VIABLE POPULATION MONITORING: RISK-BASED POPULATION MONITORING FOR THREATENED AND ENDANGERED SPECIES WITH APPLICATION TO BULL TROUT, *SALVELINUS CONFLUENTUS*

by
David Franklin Staples

A dissertation submitted in partial fulfillment
of the requirements for the degree
of

Doctor of Philosophy

in

Fish and Wildlife Management

MONTANA STATE UNIVERSITY
Bozeman, Montana
March 2006

© COPYRIGHT

by

David Franklin Staples

2006

All Rights Reserved

APPROVAL

of a dissertation submitted by

David Franklin Staples

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies

Dr. Mark L. Taper

Approved for the Department of Ecology

Dr. David Roberts

Approved for the College of Graduate Studies

Dr. Joseph J. Fedock

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor Michigan 48106, to whom I have granted "the right to distribute my dissertation in and from microform along with the right to reproduce and distribute my abstract in any format in whole or in part."

David Franklin Staples

March 2006

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF FIGURES - CONTINUED

LIST OF FIGURES - CONTINUED

ABSTRACT

Population monitoring is a vital component for managing threatened and endangered (TE) species to demonstrate recovery, or alert managers if the status is deteriorating. Common methods for analyzing monitoring data, however, have poor power to detect changes in population status and do not directly address questions about population status as defined for threatened (likely to be endangered) or endangered (in danger of extinction) species.

Population viability analysis (PVA) methods are used to estimate the risk of decline for population, and have been recommended to reconcile short-term management actions with the ultimate long-term goal of preserving the species. The research presented herein is concerned with how to use PVA for monitoring population status, with a general focus on TE species and specific application to a bull trout (Salvelinus confluentus) population.

The bull trout population in the Flathead Lake and River System of NW Montana, USA provides an example for the motivation of VPM, and will serve as a test bed for developing and applying VPM for population management. Bull trout are listed as threatened under the Endangered Species Act, and the Flathead Lake population is of special concern because of a dramatic decline in the late 1980's.

Risk estimates are constrained by the ability to estimate model parameters from data. I develop methods to accommodate sampling error in population data and temporal correlations in population growth for count-based PVA models, and evaluate the effects of model extrapolation errors on risk estimates. Further, I present model structural adequacy analyses in which model evaluation criteria are not based on a model's fit to data, but on how well the model answers the scientific question of interest concerning the population's future status. This study suggests monitoring with a Gompertz density dependent model is likely the best available means for estimating average risk of decline for the bull trout population. Data on juvenile vital rates and abundances incorporated into a relatively simple demographic model could potentially enhance the ability to foresee imminent declines in adult abundances, though risk estimates can be detrimentally affected by uncertainty in sub-adult survival rates.

CHAPTER ONE

INTRODUCTION

Conservation biologists are concerned with the science that bears on preventing extinctions of natural populations and the subsequent loss of biodiversity. Population monitoring is a vital component for managing imperiled species. Monitoring is often specified in recovery plans for threatened and endangered (TE) species under the Endangered Species Act (ESA) (2000) to determine population status, and to evaluate population response to management actions. The primary goal for monitoring TE species is assess the risk of extinction. Ideally, a monitoring program should detect small changes in status, especially if the population's status deteriorates. For TE populations, which are by definition at high risk of extinction, quick implementation of remedial measures may be critical to prevent extinction. Such management actions can be costly and controversial. An ideal monitoring program should be able to demonstrate that recovery actions are effective in reducing the risk of extinction. The primary management goal is to improve the population's status so it can be removed from ESA protection. Thus, an ideal monitoring strategy should also establish when a population has achieved recovery.

Most monitoring methods fall short of the ideal. The lack of general, biologically relevant monitoring methods or recovery criteria is a major obstacle to effective monitoring for TE species (Gerber and Hatch 2002). This often leads to confusion and disagreement over population status and delisting decisions (Doremus and Pagel 2001). Compounding these problems is that common methods used for evaluating monitoring

data, such as statistical tests for changes in abundance (Thompson et al. 1998) or trend in population growth (Campbell et al. 2002), have poor power to detect changes in population status (Gerrodette 1987) with the quantity and quality of data usually available. Drastic drops in abundance can occur before the decline can be determined significant (Maxell 1999). Even then, inference is limited to the past performance of the population. In this sense, monitoring based on these methods is largely a retrospective process.

A retrospective monitoring program focuses on past performance of a population, and in many cases, is not very informative; a listed population presumably has experienced problems in the past. A more important issue, however, is that by limiting inference to past population performance, monitoring programs do not directly address questions of most interest to managers. By inclination and by statute, managers are primarily concerned with the future of a population, either in preventing a decline or in facilitating increase. Additionally, the terms "Endangered" and "Threatened" are defined in the ESA as 'in danger of extinction', and 'likely to be endangered' respectively. These definitions inherently refer to the future risk of decline for a population. Population viability analysis (PVA) is the general term for methods used to estimate future risk of decline for population, and has been recommended for recovery planning (Morris et al. 2002, Lande et al. 2003). The use of PVA, however, has often been based on single long-term predictions that make it difficult to satisfactorily demonstrate recovered status. Attempting to reduce risk of extinction over such long time frames may seem an intangible goal for population managers necessarily working in the short-term.

Projections 100 or more years in the future, while conforming to long-term management goals, are difficult to make with any amount of useful precision (Fieberg and Ellner 2000), often making recovery criteria (e.g.,(Schultz and Hammond 2003)) based on these analyses difficult to achieve. Further, these analyses do not incorporate future management actions that may be implemented if the population experiences problems in the future. Even if the population does achieve recovery goals, there is always the chance that unexpected threats to the population may emerge.

Goodman (2002) recommended a strategy to use PVA methods to reconcile short-term management actions with the ultimate long-term goal of preserving the species. In this strategy, short-term PVA predictions in periodic recovery plan reviews are used to guide management actions. Higher estimates of short-term risk would trigger more intense intervention. This requires commitment to data collection for monitoring and to interventions when the population is at high risk. If managers commit to such a strategy, however, the long-term probability of extinction can be greatly reduced (Goodman 2002). Thus, recovery is not a fixed set of interventions, but a fixed rule set on how interventions change in response to monitoring data and short-term PVA predictions associated with periodical recovery plan evaluations. The research presented herein is concerned with how to use PVA for monitoring population status, with a general focus on TE species and specific application to a bull trout (*Salvelinus confluentus*) population. In doing so, I outline a strategy to use short-term PVA predictions to evaluate populations status periodically and initiate conservation actions when needed. This monitoring protocol uses periodic short-term PVA estimates as the monitoring 'signal' (Hellawell 1991), and

allows population monitoring relative to status as defined for threatened and endangered species. I use the term 'viable population monitoring', or VPM, as a general term for a monitoring strategy based on population viability analysis predictions.

The bull trout population in the Flathead Lake and River System of NW Montana, USA provides an example population for developing and applying viable population monitoring for population management. Bull trout, a salmonid species with an adfluvial life history, is a native top-level predator species in the Flathead basin (Shepard et al. 1984). After maturing in Flathead Lake, adult bull trout migrate up to 200km to tributaries of the North and Middle Forks of the Flathead River to spawn; juveniles spend 2-4 years in natal tributaries before immigrating to the river and lake areas. Historically, Flathead Lake was connected with two major tributary rivers, the Flathead and Swan Rivers. Construction of the Bigfork Dam blocked fish movement into the Swan basin in 1902, while the South Fork of the Flathead River (one of 3 forks) was blocked by Hungry Horse Dam in 1953. Thus, spawning habitats for the Flathead Lake population is largely restricted to the Middle and North Forks of the Flathead River. Presently, both the Swan and South Fork basins contain relatively strong adfluvial bull trout populations, but the Flathead Lake population has recently experienced a dramatic decline.

Following species introductions in the early 20th century, the Flathead Lake community was relatively stable until the mid 1980s.An unexpected community shift in Flathead Lake around 1990 following introductions of the opossum shrimp, *Mysis relicta*, into four lakes of the Flathead drainage between 1968 and 1975led to a change in food resources in Flathead Lake and a drastic increase in lake trout (*Salvelinus namaycush*).

Lake trout, a top-level predator, were introduced in Flathead Lake in 1905, and were at relatively low abundance prior to 1990. The community shift is believed to have primarily caused the decline in the bull trout population (Deleray et al. 1999), but habitat changes in the natal tributaries may also have contributed to the observed decline (Rieman et al. 1997). The Flathead population has been well studied compared to other populations, yet effective data collection can often be difficult for such a wide-ranging population in rugged terrain, and important data gaps remain. Data series on yearly spawning effort, use of Flathead Lake, and juvenile densities show that qualitatively the population is expressing the full life history, however, there remains considerable uncertainty about the population's status.

A strong monitoring program is needed for the bull trout population because of its low abundance and uncertain future. A risk-based monitoring strategy can be effective in detecting changes in status, but is dependent on model-based predictions of the risk of decline, as well as data inputs to the model. Thus, population model structure will affect monitoring inference. For example, data on spawning nests, or redds, for the Flathead population can be used to parameterize a discrete-time exponential growth model, and used to calculate the probability of a decline in redd numbers. It is unknown how risk estimates from redd data correspond to the true risk of decline, and models with more complex structure, that incorporate age class abundances for example, may provide better estimates of risk for the Flathead population. Model structural adequacy analyses presented in this doctoral research provide a means to compare potential models based on how well they address management questions in this case where traditional model

selection methods cannot be used due to lack of data. Evaluation criteria are not based on a model's fit to data, but on how well the model answers the scientific question of interest concerning the population's future status (Lindsay 2004). These analyses help elucidate the level of model complexity necessary to adequately predict risk of decline for the bull trout population given assumptions about the population's dynamics.

The research reported in this doctoral dissertation addresses three major objectives. The first was to develop a risk-based monitoring strategy for Threatened and Endangered species. I describe a monitoring framework that will allow managers to determine how risk of decline is changing over time through comparisons of periodic short-term PVA predictions, allowing managers to evaluate population status as defined for TE species. This strategy addresses the need for a proactive monitoring program that will warn of increased risk or will detect when management actions are reducing risk.

The second major objective was to address technical issues concerning the application of PVA to actual ecological data. Viable population monitoring uses PVA models to predict the future status of a population, but in application, the accuracy and precision of risk predictions are affected by the ability to estimate model parameters from available data. Sampling error in data will increase uncertainty in predictions and can bias risk estimates. I present a method to partition sampling error variance from estimates of process variance for improved estimates of population risk with a density independent PVA. I have also collaborated on developing a method for estimating parameters for a Gompertz model of density dependent population growth from time-series of abundance data corrupted by sampling error (Dennis et al. 2006). Temporal

correlations in population growth, due to environmental or intrinsic factors, can further affect PVA in unpredictable ways. I examined the feasibility of accounting for temporal correlations in population growth for count-based PVA in the presence of sampling error in abundance data. Additionally, PVA predictions may be biased when models are used to predict population dynamics at unobserved densities. I explored the effects of non-linearities in density dependence on estimates of population risk from a Ricker model (Ricker 1954) of density dependent population growth. Finally, yearly redd counts are a major data series used for monitoring bull trout. I have collaborated on developing a bias correction and constructing confidence intervals for observer error in redd counts (Muhlfeld et al. 2006). These methods will enhance the ability to estimate risk for bull trout using current data series.

The third objective was to determine model complexity necessary for effective risk predictions for the Flathead bull trout population through model structural adequacy analyses. For risk-based monitoring, a model is needed that can effectively approximate the complex life history of the bull trout and deliver useful estimates of risk of decline. It would be beneficial to evaluate how risk predictions, and thus monitoring inference, are affected by model structure. I used a systems analysis approach (Coulman et al. 1972) to develop mechanistic simulation models of the Flathead population. Simulation models were then used to develop and evaluate the adequacy of simpler models for VPM. The model structural adequacy analyses also address the effects of uncertainty about juvenile bull trout population dynamics on monitoring-based predictions and model selection, and can be used to guide future data collection for increased understanding of the population.

This thesis is organized as a collection of independent manuscripts addressing the three major research objectives. Each chapter is an independent manuscript either published, in press, or submitted for publication (Note: the final chapter on model structural adequacy analyses will be submitted after March 21, 2006). At the end of this introduction, I also briefly discuss two additional papers I have collaborated on addressing monitoring issues. These are part of my overall body of work on monitoring recovery, but are not given dissertation status, as I am not the primary author.

Chapter 2, titled *Risk-Based Viable Population Monitoring*, details a framework to use PVA for proactive population monitoring. This chapter was published in *Conservation Biology* (Staples et al. 2005). In viable population monitoring, comparisons of periodic short-term predictions of the probability that the population will decline below a biologically relevant abundance threshold are used for inference about population status. With a risk-based monitoring strategy, critical changes in a population's status are more likely to be demonstrated before a devastating decline than with common monitoring methods allocating the same total monitoring effort. Conversely, decreases in estimated risk will demonstrate when management actions are reducing the risk of decline. Recovery is viewed not as a single evaluation of viability, but as a commitment to maintaining low short-term risk of decline. Because it focuses on relative changes in estimated risk, VPM also appears robust to model structure errors. Risk predictions with a simple exponential growth model can be effective monitoring indicators for population dynamics ranging from random walk to density dependence with stable, decreasing, or increasing equilibrium

Chapter 3, *Estimating Population Trend and Process Variation for PVA in the Presence of Sampling Error*, addresses the problem of sampling error in abundance data corrupting count-based population viability analyses. This chapter was published in *Ecology* (Staples et al. 2004). Time-series of population abundance estimates often are the only data available for calculating the risk of decline. With such data, a PVA can be performed using estimates of the mean population trend and the variance of the trend, the so-called process variation. Sampling error in the data, however, may bias estimators of process variation derived by simple methods. I developed a state-space model (De Valpine and Hastings 2002) for estimating trend, process variation, and sampling error from a single time-series based on a discrete-time model of density independent growth coupled with a model of the sampling process. This method provides essentially unbiased estimators of trend, process variation, and sampling error over a range of process variation/sampling error combinations. This method is useful for PVA applications that depend on accurate estimation of process variation from time-series data.

Chapter 4, Interaction of Sampling Error and Temporal Correlations in Population Growth Present an Intractable Problem for PVA Applications, examines the feasibility of accounting for temporal correlations in growth in PVA. This chapter has been submitted to Ecology. Correlations in the data from sampling error, intrinsic factors, or environmental conditions can detrimentally affect PVA risk predictions. Recently, it has been claimed that explicitly incorporating estimates of all between-year correlations in population growth into the process variance estimate, termed the 'long-term variance',

will exactly counterbalance effects of sampling error giving unbiased risk predictions (McNamara and Harding 2004). Alternatively, I show in Chapter 3 that sampling variability can be partitioned from process variation to greatly reduce bias in parameter estimates if environmental errors are assumed to be uncorrelated. I extend the state-space model of chapter 3 to accommodate autocorrelated environmental error, thus allowing long-term variance to be estimated with one additional parameter. In comparisons with simulated correlated abundance data, however, long-term variance estimators were imprecise and biased. Ignoring correlation while only accounting for sampling error generally resulted in better estimates of the variability in population growth. This study suggests sampling error can have pervasive effects that will complicate attempts to estimate temporal correlations in population growth, and supports the use of short-term risk predictions in PVA applications to minimize the effect of extrapolation errors.

Chapter 5, *Impact of Non-Linearities in Density Dependence Above the Range of the Data on Predictions of Population Extinction Risk*, explores the effect of model extrapolation on PVA risk estimates. This chapter has been accepted for publication in the *Journal for Nature Conservation* (Staples and Taper 2006). PVA predictions can be detrimentally affected when models are used to extrapolate population dynamics at unobserved densities. For example, there may be density dependent factors that only affect the population at very high, unobserved densities. A possible biological mechanism for this phenomenon could be territorial behavior placing a limit on the maximum density of fish in a given area. Such a mechanism may serve as a limit or upper bound to maximum population density even though it does not influence 'normal'

density dependent dynamics when the population is observed at densities near equilibrium. To explore the potential consequence of a regulation mechanism affecting the population only at high densities, we placed a reflecting upper bound in simulations of a Ricker model (Ricker 1954). The upper bound prevented the population from attaining unrealistically high densities (that would have directly resulted in extinction) and so this modification to the dynamics decreased the probability of extinction, all other things being equal. This research supports the use of shorter prediction timeframes and population thresholds near the range of observed data in VPM to minimize effects of model extrapolation errors on monitoring inference.

Chapter 6, *Model Structural Adequacy: Developing Bull Trout Population Models for Viable Population Monitoring*, details an inductive model selection method using detailed simulation models. This chapter will be submitted to *Conservation Biology*. Risk-based monitoring is dependent on model-based risk predictions for inference to population status, thus a model's explanatory power about past data is ancillary to its ability to predict short-term risk of decline. Model structural adequacy analyses were used for insight into the level of model structural complexity needed for adequate risk predictions for a population with a complex, age-structured life history like adfluvial bull trout. Because the actual population dynamics for the population are unknown, I developed a range of mechanistic age-structured simulation models of an adfluvial bull trout population using available information and with input from Flathead biologists. These models incorporate hypotheses about juvenile dynamics and population vital rates, and represent the 'true' population growth process for this study. Using data simulated

with the process models, potential monitoring models are evaluated on whether they can usefully approximate the risk of decline for the more complex population dynamics. Monitoring models are supported according to how well they estimate risk for the different process models. Monitoring models included 3 count-based models that could be used with existing bull trout data: a density independent model (Dennis et al. 1991) in addition to the Ricker (Ricker 1954) and Gompertz (Dennis and Taper 1994) models of density dependent growth. Two demographic models were also evaluated for monitoring effectiveness: a 3-stage population model (Juvenile, sub-adult, adult), and an age-stage population model (age classes for age 0-4, sub-adult and adult stages). The model structural adequacy analyses will help evaluate the likelihood that simple models are adequate for monitoring (i.e. they approximate risk adequately for the various process models), or whether it is possible to improve predictions with demographic models. The Gompertz model was the only model that gave useful estimates of risk in all population simulations, suggesting it is likely the best available means for estimating average risk of decline for the population. Our study suggests further data collection toward developing an age-stage model of the population could potentially enhance the ability to foresee declines in adult abundances, but the reliability of risk predictions with demographic models could depend on survival rates for sub-adults, a very difficult vital rate to estimate for this population.

Population growth rates can be affected by abundance, and in such cases use of density independent models will be inappropriate for making risk predictions. Estimation of parameters for a density dependent growth model, however, has been constrained by

sampling error in abundance estimates. I have collaborated on developing a method to estimate parameters for a density dependent Gompertz model from time-series abundance estimates. This paper, titled *Estimating Density Dependence, Process Noise, and Observation Error*, has been accepted for publication in *Ecological Monographs* (Dennis et al. 2006). We developed a state-space model (De Valpine and Hastings 2002) that allows estimation of the strength of density dependence, process variation, and observer error from time-series of abundance estimates using common statistical software. The model contains density dependence in the form of a Gompertz-type population growth function with lognormal process variation in addition to lognormal sampling error in estimates of population abundance. In application, this method will make it easier to calculate risk estimates for populations exhibiting density dependent growth.

Redd counts are a primary means of monitoring bull trout populations, yet few studies have evaluated errors in observed counts relative to the true number of redds present. I have collaborated on developing a method to correct for bias and construct confidence intervals for observed redd counts. This paper, titled *Observer Error Structure in Bull Trout Redd Counts in Montana Streams: Implications for Inference on True Redd Numbers*, has been accepted for publication in *Transactions of the American Fisheries Society* (Muhlfeld et al. 2006). Biologists from Montana Fish, Wildlife, & Parks assessed the variability in redd counts among experienced observers in two tributaries of the Swan River in NW Montana (the Swan River flows into Flathead Lake). Observers made both errors of omission and false identification; we combined them using a binomial distribution of true redds detected and a Poisson distribution of false counts.

At low redd abundances, this error structure leads to overestimates of true redd numbers. Conversely, at high redd abundances, observed redd counts will underestimate the number of true redds present. We used this model to correct for bias and construct confidence intervals for historical redd count data. This method will improve inference to population status from redd count data, and incorporating estimates of observer error into the state-space Gompertz model (Dennis et al. 2006) may improve the reliability of risk predictions with the state-space Gompertz model (Hakoyama and Iwasa 2000).

Literature Cited

2000. U.S. Endangered Species Act of 1973. *in* 16th U.S. Code.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: A practical information-theoretic approach. Springer-Verlag, New York.

Campbell, S. P., J. A. Clark, L. H. Crampton, A. D. Guerry, L. T. Hatch, P. R. Hosseini, J. J. Lawler, and R. J. O'Connor. 2002. An assessment of monitoring efforts in endangered species recovery plans. Ecological Applications **12**:674-681.

Coulman, G. A., R. L. Tummala, and S. R. Reice. 1972. Population Modeling - Systems Approach. Science **175**:518-&.

De Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise an observation error. Ecological Monographs **72**:57-76.

Deleray, M., L. Knotek, S. Rumsey, and T. Weaver. 1999. Flathead Lake and River System Fisheries Status Report. F-78-R-1 through 5, Element 1, Project 1 and 2, SBAS Project No. 3131., Montana Fish, Wildlife, & Parks, Kalispell, Montana.

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:In Press.

Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. Ecological Monographs **64**:205-224.

Doremus, H., and J. E. Pagel. 2001. Why listing may be forever: Perspectives on delisting under the US Endangered Species Act. Conservation Biology **15**:1258-1268.

Fieberg, J., and S. P. Ellner. 2000. When is it meaningful to estimate an extinction probability? Ecology **81**:2040-2047.

Gerber, L. R., and L. T. Hatch. 2002. Are we recovering? An evaluation of recovery criteria under the US Endangered Species Act. Ecological Applications **12**:668-673.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology **68**:1364-1372.

Goodman, D. 2002. Predictive Bayesian population viability analysis: A logic for listing criteria, delisting criteria, and recovery plans. *in* S. R. Beissinger and D. R. McCullough, editors. Population Viability Analysis. The University of Chicago Press, Chicago.

Hellawell, J. M. 1991. Development of a rationale for monitoring. *in* F. B. Goldsmith, editor. Monitoring for Conservation and Ecology. Chapman and Hall, New York, NY.

Lande, R., S. Engen, and B.-E. Saether. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, Inc., New York, NY.

Lindsay, B. G. 2004. Statistical Distances as Loss Functions in Assissing Model Adequacy. *in* M. L. Taper and S. R. Lele, editors. The Nature of Scientific Evidence: Empirical, Statistical, and Philosophical Considerations. The University of Chicago Press.

Maxell, B. A. 1999. A power analysis on the monitoring of bull trout stocks using redd counts. North American Journal of Fisheries Management **19**:860-866.

McNamara, J. M., and K. C. Harding. 2004. Measurement error and estimates of population extinction risk. Ecology Letters **7**:16-20.

Morris, W. F., P. L. Bloch, B. R. Hudgens, L. C. Moyle, and J. R. Stinchcombe. 2002. Population viability analysis in endangered species recovery plans: Past use and future improvements. Ecological Applications **12**:708-712.

Muhlfeld, C. C., M. L. Taper, D. F. Staples, and B. B. Shepard. 2006. Observer error structure in bull trout redd counts in Montana streams: Implications for inference on true redd numbers. Transactions of the American Fisheries Society:In Press.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada **11**:559-623.

Rieman, B. E., D. C. Lee, and R. F. Thurow. 1997. Distribution, status, and likely future trends of bull trout within the Columbia River and Klamath River basins. North American Journal of Fisheries Management **17**:1111-1125.

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.

Shepard, B. B., K. L. Pratt, and P. J. Graham. 1984. Life histories of westslope cutthroat and bull trout in the Upper Flathead River Basin, Montana. Final Report for EPA

Under Contract Number R008224-01-5. Montana Department of Fish, Wildlife and Parks, Helena.

Staples, D. F., and M. L. Taper. 2006. Impact of non-linearities in density dependence above the range of the data on predictions of population extinction risk. Journal of Nature Conservation:In Press.

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.

Staples, D. F., M. L. Taper, and B. B. Shepard. 2005. Risk-based viable population monitoring. Conservation Biology **19**:1908-1916.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, Inc., San Diego, California, USA.

