the population will experience in the near future.

*Simulations*

To study the impact of the length of a observation period and the number of total observations, five lengths of observation periods (5, 15, 30, 45, and 60 time steps) and five different levels of total observations (15, 30, 60, 75, and 90) were considered. For each combination of observation period length and total observations, 500 simulated time series were generated from predetermined rates of decline, process error variance, and measurement error variance. For some combinations, this led to one observation at every time step, for others there were multiple observations at each time step, and for some there were time series with missing observations for some time steps. In the last case, the placement of the missing values was randomized for each simulated time series, perhaps equating to a scenario with unequal funding through time.

The parameter values were chosen based on previous studies that have estimated  $\mu$  and  $\sigma^2$  from real populations (Brook et al. 2000, Lindley 2003, Staples et al. 2004, Dennis et al. 2006, Holmes et al. 2007). The vast majority of these species were terrestrial mammals and birds. The rates of decline correspond to deterministic declines of about 30% and 50% over 20 years ( $\mu = -0.02, -0.04$ ), which would lead to listings as vulnerable and endangered under the IUCN criteria (IUCN Standards and Petitions Subcommittee 2014). We chose parameter values such that the sum of the process error and measurement error variances was always 0.1 ( $\sigma^2 + \tau^2 = 0.1$ ) and the ratio of process- to measurement-error variance ranged from 0.05 (relatively little environmental stochasticity) to 5 (relatively high environmental stochasticity). This led to process error variances in the ranges discussed by Holmes et al. (2007), Ellner and Holmes (2008) and Humbert et al. (2009). Simulations began with an initial population of  $\exp(10)$ , and were allowed to burn-in for 100 time steps before simulated observations

began. This resulted in initial population sizes of around 400 and 3000, depending on the value of  $\mu$ .

We generally equate a time step to one year, primarily because that is the timescale on which many species reproduce. For some species that reproduce faster, it may make sense to think of time steps in terms of months, or weeks, or possibly days. For some species that reproduce much more slowly, the appropriate time step may be several years, or possibly a decade. However, for most species a year is the appropriate time step, both because it is their reproductive timescale, but also because many species are affected by the seasons, and therefore it is appropriate to estimate their trend on a scale that encompasses all four seasons. Therefore, in this paper we use time step and year interchangeably.

#### *Parameter and extinction metric estimation*

Each simulated time series was analyzed to estimate the trend ( $\mu$ ), process error variance ( $\sigma^2$ ) and measurement error variance ( $\tau^2$ ), using maximum likelihood estimates obtained from the Kalman filter and EM (expectation–maximization) algorithm, implemented by the MARSS package (Holmes et al. 2014) for R software (R Development Core Team 2009). Other methods were evaluated (see Appendix A), but the MARSS algorithm performed the best. In our simulations and estimating models, it was assumed that the measurement errors within each time step were independent and unbiased (centered around 0), so the observation errors could depend on the patchiness or within-year movement of the population, or could be due to measurement error, or some combination. If within-year samples were taken from different locations (e.g., three sampling stations, each of which was visited every year, or two were visited a year in on a rotating schedule, etc.), the CSEG model can accommodate that by estimating a bias parameter for each location (relative to one of the locations). However, in an analysis focused on estimating  $\mu$  and  $\sigma^2$ , one can center all of the time series and eliminate the need to estimate extra (bias) parameters. Therefore, these results are applicable to designs that sample the same location multiple times, or multiple locations once within a time step, as long as the within-year observations can be considered independent.

For each of the 25 combinations of observation period length and total number of observations, the coefficient of variation (CV) of the parameter estimates across all simulations was calculated as a measure of precision, defined as the standard deviation of the estimates divided by the true parameter value. From these 30 combinations, we used loess smoothing to interpolate a contour plot to examine the trade-offs in precision that arise from changing the length of an observation period and changing the total number of observations. Relative and absolute errors for the parameter estimates were compared across the lengths

of observation periods and the number of total observations.

