Use of state-space population dynamics models in hypothesis testing: advantages over simple log-linear regressions for modeling survival, illustrated with application to longfin smelt (*Spirinchus thaleichthys*)

Mark N. Maunder [a,b,*], Richard B. Deriso [b], Charles H. Hanson [c]

[a] *Quantitative Resource Assessment LLC, San Diego, CA 92129, USA*
[b] *Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA*
[c] *Hanson Environmental, Inc., 446 Green View Court, Walnut Creek, CA 94596, USA*

ARTICLE INFO

ABSTRACT

Factors impacting the survival of individuals between two life stages have traditionally been evaluated using log-linear regression of the ratio of abundance estimates for the two stages. These analyses require simplifying assumptions that may impact the results of hypothesis tests and subsequent conclusions about the factors impacting survival. Modern statistical methods can reduce the dependence of analyses on these simplifying assumptions. State-space models and the related concept of random effects allow the modeling of both process and observation error. Nonlinear models and associated estimation techniques allow for flexibility in the system model, including density dependence, and in error structure. Population dynamics models link information from one stage to the next and over multiple time periods and automatically accommodate missing observations. We investigate the impact of observation error, density dependence, population dynamics, and data for multiple stages on hypothesis testing using data for longfin smelt in the San Francisco Bay-Delta.

1. Introduction

Estimation of survival, and the factors influencing survival, are vital in the research, and to the management, of natural resources. Survival is a critical component of methods used to determine sustainable yields of harvested resources (Quinn and Deriso, 1999). Managers need to know the most influential factors affecting the survival of endangered species to focus limited financial resources on research and management actions that obtain the most benefit. Anthropogenic effects have to be separated from natural impacts to determine the relative importance of restricting human activities (e.g. Deriso et al., 2008).

Survival can be estimated using a number of approaches ranging from field studies such as following individuals using radio tracking and determining their fate (White and Garrott, 1990; Skalski et al., 2010) to sophisticated statistical state-space population dynamics models that integrate multiple data types (Besbeas et al., 2003; Maunder, 2004; Schaub and Abadi, 2010). Facilitated by the availability of time series of relative abundance, log-linear modeling of the ratio of relative abundance in two life stages is a common approach to estimate relative survival and evaluate the support for different hypotheses about the factors influencing survival (e.g. Miller et al., 2012). Log-linear modeling is used because it is conveniently implemented in traditional software packages as a linear equation. However, it restricts the analysis to a subset of models that are not necessarily the most appropriate for the particular application. Log-linear modeling also aggregates process and observation error into a single term, limiting the ability to fully characterize uncertainty. Modern nonlinear modeling software such as BUGS (Lunn et al., 2009) and AD Model Builder (Fournier et al., 2012) expand the modeling options outside those covered by "fixed effects" log-linear models, allowing flexibility in model and error structure (Bolker et al., 2013).

Correctly dealing with both observation and process error is important for hypothesis testing and evaluating the data-based support for alternative hypotheses (Maunder and Watters, 2003; Deriso et al., 2008). Process error (also known as process noise or

* Corresponding author at: Inter-American Tropical Tuna Commission, 8901 La JollaShores Drive, La Jolla, CA 92037-1508, USA. Tel.: +1 858 546 7027..
E-mail addresses: mmaunder@iattc.org (M.N. Maunder), rderiso@iattc.org (R.B. Deriso), chanson@hansonenv.com (C.H. Hanson).

process variability) generally refers to stochasticity in population dynamics (but can also relate to model structure misspecification) and is hence parameterized as "random effects", and observation error refers to inaccuracy in observations (de Valpine, 2003).