CHAPTER TWO

RISK-BASED VIABLE POPULATION MONITORING

Abstract

We describe risk-based viable population monitoring in which the monitoring indicator is a yearly prediction of the probability that, within a given timeframe, the population abundance will decline below a pre-specified level. Common abundance-based monitoring strategies usually have low power to detect declines in threatened and endangered species and are largely reactive to declines. Comparisons of the population's estimated risk of decline over time will help determine status in a more defensible manner than current monitoring methods. Monitoring risk is a more proactive approach; critical changes in the population's status are more likely to be demonstrated before a devastating decline than with abundance-based monitoring methods. In this framework, recovery is defined not as a single evaluation of long-term viability, but as maintaining low risk of decline for the next several generations. Effects of errors in risk prediction techniques are mitigated through shorter prediction intervals, setting threshold abundances near current abundance, and explicitly incorporating uncertainty in risk estimates. Viable population monitoring also intrinsically adjusts monitoring effort relative to the population's true status and exhibits considerable robustness to model misspecification. We present simulations showing risk predictions made with a simple exponential growth model can be effective monitoring indicators for population dynamics ranging from random walk to density dependence with stable, decreasing, or increasing

equilibrium. In analyses of time-series data for five species, risk-based monitoring warned of future declines and demonstrated secure status more effectively than statistical tests for trend.

Introduction

Correctly evaluating the status of threatened and endangered populations is critical to detecting population declines that will further endanger the population and to verifying effectiveness of conservation actions. Monitoring to detect changes in status is frequently specified in recovery plans (Morris et al. 2002). Unfortunately, detecting changes in status with commonly used monitoring strategies such as monitoring species abundance and statistical tests (Thompson et al. 1998) for population trend can be difficult (Holmes & York 2003). Because of high variability in ecological data arising both from natural variation in population growth and from measurement errors in population estimates, a biologically devastating decline could occur before there is a reasonable chance to detect the decline. For example, at observed variation levels common in bull trout (*Salvelinus confluentus*) data, a population could decline by over 60% before the power to significantly detect such a decline reached 0.80 (Maxell 1999). For species that have presumably already declined to low levels, an inability to quickly detect further declines in population abundance is a serious problem. Ideally, monitoring would reliably detect a small (though biologically relevant) decline in a short amount of time; however, this goal is not realistic with current methods for most populations. Monitoring species with abundance-based methods also leaves managers in a reactive

posture because a decline might be demonstrated only after it has seriously diminished the population, even though the demographic problems that caused the decline occurred many years before the decline was ”significant.”

Natural variation inherent in population data also complicates the task of determining whether a population is recovering following actions designed to promote recovery. Conservation actions are usually costly and may be controversial, but many years of monitoring will often be required before the effects of these actions can be evaluated clearly through population abundance estimates. Compounding these problems is a common lack of biologically based recovery criteria in population monitoring plans. Consequently, populations are rarely de-listed (unless they go extinct) because there is no agreed upon method for evaluating population status and levels for recovery are rarely specified. Even when a species has rebounded and might be considered recovered, the lack of quantifiable recovery criteria applied to a statistically rigorous monitoring protocol prevents agencies from concluding the population is secure.

There have been recent calls to incorporate population viability analysis (PVA) in endangered species recovery plans (Morris et al. 2002; Lande et al. 2003). With PVA methods, it is possible to predict future population abundance, the time to extinction (or some other lower threshold abundance), the probability of reaching a lower threshold, or of reaching a lower threshold within a specified amount of time (Dennis et al. 1991; Morris & Doak 2002). Morris et al. (2002) give three useful functions of PVA in recovery planning. First, PVA methods can help identify particular life stages or demographic processes that should be targeted to promote recovery (Caswell 2000;

Caswell 2001). Model analyses can both focus management actions and help guide planning of data collection to improve predictive accuracy. Second, PVA can serve as a means to synthesize data for assessing recovery success. For example, monitoring data on population abundance and environmental correlates can be combined with demographic models for PVA projections to give a more comprehensive picture of population status (Morris et al. 2002; Lande et al. 2003). Finally, PVA can provide good estimates of relative risks to populations with as little as 10 years of data (McCarthy et al 2003). Measures of risk can be successfully compared among populations to provide a relative measure of how urgently recovery actions need to be implemented in specific populations, or risk can be compared for the same population under competing conservation strategies (Ellner et al. 2002).

This documented accuracy of relative risk assessments suggests that an individual population could also be compared to itself over time for an indication of how the population's risk of decline is changing over time. Monitoring risk in this fashion is also more concordant with the definitions of endangered ('in danger of extinction') and threatened ('likely to become endangered in foreseeable future') in the Endangered Species Act (U.S.ESA of 1973 [as in U.S. Code 2000]). There is also the opportunity to develop biologically relevant and adaptable recovery criteria for threatened and endangered species with a risk-based monitoring plan. As an alternative to abundance-based monitoring methods, we propose a risk-based viable population monitoring (VPM) protocol, in which yearly risk predictions are used as the monitoring indicator.

Threatened and endangered populations are likely to be at low abundance levels; consequently, two important management goals are to prevent further declines to unsustainable levels and facilitate increases to a secure abundance level. Given these objectives, an estimate of the risk of the population declining below a biologically relevant threshold is a useful monitoring indicator that consolidates relevant information about the population's status. For monitoring purposes, risk is defined as the probability of population abundance declining below a lower threshold within a given time frame. Increases in this risk over time indicate that the population's status is deteriorating and remedial actions should be implemented. Conversely, decreases in risk demonstrate that management actions are reducing the likelihood of a catastrophic decline in the near future. A population could be considered secure after a period (e.g. 10 years) of consistently low risk; however, in the VPM framework, recovery is not a one-time evaluation but the *maintenance* of low risk of a serious population decline. By comparing predicted risk over time, VPM can provide more defensible status evaluations and help facilitate better communication between agencies as to what constitutes a 'healthy' or 'threatened' population.

Methods

A common situation in monitoring is when the target population is undergoing so-called random walk population growth. In this case, the long-term population growth rate is zero but observed growth could be positive or negative in any given year due to environmental variation. At any time, a series of bad years (or even one *really* bad year)

could cause a large drop in abundance. Random walk population growth will, by definition, tend to give trend estimates that are not different from zero. This makes monitoring based on trends in abundance alone unreliable. If the population has a significant negative or positive growth rate, it will more than likely be apparent with a few years of data. To evaluate how a VPM strategy will perform for such a population, we simulated a random walk process for 15 years with an initial abundance of 1000. We then estimated the risk of the population abundance declining below 500 within 5 years (denoted PLT5) or within 10 years (PLT10) with an exponential growth population model (Dennis et al 1991; Morris and Doak 2002). The random walk simulation continued for 50 additional time-steps with updated estimates of PLT5 and PLT10 calculated at each time-step with the entire data series up to that point (i.e. risk estimated with 15 years of data, 16 years of data, 17 years of data, and so on until the end of the series). We compared the population's abundance trajectory with the series of estimated risk metrics.

The risk calculations in VPM should be based on the best available model of the target population's dynamics, however, in reality, the true population dynamics are unknown and it can be difficult to discern between competing models (e.g. density independent or dependent growth) in most data series (Zeng et al. 1998). For this reason, it is instructive to evaluate VPM when the incorrect model is used for analysis (i.e. when the actual population growth is density dependent, but risk is estimated with an exponential growth model). We used three density dependent models to generate time-series that were analyzed with VPM using an exponential growth PVA. The first model

was a population with a stable equilibrium slightly above the lower threshold, a very different process but potentially indiscernible from random walk population growth in actual data. The second was a population with a slowly declining (1% per year) carrying capacity, such as could result from gradual degradation of habitat over time. The third scenario was an increasing (3% per year) carrying capacity, which should test whether VPM can detect when a population is on the path to recovery. We compared the estimated risk from an exponential growth analysis with the true risk calculated with simulations of the density dependent model.

To see whether a risk-based monitoring strategy can warn of potential declines and demonstrate population security more reliably than testing for a significant trend in abundance, we also applied VPM in a post hoc evaluation of four historical time-series of population data: 1965-1980 surveys of California Condor (*Gymnogyps californianus*; Snyder & Johnson 1985), 1969-1989 censuses of Puerto Rican Parrot (*Amazona vittata*; Dennis et al. 1991), 1938-1993 winter counts of Whooping Crane (*Grus Americana*; USFWS 1994), and 1959-1997 adult female estimates for grizzly bear (*Ursus arctos horribilis*; Morris & Doak 2002). To prevent bias and low power resulting from sampling errors in the data, parameters for risk estimates and tests for trend were calculated using the exponential growth PVA method presented in Staples et al. (2004). Exponential growth PVA analyses and trend estimation for these data sets have been discussed previously (Dennis et al. 1991; Staples et al. 2004). For the analysis of these data sets, the lower threshold was arbitrarily set at one half the initial population estimates. Because the grizzly data are running 3-year sums, the prediction interval for

that data series was set at 20 years; for all others it was set at 5 years. All significance tests for trend were two-sided with test size at 0.10. We calculated risk estimates and tests for significant trend yearly starting after the first 5 years of observations.

We also analyzed a contemporary data set of index counts of spawning nests, or redds, for bull trout (*Salvelinus confluentus*) from the Flathead River drainage in northwest Montana (U.S.A.). Though redd count data have been collected on this population consistently since 1980, major community changes in Flathead Lake around 1990 caused an explosion in lake trout (*Salvelinus namaycush*) abundance that altered the dynamics of this bull trout population. These changes likely precipitated a large crash in the bull trout population observed between 1991 and 1992 (Deleray et al. 1999). Because PVA predictions with an exponential growth model assume a constant process (i.e., the true nature of population growth remains the same) and no catastrophic environmental events, only data since the system stabilized (approximately 1992) were deemed appropriate for the VPM analysis.

The bull trout data are not direct counts of individuals but of spawning sites, therefore it would be useful to define the analysis in terms of a biologically significant adult abundance. In this case, we set the lower threshold for adult abundance at 1000 spawning adults based on the estimated effective population size necessary to maintain genetic diversity in the populations (Rieman & Allendorf 2001). Assuming 3.2 fish/redd (Fraley & Shepard 1989), and the total redd to index redd count relationship observed by Deleray et al. (1999), 1000 spawning adults would correspond to approximately 125 index redd counts. (Much of the data for these conversions were collected before the

population crash and it is unknown whether these conversions hold for the current population.) As in the above data analysis, the estimated probability of redd counts declining below the lower threshold within 5 years was calculated yearly starting with 5 years of observations. Changes in probability of redd counts declining below 125 should then reflect the changes in the risk of declining below 1000 spawning adults subject to the accuracy of the conversions.

Results

In random walk simulations, VPM risk estimates effectively at warned of future declines in population abundance; because risk estimates are a probability of future decline, however, a high risk estimate does not necessitate a future decline. In an example of simulated random walk growth, the initial risk estimates were high because of negative trend and high variance estimates (Fig. 1). After a period of steady increases in abundance, the risk estimate dropped; however, after 2 years of decline, there was a sharp increase in estimated risk. This increase in estimated risk was due to lower abundance and a downturn in the estimated trend. As time passed, the estimated risk declined because of higher abundance and more stable estimates of both trend and process variation. The only trend estimate that was close to being significantly <0 was the first estimate, which was calculated with only five observations ($p = 0.08$; one-sided test). Trend estimates converged on the true value of 0, as expected, indicating tests for a significant trend would not be very informative for monitoring a population with random walk dynamics. Although risk estimates calculated with the 10-year prediction interval

were higher, trends in risk estimates from the 5- and 10-year prediction intervals were similar.

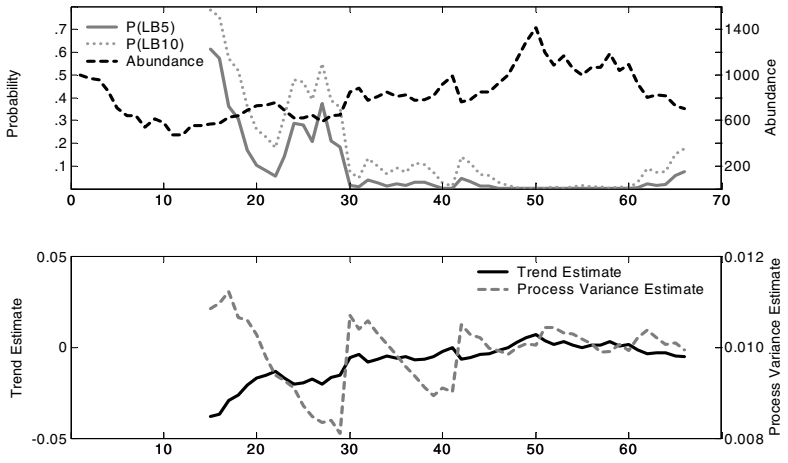


Figure 1. (top) Simulated random walk population growth with predicted probabilities for declining below a threshold value within 5 years (PLT5) and 10 years (PLT10) calculated with an exponential growth model. (bottom) Trend and process variance estimates used to calculate PLT5 and PLT10.

For density-dependent simulations with stable equilibrium abundance, VPM analysis with the exponential growth model showed biased high risk estimates for abundance levels below equilibrium and biased low risk estimates for abundances above equilibrium (Fig. 2). These biases were due to the discrepancy between the exponential growth analysis model and the true density-dependent process. An exponential growth model does not account for increases in the growth rate as the population gets further below equilibrium nor overcompensatory effects that can lead to severe crashes directly from abundances above equilibrium. We note that high true risk at abundances above equilibrium is an idiosyncrasy of the Ricker model used for the simulations (Ripa &

Lundberg 2000); other types of density dependent population growth do not necessarily have high risk of decline directly from high abundances.

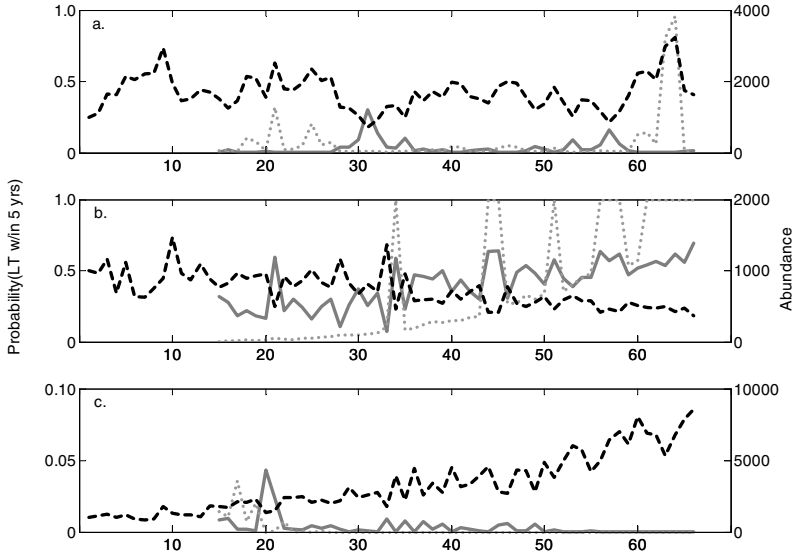


Figure 2. Estimated probability of a population declining below 500 within 5 years calculated with an exponential growth model from simulations of density-dependent population growth with equilibrium value (a) stable, (b) declining 1% per year, and (c) increasing 3% per year. Dashed line is population abundance; dotted line is true probability of declining below lower threshold (LT); solid line is estimated probability of declining below 500 within 5 years.

When the equilibrium abundance was declining over time, risk estimates generally increased over time, corresponding to decreases in abundance. These risk estimates were stochastic but never < 0.10 for consecutive estimates. In contrast, using classical trend analysis, there were no significant tests for negative trend for the entire series (lowest *p*-value = 0.32; one-sided test). In situations when the equilibrium was increasing over time, VPM generally showed low risk with estimated risk usually < 0.01 (Fig 2c). Risk increased to 0.05 once when the abundance declined at time 20. Again, no classical tests showed a trend >0 despite the long data series (lowest *p* value = 0.21; one-

sided test). In contrast to the results with classical methods, this population could be considered secure according to VPM when, for example, the risk of decline was below 0.01 for 5 consecutive years. However, true recovery should be viewed also as maintaining low risk of decline, not just a one-time evaluation of status.

In analyses of historical data sets, VPM warned of future declines and demonstrated secure status before such changes in abundance could show even a marginally significant trend (Fig. 3). Increased risk occurred after only 6 years of observations for the sharply declining California Condor population, and after 9 years of observations, the estimated probability of declining below the threshold was 0.80. In contrast, the only significant trend estimate made in the Condor series occurred after 13 years of observations ($p < 0.10$; two-sided test). For populations that were increasing, as was the case for the Puerto Rican Parrot and Whooping Crane, the VPM analysis demonstrated low risk before any significant trends in abundance were detected. For the Puerto Rican Parrot, there was a general decline in estimated risk from an initially high estimate. Predicted risk was quite low after 10 years of observations and near 0 after 14 years; there was, however, no significant trend estimate for the entire data series. In the case of the Whooping Crane, the risk estimates initially were 0 and only increased above 0.01 for 1 year. There was no significant trend in Whooping Crane abundance until there were 41 years of observations. The grizzly bear analyses showed significantly negative trend estimates for the years 1973-1980. This period was preceded by higher and more volatile risk estimates for 1967-1975. Grizzly bear abundance increased after 1980 and

subsequent risk estimates were low; there were, however, no significant trend estimates after 1980.

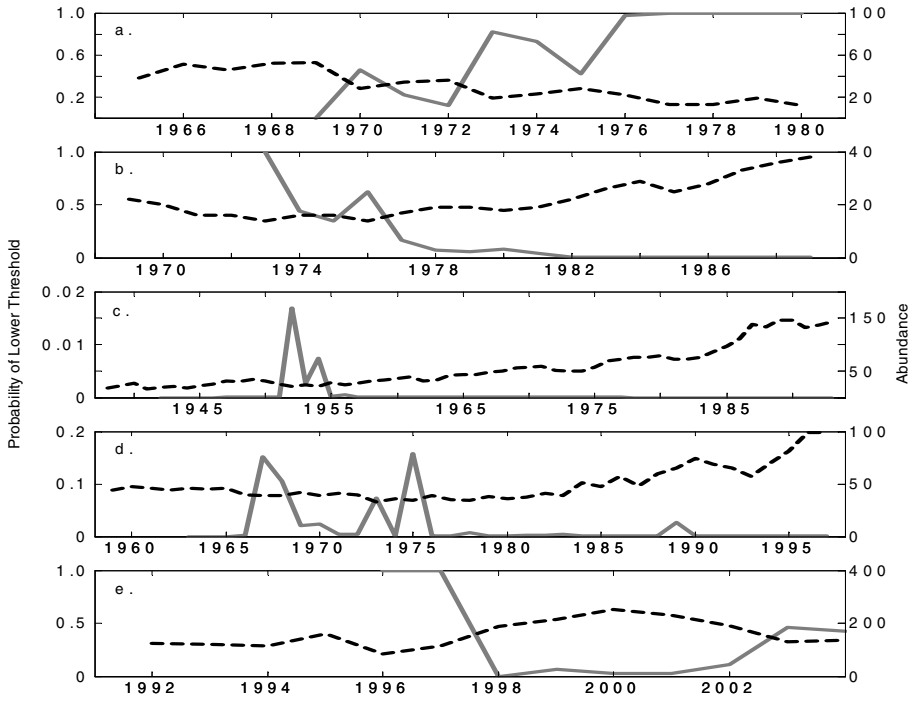


Figure 3 Viable population monitoring analysis for (a) California Condor, (b) Puerto Rican Parrot, (c) Whooping Crane, (d) grizzly bear, and (e) bull trout. Dashed line is estimated population abundance except for bull trout (e) in which case it is index redd counts. Solid line is probability of the population declining below lower threshold within 5 years (a,b,c,e) or 20 years (d). Threshold is one half the initial population estimates for a, b, c, and d; threshold is 125 redds for (e).

Between 1992 and 1996, counts of bull trout redds declined below the threshold; therefore the initial estimated risk of going below 125 redds was 1 (Fig. 3). Redd counts for the next 4 years, however, sharply increased and, by 2000, they were double the 1992 level. Consequently the estimated probability of decline below the threshold decreased to nearly 0. Unfortunately, redd counts have steadily declined since 2000, which has

resulted in increasing risk estimates. The steep increase in estimated risk is evidence that the likelihood of declining below 1000 total spawners increased since 2001 despite the uncertainty about the true risk at the present. It is too early to come to any long-term conclusions about the population because of relatively few years of observations and the coarse prediction method used for this analysis (i.e. an exponential growth model was used to predict adult numbers in an age-structured population with a long pre-reproductive period).

Discussion

Because VPM focuses on monitoring risk, it offers several improvements over traditional strategies. Management of threatened and endangered populations can be more proactive in addressing problems. Increases in risk should be evident and remedial actions could be initiated before large declines in abundance. For example, VPM showed large increases in risk for the California Condor data series well before any precise estimates of trend would have been available. In a similar manner, VPM can demonstrate improvements in population status before traditional methods. For the Puerto Rican Parrot analyses, risk estimates were zero for the last 9 years but there were no significant trend estimates for the entire series.

A risk-based monitoring strategy has an intrinsic mechanism to adjust monitoring effort in relation to the population's true status. More precise predictive ability lowers the calculated risk. Conversely, uncertainty in predicted abundance results in higher estimated risk. A stable population well above the lower bound may have high estimated

risk entirely because little is known about it (i.e. predictions about future abundance are very wide). As more is learned, population abundance predictions will be more precise, thereby reducing the estimated risk without the population's true status changing. Populations near the lower bound will likely require more precise abundance predictions to demonstrate secure status. In contrast, minimal monitoring effort will be required for populations far above the lower bound because less precise predictions will still lead to low risk estimates.

Abundance thresholds are used to lessen prediction errors and can lead to more robust monitoring in face of possible catastrophic declines. Thresholds set near observed abundance levels help prevent errors that would arise from large increases in demographic stochasticity (Lande 1993) or changes in population dynamics such as the Allee effect (Dennis 1989; Courchamp et al. 1999) that may occur at abundances near extinction. The abundance threshold may also be modified to incorporate the effects of catastrophes. It would be reasonable to raise the threshold to the level where the population would still be viable after a likely catastrophe (see Ralls et al. 1996 for an example).

Shorter prediction intervals decrease prediction errors that are problematic in PVA techniques (Ludwig 1999) and lead to bias in risk estimates. Although long-term (e.g. 100-year) predictions may not be meaningful, it might be possible to reliably estimate short-term risk (Fieberg & Ellner 2000). The ideal prediction interval will likely depend on the monitored population. Our analyses of the Whooping Crane data suggested a 5-year prediction interval was not very sensitive for monitoring purposes. A

prediction interval of 10 years or more is probably warranted for longer-lived species such as the Whooping Crane. An even better strategy may be to monitor several prediction intervals concurrently with interval selection based on the generation length of the target species. One possibility is to monitor the probability of declining below the threshold abundance within one, two, and five generations. Additionally, the nature of the data itself may necessitate a longer prediction interval. The grizzly bear data are three-year sums of adult females with cubs so it is unlikely that a 5-year prediction interval will be very sensitive to population changes, especially considering the long lifespan and generation time of the grizzly bear.

For many populations, data are inadequate to discern between competing population models (e.g. differentiate between density independent and dependent growth). Comparison of risk estimates from an exponential growth model may provide a practical method for assessing the changes in risk for a variety of populations. In our simulations, predictions made with a simple exponential growth model appeared to be useful for monitoring populations with a range of underlying growth processes. A density independent approximation correctly or conservatively estimates the risk of severe decline for a broad range of density dependent processes and is most accurate when it matters most to conservation managers, i.e. when a population is fluctuating near its carrying capacity, recovering slowly, or declining toward extinction (Sabo et al. 2004). Additionally, any effect of biases in risk estimates on the inference to population status is reduced because VPM does not rely on a single risk estimate, but the comparison of risk estimates over time.

The usefulness of this approximation depends on how well exponential growth can characterize actual population growth patterns. When the equilibrium for a density dependent population is increasing, for example, resulting abundance trajectories will have similar behavior as an exponential growth process (i.e. observed population abundance is variable around an increasing mean abundance). In our simulations, the estimated trend from the exponential growth model converged on the true rate of change for both increasing and decreasing carrying capacities. Similarly, stable density dependent growth can behave like a random walk if observed population abundance varies around the same mean abundance. In such cases, the exponential growth approximation of the distribution of population abundance can be useful, especially for shorter prediction intervals. Again, the approximation will overestimate risk for abundances below equilibrium. It may be reasonable, however, to view a density-dependent population close to a critical lower threshold as at higher risk because declining abundance could be a signal of deteriorating habitat reducing the equilibrium abundance. For strong overcompensatory density dependent growth, however, where overshoot is possible, an exponential growth analysis will underestimate risk for a population above equilibrium.