Based on the CSEG parameter estimates, a probability of quasi-extinction was calculated for each simulated time series. We examined the probability that simulated populations would decline by 30%, 50%, and 80% over 30- and 50-year horizons, which begin from the end of the observed time series. The impacts of the length of the observation period and total number of observations were similar across all three forecast horizons, so we only present results from the 30-year time horizon (see Appendix B for additional results). The precision of these quasi-extinction estimates was examined by comparing the coefficient of variation of the  $P_e$  estimates across each combination of length of the observation period and total number of observations, calculated in the same fashion as the CV of the parameter estimates. Using as a reference the case of an observation period of 30 years with one observation every year, we interpolated a contour plot to show the improvement in precision when the available data differed from that reference case. The effects of length of the observation period and total observations on the precision of quasi-extinction probabilities were similar for  $\mu = -0.02$  and  $\mu = -0.04$ , and across the three levels of potential decline, so we only present results from one group of simulations ( $\mu = -0.02$  and a 50% decline; but see Appendix B for full results).

## RESULTS

### *Parameter bias*

Estimates of the population's mean rate of decline were unbiased (centered over the true value) regardless of the length of the observation period or the total number of observations. In most simulated scenarios, estimates of process and measurement error variance were also unbiased. However, when measurement error variance was very small relative to the process error variance ( $\sigma^2/\tau^2 = 5$ ), and there was less than one observation per time step, we tended to overestimate measurement error variance, probably because the data contained little information about that parameter. This in turn led to an underestimate of the process error variance. This bias was reduced if there were one or more observations per time step. Similarly, when the process error variance was relatively quite small ( $\sigma^2/\tau^2 = 0.05$ ), we tended to overestimate the process error variance and underestimate the measurement error variance, particularly if the total number of observations was low, regardless of the length of the observation period (see Appendix B for supporting figures).

### *Parameter precision*

The precision of parameter estimates depended on different structures in the simulated data, depending on which parameter we focused on. Estimates of the population's mean rate of decline ( $\mu$ ) became more precise as the observed period became longer, but were