One approach for dealing with both observation and process error is to ignore one or the other entirely. Polacheck et al. (1993) found that ignoring process error (an observation error estimator) was superior to ignoring observation error (a process error estimator) when estimating the parameters of a simple population dynamics model, but they did not evaluate which choice was best for hypothesis testing. Ignoring process error biases likelihood ratio and Akaike information criterion (AIC; Akaike, 1973) based tests towards incorrectly accepting covariates (Maunder and Watters, 2003). Other tests such as analysis of deviance (Skalski, 1996) or randomization tests (Edgington, 1987; Deriso et al., 2008) can be used, but they are less elegant and impractical in some situations. An alternative approach is to include both process and observation error, but assume the ratio of the variances between these two sources of variation is known (e.g. Walters and Ludwig, 1981) or that one of the variances is known (e.g. Maunder and Watters, 2003). Incorrectly specifying the variance terms can bias hypothesis tests (Deriso et al., 2007).

The preferred approach is to use state-space models (e.g. Schnute, 1994; Newman, 1998; de Valpine, 2002; Buckland et al., 2004, 2007; Maunder and Deriso, 2011) that allow the estimation of the both observation and process error variances. It should be noted that state-space models are often described as random effect, hierarchical, or Bayesian models. de Valpine and Hastings (2002) found that state-space models led to lower bias and often lower variance estimates than least squares estimators that ignore either process noise or observation error. Traditionally, state-space models have been used to model demographic variability such as the binomial probability of individuals surviving given an average survival rate (Dupont, 1983; Besbeas et al., 2002). However, demographic variability is typically overwhelmed by environmental variability (Buckland et al., 2007), so environmental variability is often modeled instead of demographic variability or in addition to demographic variability (e.g. Rivot et al., 2004; Newman and Lindley, 2006). Nonlinear, non-Gaussian state-space models generally require computationally intensive high dimensional integrals that have no closed form solution (de Valpine, 2003). The implementation of state-space models in a Bayesian framework has been facilitated by the development of Markov chain Monte Carlo (MCMC) methods (Punt and Hilborn, 1997; Newman et al., 2009; Lunn et al., 2009). MCMC methods have also been adapted to implement state-space models in a classical framework (Lele et al., 2007). Alternatively, the Laplace approximation (Skaug, 2002; Skaug and Fournier, 2006) or importance sampling (Maunder and Deriso, 2003) can be used to implement the integration in a classical framework. Modern nonlinear modeling software packages such as BUGS and AD Model Builder have made state-space models practical for many applications (Bolker et al., 2013).

Log-linear models, such as generalized linear models, analysis of variance (ANOVA), and related statistical methods, do not incorporate demographic relationships between abundances through time (de Valpine, 2003). In contrast, lifecycle models link life-stages and time periods using population dynamics propagating information and uncertainty (Buckland et al., 2007; Maunder and Deriso, 2011). This link allows information related to one life-stage to inform processes influencing other life-stages and is particularly important when data are not available for all life stages for all time periods. Hypotheses that are difficult to consider with ANOVA and related methods can be simple to express using a population dynamics model (de Valpine, 2003). de Valpine

(2003) found that a population dynamics model had much higher statistical power than ANOVA, and provided greater biological insight. Even approximately correct population dynamics models had higher power than omitting demographic structure, but the rate at which Type I error occurs may increase, or the power might be reduced as the model structure becomes more incorrect (de Valpine, 2003).

Hypothesis testing is an essential part of statistical analysis and is particularly important when evaluating factors that are impacting survival. When we refer to hypothesis testing, we are more generally referring to the evaluation of the data based support for alternative configurations of a model, where each configuration could represent an alternative hypothesis. This approach is often termed model selection to differentiate it from traditional hypothesis testing that involves the rejection of a null hypothesis (Johnson and Omland, 2004). Hypothesis testing can easily become complex when analysing population dynamics because of the many factors operating on different stages under the presence of density dependence. Deriso et al. (2008) present a framework for evaluating alternative factors influencing survival, and Maunder and Deriso (2011) extended the framework to include density dependence in survival. The first step is to identify the factors to be considered, including the life stages that are impacted by each factor and where density dependence occurs. Next, a model should be developed to include these factors. Then hypothesis tests should be conducted to determine which factors are important. Finally, impact analysis (Wang et al., 2009; Maunder and Deriso, 2011) should be conducted to determine the impact of the factors on quantities useful for management.