Certainly in some monitoring situations, an exponential growth analysis, as presented above, will be of little practical use. This may be simply due to the lack of a suitable data time-series with which to estimate the risk. If there are no such data, it will likely take up to a decade or more before there is enough data for reasonably precise abundance predictions (Morris & Doak, 2002). Even when there are suitable data, the

exponential growth approximation of the true biological process may be too inaccurate for useful predictions, especially if current abundance is close to the threshold. In such cases it is possible to construct more detailed demographic models for more precise predictions in a VPM analysis. Models that are more detailed may produce relatively reliable monitoring indicators more quickly than waiting for enough yearly observations so that the exponential growth analysis can be reasonably applied. Estimating demographic parameters for such models, however, may be expensive. The bull trout population analyzed above is an example of the latter situation in that current redd counts are close to the threshold and the exponential growth approximations are imprecise due to fluctuations in age-class abundances.

Bull trout have been monitored for over 20 years in the Flathead with index redd counts, and although the above analysis demonstrates an uncertain future for the population, the process by which redd counts change over time is not well approximated by an exponential growth process (i.e. this years redds are not a direct function of last years' redds). There is the possibility for age-class interactions to generate cycles in adult abundances that confound the use of an exponential growth model. This population has a maturation time of approximately 6 years and the potential for density dependent interactions among juvenile cohorts that can cause large fluctuations in age-class abundances (Paul et al. 2000). We have found that risk predictions from an exponential growth model based on adult-only data are biased high if the population experiences severe age class fluctuations (D.F.S. & M.L.T., unpublished data). Holmes (2001) simulated adult salmon abundances with an age-based model and also found that using

adult abundance counts as a surrogate for total population size led to severe bias in estimators of trend and process variation, even with no sampling error. It is likely that structural differences between the true biological process and predictive model results in overestimates of risk in the VPM analysis above. This implies abundance levels would have to be very high relative to the threshold for this redd-based risk analysis with the exponential growth model to show low risk. We are working to improve prediction models for this population, and remedial actions are being implemented to increase bull trout population abundance.

To decrease controversy about population status and facilitate communication among management agencies, establishing *a priori* goals for recovery is a critical feature of a monitoring strategy. Schultz and Hammond (2003) recommended a recovery goal for Fender's blue butterfly (*Icarica icarioides fenderi*) based on the minimum population growth rate necessary for 100-year persistence. In this case, high variation in population growth and a long prediction interval led to the conclusion that a relatively high population growth rate would be necessary for persistence (minimum average growth rate of $\lambda = 1.55$ over 10 years). This goal may be difficult to achieve in practice, especially if there are density-dependent interactions within the population. Alternatively, it may be constructive to view recovery not as a one-time verification of long term persistence, but as maintaining low risk of a biologically relevant decline within the next few generations. Very long prediction intervals compound uncertainty and require the assumption that environmental conditions and the growth process itself remain the same far into the future, making it very difficult to verify long-term viability satisfactorily.

Recovery criteria based on maintaining low risk can be a more tangible and achievable goal for managers of threatened and endangered species. Several broad categories of recovery criteria have been used for threatened and endangered species: population size, population trend, habitat fragmentation, demographic rates, and legal or policy criteria (Gerber & Hatch 2002). There is also concern that monitoring tasks specified in recovery plans do not adequately address specific threats affecting populations (Campbell et al. 2002) or recovery criteria are poorly linked to species biology (Schultz & Hammond 2003). A risk-based strategy can consolidate these broad categories of recovery criteria into a single, biologically relevant criterion with flexibility to assess specific threats relative to their effect on the population's risk of decline. An added advantage is that risk-based recovery criteria can adapt to unanticipated threats to a population after de-listing.

A useful characteristic in VPM is that the measure of status is congruent with the definitions of 'threatened' and 'endangered' under ESA and potentially could lead to less controversy about de-listing species. The primary difficulty in de-listing is that there is no regulatory mechanism to address threats to a population except for the ESA; a population is only de-listed if there are protections for the population from all known threats under other existing laws (Doremus and Pagel 2001). It may be possible to develop a regulatory mechanism that provides a means to address and *adapt* to threats after de-listing with risk-based recovery criteria. In VPM, recovery is defined as perpetually maintaining low risk of decline for the population, or in other words, a low probability that the population would qualify as threatened or endangered. After de-

listing, management agencies could manage a population as it sees fit as long as the requirement of low risk of decline is satisfied.

Unlike conventional monitoring methods, VPM answers a fundamentally different question. The goal of VPM is to estimate how the population's status is changing (i.e. toward recovery or extinction) given current conditions instead of proving a past decline was significant. The significance value for classical tests for trend is strictly on what has happened in the past. Conversely, VPM uses prediction intervals about future abundance for inference on population status. Risk-based monitoring should show increases in risk for declining populations to signal the need to implement conservation actions. To establish recovery, low risk of decline to untenable levels must be demonstrated and maintained. Further work is needed to determine exactly what constitutes a critical increase in risk or a stable, secure population through viable population monitoring. Risk-based monitoring can be a valuable tool for managing threatened and endangered species because it is proactive in detecting potential population declines and can show how a population's status is changing relative to consistent, biologically relevant recovery criteria.

Acknowledgements: We thank fishery biologists from Montana Fish, Wildlife, & Parks, U.S. Fish and Wildlife Service, and the Confederated Salish and Kootenai tribes for providing data and helpful discussions about bull trout life history. We also thank D. Goodman, L. Gerber, and two anonymous reviewers for their constructive comments. Funding for this work provided by Montana Fish, Wildlife, & Parks (grant Z0862) and the Northwest Power and Conservation Council (grant PTMM11).

Literature Cited

Campbell, S. P., J. A. Clark, L. H. Crampton, A. D. Guerry, L. T. Hatch, P. R. Hosseini, J. J. Lawler, and R. J. O'Connor. 2002. An assessment of monitoring efforts in endangered species recovery plans. Ecological Applications **12**:674-681.

Caswell, H. 2000. Prospective and retrospective perturbation analyses: Their roles in conservation biology. Ecology **81**:619-627.

Caswell, H. 2001. Matrix population models. Sinauer Associates, Sunderland, Massachusetts.

Courchamp, F., T. Clutton-Brock, and B. Grenfell. 1999. Inverse density dependence and the Allee effect. Trends in Ecology & Evolution **14**:405-410.

Deleray, M., L. Knotek, S. Rumsey, and T. Weaver. 1999. Flathead Lake and River system fisheries status report. Montana Fish, Wildlife, & Parks, Kalispell, Montana.

Dennis, B. 1989. Allee effects: population growth, critical density, and the chance of extinction. Natural Resource Modeling **3**:481-538.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of growth and extinction parameters for endangered species. Ecological Monographs **61**:115-143.

Doremus, H., and J. E. Pagel. 2001. Why listing may be forever: Perspectives on delisting under the US Endangered Species Act. Conservation Biology **15**:1258-1268.

Ellner, S. P., J. Fieberg, D. Ludwig, and C. Wilcox. 2002. Precision of population viability analysis. Conservation Biology **16**:258-261.

Fieberg, J., and S. P. Ellner. 2000. When is it meaningful to estimate an extinction probability? Ecology **81**:2040-2047.

Fraley, J. J., and B. B. Shepard. 1989. Life history, ecology and populations status of migratory bull trout (Salvelinus confluentus) in the Flathead Lade and River system, Montana. Northwest Science **63**:133-143.

Gerber, L. R., and L. T. Hatch. 2002. Are we recovering? An evaluation of recovery criteria under the US Endangered Species Act. Ecological Applications **12**:668-673.

Holmes, E. E. 2001. Estimating risks in declining populations with poor data. Proceedings of the National Academy of Sciences of the United States of America **98**:5072-5077.

Holmes, E. E., and A. E. York. 2003. Using age structure to detect impacts on threatened populations: A case study with Steller sea lions. Conservation Biology **17**:1794-1806.

Lande, R., S. Engen, and B.-E. Saether. 1995. Optimal harvesting of fluctuating populations with a risk of extinction. American Naturalist **145**:728-745.

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

Ludwig, D. 1999. Is it meaningful to estimate a probability of extinction? Ecology **80**:298-310.

Maxell, B. A. 1999. A power analysis on the monitoring of bull trout stocks using redd counts. North American Journal of Fisheries Management **19**:860-866.

McCarthy, M. A., S. J. Andelman, and H. P. Possingham. 2003. Reliability of relative predictions in population viability analysis. Conservation Biology **17**:982-989.

Morris, W. F., P. L. Bloch, B. R. Hudgens, L. C. Moyle, and J. R. Stinchcombe. 2002. Population viability analysis in endangered species recovery plans: Past use and future improvements. Ecological Applications **12**:708-712.

Morris, W. F., and D. F. Doak2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Ralls, K., D. P. Demaster, and J. A. Estes. 1996. Developing a criterion for delisting the southern sea otter under the US endangered species act. Conservation Biology **10**: 1528-1537.

Rieman, B. E., and F. W. Allendorf. 2001. Effective population size and genetic conservation criteria for bull trout. North American Journal of Fisheries Management **21**:756-764.

Ripa, J., and P. Lundberg. 2000. The route to extinction in variable environments. Oikos **90**:89-96.

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.

Snyder, N. F. R., and E. V. Johnson. 1985. Photographic Censusing of the 1982-1983 California Condor Population. Condor **87**:1-13.

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.

Thompson, W. L., G. C. White, and C. Gowan 1998. Monitoring vertebrate populations. Academic Press, San Diego, California.

U.S. Endangered Species Act of 1973. 16 U.S. Code Section 1532 (2000).

USFWS. 1994. Whooping Crane recovery plan. U. S. Fish and Wildlife Service, Albuquerque, NM.

Zeng, Z., R. M. Nowierski, M. L. Taper, B. Dennis, and W. P. Kemp. 1998. Complex population dynamics in the real world: Modeling the influence of time-varying parameters and time lags. Ecology **79**:2193-2209.

CHAPTER THREE

ESTIMATING POPULATION TREND AND PROCESS VARIATION FOR PVA IN THE PRESENCE OF SAMPLING ERROR

Abstract

Time-series of population abundance estimates often are the only data available for evaluating the prospects for persistence of a species of concern. With such data, it is possible to perform a population viability analysis (PVA) with diffusion approximation methods using estimates of the mean population trend and the variance of the trend, the so-called process variation. Sampling error in the data, however, may bias estimators of process variation derived by simple methods. We develop a restricted maximum likelihood (REML)-based method for estimating trend, process variation, and sampling error from a single time-series based on a discrete-time model of density independent growth coupled with a model of the sampling process. Transformation of the data yields a conventional linear mixed model, where the variance components are functions of the process variation and sampling error. Simulation results show essentially unbiased estimators of trend, process variation, and sampling error over a range of process variation/sampling error combinations. Example data analyses are provided for the Whooping Crane (*Grus americana*), grizzly bear (*Ursus arctos horribilis*), California Condor (*Gymnogyps californianus*) and Puerto Rican Parrot (*Amazona vittata*). This REML-based method is useful for PVA methods that depend on accurate estimation of process variation from time-series data.

Introduction

Conservation or management policies generally require assessment of a population's prospects for persistence. Unfortunately, these assessments often must be made with little data. Commonly, the most extensive data available are count-based data in which the entire population or a subset of the population is sampled over multiple years. Count-based data may include time-series of population abundance estimates, catch-per-unit effort, or samples of a portion of the life cycle, e.g. spawning redds, nesting adults or mother-cub pairs. Dennis et al. (1991) presented methods for fitting an exponential growth model to count-based data and techniques to perform a population viability analysis (PVA) with the fitted model parameter estimates using diffusion approximations (DA). These PVA techniques include the calculation of risk metrics such as the mean time to extinction, distribution of extinction times, probability of hitting a lower threshold (not necessarily extinction) or the distribution of the population density in the future. The methods presented in Dennis et al. (1991) have been used to calculate risks for a variety of species (Dennis et al. 1991, Nicholls et al. 1996, Gerber et al. 1999, Morris and Doak 2002).

Count-based PVA using diffusion approximation (DA) methods are useful for analyzing risk to a population. However, it has been noted these methods are vulnerable to sampling or measurement error in the data. Ludwig (1999) found that errors in estimated abundance could increase confidence intervals for quasi-extinction to the point where estimates may become meaningless. Meir and Fagan (2000) found that sampling error had its largest effect on estimates of extinction risk for cases in which it was

uncertain whether the population would go extinct, presumably a situation applicable to many species of conservation interest.

PVA techniques with the Dennis et al. model depend on the estimates of two population parameters in the exponential growth model: the long-term geometric mean growth rate and the variance of the realized growth due to environmental effects. Estimation of the latter parameter, termed process variation, is confounded by sampling error in the data. In this paper, the term "sampling error" is used to indicate any deviation of the observed count from the true population size. This includes error from sampling only a portion of the population (traditional definition of sampling error) in addition to error from incorrect measurements within a chosen sampling unit (conventionally called observation, or measurement, error). Sampling error adds to the variability in the data leading to positively biased estimators of process variation in the model of Dennis et al. This bias leads to poor predictive precision and pessimistic PVA projections. An ability to partition variation in the data due to sampling error from variation due to environmental effects would greatly improve the quality of count-based PVA.

Holmes (2001) described a regression method for estimating process variation and sampling error from a single time-series of population estimates. Holmes' method, however, does not necessarily produce parameter estimates with desirable statistical properties. A particularly worrisome feature of the Holmes method is that estimators of process variation can be biased high or low and the magnitude and direction of bias is

unknown in actual applications. For more detailed discussions of the Holmes method see Morris and Doak (2002) and Holmes and Fagan (2002).

We present a new restricted maximum likelihood (REML)-based method for estimating trend, process variation, and sampling error variance from a single time-series of count-based data that can be performed in statistical computing packages such as SAS PROC MIXED (Littell et al. 1996). These parameter estimates are obtained by embedding a model of the sampling process within a model of population growth. This allows estimation of process variation while accounting for variation in the data due to sampling error. The method presented, however, has stricter data requirements than the Dennis et al. method in that the REML-based method requires uniformly spaced population estimates in time with no missing data points.

Methods

The underlying change in population density over time is modeled with a stochastic, discrete-time model of exponential growth. This model represents an assumption of density-independent growth for the population, a reasonable assumption for depleted populations. The population density at time (t+1) is thus

$$\boldsymbol{N}(t+1) = \boldsymbol{N}(t)\exp(\mu)\boldsymbol{E}(t) \qquad (1)$$

for $t = 1,2,\ldots, T$; where $\boldsymbol{N}(t)$ represents the actual population density at time (t) and $\exp(\mu)$ is the deterministic per-unit-abundance growth rate in population density. Natural population growth is not constant from year to year, so a multiplicative term, $\boldsymbol{E}(t)$, is included to represent the process variation or deviation from the long-term trend at time

(t) due to environmental effects. The random variables $\boldsymbol{E}(1)$, $\boldsymbol{E}(2)$, … are assumed to be independent and identically distributed (iid) as lognormal(0, σ_p^2), that is, ln($\boldsymbol{E}(t)$) has a normal distribution with a mean of 0 and variance σ_p^2. We write this as $\boldsymbol{E}(t)$~lognormal(0, σ_p^2) and ln($\boldsymbol{E}(t)$)~normal(0, σ_p^2). Dennis et al. (1991) used this model as an approximation of a more complicated demographic population model under the assumption of constant trend (μ) and process variation (σ_p^2) in the absence of density-dependence. The parameters μ and σ_p^2 are the main quantities of interest when using DA methods for PVA. If this model is used for abundance estimates, a discrete random variable, as opposed to density, a continuous random variable, then caution may be required should the population abundance not be large enough to justify the continuous approximation with the exponential model.

If process error is the sole source of error, the estimates of the parameters μ and σ_p^2 given in Dennis et al. (1991) are unbiased. However, estimation is complicated when observed variability in the data reflects sampling error as well as process variation. Holmes (2001) proposed a model in which the population growth model (1) is embedded in a model that includes sampling error. The observed data are modeled as

$$\boldsymbol{O}(t) = \boldsymbol{N}(t)\boldsymbol{Z}(t), \qquad (2)$$

where $\boldsymbol{O}(t)$ is the observed estimate of population density at time (t) and $\boldsymbol{Z}(t)$ represents sampling error. The random variables $\boldsymbol{Z}(1)$, $\boldsymbol{Z}(2)$, … are assumed to be iid lognormal(μ_s, σ_s^2) and independent of $\boldsymbol{E}(t)$. Bias in measurement is represented by μ_s. By combining the population growth and sampling models, the observation at time (t+1) is

$$\boldsymbol{O}(t+1) = \boldsymbol{N}(t)\exp(\mu)\boldsymbol{E}(t)\boldsymbol{Z}(t+1). \qquad (3)$$

Define $\boldsymbol{M}(t) = \ln(\boldsymbol{N}(t))$, $\boldsymbol{\varepsilon}(t) = \ln(\boldsymbol{E}(t))$, and $\boldsymbol{\varphi}(t) = \ln(\boldsymbol{Z}(t))$. Note that $\boldsymbol{\varepsilon}$~normal$(0, \sigma_p^2)$ and $\boldsymbol{\varphi}$~normal(μ_s , σ_s^2). Log transformation of the observed values yields:

$$\boldsymbol{Y}(t) = \ln(\boldsymbol{O}(t)) = \boldsymbol{M}(t) + \boldsymbol{\varphi}(t), \text{ and} \qquad (4)$$

$$\boldsymbol{Y}(t+1) = \boldsymbol{M}(t) + \mu + \boldsymbol{\varepsilon}(t) + \boldsymbol{\varphi}(t+1). \qquad (5)$$

Note the entries of $\boldsymbol{Y}$ are autocorrelated because the effect of each realization of the process variation, $\boldsymbol{\varepsilon}(t)$, is passed on to later population densities through the exponential growth of the population, e.g. $\boldsymbol{Y}(t+2) = \boldsymbol{M}(t) + 2\mu + \boldsymbol{\varepsilon}(t) + \boldsymbol{\varepsilon}(t+1) + \boldsymbol{\varphi}(t+1)$ and $\boldsymbol{Y}(t+3) = \boldsymbol{M}(t) + 3\mu + \boldsymbol{\varepsilon}(t) + \boldsymbol{\varepsilon}(t+1) + \boldsymbol{\varepsilon}(t+2) + \boldsymbol{\varphi}(t+2)$ and so on. The differences $\boldsymbol{W}(t) = \boldsymbol{Y}(t+1) - \boldsymbol{Y}(t)$ for $t = 1,2,..,T-1$ give the growth rate plus error at each individual time step due to process variation at time t and sampling error at times t and $t+1$. The mixed model representation of $\boldsymbol{W}(t)$ and $\boldsymbol{W}(t+1)$ are

$$\boldsymbol{W}(t) = \mu + \boldsymbol{\varepsilon}(t) + \boldsymbol{\varphi}(t+1) - \boldsymbol{\varphi}(t), \text{ and} \qquad (6)$$

$$\boldsymbol{W}(t+1) = \mu + \boldsymbol{\varepsilon}(t+1) + \boldsymbol{\varphi}(t+2) - \boldsymbol{\varphi}(t+1). \qquad (7)$$

The vector $\boldsymbol{W}$ is a series of the observed log population growth rate at each of the $(T-1)$ one-step time intervals in the time-series of observations. The distribution of $\boldsymbol{W}$ is multivariate-normal with common mean μ and variance-covariance matrix $\boldsymbol{\Sigma}$. The empirical growth observations in $\boldsymbol{W}$ can be expressed as a conventional linear mixed model;

$$\boldsymbol{W} = \mathbf{X}\,\mu + \mathbf{v}, \qquad (8)$$

where $\mathbf{X}$ is a column of ones with length $(T-1)$ and $\mathbf{v}$~$(\underline{0}, \boldsymbol{\Sigma})$. Autocorrelation in $\boldsymbol{Y}$ due to exponential population growth is removed in the creation of $\boldsymbol{W}$ by the differencing step. The entries of $\boldsymbol{W}$ are still correlated, however, because a single realization of the

sampling error will be shared by two successive entries of $\boldsymbol{W}$, e.g. $\boldsymbol{\varphi}(t+1)$ appears in both $\boldsymbol{W}(t)$ and $\boldsymbol{W}(t+1)$. This creates a one-step covariance in $\boldsymbol{W}$. Accordingly, the variance-covariance matrix for $\boldsymbol{W}$ has a structure composed of the variance of individual entries of $\boldsymbol{W}$ on the main diagonal and the one-step covariance on the diagonals above and below the main with zeroes elsewhere (this is known as a banded Toeplitz(2) structure (Graybill 1983)). Let σ_1 and σ_2 represent the variance and one-step covariance of $\boldsymbol{W}$ respectively, then;

$$\underset{(T-1)\times(T-1)}{\Sigma} = \begin{bmatrix} \sigma_1 & \sigma_2 & 0 & 0 & . & . & . & 0 \\ \sigma_2 & \sigma_1 & \sigma_2 & 0 & & & & . \\ 0 & \sigma_2 & \sigma_1 & \sigma_2 & 0 & & & . \\ 0 & 0 & \sigma_2 & . & . & . & & . \\ . & & 0 & . & . & . & . & . \\ . & & & . & . & . & . & 0 \\ . & & & & . & . & . & \sigma_2 \\ 0 & . & . & . & . & 0 & \sigma_2 & \sigma_1 \end{bmatrix}. \qquad (9)$$

The parameters in $\boldsymbol{\Sigma}$ are functions of the process variation and sampling error. The variance of an individual entry of $\boldsymbol{W}$ is $\sigma_1 = 2\sigma^2_s + \sigma^2_p$. The one-step covariance of the entries of $\boldsymbol{W}$ is $\sigma_2 = -\sigma^2_s$.

In simulations, maximum likelihood estimators of μ, σ_s^2, and σ_p^2 were strongly biased. Biased estimators are a common problem in variance component problems when using maximum likelihood, e.g. the maximum likelihood estimate of a sample variance is biased as it ignores the difference of the sample mean from the true mean. This bias is overcome by using restricted maximum likelihood (REML) estimation (Searle et al. 1992) in which a transformation of the data is used to eliminate the fixed effect (μ) before

estimating σ_1 and σ_2. The data transformation used in this case is again a one-step difference of the ***W*** series;

$$\boldsymbol{J}(t) = \boldsymbol{W}(t+1) - \boldsymbol{W}(t), \qquad (10)$$

where $\boldsymbol{J}(t)$ is the transformed data vector of length (T-2). This transformation can be constructed as $\boldsymbol{J}=\mathbf{U}'\boldsymbol{W}$, where $\mathbf{U}$ is a (T-1)x(T-2) contrast matrix with negative 1's on the main diagonal and 1's on the diagonal below the main. The parameters σ_1 and σ_2 are then estimated with maximum likelihood from the transformed data. Because the covariance matrix of ***J*** is related to that of ***W*** by the equation $\boldsymbol{\Theta} = \mathbf{U}'\boldsymbol{\Sigma}\mathbf{U}$, the covariance matrix $\boldsymbol{\Theta}$ is also a function of σ_1 and σ_2. The values $\hat{\sigma}_1$ and $\hat{\sigma}_2$ that minimize the negative log-likelihood function,

$$L(\sigma_1, \sigma_2 | \boldsymbol{J}) = \tfrac{1}{2}(T-2)\ln(2\pi) + \tfrac{1}{2}\ln(|\boldsymbol{\Theta}|) + \tfrac{1}{2}\boldsymbol{J}'\boldsymbol{\Theta}^{-1}\boldsymbol{J}, \qquad (11)$$

are then the REML estimates of σ_1 and σ_2. Good initial estimates of σ_1 and σ_2 for the minimization can be obtained by the method-of-moments, i.e. the initial estimates of σ_1 and σ_2 are the sample variance and the sample one-step covariance of ***W*** respectively. The estimate, $\mathbf{V}$, of the covariance matrix, $\boldsymbol{\Sigma}$, is obtained by inserting $\hat{\sigma}_1$ and $\hat{\sigma}_2$ into equation (9).