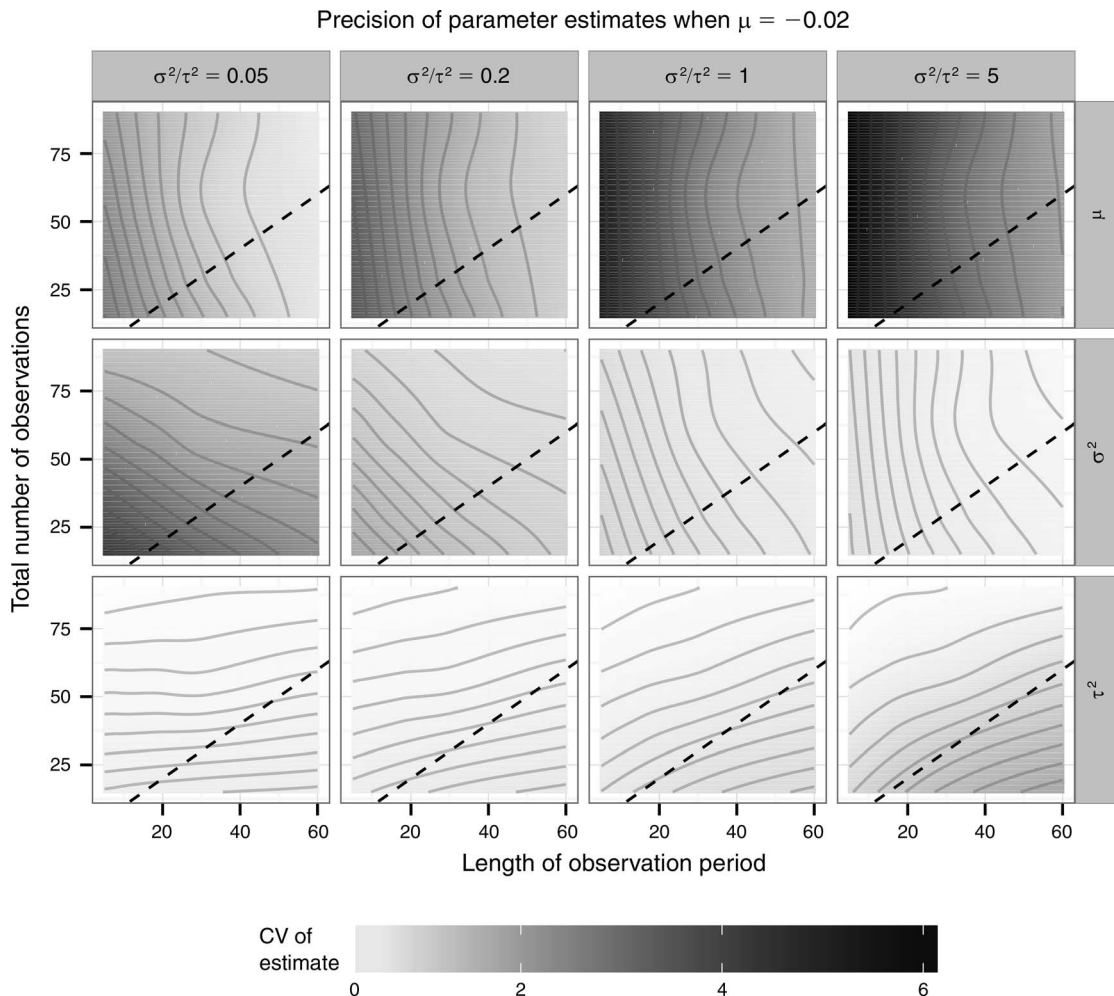


FIG. 1. Contour plots of the coefficient of variation (CV) of the parameter estimates for rate of decline (top row,  $\mu$ ), process error variance (environmental stochasticity; middle row,  $\sigma^2$ ), and measurement error variance (observation error; bottom row,  $\tau^2$ ), when the true value of  $\mu$  is  $-0.02$ . Each column corresponds to a different process- to measurement-error variance ratio ( $\sigma^2/\tau^2$ ). The dashed line corresponds to scenarios where one observation is taken every time step. There are no units for the  $x$ -axis observation period; although a year is the appropriate time step for most species, modeling also applies to shorter time steps (e.g., weeks for phytoplankton).

not affected by the number of total observations, as seen in the nearly vertical contour lines of the top row in Fig. 1. On the other hand, the precision of measurement error variance estimates ( $\tau^2$ ) improved primarily when the total number of observations was increased, regardless of the length of the observation period, as seen in the horizontal contour lines of the bottom row of Fig. 1. The improvements in the precision of estimates of  $\mu$  and  $\tau^2$  show diminishing returns as the length of the observation period or the total number of observations, respectively, were increased. In contrast, the precision of process error variance estimates was improved by both a longer observation period and an increasing number of total observations, illustrated by the diagonal contour lines of the middle row of Fig. 1. As process error variance increases, the contour lines shift from closer to horizontal to closer to vertical, depicting less improve-

ment in precision with an increasing number of observations, although some still exists.

#### *Probability of quasi-extinction*

The probabilities of quasi-extinction were estimated with little bias for all three levels of decline and all combinations of observation period length and number of total observations (see Appendix B for relevant plots). Estimates that were biased low, such as the probability of an 80% decline when the process error variance was relatively high and the total number of observations was low, correspond to cases in which the process error variance was underestimated (often accompanied by overestimates of the measurement error variance). This illustrates the importance of accurate estimates of process error variance when conducting this type of PVA. The precision improved for all three levels of

possible decline as the observation period length was increased, and there was a much smaller effect on the precision due to the total number of observations.

However, the relative effects of more total observations and longer observation periods on the precision of  $P_e$  depended on the ratio of process- to measurement-error variance. When that ratio was small (relatively little process error variance), the precision of  $P_e$  estimates improved primarily as a result of increasing the observation period. As that ratio grew larger (i.e., relatively more variation due to environmental stochasticity), the same precision could be obtained with a shorter observation period, provided additional samples were taken (Fig. 2). For instance, if this ratio were high (i.e.,  $\sigma^2/\tau^2 = 5$ ), then compared to the reference case of a 30-year time series with one observation a year, the precision of estimating the risk of a 50% decline was improved by 20% if observations were taken twice a year. Although there are twice as many data, they do not contain twice as much information about the population dynamics, but this does increase the confidence that a manager should have in the results by improving the estimate of measurement error variance and, thus, of the process error variance. However, increasing the number of observations to three per year resulted in very little additional improvement in precision, illustrating the diminishing returns that come with more repeated measurements. Looking at these results another way, the same precision as the reference case could be obtained by increasing the total number of observations from 30 to 40 and collecting observations twice a year for 20 years rather than once a year for 30 years (Fig. 2).