Density dependence is an important factor in the dynamics of many populations (Brook and Bradshaw, 2006) and can occur in multiple life stages (e.g. Ciannelli et al., 2004). It is important to consider density dependence when carrying out model selection because it can modify the impact of factors (Rose et al., 2001; Maunder and Deriso, 2011). Environmental conditions can also have a large impact on population dynamics. Environmental factors can directly affect survival through processes such as temperature tolerance or can interact with density dependence through affecting density limiting processes such as habitat or prey availability. Environmental factors and density dependence have been identified as impacting population dynamics in numerous studies either independently or in combination (e.g. Sæther, 1997; Brook and Bradshaw, 2006; Ciannelli et al., 2004; Deriso et al., 2008; Maunder and Deriso, 2011). Density dependence can easily be integrated into state-space models (e.g. de Valpine and Hastings, 2002; Maunder and Deriso, 2011).

Data from longfin smelt (*Spirinchus thaleichthys*) in the San Francisco Bay-Delta are used to illustrate the development and advantages of using state-space population dynamics models over simple log-linear regressions for modeling survival. The models are implemented in AD Model Builder using the Laplace approximation for random effects (Skaug and Fournier, 2006) under a classical (frequentist) framework. Longfin smelt is of conservation concern because it is exposed to a variety of anthropogenic factors (e.g. habitat modification, sewage outflow, farm runoff, and water diversions) and survey data have shown a decline in abundance. Longfin smelt was listed as threatened under the California Endangered Species Act in 2009. The U.S. Fish and Wildlife Service also evaluated the status of the Bay-Delta longfin smelt population and concluded in 2012 that although the species warranted protection under the federal Endangered Species Act, staff limitations precluded listing the species as of that time. Several other species in the San Francisco Estuary have also experienced declines (e.g., Bennett, 2005; Sommer et al., 2007; Mac Nally et al., 2010; Thomson et al., 2010; Maunder and Deriso, 2011), but the declines have yet to be fully explained.

2. Theory

State-space models appropriately accommodate both observation and process error. de Valpine (2002, 2003) provides a useful description of state-space models in the context of population dynamics models. Here we illustrate state-space models using a simple population dynamics model where the abundance in the next time period is simply those that survive from the previous time period:

$$E[X_{t+1}|X_t] = \mu_s X_t \tag{1}$$

where X_t is the number of individuals at time t, which are the states; and μ_s is the mean survival rate. The observations of the population are estimates of absolute abundance and the sampling variation in these estimates is assumed to be normally distributed:

$$Y_t \sim N(X_t, \sigma^2) \tag{2}$$

where Y_t is the estimate of absolute abundance at time t and σ^2 is the sampling variance.

State-space population dynamics models have three main components: (1) states (X), (2) parameters (θ), and (3) observations (Y). The states represent the population such as the abundance in a life stage at a given time. The parameters describe the average (or sometimes the exact) relationship (transition) between the states (e.g. the average survival rate), but also include the initial state (e.g. X_1) and the variance parameters (e.g. σ). The observations are measurements of the states, or some function of the states. The states and parameters are unknown and they, or a function of them, are the quantities of interest. The observations, which are known, are used to provide information about the states and parameters. Observations are generally not a census of the population, but a sample of the population and therefore contain sampling error (e.g. if a line transect or trawl survey is used to estimate the abundance of a population). This sampling error is the observation error and is generally represented by the likelihood function. In other words, the observation is known, but there is uncertainty in how the observation relates to the true abundance. There may also be additional observation error over and above the sampling variability, but for illustrative purposes we ignore this.

In traditional maximum likelihood estimation, the parameters of the model are estimated by finding the parameter values that, conditional on these values, give the highest probability (likelihood) that the observations came from the model. Since the states (X) are a direct function of the parameters (θ), for known observations and given parameter values, the probability function described in Eq. (2) can be evaluated and maximized. To better illustrate state-space models, let

$$f(\theta, Y) = f(X, Y) \tag{3}$$

be the joint distribution of the data and parameters, since the parameters determine the states, and