To estimate μ, a generalized least squares estimator is used. This is an extension of ordinary least squares, i.e. a straight average of the entries in ***W***, and takes advantage of non-zero covariance in the data. The generalized least squares estimator of the trend is

$$\hat{\mu} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{W}. \qquad (12)$$

Kackar and Harville (1984) showed that generalized least squares estimators with estimated weights are unbiased. Inference on $\hat{\mu}$ requires an estimate of its variance. If

the data have no sampling error, this variance is $\mathbf{V} = \text{var}(\boldsymbol{W})^* \mathbf{I}_{(T\text{-}1)}$ where $\mathbf{I}_{(T\text{-}1)}$ is an (T-1) identity matrix and the estimator for trend is the same as presented in Dennis et al. (1991). If $\boldsymbol{\Sigma}$ is known, then $\hat{\mu} = (\mathbf{X}' \boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}' \boldsymbol{\Sigma}^{-1}\boldsymbol{W}$ and the variance of $\hat{\mu}$ is exactly $\Phi = \text{Var}(\hat{\mu}) = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$. Because $\boldsymbol{\Sigma}$ is unknown, the estimated covariance matrix must be used as in (12), resulting in a biased estimator of $\text{Var}(\hat{\mu})$, namely

$$\text{Var}(\hat{\mu}) = \hat{\Phi} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \qquad (13)$$

Kenward and Roger (1997) presented a technique to remove the bias in $\hat{\Phi}$, however, application of this technique is not easily implemented in SAS and simulation results suggest that the bias in $\hat{\Phi}$ is small, at least for the larger data sets examined. The Kenward and Roger adjustment to $\hat{\Phi}$ was therefore not used for this analysis.

The behavior of estimators for the REML-based procedure was explored with analysis of simulated time-series data at nine combinations of process variation and sampling error. For each simulation, 2000 time-series of length 50 were generated by first simulating the actual population growth with process variation and then 'observing' the population densities with sampling error. Process variance and sampling error variance were each set at three levels: L = 0.09, M = 0.0225, and S = 0.0025. For each of the nine variation combinations, simulations with and without a true non-zero underlying trend were analyzed. The simulations with zero trend, random walks, consisted of population growth that had random fluctuations due to process variation only. Simulations with a non-zero trend had μ = -0.05, a decline in total population density of ~5% per year.

The performance of the REML-based procedure is compared with the Dennis et al. model in the simulations described above and in example analyses of data on the Whooping Crane (*Grus americana*), grizzly bear (*Ursus arctos horribilis*) in Yellowstone National Park, California Condor (*gymnogyps califonianus*) and Puerto Rican Parrot (*Amazona vittata*). The Whooping crane data are winter population counts conducted 1938-1993 in the Aransas National Wildlife Refuge (USFWS 1994). The grizzly bear data are yearly estimated number of adult females in the Greater Yellowstone population for 1959-1997 (Morris and Doak 2002). The California condor data are the October Surveys conducted 1965-1980 that have been described as problematic due to inconsistent sampling efforts (Snyder and Johnson 1985). The Puerto Rican Parrot data are 1969-1989 censuses for an intensively managed population in Luquillo Forest of Puerto Rico (Dennis et al. 1991).

As an example of behavior of extinction metrics using estimators from the Dennis et al. and REML-based methods, predictions of the probability of reaching a threshold of 75% of current population size (so-called quasi-extinction) are compared for simulated data with a slight positive trend. For these simulations, the true underlying trend is a 2% per year increase in population size with the process variation and sampling variance taking on the same nine combinations as the previously mentioned simulations.

Results

In simulations, the REML-based method provided an unbiased estimator of trend and only slight bias in estimators for variance components. Performance of the REML-

based method was similar for simulations with and without a non-zero trend. Increases in sampling error and/or process variation resulted in more variable estimates of sampling error. The level of sampling error affected the distribution of process variation estimates similarly for the three levels of process variation examined and for simulations with and without a non-zero trend. The distributions of process variation estimates for the REML-based method and the Dennis et al. model are compared in Fig. 4 over the range of sampling error for simulations with medium level process variation and ~5% per year decline. In addition to decreasing bias in process variation estimates, the REML-based method has less variability in process variation estimates. In data sets with large sampling error, however, estimates of process variation occasionally are extremely high with the REML-based method. These high estimates are typically higher than the total variation observed in the time-series but occur less than 1% of the time in the simulated data. The sampling error estimates are zero or very close to zero for the time-series with high process variation estimates. In practice, a near zero sampling error estimate accompanied by a process variation estimate that is larger than the variance in the time series (specifically the variance of the $\boldsymbol{W}$ vector) should warn the user of unreliable variance component estimates. In such cases, we would recommend using the Dennis et al. model process variation estimator while acknowledging that it is likely to be biased high.

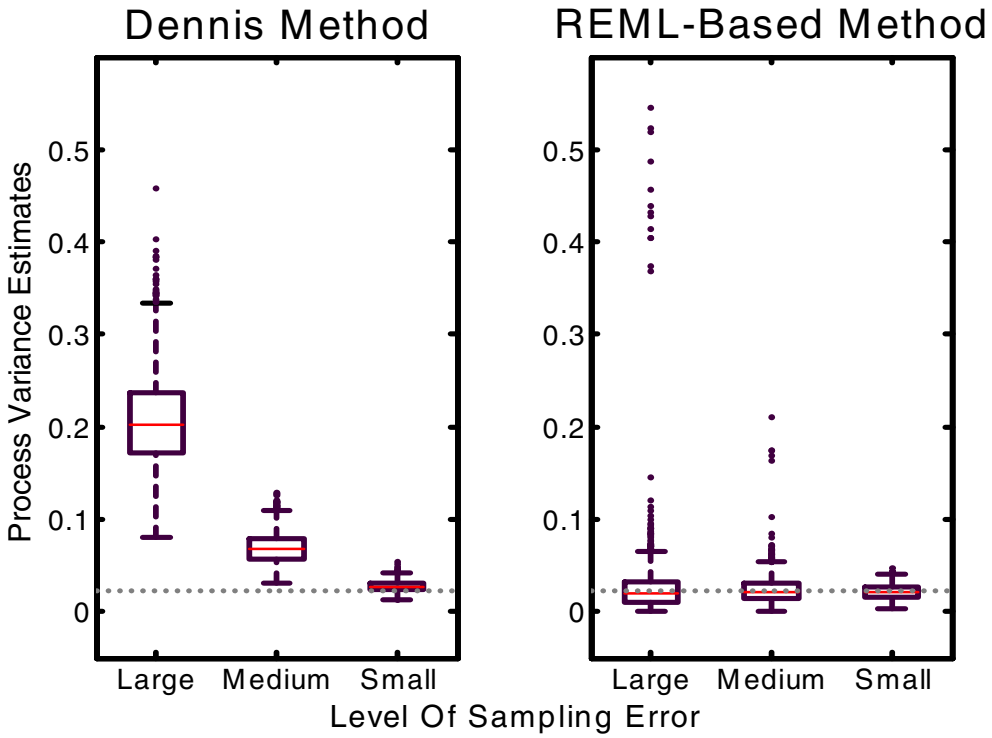


Figure 4. Comparison of Dennis et al. model and REML-based estimates of process variation over range of sampling error. 2000 simulations at medium level of process variation with 5% decline per year. Dotted line represents true process variation value.

Bias and variance in predictions of quasi-extinction depend on the true probability of reaching of quasi-extinction and the level of sampling error in the data. For high process variation, when quasi-extinction was highly likely, the higher variance of predictions for the REML-based method made it a poorer choice than the Dennis et al. method despite its bias. When the true probability of quasi-extinction was near zero (when process variation is small), the REML-based method is better due to the large bias in the Dennis et al. method attributable to sampling error in the data. The distributions of predictions for both estimation methods over the nine variation combinations are shown in figure 5.

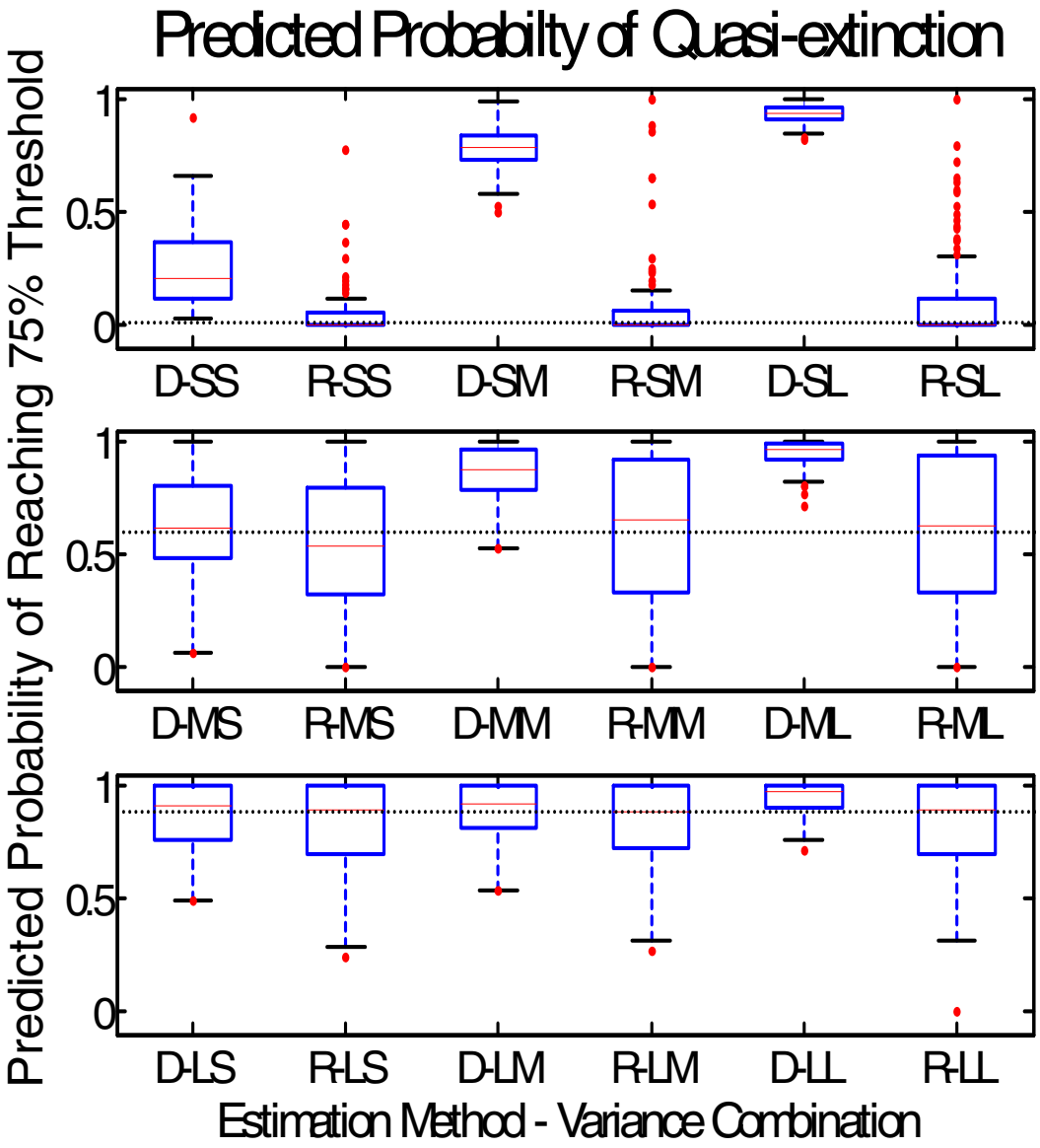


Figure 5. Predicted probability of reaching quasi-extinction level of 75% of current population size with Dennis et al. model and REML-based model. Population growth is approximately a 2% increase per year. SS,SM,SL,…,LL represent respectively the level of process variance and sampling error variance. The D and R in front of the variance combination represent the Dennis et al. and REML-based methods. True probability of quasi-extinction is given by the horizontal dotted line.

Animal	years	Analysis Method	Trend (SE)	Process Variance	Sampling Variance	P(Lower Threshold) †
GB	39	Dennis	0.0213 (0.0185)	0.0131		0.40
		REML	0.0211 (0.0148)	0.0082	0.0023	0.24
WC	56	Dennis	0.0377 (0.0187)	0.0194		0.32
		REML	0.0372 (0.0159)	0.0137	0.0028	0.21
CC	16	Dennis	-0.0768 (0.0885)	0.1176		1.00
		REML	-0.0948 (0.0131)	0*	0.0579	1.00
PP	21	Dennis	0.0273 (.0275)	0.0151		0.33
		REML	0.0273 (.0275)	0.0151	0*	0.33

* estimates at boundary

† Probability of population size reaching lower threshold equal to 0.75 of last population size.

Table 1. Parameter estimates and risk metric comparisons for Dennis et al. and REML-based methods.

Parameter estimates and the probability of quasi-extinction for the example data analyses are given in Table 1. The trend estimates for the Grizzly bear and Whooping crane are similar for the Dennis et al. and REML-based approaches, however, the REML-based model gives a smaller standard error for the trend estimate in both cases. For both analysis methods, the estimates of trend, standard error of the trend estimate and the estimated process variation are the same for the Puerto Rican Parrot data because the estimated sampling error is less than zero. In contrast, the California Condor data analysis resulted in an infinite likelihood in the SAS analysis (code available in

Ecological Archives) due to the estimate of process variation being less than zero (see discussion below). Therefore the results given for the Condor were calculated with the MATLAB code available in *Ecological Archives*.

Discussion

The REML-based method may be performed in standard statistical packages such as SAS PROC MIXED. Because SAS places no boundaries on variance component estimates, there are occasional instances when the sampling error estimate is negative or the procedure fails to converge on a solution. The former is due to SAS allowing $\hat{\sigma}_2$ to be larger than zero (recall that the sampling error is the negative of $\hat{\sigma}_2$). In this case we recommend setting the sampling error estimate to the boundary value of zero thereby recovering the Dennis et al. method; see discussion of the Puerto Rican Parrot below. The latter occurs when the likelihood function is infinite in the unbounded optimization. An ad hoc solution is to set the process variation estimate to the minimum boundary value of zero; see discussion of California Condor below. The MATLAB code provided in *Ecological Archives* automatically makes these adjustments.

A t-test for significant trend may be performed in SAS if non-boundary parameter estimates exist, i.e. a minimum is found for the likelihood function and the sampling error estimate is greater than zero. The correct degrees of freedom for the test, however, is not clear for all circumstances. We suggest using the Kenward and Roger (1997) adjustment option in PROC MIXED for the degrees of freedom. However, this test is problematic if the estimated degrees of freedom are less than four (D.F. Staples and M.L. Taper,

personal observation). In such cases, we recommend using the same degrees of freedom as the Dennis et al. method minus one for the additional parameter being estimated (T-4), while being aware this is likely to be a liberal test (D.F. Staples, *unpublished data*).

The unbiased estimator of process variation rectifies a significant liability when using the Dennis et al. model for PVA, i.e. the vulnerability of conclusions to sampling error in the data. It does, however, retain other assumptions for PVA with the Dennis et al. model. Both methods assume there are no large environmental events like catastrophes or bonanzas that dramatically change the population size. Dennis et al. (1991) presented model diagnostic procedures that can be used to detect extreme events in the data. We would like to note that the detection of outliers does not mean that the point should be dropped for the REML-based method as the 'extreme' event may be due to sampling error. If due to sampling error, the REML-based method can use the information for better estimates of process variation thereby avoiding the inflation of the process variation estimate by the extreme event.

The process variation (or environmental error) is also assumed to be temporally uncorrelated for both methods. A two-tailed Durbin-Watson test [Sen, 1990 #333] on the residuals of the Dennis et al. model can be used to evaluate if errors are correlated. If there is a positive correlation, neither model may be appropriate. A negative correlation supports the assumption of the REML-based method that one-step correlation in the observed growth rates (the $\boldsymbol{W}$ series) is due to sampling error. Preliminary simulation results suggest REML-based sampling error estimators are biased low and process variation estimators are biased high in the presence of positively correlated

environmental error (*unpublished data* DFS). The magnitude of the bias depends on the magnitude of the correlation. Development of improved tests for correlated environmental error and methods to incorporate it into the REML-based method are the focus of ongoing research.

Another assumption is constant trend and process variation throughout the time-series. Violations of this may occur due to demographic stochasticity due to small population size or other factors that change the growth rate of the population, e.g., density dependence or the Allee effect. Demographic stochasticity can change the variance in growth rate, especially when population estimates vary from low to high densities. Density Dependence may be tested using methods presented in Dennis and Taper (1994) or, alternatively, model identification techniques may be used as in Taper and Gogan (2002).

In addition to the smaller standard error in the trend estimates, another benefit of the REML-based method for the Grizzly Bear and Whooping Crane data sets is the removal of sampling error from the observed variation in the data resulting in smaller estimates of process variation. This decreases bias in the process variation and results in a less biased analysis of the risk to the populations. For the Puerto Rican Parrot data, the estimate of sampling error is at the boundary value of zero indicating that sampling error is little of the variability observed in the data. This is reasonable considering the data come from a small, intensively managed population and represent full population censuses. With the sampling error set to zero, the REML-based method is identical to the Dennis et al. (1991) method.

The California Condor data represent the opposite end of the parameter spectrum compared to the Puerto Rican Parrot. In contrast to a zero estimate of sampling error, the REML-based method gives an estimate of process variation that is at the boundary value of zero for the condor data, indicating that the variation in the data is mainly due to sampling error. This again is a reasonable conclusion considering the data come from the 'October Surveys' that have been described as problematic due to sampling irregularities (Wilbur 1980) and the population was undergoing a severe crash at the time and likely had little variability in the true negative growth rate. If the process variation is set to zero and a significance test is performed, the REML-based method does give a significantly negative growth rate ($P<.001$) in contrast to a non-significant trend in the Dennis et al. method ($P=0.40$). This case is, however, an instance where the degrees of freedom estimate from the Kenward and Roger (1997) method is extremely small (0.07) and therefore T-4 degrees of freedom was used for the test. While inference in this case is questionable due to uncertainty about degrees of freedom, it is reassuring to get a significant negative trend in a population known to have experience such a severe decline.

The REML-based method is a useful addition to PVA techniques using diffustion approximation methods. It allows partitioning out variability in the data due to sampling error for more precise estimates of trend and less bias in the process variation estimator. When sampling error is small in the data, as in the Puerto Rican Parrot, the REML-based method recovers the estimators in the Dennis et al. model. The REML-based method does, however, have more strict data requirements than the Dennis et al. model at the

present. Data must come from uniform sampling intervals with no missing points and the REML-based method is likely to require more data points because of the additional parameter to be estimated. In cases with too few data points, an estimate of the trend, which is relatively straightforward and unbiased, may be the best metric available for assessing risk to the population. We caution that these models or any other PVA model are only tools for a manager and results must be conditional on validity of assumptions in the model and the nature of the population itself.

Acknowledgements: We would like to thank Robert J. Boik, John J. Borkowski, and James F. Robison-Cox for their help in developing the method presented. We would also like to thank reviewers Ben Bolker, Aaron M. Ellison, Lev Ginzburg, and Jay Ver Hoef for their constructive comments. Funding for this work was provided by Montana Fish, Wildlife, & Parks grant # Z0862 and EPA STAR grant # G9A10014.

Literature Cited

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., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. Ecological Monographs **64**:205-224.

Gerber, L. R., D. P. DeMaster, and P. M. Kareiva. 1999. Gray whales and the value of monitoring data in implementing the US Endangered Species Act. Conservation Biology **13**:1215-1219.

Graybill, F. A. 1983. Matrices with applications in statistics, 2nd ed. edition. Wadsworth International Group, Belmont, CA.

Holmes, E. E. 2001. Estimating risks in declining populations with poor data. Proceedings of the National Academy of Sciences of the United States of America **98**:5072-5077.

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

Kackar, R. N., and D. A. Harville. 1984. Approximations for standard errors of estimators of fixed and random effects in mixed linear models. Journal of the American Statistical Association **79**:853-862.

Kenward, M. G., and J. H. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics **53**:983-997.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. SAS System for Mixed Models. SAS Institute, Cary, NC.

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.

Morris, W. F., and D. F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Nicholls, A. O., P. C. Viljoen, M. H. Knight, and A. S. vanJaarsveld. 1996. Evaluating population persistence of censused and unmanaged herbivore populations from the Kruger National Park, South Africa. Biological Conservation **76**:57-67.

Searle, S. R., G. Casella, and C. E. McCulloch. 1992. Variance components. John Wiley & Sons, Inc., New York, NY.

Sen, A., and M. Srivastava. 1990. Regression Analysis: Theory, methods, and applications. Springer-Verlag New York Inc., New York, NY.

Snyder, N. F. R., and E. V. Johnson. 1985. Photographic Censusing of the 1982-1983 California Condor Population. Condor **87**:1-13.

Taper, M. L., and P. J. P. Gogan. 2002. The Northern Yellowstone elk: Density dependence and climatic conditions. Journal of Wildlife Management **66**:106-122.

USFWS. 1994. Whooping crane recovery plan. U. S. Fish and Wildlife Service, Albuquerque, NM.

Wilbur, S. R. 1980. Estimating the size and trend of the California condor population. California Fish and Game **66**:40-48.

CHAPTER FOUR

INTERACTION OF SAMPLING ERROR AND TEMPORAL CORRELATIONS IN POPULATION GROWTH PRESENT AN INTRACTABLE PROBLEM FOR PVA APPLICATIONS

Abstract

Estimates of a population's average growth rate and process variance can be calculated from time-series of abundance estimates, and used to estimate probability of decline for a population. Correlations in the data from sampling error, intrinsic factors, or environmental conditions can detrimentally affect risk predictions. It has been claimed that explicitly incorporating estimates of all between-year correlations in population growth into the process variance estimate, termed the 'long-term variance', will exactly counterbalance effects of sampling error giving unbiased risk predictions. This approach potentially could accommodate correlations in the data from any source, but is hampered by inadequate estimates of the population growth correlations for real data. Alternatively, we have shown previously that sampling variability can be partitioned from process variation to greatly reduced bias in parameter estimates if the process noise is uncorrelated. In this paper, we extend the method to accommodate autocorrelated process noise, allowing long-term variance to be estimated with one additional parameter. In comparisons with simulated correlated abundance data, long-term variance estimators were imprecise and biased. Surprisingly, ignoring correlation and only explicitly accounting for sampling error generally resulted in better estimates of the variability in

population growth. This study suggests sampling error can have pervasive effects that will complicate attempts to estimate temporal correlations in population growth.

Introduction

Estimates of a population's growth rate and its variability, commonly denoted trend and process variance, are often used to calculate risk metrics such as the probability of extinction (Dennis et al. 1991, Morris and Doak 2002). Temporal correlations in population growth, however, can lead to bias in these estimated parameters resulting in biased risk estimates (Ripa and Lundberg 1996, Morales 1999, Holmes and Fagan 2002). Correlations in observed population growth may be due to temporal correlations in environmental factors that affect yearly growth rates, or intrinsic population factors, such as age structure, and their interaction. Though not usually thought of as a correlation issue, sampling error in abundance estimates generates a negative one-step correlation in the observed yearly growth rates [McNamara, 2004 #370], and if ignored, will inflate process variance estimates, leading to pessimistic risk estimates (Ludwig 1999, Meir and Fagan 2000). (We will refer to all errors from the sampling protocol and measurement error collectively as sampling error). If process noise is uncorrelated, the correlation generated by sampling error can be used to estimate the sampling variability in abundance estimates, resulting in greatly reduced bias for process variance estimates (Staples et al. 2004). Sampling error will increase uncertainty in process variance estimates, however, even when the sampling variance is known (Ferrari and Taper 2006).

An alternative approach for dealing with sampling error based on estimating a 'long-term variance' was proposed by McNamara and Harding (2004). The long-term variance parameter, ψ, includes effects of both the variability and correlations in population growth on future abundances. Because the process variance parameter governs many risk metrics of interest in population viability analysis (PVA), an uncorrelated error model using ψ as the estimate of process variance can provide useful estimates of the future abundance for populations with correlated growth rates. Though the McNamara and Harding (MH) method can accommodate any source of correlation among yearly observations of population growth, it does assume the correlations have second order stationarity (i.e. correlation is only a function of time between observations) and quickly converge to zero as time between observations increases. The MH results have been misinterpreted to show that sampling error cancels out for large samples (Almaraz and Amat 2004). That is certainly not the case; MH's claim is that only by explicitly incorporating the covariances in population growth rates does sampling error cancel out in the long-term variance estimate.