#### DISCUSSION

The three parameters in the CSEG model are affected differently by different structures in the data. Under almost all of the scenarios that we tested, the parameter estimates were unbiased, but the length of the observation period and the total number of observations influenced the precision of the parameter estimates in contrasting ways. The rate of population decline is estimated more precisely with a longer observation period, whereas the total number of observations is almost irrelevant. This results supports previous simulation studies that have shown estimates of population trends to be more precise with longer time series, but that precision is unaffected by the number of missing observations in a given time series length (Humbert et al. 2009).

The effect of data structure on estimates of process error, or environmental stochasticity ( $\sigma^2$ ), changes as the ratio of process- to measurement-error variance shifts (middle row of Fig. 1). When this ratio is low, precision can be improved appreciably with more total observations—via either more observations per time step or one observation per time step and a longer observation period. However, as this ratio grows, the length of the observation period becomes more important, particu-

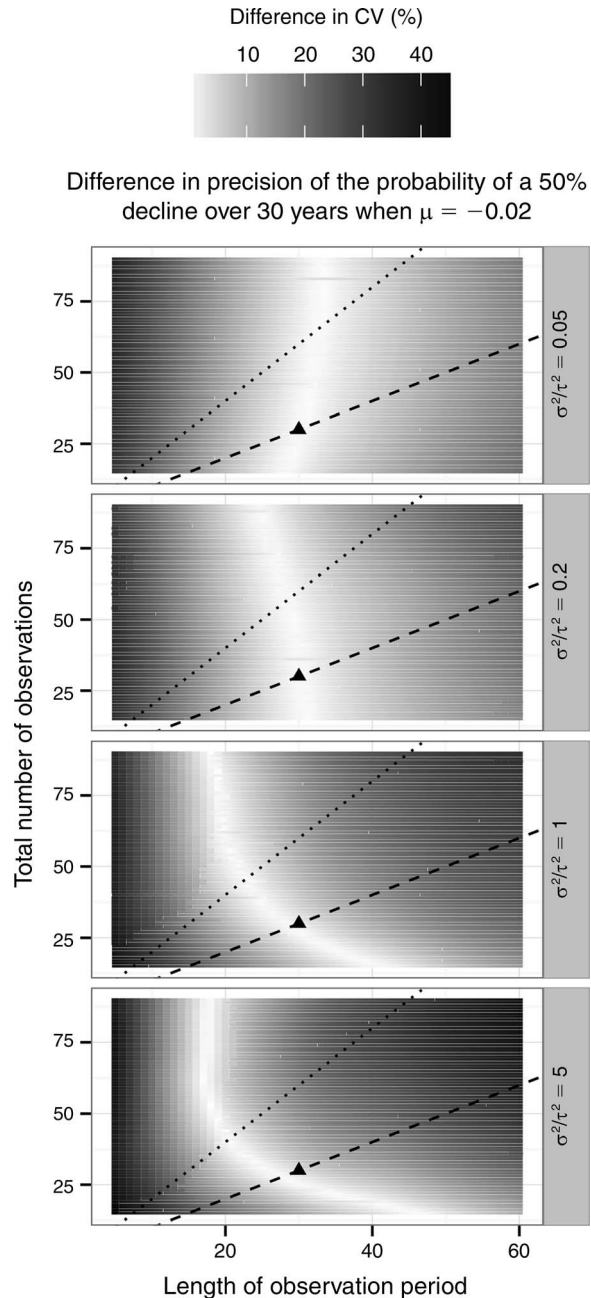


FIG. 2. Contour plots of the percentage improvement to the coefficient of variation (CV) of the estimates of the probability of a 50% decline in 30 years, compared to a reference case of a 30-year observation period with one observation each year (marked with a solid triangle). The white contour area corresponds to the same level of precision as the reference case. Areas to the right of that white contour have a more precise estimate of  $P_e$  (probability of quasi-extinction) than the reference case, whereas areas to the left have a less precise estimate. The dashed and dotted straight lines correspond to scenarios where one and two observation(s) are taken every time step, respectively. Each facet corresponds to a different ratio of process error to measurement error variance ( $\sigma^2/\tau^2$ ).

larly when the ratio of process to measurement error variance is 1 or greater. If the observation period is long enough, more total observations can still improve the precision of  $\sigma^2$  estimates, but not as markedly as when the  $\sigma^2/\tau^2$  ratio is smaller.

Finally, the measurement error variance is estimated more precisely as the total number of observations is increased. For higher ratios of process error to measurement error variance, precision of the estimate of measurement error variance begins to deteriorate rapidly in longer time series once observations are taken less than once per time step, and the estimates are more biased (see relevant figures in Appendix B). Because measurement errors do not feed back into the true state of the system, longer time series provide no additional information for estimating the measurement error variance. In fact, the most precise estimates of measurement error variance can be obtained from a very short time series with a large number of (independent) observations taken at each time step (bottom row of Fig. 1). This suggests a combination strategy of deploying observations through time, which will be discussed further.

It is encouraging to note that the quasi-extinction estimates are unbiased, but in practice a biologist will be assessing one data set, and making one quasi-extinction estimate. Thus precision is most important, and the data structure that provides a more precise estimate will allow the biologist to have more confidence in the result. Consistent with results from previous studies (Holmes et al. 2007), we found that the precision of quasi-extinction estimates deteriorates rapidly when time series are less than 20 time steps for all simulated scenarios. This 20-time-steps threshold probably arises from the precision of a population's estimated rate of decline ( $\mu$ ), which also drops off precipitously with less than 20 time steps of observations, regardless of the ratio of process error to measurement error variance (top row of Fig. 1). This threshold of observation length almost certainly depends on  $\sigma^2$ ; we used  $\sigma^2$  values in our simulations that are consistent with estimates from small to large terrestrial vertebrates. If  $\sigma^2$  is larger than our simulated values, we expect the 20-year threshold to increase, whereas if  $\sigma^2$  is smaller, the threshold will probably shrink as well. It is important to note that estimates of quasi-extinction will always become more precise with a longer time series, so if quasi-extinction risk is important, a long-term monitoring program will be needed.

However, for certain populations, repeated observations within, and potentially beyond, those 20 years can improve the precision of quasi-extinction estimates. Given at least a 20-year observation period, how the precision of quasi-extinction estimates changes with the length of the observation period and total number of observations depends on the relative strengths of process and measurement error variances. When the risk of quasi-extinction is dominated by chance events, such as when  $V = a/\sigma\sqrt{T}$  (from Eq. 4) is quite small, an estimate

of  $P_e$  depends less on the estimate of the population trend,  $\mu$ , and more on the estimate of process error variance,  $\sigma^2$  (Fieberg and Ellner 2000). In our simulations, these scenarios occur when the level of decline is moderate (30–50%) and process error variance is relatively high ( $\sigma^2/\tau^2 \geq 1$ ). These are also the scenarios in which multiple observations within a time step can improve the precision of quasi-extinction estimates by providing more precise estimates of the measurement error variance and, consequently, the process error variance. Projecting forward 30 years, this corresponds to the lower two panels in Fig. 2. When  $\sigma^2/\tau^2 = 5$ , similar precision in quasi-extinction estimates can be obtained with 20 years of data, if two observations are taken each year, compared to a sampling regime of one observation each year for 30 years. Obtaining the same precision requires 10 more observations (40 vs. 30), but an analysis performed 10 years sooner could lead to conservation actions being taken a decade earlier, which could make a large difference for the future health of the population.