In this paper, we present an extension of the density independent sampling error (DISE) model developed in Staples et al. (2004) to accommodate one-step autoregressive (AR1) process noise and sampling error in an exponential growth PVA. While long-term variance can be estimated with only one additional parameter, this method does assume a specific form of correlation in population growth. We compare the long-term process variance estimates from the DISE-AR1 and MH methods with process variance estimates from a density independent (DI) and DISE models using simulated population abundance

data with sampling error. Data were simulated with an exponential growth model with AR1 process noise, and with an age structured model with independent environmental errors.

Methods

Stochastic exponential population growth is given by $N_{i+1}=N_i*\exp(r + \varepsilon_i)$, where N_i is the population abundance at time $i = 0,1,2,..,t$. The parameter r is the trend in population growth, and ε_i is random variation in population growth, i.e. the process noise. In the presence of AR1 process noise with correlation strength ρ, ε_i is distributed as $N(0,\sigma^2 \Omega)$ where Ω is a matrix with ones on the main diagonal and ρ^i on the i^{th} off-diagonal. As true population abundances are rarely known, estimates of abundance must be used in application. Let $O_i=N_i*Z_i$ be the estimated population abundance with Z_i representing multiplicative sampling error in abundance estimates assumed to be distributed as $\text{LogNormal}(0,\tau^2)$. As in the DISE model, we will use the difference of the natural log of successive observed abundances, i.e. the observed yearly growth rates denoted as $W_i= \log(O_{i+1})-\log(O_i)$, to estimate model parameters. Sampling error in abundance estimates inflates the variance of W_i by $2\tau^2$, and introduces a one-step correlation equal to $-\tau^2$ (i.e. a banded Toeplitz covariance structure (Staples et al. 2004)). Because environmental errors and sampling errors are independent, incorporating the correlated error into the DISE model is accomplished by adding the AR1 covariance matrix to the Toeplitz covariance matrix that represents the effect of sampling error in abundance estimates. The expected value and variance of each observation of **W** is

$E(W_i) = r$ and $V(W_i) = 2\tau^2 + \sigma^2$, while the one-step covariance is $Cov(W_i,W_{i+1}) = \sigma^2\rho - \tau^2$, and higher order covariances are $Cov(W_i,W_{i+k}) = \sigma^2\rho^k$ for k>1. In this formulation, W has a multivariate normal distribution with common mean *r* and variance-covariance structure as described above. The parameters σ^2, τ^2, and ρ are first estimated with restricted maximum likelihood; the estimated covariance matrix is then used to estimate the population trend with generalized least squares.

McNamara and Harding (2004) proposed an approximation for the variability in population growth, the 'long-term variance', based on the distribution of population abundance *t* time steps in the future. For a series of population abundances, $\boldsymbol{N}=\{N_0, N_1, \ldots, N_t\}$, let $R_i = \log(N_{i+1})-\log(N_i) = r + e_i$; R_i is the true log-scale growth rate between time *i* and *i*+1. On the log-scale, the future abundance, $\log(N_t)$, is the current abundance, $\log(N_0)$, plus the sum of the yearly growth increments, $\sum_{i=0}^{t-1} R_i$. When process noise is uncorrelated, the variance of the future abundance distribution is $Var(\log(N_t)) = Var(\sum_{i=0}^{t-1} R_i) = \sigma^2 t$, i.e. a linear function of time, with σ^2 representing the process variance, the parameter used in exponential growth PVA to represent yearly variation in population growth. In their development of the long-term variance parameter, ψ, McNamara and Harding claimed that for correlated process noise long-term average variance, $Var(\sum_{i=0}^{t-1} R_i)/t$, is approximately equal to ψ if *t* is large, where $\psi = \left[\sigma^2 + 2\sum_{i=1}^{t} c_i\right]$, and c_k denotes the covariance of R_i and R_{i+k}. MH do not verify or prove this result; they

merely state that it is "easy to see". The true average long-term variance Var($\sum_{i=0}^{t-1} R_i$)/t = $\left[\sigma^2 + 2\sum_{i=1}^{t-1} c_i - 2\sum_{i=1}^{t-1} \frac{i}{t} c_i\right]$. For the MH approximation to be correct, the term $2\sum_{i=1}^{t-1} \frac{i}{t} c_i$ must go to zero as t gets large, thus the c_k's must go to zero rather quickly as k gets large. For example, if $c_k = \frac{1}{2k}$, the sum will not converge to zero but to $\frac{1}{2}$. If yearly growth has an AR1structure as described above, the term $2\sum_{i=1}^{t-1} \frac{i}{t} c_i$ does converge to zero.

The total long-term variance, ψt, has a similar form as in the uncorrelated model, i.e. variance of future abundance is a linear function of time. The ψ parameter is a measure of the yearly variation in population growth adjusted by the strength of correlations in population growth, and can be used in place of σ^2 in an uncorrelated PVA to model correlated population growth. McNamara and Harding also claimed thant the average variance of the *estimated* growth increments, i.e. Var($\sum_{i=0}^{t-1} W_i$)/t, is $\left[\tilde{\sigma}^2 + 2\sum_{i=1}^{t} \tilde{c}_i\right]$, where $\tilde{\sigma}^2$ is the variance of W_i, and $\tilde{c}_k$ is the covariance of W_i and W_{i+k}. As above, process noise and observation errors are assumed independent, thus $\tilde{\sigma}^2 = 2\tau^2 + \sigma^2$, and $\tilde{c}_1 = c_1 - \tau^2$, and $\tilde{c}_k = c_k$ for $k > 1$. In this formulation, the terms with τ^2 cancel out so that $\left[\tilde{\sigma}^2 + 2\sum_{i=1}^{t-1} \tilde{c}_i\right] = \left[\sigma^2 + 2\sum_{i=1}^{t-1} c_i\right]$, i.e. long-term variance estimation is not affected by sampling error in abundance. MH state that ψ can then be used to calculate unbiased quasi-extinction risk estimates (McNamara and Harding 2004).

A fundamental problem with this approach is that the expression McNamara and Harding give for the long-term variance, ψ, is written in terms of the true population parameters ($\tilde{\sigma}^2, \tilde{c}_i$), but they are not specific on how these parameters should be estimated, nor do they give any examples of its use. Their derivation leaves the impression that investigators could use the sample variance and covariances of the W's as estimators for $\tilde{\sigma}^2$ and $\tilde{c}_i$; we investigate the performance of the MH method when the variance and covariance are estimated with standard methods. For the DISE-AR1 model, the true long-term variance in population growth is

$\psi = \sigma^2 + 2\sigma^2\left[\frac{1-\rho^t}{(1-\rho)} - 1\right] + \frac{\rho^t(1-\rho) + (1-\rho^t)/t}{(1-\rho^2)} + \frac{2\tau^2}{t}$. We use the MH approximation (i.e. drop terms that go to zero as *t* gets large) to estimate long-term variance with the AR1 correlation used in place of individually estimated covariances:

$\hat{\psi} = \hat{\sigma}^2 + 2\hat{\sigma}^2\left[\frac{1}{(1-\hat{\rho})} - 1\right]$.

We used simulated autocorrelated time-series data with sampling error in abundance estimates to compare the performance of the DI, DISE, DISE-AR1, and MH methods in estimating variability of population growth. We generated 1000 time-series of length 50 with a discrete exponential growth model (trend = 0) with AR1 process noise using two levels of sampling variance (0.01 and 0.05) and correlation (0.2 and 0.6). AR1 process error was generated with the formula $\varepsilon_i = \rho\varepsilon_{i-1} + z_i$ where $z_i \sim N(0, 0.05)$. This results in $\sigma^2 = \frac{0.05}{(1-0.2^2)}$ and $\sigma^2 = \frac{0.05}{(1-0.6^2)}$ for correlation equal 0.2 and 0.6

respectively. Process variance was estimated from each series with the DI and DISE models, and the long-term variance was estimated with the DISE-AR1 and the MH methods. Reducing the number of estimated parameters should improve the precision of long-term variance estimates. To evaluate this, we estimated long-term variance with the DISE-AR1 model under the assumption that the true sampling variance is a known constant. For the MH method, we used up to lag-10 covariances to estimate ψ because preliminary analyses showed that fewer lags gave the best performance. Fewer lags performed better because correlation in growth decreases rapidly to zero with increasing lags in an AR1 process. Therefore, attempting to estimate the higher lag covariances (that are in truth close to zero) tended to increase bias and variability of the long-term variance estimate with the MH method.

Correlations among changes in observed abundance also can arise from population age-structure. We simulated data from an age-structured population to evaluate how effectively the models could estimate variability in growth from data with this type of intrinsic correlations in addition to sampling errors. Though the DI, DISE, and DISE-AR1 models obviously will have some model misspecification error, the MH model should accommodate this correlation structure. The simulated population had three age classes with initial abundance of 4000, 750, and 50 for age classes 1, 2, and 3. Fecundity equaled 8 and 11 for age class 2 and 3 respectively. Survival rates were 0.12 for age class 1, 0.04 for age class 2, and 0 for age class 3. All survival and fecundity rates had a coefficient of variation (CV) equal to 0.1 and were uncorrelated both within and among years. The data were the sum of the two oldest age classes collected with

sampling error (CV = 0.10). We estimated the 'true' long-term variance from a simulated 50,000-year time-series. We compared process variance estimates from the DI and DISE models with long-term variance estimates from the MH and DISE-AR1 methods for 500 time-series of length 50. We used up to lag-45 covariances for the MH method.

Results

For the simulated AR1 data, the DI model gave biased estimates of process variance (fig. 6) compared to the true long-term variance, i.e. the slope of the relationship between Var(log(N_t)) and time. The direction and magnitude of the bias, however, depended on the sampling error variance and the strength of correlation. Process variance estimates were biased low for series with low sampling error and high correlation; conversely, process variance estimates were biased high for series with high sampling error and low correlation. The DISE model partially corrected the bias direction due to sampling error, and gave relatively precise estimates of process variance. Nonetheless, process variance estimates were consistently biased low compared to the calculated long-term variance. The magnitude of the bias varied with the strength of autocorrelation in the series; for the smaller correlation value (0.2), the bias was relatively small.

All methods that attempted to estimate ψ from the AR1 data performed poorly (fig.6). The DISE-AR1 model rendered positively skewed and imprecise estimates of long-term variance. Including the true sampling error variance into the DISE-AR1 model as a fixed parameter only gave marginal improvement to estimates of long-term variance.

The MH method also produced highly variable estimates of long-term variance, and commonly gave negative variance estimates (which are of no use for estimating extinction risk). Long-term variance estimates with the MH method were biased low, especially for the higher level of environmental correlation.

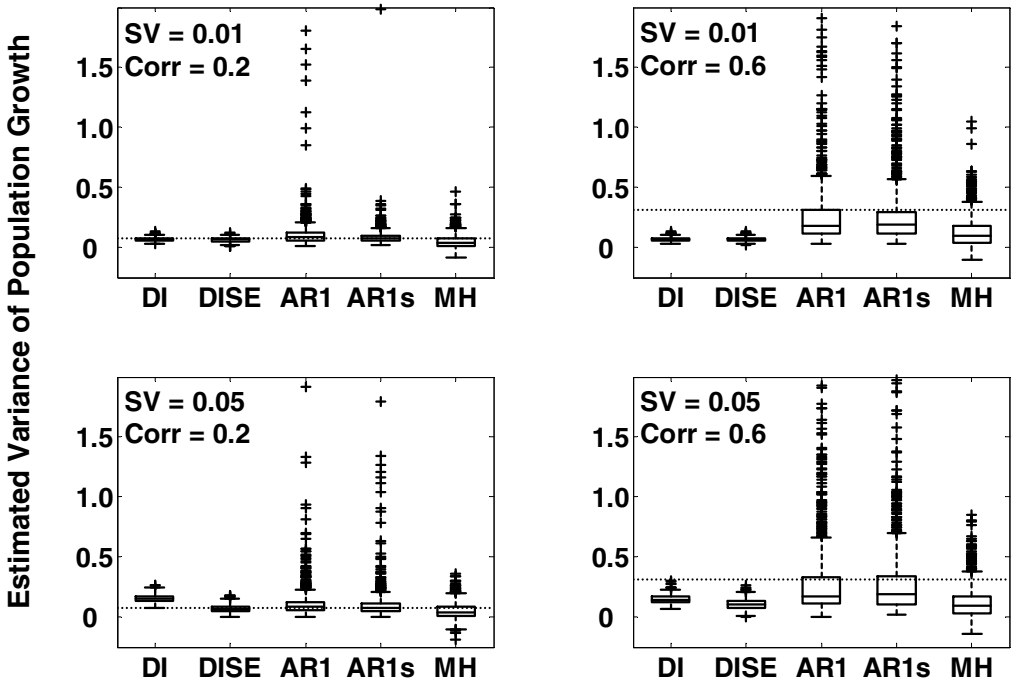


Figure 6. Estimates of variation in growth (either process variance or long-term variance-see text) for simulated exponential population growth data with one-step autoregressive environmental errors and sampling error in abundance estimates. Dotted lines represent true long-term variance of population abundance. SV = sampling error variance; Corr = AR1 correlation. DI = exponential growth model, DISE = exponential growth with sampling error model, AR1 = DISE model with AR1 environmental errors, AR1s = AR1 model when sampling variance is known, MH=McNamara and Harding model.

In analyzing the age-structured data with the DI model, fluctuations in abundance due to age structure and sampling error in the data lead to overestimates of process variance compared to the true long-term variability in abundance (fig.7). The DISE

model had only a slight bias, and gave the most precise prediction of variability in growth compared to all other methods. We note in this case the age structure generated a negative one-step correlation similar to the effect of sampling error for which this model was designed; the performance of the DISE analysis may or may not perform as well for other forms of intrinsic correlations. The DISE-AR1 model gave imprecise and biased estimates, not surprising considering the data were generated with much different process than the assumed autocorrelated exponential growth. This example should be a case particularly suited to the MH method because of its ability to accommodate any correlation structure. Unfortunately, the MH method also gave imprecise estimates for the long-term variance. Many MH long-term variance estimates were negative.

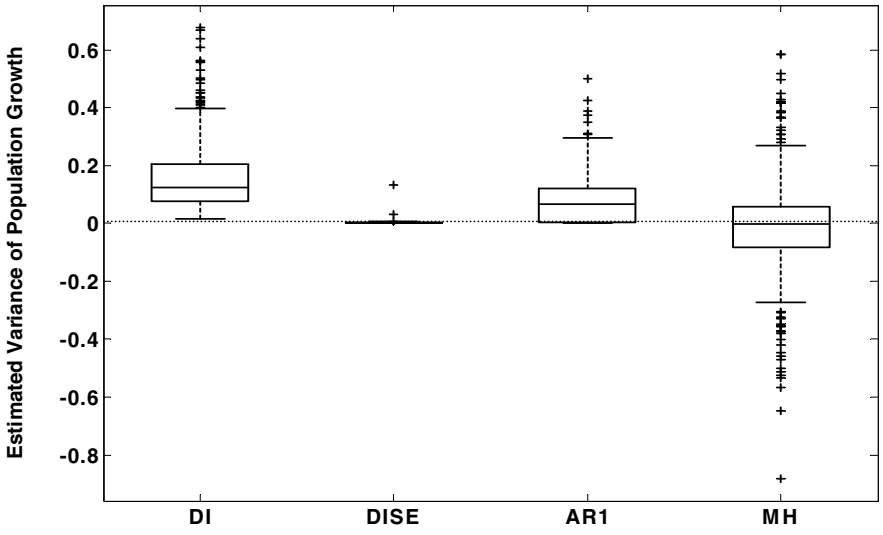


Figure 7. Figure 2. Estimates of variation in growth for simulated age class data with sampling error. Dotted lines represent best estimate of true long-term variance. DI = exponential growth model, DISE = exponential growth with sampling error model, AR1 = DISE model with AR1 environmental errors, MH=McNamara and Harding model

Discussion

The Durbin-Watson test has been proposed to detect first order correlation in environmental errors when fitting an exponential growth model (Dennis et al. 1991). However, this test may not be reliable in population data series that are relatively short. Sampling error in abundance also can detrimentally affect the ability to detect autocorrelation in population growth. In simulations of AR1 population growth, we have found that the Durbin-Watson test can detect a 0.6 correlation in growth only about 50% of the time in a 50-year time-series with our low sampling error (0.01); at our high levels of sampling error (0.05), correlation is detected less than 1% of the time. The decrease in power for the Durbin-Watson test is because the negative correlation in observed population growth generated by sampling error cancels out the true positive correlation in population growth. It may be possible to adjust for sampling error in a test for autocorrelation if the variance of the sampling error is known. We note that whether or not positive autocorrelation can be detected in ecological time-series may be irrelevant as such data may also have intrinsic correlations driven by life history that will complicate any correlation analysis. The possibility of such intrinsic correlations interacting with environmental factors (that may or may not be correlated) further complicates estimation of correlations in population growth. It is difficult to estimate correlations in population growth with classical time-series analysis with small data sets (Lande et al. 2003). The problem is only exacerbated with the inclusion of sampling error in abundance data.

Our results suggest the MH method will be hampered by the problem of estimating temporal covariances in population growth. The W_i observations are not

independent, consequently, the standard estimates of Var(W_i) and Cov(W_i , W_{i+k}) are biased to an unknown degree (see for instance, (Ceylan and Schmeiser 1993)). Further, covariance estimates have notoriously large variances. Despite the relatively long 50-year data series, the MH estimate for ψ does not provide a better estimate for variability in growth than is produced by the DISE estimates that simply ignore all environmental correlations. The effect that high variability in variance estimates would have on estimated quasi-extinction risk depends on the threshold used and the underlying population growth rate. In general, positive bias and higher variation in process variance estimates will lead to positive bias and higher variability in risk estimates.

For the MH method, covariances were estimated individually at each lag. A more effective method for calculating a long-term variance may be to assume a particular form of correlation to reduce the number of parameters. For example, estimation of only a single correlation parameter is required for the DISE-AR1 method presented here. Unfortunately, this model also yields highly variable estimates of long-term variance, even when sampling variance is known. Our analysis suggests the DISE model is a reasonable starting point for count-based PVA combining simplicity of the exponential growth model while accommodating sampling error in abundance estimates. The DISE method gives relatively precise estimates of variation, and has consistent negative bias for autocorrelated environmental errors. The DISE model has been extended to include Gompertz-type density dependent population growth (Dennis et al. 2006).

Realistic correlations in environmental conditions can have substantial effects on the long-term variability in population growth. These effects can translate into

substantial effect on estimated risk. Properly incorporating correlation structure in estimates of process variance in the presence of sampling error must be the focus of future research if PVA analyses are to be useful management tools. This problem, however, may not be resolved with statistical techniques to estimate long-term variance. Correlations in the data reduce the amount of independent information in the data series, in effect reducing the true degrees of freedom in the data. There simply is insufficient information in data sets of realistic length to estimate the necessary parameters. In our DISE-AR1 example, even a maximum likelihood solution with the correct model is inadequate. An alternative strategy may be to develop management tools more sensitive to uncertainty in PVA predictions. For example, viable population monitoring (Staples et al. 2005) can mitigate the effect of negative bias with the DISE method by using comparisons of yearly short-term PVA risk estimates instead of any single long-term risk prediction.

Acknowledgements: We thank D. Goodman for discussions that improved this paper. Funding for D.F.S. and M.L.T. provided by Montana Fish, Wildlife, & Parks (grant Z0862) and the Northwest Power and Conservation Council (grant PTMM11). DFS acknowledges the Kenneth David Lorang Memorial Award for support.

Literature Cited

Almaraz, P., and J. A. Amat. 2004. Multi-annual spatial and numeric dynamics of the white-headed duck Oxyura leucocephala in southern Europe: seasonality, density dependence and climatic variability. Journal of Animal Ecology **73**:1013-1023.

Ceylan, D., and B. W. Schmeiser. 1993. Interlaced variance estimators. *in* G. W. Evans, M. Mollaghasemi, E. C. Russell, and W. E. Biles, editors. Proceedings of the 1993 Winter Simulation Conference.

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:In Press.

Ferrari, M., and M. L. Taper. 2006. Using auxilary data on measurement error to improve estimates of process variation in ecological time series. Journal of Environmental and Ecological Statistics:in press.

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

Lande, R., S. Engen, and B.-E. Saether. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, Inc., New York, NY.

Ludwig, D. 1999. Is it meaningful to estimate a probability of extinction? Ecology **80**:298-310.

McNamara, J. M., and K. C. Harding. 2004. Measurement error and estimates of population extinction risk. Ecology Letters **7**:16-20.

Meir, E., and W. F. Fagan. 2000. Will observation error and biases ruin the use of simple extinction models? Conservation Biology **14**:148-154.

Morales, J. M. 1999. Viability in a pink environment: why "white noise" models can be dangerous. Ecology Letters **2**:228-232.

Morris, W. F., and D. F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Ripa, J., and P. Lundberg. 1996. Noise colour and the risk of population extinctions. Proceedings of the Royal Society of London Series B **263**:1751-1753.

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.

Staples, D. F., M. L. Taper, and B. B. Shepard. 2005. Risk-based viable population monitoring. Conservation Biology **19**:1908-1916.

CHAPTER FIVE

IMPACT OF NON-LINEARITIES IN DENSITY DEPENDENCE ABOVE T HE RANGE OF THE DATA ON PREDICTIONS OF POPULATION EXTINCTION RISK

Abstract

Population viability analysis (PVA) is often used to predict performance of a population in the future. However, a potential bias in PVA exists when population dynamics are predicted for unobserved population densities. The Allee effect, a change in density dependence at low population densities, is one such factor that can have profound effects on PVA. In contrast to the Allee effect, we explore potential effects of a change in density dependence dynamics at unobserved high densities. As a heuristic example, we parameterized a stochastic Ricker model to a time-series of cutthroat trout (*Oncorhynchus clarki*). In simulations with the model, the majority of extinctions were directly from population densities higher than observed in the data. To explore the potential consequence of a regulation mechanism affecting the population only at high densities, we placed a reflecting upper bound in simulations of the Ricker model as fit to the data. The upper bound prevented the population from attaining unrealistically high densities that directly resulted in extinction leading to a decrease in the probability of extinction and an increase in the expected time to extinction. This demonstrates a possible error in PVA when population models are used for predictions at high, unobserved densities.

Introduction

Population viability analysis (PVA) is the use of stochastic quantitative population models to predict the future status of a population. With PVA methods, it is possible to predict future population size, the time to extinction (or some other lower threshold), the probability of ever reaching a lower threshold, or the probability of reaching a lower threshold within a specified amount of time (Dennis, Munholland et al. 1991; Morris and Doak 2002). A useful characteristic of these methods is that these risk metrics can be calculated with count-based models that can be parameterized with time-series of population estimates, some of the most common data available for threatened species. However, a potential bias exists in PVA when data are used to predict population behavior at unobserved densities. Inverse density dependence at low densities, the so-called Allee effect, is one phenomenon that may have implications in PVA (Stephens, Sutherland et al. 1999). Allee effects (Allee, Emerson et al. 1949), non-linearities in the density dependence function at low population densities, can have profound impacts on both the probabilities of extinction and on persistence times (Dennis 1989; Dennis 2002). Ignorance of an Allee effect can affect the estimation of model parameters such as growth rate (Stephens and Sutherland 1999) and consequently result in misestimating the population viability (Brassil 2001). In contrast, our purpose in this note is to demonstrate that unobserved non-linearities in population growth rate at high population densities may also have important influences on the extinction dynamics of populations and can greatly affect predictions of population risk.

When performing PVA using count-based data, it is first necessary to choose a model of population growth that best fits the data. Dennis and Taper (1994), Morris and Doak (2002), and Taper and Gogan (2002) give testing and model selection procedures for count-based data to discern between density independent and density dependent population growth models. The density dependent Ricker model (Ricker 1954) is a candidate model in all of these selection methods; however, Ripa and Lundberg (2000) found that in simulations with the Ricker model, extinctions were directly from population densities above equilibrium and described the Ricker model as having a 'pathological' extinction behavior. This behavior is the result of stochastic fluctuations in population growth pushing the population to densities far above equilibrium. Once at these extremely high densities, the Ricker model dictates a direct crash to extinction. But does this behavior represent the dynamics of real populations? Data suggest most real population extinctions occur after periods of negative trend, eventually bringing the population to small numbers (Ripa and Lundberg 2000).

In contrast to the Allee effect, we hypothesized there could be density dependent factors that only affect the population at very high densities. Such a mechanism may serve as a limit or upper bound to maximum population density though it does not influence 'normal' density dependent dynamics when the population is observed at densities near equilibrium. Such an upper bound would prevent the population from reaching unrealistically high densities predicted by the Ricker model. A possible biological mechanism for this phenomenon could be territorial behavior placing a limit on the maximum density of fish in a given area. Territory size has been shown as a

possible predictor of an upper limit to population density in salmonids (Grant and Kramer 1990). Territorial behavior can be conceptualized as a form of contest competition (Hassell 1975) that may ultimately limit population density; however, it may not be a prominent factor in population dynamics at observed population densities near equilibrium. A factor that affects population growth only at high densities may be difficult to detect in time series data but could lead to error in PVA predictions as population dynamic behavior at high densities will influence predictions both in simulation and analytic approaches. To evaluate the theoretical consequence of an upper bound on PVA procedures, we use a time series of population estimates to parameterize the Ricker model (Dennis and Taper 1994). PVA metrics are then calculated with simulations of the model with and without an upper bound on population density.

Methods

We used a stochastic generalization of the Ricker model relating N_{t+1} to N_t (Dennis and Taper 1994):

$$N_{t+1} = N_t \exp [a + bN_t + \sigma Z_t], \qquad (1)$$

where N_t represents population density at time t for $t = \{0,1,2\ldots\}$, a and b are constants representing the intrinsic growth rate and strength of intraspecific competition respectively, and σ represents variation in the observed growth rate due to random environmental shocks (commonly denoted as process variation). Z_t is assumed to have independent standard normal distribution. The model was fit to a time series of population estimates of cutthroat trout (*Oncorhynchus clarki*) abundances (Hamm and

Pearsons 2000). We note that for heuristic purposes, the Ricker model was purposely chosen as the 'best' model; a suite of models should be compared for any actual PVA application to determine which model has the best fit. As is common in ecological data, the time series is short, consisting of only nine observations, and is very stochastic, ranging from 15 to 251 fish/km. Due to high variation in the observed abundances, the estimate of process variation is very large (1.018). In model simulations, this high variation in the growth rate causes the population to reach extremely high densities that lead to a crash to extinction. To explore the potential consequence of a regulation mechanism outside the range of data, we added a reflecting upper bound on population density in simulations of the Ricker model as fit to the data. The upper bound is added as an approximation of effects a limit on maximum density, such as territoriality can have on population dynamics. In this case, the upper bound was arbitrarily chosen to be 1.2 times the maximum observed population density or 300 in this example. If any iteration in the simulation resulted in a population density above the upper bound, the population density was set to equal the upper bound. We note this theoretical modeling exercise is not intended to prove or disprove the existence of an upper bound in population density due to territoriality, or any other factor, in this population.

A lower limit of 2 individuals was set as a quasi-extinction level for model simulations. If any population density was below the quasi-extinction level, the population was considered to have gone extinct. To estimate the probability of quasi-extinction in 100 years, ten thousand simulations of 100 years were performed with and without the upper bound to population density and the percent of populations going

below the quasi-extinction level was recorded. In addition, simulations were performed until all populations went below the quasi-extinction level to examine the full distribution of time to extinction and distribution of population density one time-step prior to extinction.

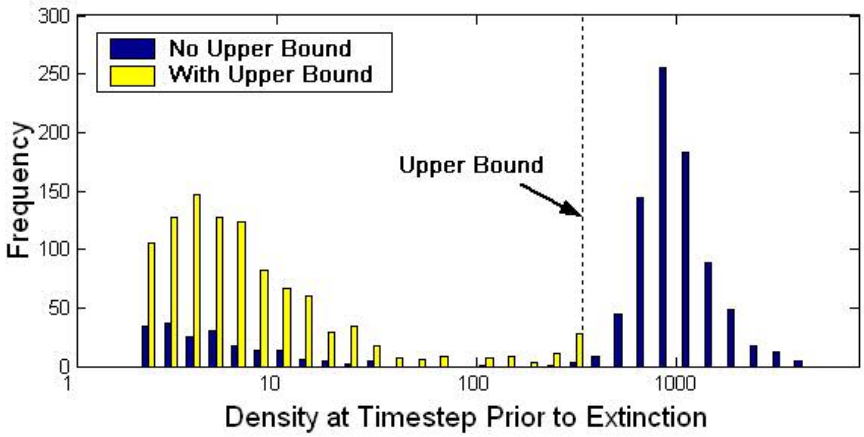
Results

The upper bound prevented the population from reaching extremely high densities in model simulations. Though the calculated equilibrium density was 106 fish/km, the population density could reach values as high as 5500 fish/km in simulations without an upper bound. Once the population reached extremely high densities, the population would then deterministically crash below the quasi-extinction level; indeed, the majority of extinctions in simulations without an upper bound occurred directly from densities far above the population equilibrium, typically over 500 fish/km. When the upper bound was incorporated in the simulations, extinction events were typically the result of a gradual decline in population density from an extended period of low realized discrete growth rates; the population density immediately prior to extinction was typically less than 10 fish/km. The distributions of population density immediately prior to extinction for simulations with and without an upper bound are given in figure 8a.

The presence of an upper bound decreased the probability of extinction and changed the distribution of time to extinction. The probability of extinction in 100 years decreased from 0.85 to 0.21 when an upper bound was included in model simulations. In the absence of an upper bound to population density, the majority of the populations

reached quasi-extinction threshold within 35 years with few persisting past 150 years; see Figure 8b. When the upper bound was added, the majority of populations lasted more than 250 years with some persisting more than 1000 years.

A.



B.

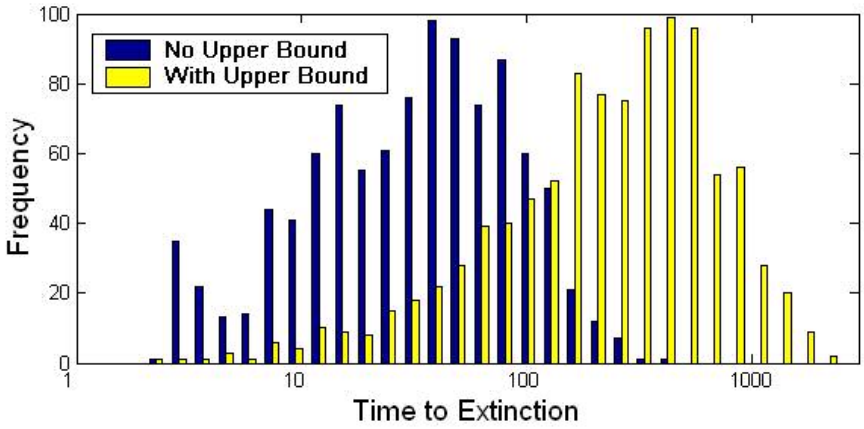


Figure 8. A) Distribution of population density at time step immediately prior to quasi-extinction and B) Distribution of time to extinction for simulations of Ricker model with and without an upper bound to population density. Model Parameter Values: $a = 0.9612$, $b = -0.00911$, $\sigma = 1.018$, Upper Bound = 300.

Discussion

In this analysis we have compared two Ricker-like models, however, we stress the importance of the model selection process in actual PVA. We assumed a Ricker model in this example for heuristic purposes; and make no claim this is necessarily the best model of this population. It is possible that a more complex model would better represent actual population dynamics if an upper bound actually existed for this population. In many conservation problems, however, limited data may not allow estimation of necessary parameters for complex models (Hilborn and Mangel 1997). In such cases, a simpler model can be used, but results may be biased because important factors in the population's dynamics are left out (Sen and Srivastava 1990). A simpler model can be useful, but it is important to recognize the limited inference the models provide; specifically, they may be good approximations of observed dynamics, but may not provide good predictions of population behavior outside the range of observed data or over long time intervals.

Fagan *et al*. (1999) used 'variation thresholds' to place species in broad classes of ecological dynamics to differentiate among species whose extinction risk is based on features such as carrying capacity, dispersal dynamics, or chance catastrophe. The variation thresholds were based on maximum amounts of process variation in model simulations that would allow the population to persist for a given intrinsic growth rate. Populations with a high intrinsic growth rate and high process variation were termed as 'dispersal dependent'; overcompensatory effects from high densities led to high extinction rates in such populations. The authors hypothesized dispersal dependent

populations could persist by breaking model assumptions about population dynamics at high densities (e.g. through lack of well mixedness or individual variation). The Ricker model was found to correctly identify dispersal dependent dynamics > 80% of the time with only 5-year time series (Fagan *et al.* 1999); our analysis of the cutthroat data places the population in the dispersal dependent category. However, there is an alternative hypothesis about how such a population can persist; instead of population dynamics departing from model assumptions in the response to high density (i.e. doesn't crash when the population density is high), there could be a biological limit to abundance that prevents the population from reaching high densities predicted by the model. Model predictions of population behavior at extremely high densities are questionable because 1) the population's response to such a high density has never been observed, and 2) densities this high have never been observed *in the first place*. This may lead to investigation of why a population may exhibit stochastic density dependent growth, but not reach densities predicted by the model. In contrast to density independent growth below a ceiling (Foley 1994), a population could undergo density dependent growth below the ceiling. Furthermore, the ceiling could range from a density at equilibrium to a density so far above equilibrium that it does not ever affect population dynamics. Our simulations with a density dependent model limited by a range of upper bounds point to possibilities of modeling population growth with decreasing variance, or even negative bias in the error, at high population densities.

Density dependent factors that only affect the population at high densities, if present, can greatly affect predictions of a PVA. In our simulations, the time to

extinction and the final trajectory to extinction were changed by an upper bound to population density. It is certain that extreme events or important factors in the population's dynamics can be missed in situations with very limited data. Models fitted with the data are then approximations of the observed dynamic. Subsequent analyses are susceptible to error when these models are used for predictions of population dynamics at unobserved densities. We have demonstrated that if a population threshold or similar dynamical phenomenon exists above the data range, the extinction probabilities calculated with a Ricker model can change importantly. Error resulting from uncertainty in model structure can be very serious but is often ignored; robustness to model structure is important as model selection error can lead to biases in the analysis (Chatfield 1995). Even when the correct model structure is known a priori, accurate long-term extinction predictions will require large amounts of data (Fieberg and Ellner 2000). Uncertainty in model structure further increases error in the predictions. Dennis and Taper (1994) successfully use an unaltered Ricker model; however, testing for density dependence in the dynamics of populations near equilibrium is a very different problem from selecting a model for population viability analysis. As Taper and Lele (2004) and Lindsay (2004) emphasize, the adequacy of a model depends entirely on the use to which it will be put.

One strategy to minimize effects of model selection error on PVA predictions is to limit the timeframe of predictions and the range of allowable population density. Previous studies have found short-term predictions to be more useful than long-term predictions (Dennis, Munholland et al. 1991; Foley 1994; Ludwig 1999; Fieberg and Ellner 2000). Short term predictions, e.g. 5-20 years, not only have lower variance than a

100 year prediction but the predicted population densities are much more likely to be in or near the range of observed data. This allows comparison of predictions to actual outcomes permitting investigation of sources of prediction error. Ideally this will lead to refinement over time of the model used for analysis. Another approach to increase the robustness of PVA is to predict the probability of reaching a lower threshold higher than extinction. A threshold near the lower values of the data used to develop the model would minimize error due to unforeseen changes in the dynamics at low densities. Another alternative would be to predict the probability of reaching a lower threshold before reaching an upper threshold (Dennis 2002). Threshold levels near the upper and lower observed values minimize possible error from changes in population growth dynamics at lower or higher densities that can't be discerned from the data.

Acknowledgements: We would like to thank Brian Dennis, Subhash Lele, and Brad Shepard for their helpful comments. Funding for this work was provided by Montana Fish, Wildlife, & Parks grant # Z0862 and Northwest Power and Conservation Council grant # PTMM11.

Literature Cited

Allee, W. C., A. E. Emerson, et al. (1949). Principles of animal ecology. Philadelphia, PA, Saunders.

Brassil, C. E. (2001). "Mean time to extinction of a metapopulation with an Allee effect." Ecological Modelling 143(1-2): 9-16.

Chatfield, C. (1995). "Model Uncertainty, Data Mining and Statistical Inference." Journal of the Royal Statistical Society A 158: 419-466.

Dennis, B. (1989). "Allee effects: population growth, critical density, and the chance of extinction." Natural Resource Modeling 3: 481-538.

Dennis, B. (2002). "Allee effects in stochastic populations." Oikos 96(3): 389-401.

Dennis, B., P. L. Munholland, et al. (1991). "Estimation of growth and extinction parameters for endangered species." Ecological Monographs 61(2): 115-143.

Dennis, B. and M. L. Taper (1994). "Density dependence in time series observations of natural populations: estimation and testing." Ecological Monographs 64(2): 205-224.

Fagan, W. F., E. Meir and J. L. Moore (1999). "Variation thresholds for extinction and their implications for conservation strategies." American Naturalist **154**(5): 510-520.

Fieberg, J. and S. P. Ellner (2000). "When is it meaningful to estimate an extinction probability?" Ecology 81(7): 2040-2047.

Foley, P. (1994). "Predicting Extinction Times from Environmental Stochasticity and Carrying Capacity." Conservation Biology 8(1): 124-137.

Grant, J. W. A. and D. L. Kramer (1990). "Territory size as a predictor of the upper limit to population density of juvenile salmonids in streams." Canadian Journal of Fisheries and Aquatic Sciences 47: 1724-1737.

Ham, K. D. and T. N. Pearsons (2000). "Can reduced salmonid population abundance be detected in time to limit management impacts?" Canadian Journal of Fisheries and Aquatic Sciences 57: 17-24.

Hassell, M. P. (1975). "Density-Dependence in Single-Species Populations." Journal of Animal Ecology 44(1): 283-295.

Hilborn, R. and M. Mangel (1997). The ecological detective. Princeton, NJ, Princeton University Press.

Lindsay, B. G. (2004). Statistical Distances as Loss Functions in Assissing Model Adequacy. The Nature of Scientific Evidence: Empirical, Statistical, and Philosophical Considerations. M. L. Taper and S. R. Lele, The University of Chicago Press.

Ludwig, D. (1999). "Is it meaningful to estimate a probability of extinction?" Ecology 80(1): 298-310.

Morris, W. F. and D. F. Doak (2002). Quantitative conservation biology: theory and practice of population viability analysis. Sunderland, Massachusetts, USA, Sinauer Associates, Inc.

Ricker, W. E. (1954). "Stock and recruitment." Journal of the Fisheries Research Board of Canada 11(5): 559-623.

Ripa, J. and P. Lundberg (2000). "The route to extinction in variable environments." Oikos 90: 89-96.

Sen, A. and M. Srivastava (1990). Regression Analysis: Theory, methods, and applications. New York, NY, Springer-Verlag New York Inc.

Stephens, P. A. and W. J. Sutherland (1999). "Consequences of the Allee effect for behaviour, ecology and conservation." Trends in Ecology & Evolution 14(10): 401-405.

Stephens, P. A., W. J. Sutherland, et al. (1999). "What is the Allee effect?" Oikos 87(1): 185-190.

Taper, M. L. and P. J. P. Gogan (2002). "The Northern Yellowstone elk: Density dependence and climatic conditions." Journal of Wildlife Management 66(1): 106-122.

Taper, M. L. and S. R. Lele (2004). The Nature of Scientific Evidence: A Forward Looking Synthesis. The Nature of Scientific Evidence: Empirical, Statistical, and Philosophical Considerations. M. L. Taper and S. R. Lele, The University of Chicago Press.

CHAPTER SIX

MODEL STRUCTURAL ADEQUACY: DEVELOPING BULL TROUT POPULATION MODELS FOR VIABLE POPULATION MONITORING

Abstract

Conservation decisions are often based on simple models that have been found to 'best' describe existing data. The suitability of any model, however, depends on how well it addresses user objectives. We present a model structural adequacy analysis to evaluate potential models for viable population monitoring of a threatened adfluvial bull trout (*Salvelinus confluentus*) population. Model evaluation criteria are not based solely on the model's fit to data, but also on how well the model is able to answer the scientific questions of interest concerning the population's risk of decline. While much is known about this bull trout population's biology, there is uncertainty about critical population processes and vital rates that could have large effects on adult abundances. We used available data and expert opinion to construct an array of mechanistic simulation models representing a range of hypotheses about the population's dynamics. The models were used to simulate population abundance data with which potential monitoring models were evaluated for how well they approximated the 'true' risk of decline for the simulated population. The measure of support for a monitoring model is how well it can estimate true risk over the range of population dynamics examined. The model structural adequacy analyses give insight into both model structure and data needed for adequately

accurate estimation of risk for a bull trout population. Our study suggests a density dependent Gompertz model parameterized with redd count data is likely to be overall the most effective available means for estimating average risk of decline for the bull trout population over a time horizon of several generations. Alternatively, a relatively simple demographic model incorporating data on juvenile vital rates and abundances potentially could enhance the ability to foresee imminent declines in adult abundances, albeit at considerably greater data costs.

Introduction

Conservation biology is often a discipline of crisis. In order to preserve threatened species, management decisions must be made with little available population data, and inadequate means or time to collect necessary data (Doak and Mills 1994). Lacking concrete information, decisions often are based on relatively simple models that, through a model selection procedure, have been found to 'best' describe existing data. Available data usually limit the ability to develop complex models that contain processes known to be important to the population. This results in uncertainty as to whether simple models provide useful predictions concerning the population of interest. The suitability of any model in a conservation context depends on how well it addresses the management question of interest and the relevant temporal and spatial framework (Pascual et al. 1997). For example, managers are often interested in a model's predictive capabilities, especially for population viability analysis (PVA) models used in many conservation biology applications.

Viable population monitoring is a risk-based monitoring framework in which comparisons of periodic PVA risk estimates are used for inference on population status (Staples et al. 2005). Because this monitoring strategy aims to detect problems early enough for managers to mitigate them, a model's explanatory power about past data is ancillary to its ability to predict short-term risk of decline. For many populations there is uncertainty about model complexity or data necessary to effectively estimate the risk of decline. Even if simple count-based PVA models (Morris and Doak 2002) can be used to estimate risk of decline, there will remain uncertainty about how risk estimates with the simple model compare to the population's unknown true risk of decline (Staples et al. 2005).

Two things are required for effective risk-based monitoring. First, a population model is needed that can adequately estimate risk of decline. And second, the requisite data for the model must be obtainable at reasonable cost. Common methods of model selection such as likelihood ratio tests (Neter et al. 1996) or Akaike's information criteria (de Leeuw 1992, Johnson and Omland 2004) are not helpful in this case because adequate data are not available. In this paper, we present a model structural adequacy analysis to guide model selection for risk-based monitoring of a threatened bull trout population. The goal of this analysis is neither an exact model of the population nor a prediction of risk, but insight into the structural complexity needed in a monitoring model and the data necessary to adequately predict risk of decline. Model evaluation criteria are not based just on fit to data, but on how well the model answers the scientific question of interest concerning the population's future status (Lindsay 2004).

As the actual risk for the Flathead population is unknown, we used mechanistic models of an adfluvial population to represent the 'true' population growth process for our study. Managers and biologists often know much about the population structure and life history of the species of interest even though there may not be sufficient data available to support or parameterize the complex population models implied by this knowledge. We developed a range of age-structured, spatially explicit process models based on information available in the literature, government reports, university theses, and expert opinion of Flathead bull trout field biologists (see acknowledgements) educed in a series of meetings. The process models, however, were not used for predicting the risk of decline for the Flathead population, but were used to simulate population data with which potential candidate models were evaluated for monitoring effectiveness (i.e. how well the relatively simple candidate models approximated the risk of decline for the 'true' complex, age-structured simulation models). A potential monitoring model is supported according to how well it estimates risk for the various process models.

A systems analysis approach, first established in engineering sciences (Forrester 1961), was used to develop the full process models. Systems analysis is a procedure for studying complex systems like adfluvial population dynamics through the construction and analysis of mechanistic simulation models (Hall and Day Jr. 1977). Life-history analyses have used this approach (Coulman et al. 1972, Ruesink 1976, Logan 1988), and Logan (1994) argues such "big ugly models" offer several advantages for understanding ecological systems. Simulation models provide flexibility for testing hypotheses regarding mechanistic pathways and interactions, and allow researchers to evaluate

effects of multiple causal pathways. More importantly for monitoring purposes, mechanistic simulation models can be used to assess the behavior of the population over time resulting from interactions among age classes in the adfluvial life history. Practically speaking, simulation models can elucidate emergent properties of abundance data arising from interactions of juvenile cohorts and their environments. Further, simulated data from relatively complex process models can be used to evaluate whether simpler models can adequately estimate the true risk of decline derived from the complex process models.

The utility of complex simulation models, compared to simpler analytic models, has been questioned (Berryman 1991, Liebhold 1994). An important reason for disenchantment with complex models (Logan 1994) is they often were practically unparameterizable, and never came close to meeting expectations generated by early proponents. To counter this, goals and limitations of a simulation analysis must be clear and realistic. Simulation models also have lacked expected predictive powers. While this may be due to flaws in the modeling process as Berryman and Liebhold contend, Logan (1994) suggested that the lack of predictive power in complex models of ecological systems might be a manifestation of the unpredictability of nature itself. The most serious problem with simulation modeling efforts has been limited application, testing, and revision of these models. Many simulation models have been the result of short-term investments with the model as the goal. Experience indicates, however, that the most useful models have undergone an evolution of thought, and have a team of modelers and ecologists who are committed to understanding the system rather than just

developing the model per se (Logan 1994). Thus, we strongly stress the importance of data collection and further model development to increase understanding about bull trout population dynamics and provide better inference regarding the population's status over time.

While much is known about this bull trout population, there is uncertainty about some critical processes in the population that potentially could have large effects on adult abundances. In the adfluvial life history, juvenile age-classes, or cohorts, spend 2-4 years in natal tributaries before immigrating to habitats in the Flathead River and Lake (Shepard et al. 1984). Inter-cohort density dependence among juveniles in natal tributaries has been demonstrated for an adfluvial bull trout population in Alberta, Canada (Paul 2000, Paul et al. 2000), but the exact nature of these interactions is uncertain. The accuracy of risk estimates may critically depend on the way density dependence is modeled, and it has been recommended that studies should investigate a series of scenarios of density dependence appropriate to the species of interest in order to understand the role of density dependence on the population (Henle et al. 2004). A useful feature of this simulation approach is that hypotheses about juvenile dynamics can be evaluated in terms of their ultimate effect on model selection and monitoring inferences.

We used three different sub-models of contest and scramble competition (Hassell 1975) among juvenile cohorts in the process models to evaluate the effects of density dependence on monitoring inference. Contest competition results when resources are monopolized by a fixed number of individuals with remaining individuals receiving no

resources. Scramble competition results when resources are partitioned equally among individuals; higher abundance lowers the mean survival rate for all individuals. The process models represent a range of juvenile dynamics that may be present in the Flathead, and simulations with these models should help determine whether differences in juvenile dynamics can have important effects on patterns of adult abundance, model selection, or inference to population status.

We used 3 count-based and 2 demographic age and stage models as candidate monitoring models. These models were used to estimate the probability of decline with simulated abundance data from the process models. Candidate model risk estimates were then compared to the 'true' probability of decline calculated with the process model. The count-based models can be used with existing data series in the Flathead, so determining the effectiveness at approximating risk would improve current inference to population status. The demographic models utilize vital rate information like survival and fecundity in addition to age or stage abundances to predict future adult abundance. Techniques derived from composite modeling (Logan 1989) were used to develop demographic candidate monitoring models from the more detailed process models. Composite modeling utilizes mathematical procedures and ecological intuitions for developing simpler models that capture the dynamics, or essence, of the system of interest. In this manner, simpler monitoring models can be linked to the biologically realistic process models of the bull trout population.