For managers tasked with prioritizing resources for monitoring programs, an understanding of the relative strengths of environmental stochasticity vs. other variability can aid in designing the most efficient monitoring scheme. Measurement error will most likely not be eliminated (Holmes 2001), but if the process error variance is comparatively large, managers should consider implementing a monitoring design with multiple observations taken each year. Due to the diminishing returns, a sampling regime of two or three observations per year may make the most sense from a cost–benefit perspective, depending on the cost of obtaining additional observations. Whether those multiple observations should come from the same location or different locations should be determined by discussions among the field biologists as to what design will generate the most independence between the multiple observations.

When little is known about the relative strengths of process and measurement errors, we recommend a combination approach that involves taking multiple samples in each of the first few years. An analysis at that point should yield precise estimates of the measurement error variance. Based on estimates of the process error variance, which may be quite imprecise at that point, as well as knowledge of process error variances for similar species, the ratio of process error to measurement error variance can be estimated, and a more efficient monitoring design can be implemented going forward. If process errors are relatively small, a monitoring program with a single observation per year may be sufficient. If process errors are on the same scale as, or are much larger than, measurement errors, one of two monitoring designs may be useful. The first would be continuing to collect at least two observations each year. The second would be to collect a single observation in most years, but periodically collect several observations in a single year. We did not test this approach in our

simulations, but it should provide biologists with information about the measurement error variance, which will then inform estimates of process error variance and therefore the risk of quasi-extinction.

One factor that was not considered was the possibility of correlation among the observations at each time step, which reduces the amount of information contained in each repeated observation. This might occur, for example, if an environmental driver such as temperature or time of day is affecting the detection probability, such that all observations in a given year are affected similarly, either because it is a particularly cold year, or the first survey is always done in the morning and the second in the evening. If such correlation exists, it is difficult to estimate from count data alone. If one suspects that this is occurring, including the most likely covariate(s) in the model to explain the observed bias in detectability can remove the correlation between the samples. Such scenarios were beyond the scope of this paper, but should be considered by managers and biologists on a case-by-case basis. The model framework that we used (MARSS models) can easily incorporate observation error covariates or they can be included in the step during which the raw surveys are converted to a population estimate.

In a world with limited resources to direct toward conservation actions, reliable and timely estimates of quasi-extinction risks are crucial. Designing efficient monitoring programs is essential for species of conservation concern, so as to efficiently use scarce resources for data collection. We have shown how, for some species and populations, designs with repeated measurements can reduce the length of time spent observing the population before identifying it as at risk. This can allow appropriate conservation actions to be taken sooner, potentially influencing how successful those actions are and whether the species or population will recover.

#### ACKNOWLEDGMENTS

The authors would like to thank Jennifer Ruesink, André Punt, Ray Hilborn, and Eric Ward for their valuable ideas, reviews and input. We would also like to thank Kevin Gross and two anonymous reviewers for comments that greatly improved the manuscript. This research was funded in part by a grant from the Washington Sea Grant Program, University of Washington, pursuant to National Oceanic and Atmospheric Administration Award No. NA08OAR4170761, Project E/I-15. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its sub-agencies.

#### LITERATURE CITED

- Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, H. R. Akcakaya, and R. Frankham. 2000. Predictive accuracy of population viability analysis in conservation biology. *Nature* 404:385–387.
- Buckland, S. T., K. B. Newman, C. Fernandez, L. Thomas, and J. Harwood. 2007. Embedding population dynamics models in inference. *Statistical Science* 22:44–58.
- Buckland, S. T., K. B. Newman, L. Thomas, and N. B. Koesters. 2004. State-space models for the dynamics of wild animal populations. *Ecological Modelling* 171:157–175.
- Connors, B., A. Cooper, R. Peterman, and N. Dulvy. 2014. The false classification of extinction risk in noisy environments. *Proceedings of the Royal Society B* 281:20132935.
- Dail, D., and L. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. *Biometrics* 67:577–587.
- 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. *Bulletin of Marine Science* 70:455–471.
- De Valpine, P. 2012. Frequentist analysis of hierarchical models for population dynamics and demographic data. *Journal of Ornithology* 152:393–408.
- De Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. *Ecological Monographs* 72:57–76.
- De Valpine, P., and R. Hilborn. 2005. State-space likelihoods for nonlinear fisheries time-series. *Canadian Journal of Fisheries and Aquatic Sciences* 62:1937–1952.
- Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. *Ecological Monographs* 61:115–143.
- Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. 2006. Estimating density dependence, process noise, and observation error. *Ecological Monographs* 76:323–341.
- Dennis, B., J. M. Ponciano, and M. L. Taper. 2010. Replicated sampling increases efficiency in monitoring biological populations. *Ecology* 91:610–620.
- Ellner, S., and E. Holmes. 2008. Commentary on Holmes et al. (2007): resolving the debate on when extinction risk is predictable. *Ecology Letters* 11(8):E1–E5.
- Fieberg, J., and S. P. Ellner. 2000. When is it meaningful to estimate an extinction probability? *Ecology* 81:2040–2047.
- Hefley, T. J., A. J. Tyre, and E. E. Blankenship. 2013a. Fitting population growth models in the presence of measurement and detection error. *Ecological Modelling* 263:244–250.
- Hefley, T. J., A. J. Tyre, and E. E. Blankenship. 2013b. Statistical indicators and state-space population models predict extinction in a population of bobwhite quail. *Theoretical Ecology* 6:319–331.
- Hinrichsen, R. A. 2009. Population viability analysis for several populations using multivariate state-space models. *Ecological Modelling* 220:1197–1202.
- Hinrichsen, R. A., and E. E. Holmes. 2010. Using multivariate state-space models to study spatial structure and dynamics. Pages 145–166 in S. Cantrell, C. Cosner, and S. Ruan, editors. *Spatial ecology*. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Holmes, E. E. 2001. Estimating risks in declining populations with poor data. *Proceedings of the National Academy of Sciences USA* 98:5072–5077.
- Holmes, E. E. 2004. Beyond theory to application and evaluation: diffusion approximations for population viability analysis. *Ecological Applications* 14:1272–1293.
- Holmes, E. E., W. F. Fagan, J. J. Rango, J. A. Folarin, J. A. Sorensen, J. E. Lippe, and N. E. McIntyre. 2005. Cross validation of quasi-extinction risks from real time series: An examination of diffusion approximation methods. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-67.
- Holmes, E. E., J. L. Sabo, S. V. Viscido, and W. F. Fagan. 2007. A statistical approach to quasi-extinction forecasting. *Ecology Letters* 10:1182–1198.
- Holmes, E. E., and B. Semmens. 2004. Population viability analysis for metapopulations: a diffusion approximation approach. Pages 565–597 in I. Hanski and O. Gaggiotti, editors. *Ecology, genetics, and evolution of metapopulations*. Elsevier Academic Press, Amsterdam, The Netherlands.
- Holmes, E. E., E. J. Ward, and M. Scheuerell. 2014. Analysis of multivariate time-series using the MARSS package. Version 3.9 edition. NOAA Fisheries, Northwest Fisheries Science