Our goal is to determine how to estimate risk for a population whose dynamics are a relatively unknown complex process. Unfortunately, lack of data limits our ability

to fit and select population models with traditional means. The model structural adequacy analyses are part of an inductive model selection process in which candidate monitoring models are evaluated on whether they can usefully approximate the risk of decline for a complex, age-structured population growth process. If a candidate model gives adequate risk predictions, then that is support for that model being useful in application. It certainly is not a guarantee, however, that the model will work because the actual bull trout population dynamics are unknown. To build further support for the monitoring model, it must give adequate risk estimates for different representations of the bull trout population. A well-supported model will be useful for estimating risk for a wide range of potential bull trout dynamics. If a model fails to give adequate estimates, that is evidence the model may not be useful for monitoring use, or at least there are circumstances in which it will be ineffective for estimating risk. The model structural adequacy analyses will help evaluate the likelihood that simple models are adequate for monitoring (i.e. they work well for the range of process models), or whether it is possible to improve risk estimates with more complex models. The MSA analyses are not intended to elucidate the exact processes for the Flathead population, but to give insight on what model structure and data are needed for adequate approximations of risk for adfluvial bull trout populations.

Methods

Stochastic matrix models patterned after known features of the Flathead bull trout population were used to simulate adult abundance data, i.e. summed number of

individuals 6-9 years of age. Matrix models are covered in depth by Caswell (2001), while Morris and Doak (2002) provide a more application-oriented introduction. In this simulation analysis, all risk predictions were the probability of adult abundances declining below an a priori lower threshold of 100 within 5 (PLT5) or 10 years (PLT10). Adult bull trout represent the portion of the population that has completed the full adfluvial life history, and have the highest reproductive values in the population. Estimates of effective population size (Caballero 1994) for bull trout have been based on the number of adults reproducing each year (Rieman and Allendorf 2001). More practically, adults are the segment of the population represented by redd count data. Redd count data have been collected consistently for over 25 years for this bull trout population and redd counts represent an ideal way to monitor bull trout populations because counting redds is relatively easy and non-intrusive. Additionally, redd numbers can be estimated with moderately high precision (Muhlfeld et al. 2006), and may be more indicative of trends in adult numbers than other available data series such as lake netting data (Staples et al. 2004b).

The Flathead population, which includes the entire basin above Flathead Lake with exception of the South Fork of the Flathead River, is a combination of several sub-populations (Shepard et al. 1984). The basic process model is a spatially explicit, age-structured matrix model of one such adfluvial bull trout sub-population. The model (fig.9) contains ten age classes (0-9) with juveniles (age 0-3) located in the natal tributary, except for juveniles that have emigrated from their natal tributary. All older age classes reside in the river/lake portion of the basin. Because little is known about the

true correlation structure among vital rates in the Flathead, we used uncorrelated vital rate distributions (Table 2) for our simulations.

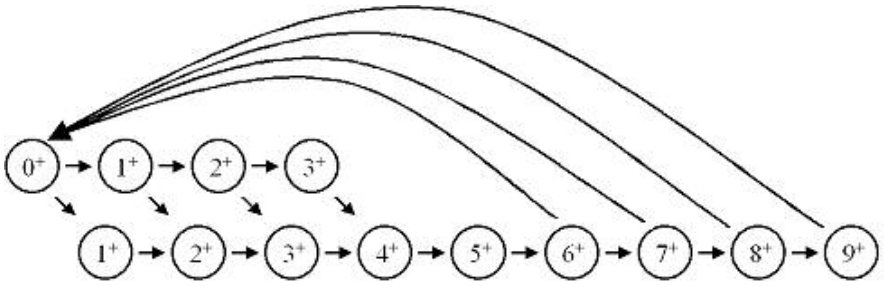


Figure 9. Basic age-structured process model of the Flathead bull trout population. Circles represent age-classes. . Upper row represents juveniles in natal tributary habitat; lower row represent juveniles and adults in River/Lake habitat.

Three models of density dependent interactions among juvenile cohorts were incorporated into the basic process model to evaluate effects of hypothesized dynamics on monitoring inference. All density dependent functions used the sum of the squared fork-lengths (distance from the tip of the snout to the center of the fork in the tail, denoted FL^2) of all individuals in the tributary as an index of abundance. This measure compensates for increases in resource consumption for larger individuals in density dependent interactions (Walters and Post 1993, Post et al. 1999). Average lengths for age 0-3 juveniles (Shepard et al. 1984) were used for calculating FL^2. The different density dependence sub-models in the process models were designed to see how changes in juvenile dynamics will affect abundance patterns, and whether that affects risk estimates with the candidate models. One sub-model had contest competition for young-of-the-year (YOY) recruitment based on tributary occupancy in which YOY filled the tributary to capacity, and once capacity had been reached, all additional YOY were assumed to emigrate from the tributary to river and lake habitats with very low YOY survival

probabilities. The survival rates of age 1, 2, and 3 juveniles were density independent. We denote this process model as YC-JI. The second process model, denoted YC-JS, also used contest competition for YOY recruitment, but older juveniles had scramble competition that altered mean survival rates while maintaining constant variance in the vital rate distribution. The third process model, denoted YS-JS, had scramble competition affecting mean rates for both YOY recruitment and juvenile survival rates.

We evaluated the ability of five candidate models to estimate the probability that adult abundance will fall below a lower threshold of 100 within 5 (denoted as PLT5) and 10 years (PLT10) using data simulated from the three age-structured process models. We deliberately used candidate models with different structure from the process models for insight into how well we could estimate risk with imperfect knowledge of the complex population dynamics. Three of the candidate models allowed estimates of risk to be calculated directly from simulated time-series of adult abundances. These models are discrete-time versions of the density independent exponential growth model (DI) (Dennis et al. 1991), a Ricker density dependent model (RD), and a Gompertz density dependent model (GD). Parameter estimation from time-series data for all three models is described in (Dennis and Taper 1994).

Vital Rate	Process	SG	AS
YOY emergence	beta(7,7)	beta(7,7)	beta(7,7)
YOY survival, Trib	beta(5,7)	beta(5,15)	
YOY survival, R/L	beta(.5,20)		
Age 1-3 Survival, Trib	beta(5.5,5.5)	beta(5.5,5.5)	beta(5.5,5.5)
Age 1 emigration	beta(5,17)	beta(5,17)	

Age 2 emigration	beta(6,4)	beta(6,4)	
Age 3 emigration	beta(9,.1)	beta(9,.1)	
Age 1 Survival, R/L	beta(1,17)		
Age 2 & 3 Survival, R/L	beta(3,7)		
Age 4 & 5 Survival	beta(5,7)	beta(5,7)*	beta(5,7)*
Shifted Low	beta(4,10)	beta(4,10)*	beta(4,10)*
Shifted High	beta(12,4)	beta(12,4)*	beta(12,4)*
Age 6 Survival	beta(10,7)	beta(16,11)*	beta(16,11)*
Age 7 Survival	beta(16,7)		
Shifted Low	beta(10,20)	beta(10,20)*	beta(10,20)*
Shifted High	beta(20,10)	beta(20,10)*	beta(20,10)*
Age 8 Survival	beta(16,5)		
Spawner Survival	beta(20,13)	beta(20,13)	beta(20,13)
P(Spawn) age 6	beta(15,15)	beta(15,15)*	beta(15,15)*
P(Spawn) age 7	beta(13,6)		
P(Spawn) age 8	beta(15,15)		
P(Spawn) age 9	beta(15,15)		
Fecundity age 6	N(2318,400)	N(4700,400)*	N(4700,400)*
Fecundity age 7	N(3589,500)		
Fecundity age 8	N(5454,600)		
Fecundity age 9	N(7631,600)		
Tributary capacity	1.0 x 10^9 (FL2)	15000 (indiv)	1.0 x 10^9 (FL2)
proportion Female	0.5	0.5	0.5

Table 2. Vital rate distributions for process models and demographic candidate monitoring models. * indicates stage-specific vital rates in SG and AS models. FL2 denotes tributary capacity is measured by the sum of the FL^2 for all juveniles in tributary.

The remaining two candidate models, a stage (SG) and an age-stage model (AS), used demographic information on vital rates and annual age class abundance estimates to make population predictions. These models were developed by aggregating components of the basic age-structured model through elasticity and life stage simulation analyses

(Wisdom et al. 2000, Caswell 2001) guided by biological intuition. Vital rate distributions for the SG and AS models are given in Table 2. The stage model (fig.10) had three stages representing juveniles age 1-3, sub-adults age 4-5, and adults age 6-9. The SG model assumed a pre-birth census, thus, recruitment and survival rates for YOY were incorporated into the Adult – Juvenile transition (individuals show up in the juvenile stage immediately prior to their first birthday). Population abundance in the SG model was limited by an upper bound placed on the juvenile stage, simulating a contest type competition affecting juvenile recruitment.

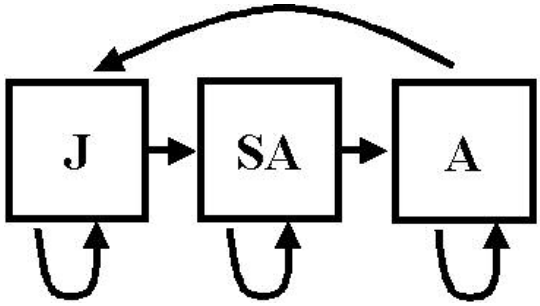


Figure 10. Stage (SG) candidate monitoring model. Boxes represent life-history stages: J = Juvenile (age 1-3), S = Sub-adult (age 4-5), A = Adult (age 6-9).

The AS model (fig.11) used a post-emergence census, and had explicit age classes for age 0-3 in the tributary with stages for sub-adults and adults. Mean YOY recruitment rates varied according to the abundance of age 1-3 juveniles in the AS model. The AS model did not use data on age 1-3 juveniles in the river/lake basin, only juveniles in the tributary were used for risk estimates. Vital rate distributions for the AS model are given in Table 2.

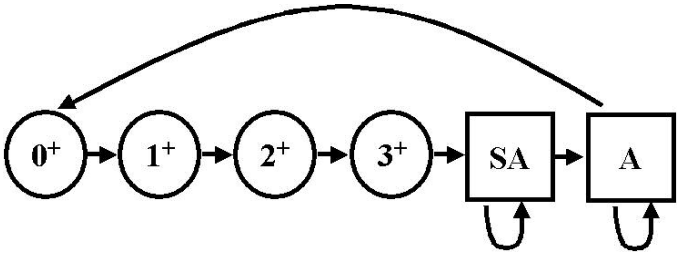


Figure 11. Age-Stage (AS) candidate monitoring model. Circles represent juveniles in natal tributary, squares represent sub-adults (SA) and adult (A) stages.

The count-based candidate models have several sources of variation in PLT10 estimates: 1) errors resulting from structural differences between the model and the true process, 2) sampling error in model parameter estimates, and consequently predictions, due to observing the population for a finite number of time periods, and 3) variation due to measurement error in abundance data. We focused on the first two sources of variation in count-based model predictions for this study. Methods to accommodate measurement error in abundance data for the DI and GD are described in Staples et al. (2004) and Dennis et al. (2006); measurement error in the RD model is described in (De Valpine and Hastings 2002). To evaluate structural error for count-based models (DI, RD, GD), each of the process models were used to generate 1000 time-series. The length of each time-series was 1000 to minimize sampling error in parameter estimation. The fitted candidate models were used to estimate PLT5 and PLT10. From the end of each data series, the process model was restarted 1000 times to accurately estimate the 'true' PLT5 and PLT10. To evaluate how the count-based models would perform with more realistic data series, the process models were used to generate 1000 time-series of adult abundance (without measurement error) of length 20 with which the parameters for the count-based models were estimated, and the PLT5 and PLT10 calculated. The 20-year series with no

measurement error somewhat represents a “best-case” scenario for how well the count-based models perform in application. Measurement error in abundance data would introduce additional variation into PLT estimates. As above, the process models were restarted from the end of each series for the ‘true’ probability reaching the lower threshold.

Variation in candidate demographic model estimates is caused by: 1) structural differences between the model and the true process, 2) measurement error in initial age/stage abundance estimates, and 3) uncertainty in candidate model vital rate estimates. The candidate demographic models were started with abundances from the last time-step of the simulated data series, and run 1000 times to estimate the PLT5 and PLT10. To evaluate structural errors in estimates, initial age and stage abundances in the demographic models were the exact abundances from the simulated data series. To represent sampling error in abundance estimates, initial stage and age abundances were contaminated with an error term resulting in coefficient of variation (CV) of 0.25 for initial abundances. As the goal for this study was to evaluate the effects of structural differences on demographic model estimates, we did not evaluate the effect of error in candidate model vital rate estimates. In general, higher uncertainty in vital rates will increase the estimated risk of decline (Staples et al. 2004b).

Though we did not check vital rate estimation errors for demographic models, we evaluated how large changes in the assumed vital rate distributions in process models affected the performance of candidate models. We used broad distributions of vital rates in the demographic models representing a range of plausible values because vital rates for

the Flathead population are poorly known. Error in the specification of vital rate distributions for the process models could potentially affect the qualitative results of this model adequacy analysis. For insight into how candidate model predictions could be affected by large changes in vital rates, we altered the distributions of two vital rates, sub-adult and adult survival, that sensitivity analyses have shown were influential on the probability of reaching the lower threshold. Two alternative distributions were specified for each of the survival rates, one with higher mean and one with lower mean (Table 2). Both distributions had reduced variance. Data were simulated with the three process models using four combinations of the modified survival rate distributions (i.e. both survival rates high, both low, one high and one low, and vice-versa). The simulated data were used to evaluate candidate model risk estimates as described above except the demographic models used the correct alternative survival rate distributions.

The above analyses evaluated differences in the distributions of candidate model risk estimates from the true risk for the simulated age-structured populations. Inference to population status in VPM, however, is through comparisons of a single population's risk estimates over time. The DI, GD, SG, and AS models were selected for further VPM simulation analyses. To evaluate how monitoring with the candidate models might perform in application, we simulated time-series with the process models for 10 years, and calculated the PLT5 and PLT10 with the candidate models. The process model simulation continued for 40 additional time-steps with updated estimates of the PLT5 and PLT10 calculated with the count-based models at each time-step with the entire data series up to that point. Initial abundance estimates had CV = 0.25 for the SG and AS

models. For every year of the last 40, we estimated the 'true' process model PLT5 and PLT10 by restarting the process model at that point 1000 times and continuing the simulations for 10 years. We compared the series of candidate model risk estimates with those calculated with the process model. These simulations demonstrated how candidate model risk estimates could change over time compared to the 'true' risk, and how VPM might signal an increased risk of decline in abundance.

Results

The distributions of thePLT5 and PLT10 for process and candidate models using the basic vital rate distributions are given in figures 12 and 13 respectively. For all three process models, the distribution of risk had a strong positive skew. The density independent model (DI) had a symmetric distribution of risk estimates, and tended to overestimate the risk of decline. Risk predictions from the DI model also had low correlation to true risk from process models (table 3). While the GD model estimated true risk much better than the RD model for the YC-JI sampling simulations, both density dependent count-based models gave similar PLT estimates for the YC-JS and YS-JS sampling simulations. PLT estimates from the GD model had higher correlation to true risk than those from the RD model in all simulations. Sampling error in parameter estimation with the more realistic 20-year time-series increased the variability of risk estimates for the RD and GD models.

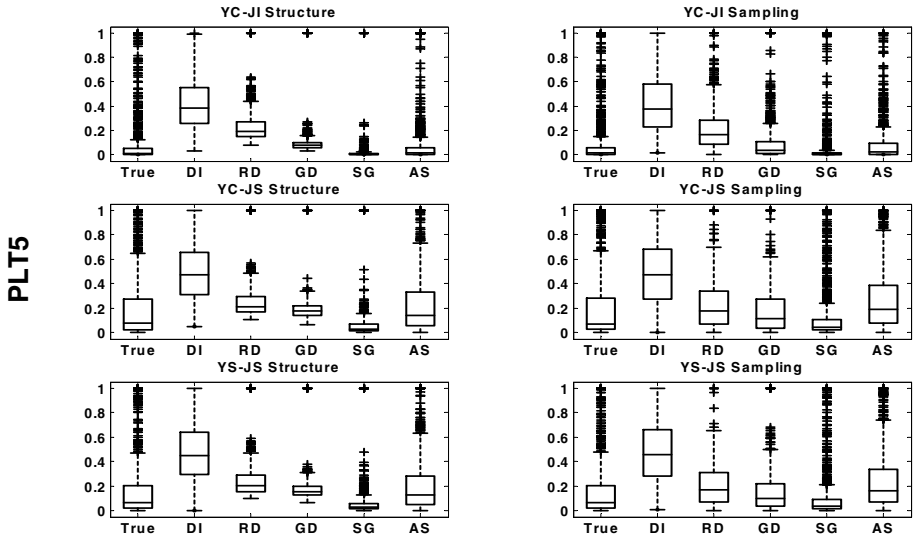


Figure 12. Estimated PLT5 for structural and sampling simulations. Models: True – Risk of decline calculated with process model, DI- density independent, RD- Ricker density dependent, GD- Gompertz density dependent, SG- Stage, AS- age-stage.

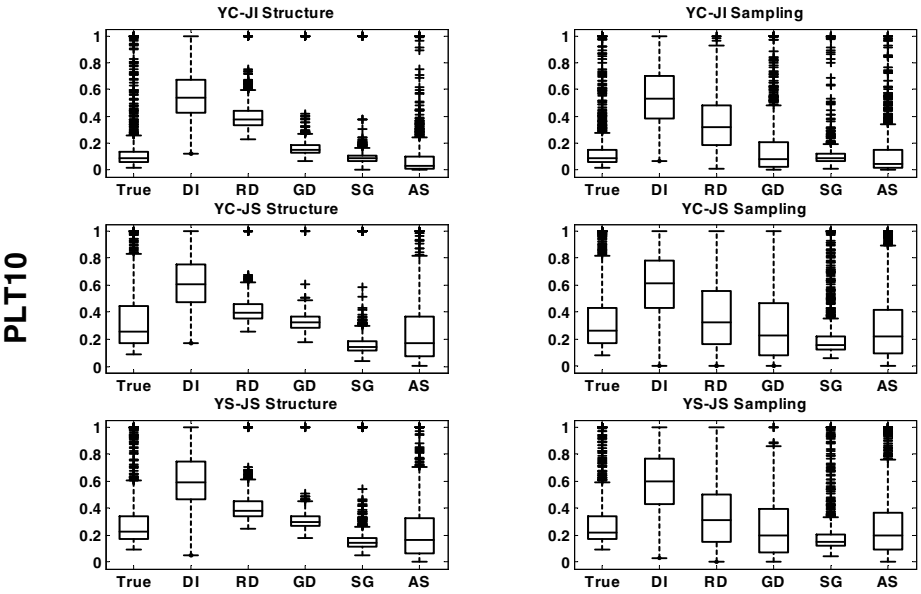


Figure 13. Estimated PLT10 for structural and sampling simulations. Models: True - process model, DI- density independent, RD- Ricker density dependent, GD- Gompertz density dependent, SG- Stage, AS- age-stage.

		CORRELATION WITH TRUE PLT10				
		DI	RD	GD	SG	AS
YC-JI	Structure	0.27	0.46	0.55	0.59	0.68
	Sample	0.31	0.35	0.45	0.55	0.63
YC-JS	Structure	0.39	0.58	0.59	0.62	0.75
	Sample	0.38	0.45	0.47	0.60	0.73
YS-JS	Structure	0.40	0.62	0.64	0.64	0.72
	Sample	0.37	0.43	0.44	0.62	0.71

Table 3. Correlation between candidate model risk predictions and true PLT10 from process models.

PLT estimates from the SG model had positive skew like the true risk distribution (figs. 12 & 13), though they tended to be biased low and less variable than the true risk. SG estimates had higher correlation with the true risk than any of the count based models, but had lower correlation than the AS model (table 3). The distribution of estimates of PLT5 and PLT10 from the AS model were close to the true distribution of risk for the process model simulations, and had the highest correlation to the true risk of decline. Sampling error in abundance estimates had little effect on risk estimates from the SG and AS models.

Changes in the assumed sub-adult and adult survival rate distributions affected the adequacy of candidate model predictions. When both survival rates were lowered, the true PLT10 was practically 1 for all process models, and all candidate models correspondingly showed high risk for decline. When sub-adult survival rates were decreased and adult survival rates increased, the true PLT10 was near 1 but with a wide negative skew, and all candidate models performed adequately (fig. 14). For simulations in which sub-adult survival rates were increased, however, the SG and AS models

underestimated the PLT10 regardless of the adult survival rate. When both survival rates distributions were high, the true risk of decline was low but positively skewed (fig. 15). The GD model slightly underestimated risk in these simulations, but the SG and AS models rarely showed any risk of decline. If sub-adult survival was high, but adult survival was low, the true PLT distribution was wide and positively skewed (fig. 16). The GD model gave adequate estimates of risk, though they were less variable than the true risk of decline. Again, the demographic candidate models rarely showed any measurable risk.

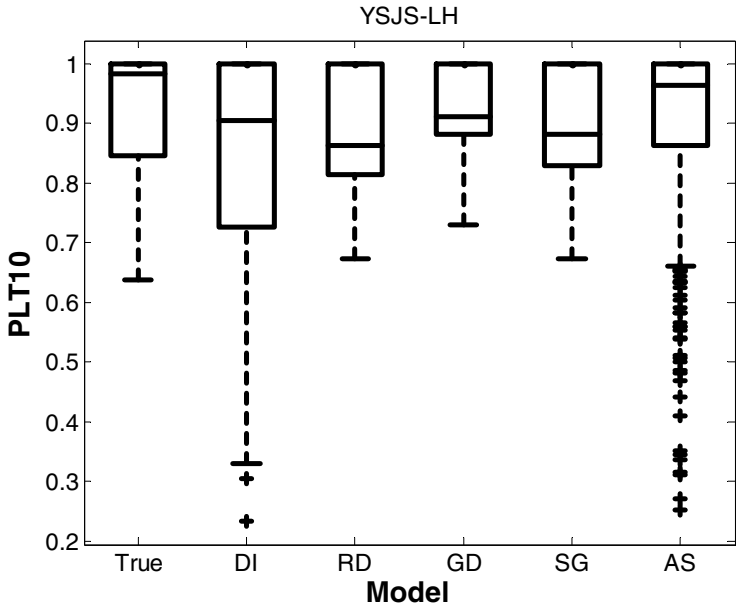


Figure 14. Distribution of PLT10 estimates from YSJS process model with sub-adult survival rates decreased and adult survival rates increased. True = risk calculated with process model, DI = density independent model, RD = Ricker density dependent model, GD = Gompertz density dependent model, SG = Stage model, AS = Age-Stage model

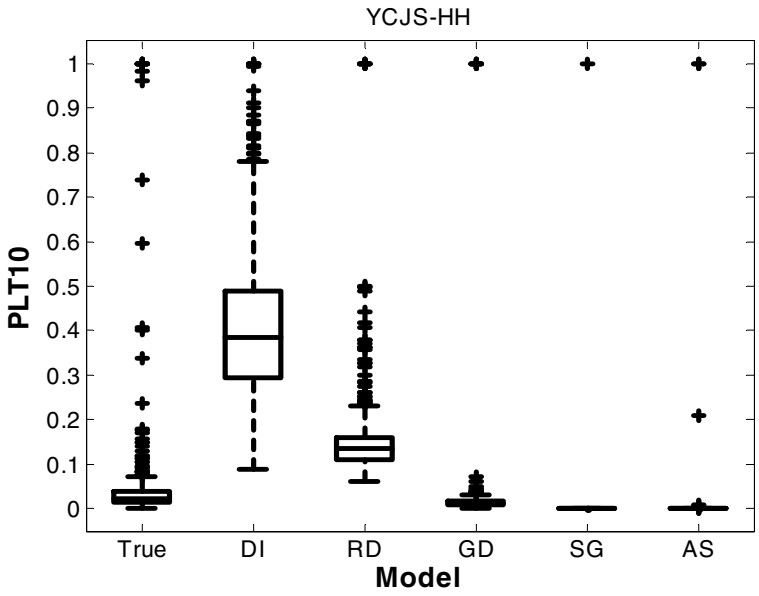


Figure 15. Distribution of PLT10 estimates from YCJS process model with both sub-adult and adult survival rates increased. True = risk calculated with process model, DI = density independent model, RD = Ricker density dependent model, GD = Gompertz density dependent model, SG = Stage model, AS = Age-Stage model

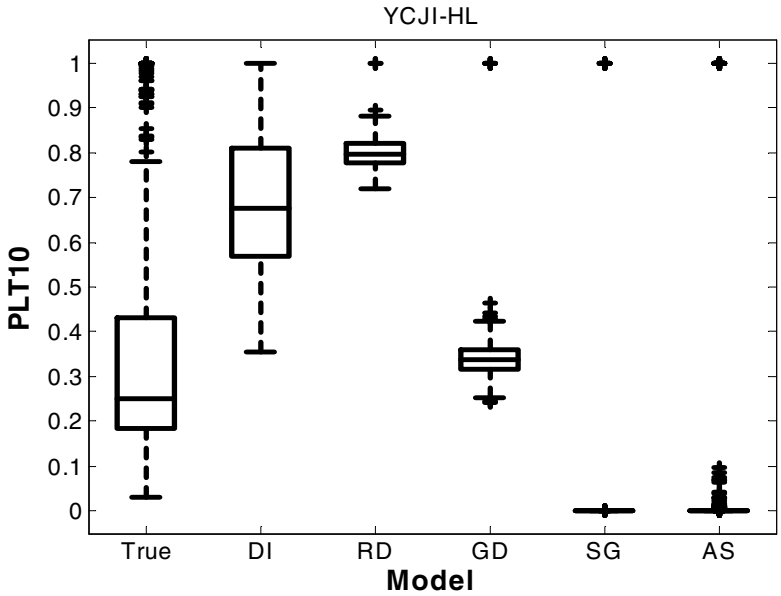


Figure 16. Distribution of PLT10 estimates from YCJI process model with sub-adult survival rates increased, and adult survival rates decreased. True = risk calculated with process model, DI = density independent model, RD = Ricker density dependent model, GD = Gompertz density dependent model, SG = Stage model, AS = Age-Stage model

Time-series simulated with the YC-JI process model showed large fluctuations in adult abundances (figs 17 & 18). The addition of density dependence among older cohorts moderated the fluctuations in adult abundance for the YC-JS and YS-JS process models. All process models showed relatively low and stable risk for long periods, but also had sharp increases in the PLT in a short amount of time. The DI model tended to show high risk of decline throughout the time-series despite the level of true risk. In the YS-JS time-series simulation, the GD model overestimated risk at the beginning due to the initial low abundance (around time-step 6), but gave a reasonably good estimate of average risk of decline for the other simulations. The GD model tracked the true risk as well as the SG demographic model, but both models failed to clearly signal increased risk before observed declines. For these simulated time-series, the AS model tracked the true risk of decline, and provided some warning of increased risk prior to declines. In the YC-JI time-series simulations, the AS showed a clear spike at time 38 prior to the decline below the lower threshold at time 43, however, risk estimates decreased for time-steps 39-42. The GD and SG models also showed a slight spike in PLT at time 38 and decreased in a similar manner. Unlike the DI and GD models, which primarily showed higher risk when current adult abundance was low, the AS model showed increased risk of decline in adult numbers when current abundances were high (e.g. time step 37 in YS-JS simulation).