- Center, Seattle, Washington, USA. <http://cran.r-project.org/web/packages/MARSS/>
- Humbert, J. Y., L. S. Mills, J. S. Horne, and B. Dennis. 2009. A better way to estimate population trends. *Oikos* 118:1940–1946.
- IUCN Standards and Petitions Subcommittee. 2014. Guidelines for using the IUCN red list categories and criteria. Version 11. International Union for Conservation of Nature, Gland, Switzerland. <http://www.iucnredlist.org/documents/RedListGuidelines.pdf>
- Ives, A. R., B. Dennis, K. L. Cottingham, and S. R. Carpenter. 2003. Estimating community stability and ecological interactions from time-series data. *Ecological Monographs* 73:301–330.
- Jamieson, L., and S. Brooks. 2004. Density dependence in North American ducks. *Animal Biodiversity and Conservation* 27:113–128.
- Kery, M., R. M. Dorazio, L. Soldaat, A. van Strien, A. Zuur, and J. A. Royle. 2009. Trend estimation in populations with imperfect detection. *Journal of Applied Ecology* 46:1163–1172.
- Knape, J. 2008. Estimability of density dependence in models of time series data. *Ecology* 89:2994–3000.
- Knape, J., P. Besbeas, and P. de Valpine. 2013. Using uncertainty estimates in analyses of population time series. *Ecology* 94:2097–2107.
- Knape, J., and P. de Valpine. 2011. Are patterns of density dependence in the global population dynamics database driven by uncertainty about population abundance? *Ecology Letters* 14:17–23.
- Lande, R., S. Engen, and B.-E. Sæther. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, Oxford, UK.
- Larson, E. I., B. High, and C. L. Sullivan. 2013. Status updates for Yellowstone cutthroat trout, redband trout, and bull trout in Idaho. Technical Report Number 13-18, Idaho Department of Fish and Game, Boise, Idaho, USA. <https://collaboration.idfg.idaho.gov/FisheriesTechnicalReports/Res13-18Meyer2014%20salmonid%20status%20report.pdf>
- Lele, S. R., B. Dennis, and F. Lutscher. 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecology Letters* 10:551–563.
- Lindley, S. T. 2003. Estimation of population growth and extinction parameters from noisy data. *Ecological Applications* 13:806–813.
- Meyer, R., and R. B. Millar. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences* 56:37–52.
- Millar, R. B., and R. Meyer. 2000. Bayesian state-space modeling of age-structured data: fitting a model is just the beginning. *Canadian Journal of Fisheries and Aquatic Sciences* 57:43–50.
- Nadeem, K., and S. R. Lele. 2012. Likelihood based population viability analysis in the presence of observation error. *Oikos* 121:1656–1664.
- Newman, K. B. 2000. Hierarchic modeling of salmon harvest and migration. *Journal of Agricultural, Biological, and Environmental Statistics* 5:430–455.
- Newman, K. B., S. T. Buckland, S. T. Lindley, L. Thomas, and C. Fernandez. 2006. Hidden process models for animal population dynamics. *Ecological Applications* 16:74–86.
- Pasinelli, G., M. Schaub, G. Häfliger, M. Frey, H. Jakober, M. Müller, W. Stauer, P. Tryjanowski, J.-L. Zollinger, and L. Jenni. 2011. Impact of density and environmental factors on population fluctuations in a migratory passerine. *Journal of Animal Ecology* 80:225–234.
- R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Russell, R. E., D. A. Schmetterling, C. S. Guy, B. B. Shepard, R. McFarland, and D. Skaar. 2012. Evaluating a fish monitoring protocol using state-space hierarchical models. *Open Fish Science Journal* 5:1–8.
- Sabo, J. L., and L. R. Gerber. 2007. Predicting extinction risk in spite of predator–prey oscillations. *Ecological Applications* 17:1543–1554.
- Sabo, J. L., E. E. Holmes, and P. Kareiva. 2004. Efficacy of simple viability models in ecological risk assessment: Does density dependence matter? *Ecology* 85:328–341.
- Schultz, C. B., and P. C. Hammond. 2003. Using population viability analysis to develop recovery criteria for endangered insects: case study of the Fender’s blue butterfly. *Conservation Biology* 17:1372–1385.
- Snover, M. L., and S. S. Heppell. 2009. Application of diffusion approximation for risk assessments of sea turtle populations. *Ecological Applications* 19:774–785.
- Sólymos, P., S. Lele, and E. Bayne. 2012. Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. *Environmetrics* 23(2):197–205.
- Staples, D. F., M. L. Taper, and B. Dennis. 2004. Estimating population trend and process variation for PVA in the presence of sampling error. *Ecology* 85:923–929.
- Williams, C. K., A. R. Ives, and R. D. Applegate. 2003. Population dynamics across geographical ranges: time-series analyses of three small game species. *Ecology* 84:2654–2667.
- Wilson, H. B., B. E. Kendall, R. A. Fuller, D. A. Milton, and H. P. Possingham. 2011. Analyzing variability and the rate of decline of migratory shorebirds in Moreton Bay, Australia. *Conservation Biology* 25:758–766.

## SUPPLEMENTAL MATERIAL

### Ecological Archives

Appendices A and B and the Supplement are available online: <http://dx.doi.org/10.1890/14-2003.1.sm>