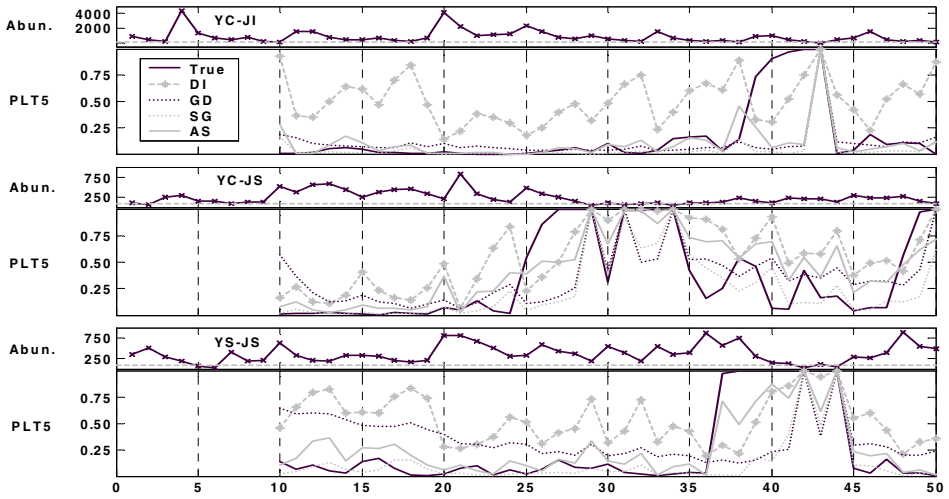


Figure 17. PLT5 time-series simulations. YC-JI, YC-JS, and YS-JS graphs: Adult abundance simulated with process models; dashed gray line is lower threshold (100). PLT5 graphs are risk of decline for time series immediately above calculated with: True= process model, DI= density independent, GD= Gompertz density dependent, SG= stage, AS= age-stage model.

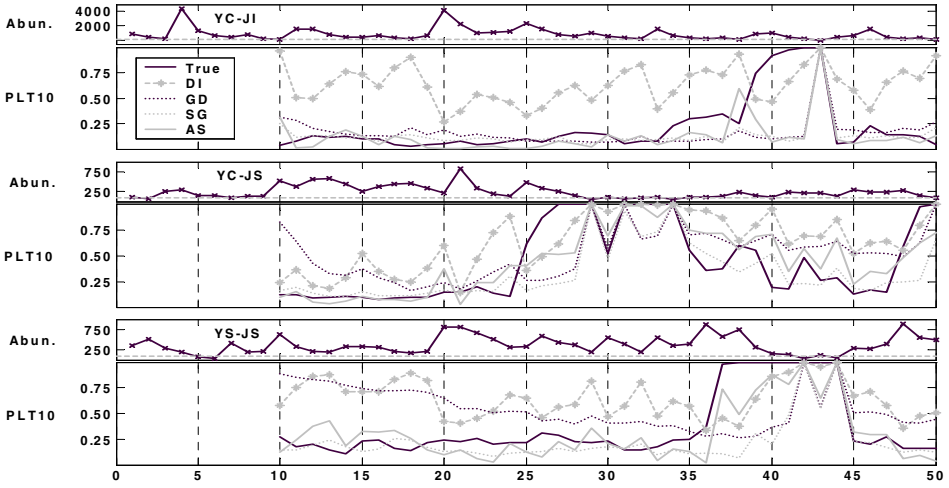


Figure 18. PLT10 time-series simulations. YC-JI, YC-JS, and YS-JS graphs: Adult abundance simulated with process models; dashed grey line is lower threshold (100). PLT10 graphs are estimated risk of decline for time series immediately above calculated with: True= process model, DI= density independent, GD= Gompertz density dependent, SG= stage, AS= age-stage model.

Discussion

A fundamental goal for this study was to evaluate the inference on population status provided by a range of simple candidate models. To address this, we used a range of mechanistic simulation models to test whether simple models provided adequate estimates of the risk of decline. Candidate models are supported based on how well they estimate risk for a range of hypothesized dynamics. Interpretation of the results of this study, especially as they apply to bull trout monitoring in the Flathead, should consider how well the mechanistic process simulation models represent a true adfluvial population. The process models assumed uncorrelated vital rates; though it is certainly possible vital rates are correlated for the Flathead population. Cross correlations among vital rates can have important effects on population dynamics (Fieberg and Ellner 2001). Age-structure and inter-cohort dynamics, however, may have larger effects on fluctuations in adult abundances. Preliminary analyses of within-year vital rate correlations in the process models showed little effect of cross correlations in vital rates on the probability of decline. For example, perfectly correlated survival, emigration, and fecundity rates had little effect on risk of decline for the YC-JI and YC-JS models (Staples et al. 2004b). Nonetheless, the effects of potential correlations among vital rates should be evaluated once more data are available for the Flathead population.

Though differences in juvenile dynamics have important effects on patterns of adult abundance and risk dynamics, this study offers general insights into how the candidate models might perform for monitoring an age-structured population like bull trout. The GD model gave good estimates of the average probability of decline over the

broad range of process models examined, and thus may be useful for monitoring populations where time-series data are available. The RD model overestimated the probability of decline and gave imprecise estimates of risk, especially for the YC-JI process model. An upper limit to abundance may reduce bias in risk estimates from the RD model (Staples and Taper 2006), but it cannot easily accommodate sampling error in abundance data so it may have limited monitoring use (see DeValpine and Hastings (2002)). While the DI and GD models can accommodate sampling error in abundance estimates (Staples et al. 2004a, Dennis et al. 2006), uncertainty in model parameter estimates will be increased by sampling error in abundance data (Ferrari and Taper 2006). Additional uncertainty in model parameter estimates may be problematic considering the short nature of typical time-series; however, auxiliary information about model parameters may improve the reliability of risk estimates from short data series (Hakoyama and Iwasa 2000, Ferrari and Taper 2006).

Models incorporating age-structure for juveniles potentially could increase the ability to detect problems in bull trout populations, and may be useful for populations for which there are no historical time-series data available. Fortunately, these models appear robust to sampling error in estimates of age-class abundances. Thus, useful age-class abundance estimates could probably be calculated from length-frequency information rather than using scale samples, which are unreliable, or some other hard body structure that would require sacrificing fish. While the AS model performed well for simulations with average vital rate distributions, large shifts in the sub-adult and adult survival rate distributions detrimentally affected the adequacy of demographic model risk predictions.

We found higher sub-adult survival rates resulted in the SG and AS candidate models greatly underestimating the PLT10 for all three process models, likely because the single sub-adult stage with high survival rates keeps the adult stage abundance artificially large. Thus, predictions from the demographic models, using the 'true' vital rate distribution, were inadequate due to structural differences between the model and true process. Splitting the sub-adult stage into two age classes, while using the same survival rate, may improve the AS model abilities to anticipate declines in adult abundance. The sensitivity of risk predictions to sub-adult survival rates suggests caution is warranted in implementing demographic models for monitoring the population. Ultimately, whether an age-structured model is worth the additional effort will also depend on if managers have the resources to effectively intervene when age-class imbalances foretell a near term decline in adult abundance.

In the time-series simulations, all models showed some level of risk prior to the population declining. These are estimates of a probability of decline, and there is no guarantee of future abundance. Considering the short-term nature of monitoring predictions, however, significant probability of decline within 10 years could certainly be cause for conservation actions or research toward understanding the true risk facing the population. Levels of estimated risk may be reduced by improving population status, or by improving precision of risk predictions (or both). For example, the DI VPM analysis for the bull trout population (Staples et al. 2005) described above indicated a high risk of decline. The present study demonstrates DI risk predictions can overestimate risk for an

age-structured population, but supports the use of the GD model over a wide range of hypothesized dynamics.

The GD model, when applied to the Flathead index redd data series with the Dennis et al. (2006) method, also indicates a high probability of declining below the lower threshold of 125 within 10 years (fig. 19). The threshold of 125 roughly corresponds to 1000 adults spawning in the system (Staples et al. 2005), the estimated abundance necessary for bull trout populations to maintain adaptive genetic variation (Rieman and Allendorf 2001). Note this data series is a composite of redd counts in 8 index tributaries, not a single population like the process models in this study. Nonetheless, given the high estimates of PLT10 and that redd counts are currently near the lower threshold, there is certainly cause for management concern about the future of this population.

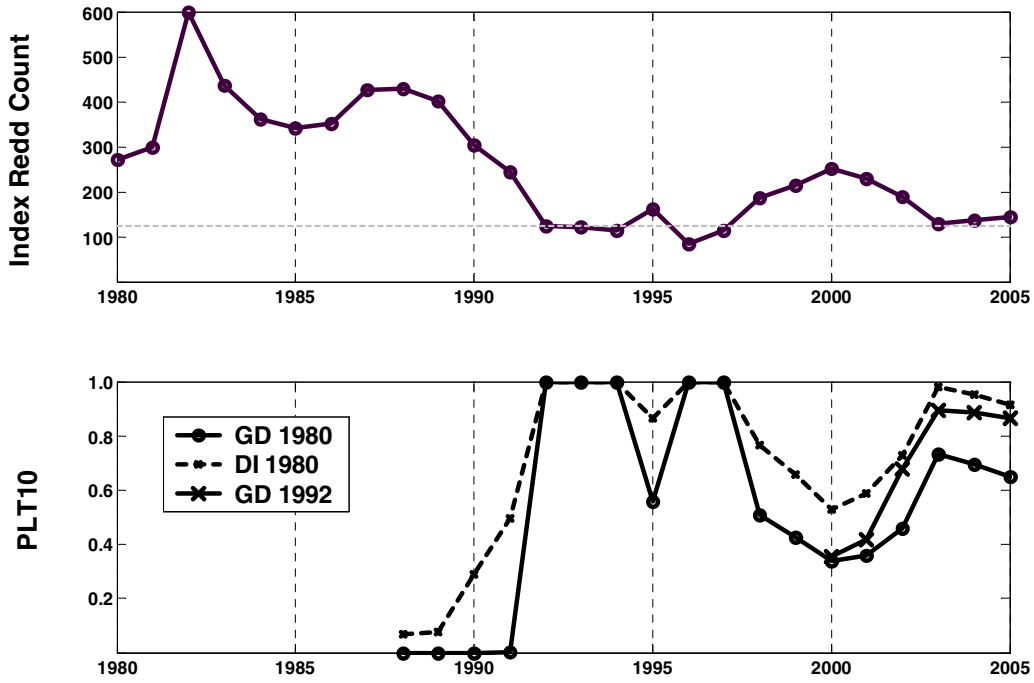


Figure 19. PLT10 estimates for the Flathead bull trout calculated with DI and GD models using index redd count data. Top panel: solid line = index redd count, dashed line = lower threshold of 125. Lower panel: GD and DI 1980 = PLT10 estimates with GD and DI models using total data series, GD 1992 = PLT10 estimates with GD model using data from 1992.

Further research on juvenile dynamics is being conducted in the Flathead due to the importance of this segment of the life history on overall population growth rates, abundance fluctuations, and the variability of risk of decline. Data on vital rates and the mechanisms of cohort density dependent interactions will improve understanding of the true risk facing the population. Sensitivity analyses of bull trout population models developed during this project suggest the viability of the population largely depends on tributary capacities and survival rates for juveniles and sub-adults (Staples et al. 2004b). These survival rates not only affect overall abundance, as the tributary capacity will, but

also strongly affect the density independent growth rate, potentially limiting the population's ability to rebound from low abundances.

Data on juvenile dynamics will be used to refine both simulation model structure and parameter estimates to better represent Flathead population. With improved process models, it will be possible to evaluate the effects of factors such as vital rate correlations or metapopulation dynamics on the population's risk of decline. Covariates to vital rates might be used in the future to reduce sampling effort necessary for effectively modeling individual sub-populations of this wide-ranging population. Estimating vital rates across the Flathead basin using spatial replication may provide a means to tie vital rates to environmental conditions using natural variation over space to increase sample size to overcome time constraints regarding data collection using space for time substitution (Dennis et al. 1998, Fieberg and Ellner 2001). Refined models could also be used to examine the effects of sampling error on vital rates on risk predictions. Preliminary work on vital rate sampling error suggests that estimating the variability in vital rates will be more important than estimating mean survival rates accurately (Staples et al. 2004b).

The model structural adequacy analyses brought theoretical and quantitative ecologists together with experienced field biologists and fishery managers for the purpose of conserving a threatened bull trout population. The collaborative process of building population models for these analyses has increased our understanding of bull trout population dynamics, which in turn has guided management actions and data collection efforts. Currently, monitoring with the GD model is likely the best means for approximating the risk of decline for the population, while data on juvenile vital rates and

abundances potentially could enhance the ability to foresee imminent declines in adult abundances. More generally, the MSA analyses presented here provide a means to compare potential models based on how well they address management questions in cases where traditional model selection methods cannot be used. In doing so, we evaluated the effect of hypotheses about the population on model selection and monitoring inference. Information from these analyses should be used to guide further research and model development to improve monitoring inference.

Acknowledgements: We thank the biologists who participated in bull trout model development and management meetings in Kalispell, MT in Spring 2002 and Fall 2003. These include: M. Deleray, J. Dunnigan, M. Hensler, B. Marotz,, C. Muhlfeld, S. Rumsey, B. Snyder, and T. Weaver from Montana Fish, Wildlife, & Parks in addition to W. Fredenberg from U.S. Fish and Wildlife Service and L. Everts, B. Hanson, and D. Rockwell from the Confederated Salish and Kootenai Tribes. DFS was supported by grants from Montana Fish, Wildlife, & Parks and the Northwest Power and Conservation Council.

Literature Cited

Berryman, A. A. 1991. Population theory: an essential ingredient in pest prediction, management, and policy making. American Entomologist **37**:138-142.

Caballero, A. 1994. Developments in the Prediction of Effective Population-Size. Heredity **73**:657-679.

Caswell, H. 2001. Matrix Population Models, 2nd edition. Sinauer Associates, Inc.

Coulman, G. A., R. L. Tummala, and S. R. Reice. 1972. Population Modeling - Systems Approach. Science **175**:518-&.

de Leeuw, J. 1992. Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. Pages 599-609 *in* S. Kotz and N. L. Johnson, editors. Breakthroughs in statistics. Springer-Verlag, London.

De Valpine, P., and A. Hastings. 2002. Fitting population models incorporating process noise an observation error. Ecological Monographs **72**:57-76.

Dennis, B., W. P. Kemp, and M. L. Taper. 1998. Joint density dependence. Ecology **79**:426-441.

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:In Press.

Dennis, B., and M. L. Taper. 1994. Density dependence in time series observations of natural populations: estimation and testing. Ecological Monographs **64**:205-224.

Doak, D. F., and L. S. Mills. 1994. A Useful Role for Theory in Conservation. Ecology **75**:615-626.

Ferrari, M., and M. L. Taper. 2006. Using auxilary data on measurement error to improve estimates of process variation in ecological time series. Journal of Environmental and Ecological Statistics:in press.

Fieberg, J., and S. P. Ellner. 2001. Stochastic matrix models for conservation and management: a comparative review of methods. Ecology Letters **4**:244-266.

Forrester, J. W. 1961. Industrial Dynamics. MIT Press, Cambridge.

Hakoyama, H., and Y. Iwasa. 2000. Extinction risk of a density-dependent population estimated from a time series of population size. Journal of Theoretical Biology **204**:337-359.

Hall, C. A. S., and J. W. Day Jr., editors. 1977. Ecosystem modeling in theory and practice: an introduction with case histories. Wiley, New York.

Hassell, M. P. 1975. Density-Dependence in Single-Species Populations. Journal of Animal Ecology **44**:283-295.

Henle, K., S. Sarre, and K. Wiegand. 2004. The role of density regulation in extinction processes and population viability analysis. Biodiversity and Conservation **13**:9-52.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology & Evolution **19**:101-108.

Liebhold, A. M. 1994. Use and abuse of insect and disease models in forest pest management: past, present, and future. *in* W. W. Covington and L. F. DeBano, editors. Sustainable ecological systems: implementing an ecological approach to land management. U.S. Forest Service Technical Report RM-247.

Lindsay, B. G. 2004. Statistical Distances as Loss Functions in Assissing Model Adequacy. *in* M. L. Taper and S. R. Lele, editors. The Nature of Scientific Evidence: Empirical, Statistical, and Philosophical Considerations. The University of Chicago Press.

Logan, J. A. 1988. Toward an expert system for development of pest simulation models. Environmental Entomology **17**:359-376.

Logan, J. A. 1989. Derivation and analysis of composite models for insect populations. *in* L. L. McDonald, B. F. J. Manly, J. A. Lockwood, and J. A. Logan, editors. Estimation and analysis of insect populations. Springer, New York.

Logan, J. A. 1994. In defense of big ugly models. American Entomologist **40**:202-207.

Morris, W. F., and D. F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.

Muhlfeld, C. C., M. L. Taper, D. F. Staples, and B. B. Shepard. 2006. Observer error structure in bull trout redd counts in Montana streams: Implications for inference on true redd numbers. Transactions of the American Fisheries Society:In Press.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. Applied linear statistical models, fourth edition. WCB McGraw-Hill, Boston, MA.

Pascual, M. A., P. Kareiva, and R. Hilborn. 1997. The influence of model structure on conclusions about the viability and harvesting of Serengetti Wildebeest`. Conservation Biology **11**:966-976.

Paul, A. J. 2000. Recruitment dynamics in Bull trout (*Salvelinus confluentus*), Linking theory and data to species management. Doctoral Thesis. The University of Calgary, Calgary, Alberta.

Paul, A. J., J. R. Post, G. L. Sterling, and C. Hunt. 2000. Density-dependent intercohort interactions and recruitment dynamics: models and a bull trout (Salvelinus confluentus) time series. Canadian Journal of Fisheries and Aquatic Sciences **57**:1220-1231.

Post, J. R., E. A. Parkinson, and N. T. Johnston. 1999. Density-dependent processes in structured fish populations: Interaction strengths in whole-lake experiments. Ecological Monographs **69**:155-175.

Rieman, B. E., and F. W. Allendorf. 2001. Effective population size and genetic conservation criteria for bull trout. North American Journal of Fisheries Management **21**:756-764.

Ruesink, W. G. 1976. Status of the systems analysis approach to pest management. Annual Review of Entomology **21**:27-44.

Shepard, B. B., K. L. Pratt, and P. J. Graham. 1984. Life histories of westslope cutthroat and bull trout in the Upper Flathead River Basin, Montana. Final Report for EPA Under Contract Number R008224-01-5. Montana Department of Fish, Wildlife and Parks, Helena.

Staples, D. F., and M. L. Taper. 2006. Impact of non-linearities in density dependence above the range of the data on predictions of population extinction risk. Journal of Nature Conservation:In Press.

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

Staples, D. F., M. L. Taper, and B. B. Shepard. 2004b. Monitoring strategies for Flathead Lake bull trout populations. PTMM11, Northwest Power and Conservation Council grant #PTMM11.

Staples, D. F., M. L. Taper, and B. B. Shepard. 2005. Risk-based viable population monitoring. Conservation Biology **19**:1908-1916.

Walters, C. J., and J. R. Post. 1993. Density-Dependent Growth and Competitive Asymmetries in Size-Structured Fish Populations - a Theoretical-Model and Recommendations for Field Experiments. Transactions of the American Fisheries Society **122**:34-45.

Wisdom, M. J., L. S. Mills, and D. F. Doak. 2000. Life stage simulation analysis: Estimating vital-rate effects on population growth for conservation. Ecology **81**:628-641.

CONCLUSION

The goal for population monitoring is to have some measure that reflects population status. Common monitoring methods are successful in indicating when a population has experienced problems in the past. For example, the decline in redd numbers and catch-per-unit effort in Flathead lake demonstrated the bull trout population declined in the late 1980's. Once a decline has been observed and the population is listed as threatened or endangered, however, proving poor performance in the past is not very useful. Threatened and endangered species are defined in terms of their risk of extinction, thus it would be useful to measure population status in terms of risk of decline in the future. A forward-looking monitoring strategy will enhance the ability to anticipate declines and address threats proactively.

Toward this goal, I have outlined a strategy to use short-term population viability analyses to determine population status in terms of risk of a biologically significant decline. Shorter prediction timeframes comport to management outlook and actions, and help mitigate sources of errors in long-term risk predictions such as model extrapolation error and correlations in population growth. As risk predictions necessarily use models to predict future status, the ability to apply such models to available data is paramount for a risk-based monitoring strategy to be useful. This research has addressed this through developing population viability analysis methods to accommodate sampling error in abundance data for count-based analyses. Unfortunately, this research suggests that it will be difficult to extend these methods to accommodate correlations in growth. Further, model evaluation and selection can be difficult in conservation applications with little

data. The model structural adequacy analyses provide a means for inference about model structure and data needed for useful risk predictions, and can be used to evaluate the effect of uncertainty concerning the true population dynamics on model selection and inference.

This research has improved monitoring inference to bull trout population status. This research suggests the Gompertz model parameterized with adult abundance data can give useful risk estimates for an age-structured population with inter-cohort density dependence. Current risk predictions basically quantify what is apparent from visual inspection of the data series, i.e. there is cause for concern because the population is close to the lower threshold. Additional information on juvenile dynamics will increase understanding of the population and means to effectively improve natal habitat. Increases in tributary output could ameliorate community factors affecting the population in Flathead Lake. Improved tributary habitat likely would also increase the potential to maintain the population should it shift more toward a resident life history in response to lake conditions. There are questions, however, as to whether estimates of tributary survival rates will markedly improve the ability to detect increased risk with demographic monitoring models. Sub-adult survival rates could have a large effect on the ability to precisely predict risk of decline with demographic models. Considering another major community upheaval is possible and the influence sub-adult survival rates have on population viability, monitoring the Lake community in addition to natal tributaries is certainly warranted.