$$f_\theta(Y), \tag{4}$$

be the likelihood function evaluated at the parameter values θ. Traditional maximum likelihood assumes that there is a single true value for each parameter. State-space population dynamics models implicitly assume that the values of the parameters representing some population processes may change over time. This is the process error. Before describing state-space models, consider the survival in each time period as a separate model parameter s_t:

$$E[X_{t+1}|X_t] = s_t X_t \tag{5}$$

In this case, the likelihood function can be denoted $f_{\theta,s}(Y)$, and traditional maximum likelihood assumes that there is a single true value for survival probability in each time period and for the other

model parameters (note that the average survival parameter is replaced with a set of survival parameters, one for each time period) and the survival parameters are estimated along with the other model parameters by maximizing the likelihood function. However, there is now one survival parameter for each observation and each survival will be estimated to exactly match the observation. No other parameters can be estimated (e.g. the observation error variance), and the process error cannot be separated from the observation error.

Intuitively, the estimation procedure could be improved by adding information based on the form of the process error probability distribution (e.g. if the temporal variability in survival is known to be low, a survival parameter in one time period that is very different from the survival in the other time periods is unlikely) and can be conceptualized as placing an informative prior, in the Bayesian sense, on the process error (except that the mean and variance of the prior are unknown) (e.g. $s_t = \mu_s \exp(\varepsilon_t)$, where $\varepsilon_t \sim N(0, \nu^2)$, which parallels the random effects approach in generalized linear mixed models (GLMMs), or in alternative notation $\ln(X_{t+1}) \sim N(\ln(\mu_s X_t), \nu^2)$). In this case, $f_{\theta,s}(Y) = f(Y|s, \theta) f(s|\theta) = f(Y|\varepsilon, \theta) f(\varepsilon|\theta)$, where $f(\varepsilon|\theta)$ is the process error probability distribution, and the resulting likelihood is often referred to as a penalized likelihood. The penalized likelihood combines the sampling probability distribution of the observations with the probability distribution of the states (recall that the parameters determine the state and similarly the process error probability distribution also defines the state probability distribution). These methods estimate the process errors (or states) along with the other model parameters while maximizing the joint probability distribution of the process error and the observations. However, the MLE of the process error variance is not statistically consistent (Seber and Wild, 1989) and the likelihood function is degenerative towards zero variance (Maunder and Deriso, 2003). There is often a negatively biased local maximum that has been used for inference, but the global maximum is at zero process error variance (Maunder and Deriso, 2003).

The process error variance will decrease as covariates are added and therefore the process variance should be reduced, which can only be practically achieved if the process variance is estimated. In contrast to penalized maximum likelihood, state-space models treat the process error (or states) as random variables rather than as parameters and when the process error is integrated out they produce a marginal likelihood or "true likelihood" function that is used for inference (e.g. Eq. (4) becomes $\int f_\theta(Y, \varepsilon) d\varepsilon$ or equivalently $\int f_\theta(Y, X) dX$. Intuitively, this can be thought of as summing up the likelihood of the observations for each possible state weighted by the probability of that state (conditioned on the parameter values). Each possible survival will lead to different population abundance (state). Hence, the derivation of "state-space", which refers to the whole range of possible trajectories through time of the population states (de Valpine, 2002). Integrating out the process error takes advantage of properties of random variables (e.g. the marginal distribution), which has the advantage that it provides a consistent non-degenerative MLE for the process error variance.

Pawitan (2003) appropriately summarizes state-space models/random effects as a convenient way to deal with many parameters. In a Bayesian framework (Punt and Hilborn, 1997), parameters are also treated as random variables and integrated out (e.g. Eq. (4) becomes $\int \int f(Y, \theta, \varepsilon) d\varepsilon d\varphi$ or equivalently $\int \int f(Y, \theta, X) dX d\varphi$, where φ are the parameters that are not of interest) and the probability distribution is used for inference rather than the likelihood function. One advantage of the state-space modeling approach over penalized maximum likelihood is that the marginal likelihood is consistent with AIC theory, which can be used for hypothesis testing and model selection.

3. Methods

3.1. Models

3.1.1. Log-linear regression

A common approach to model survival from one life-stage to the next as a function of explanatory variables is a log-linear regression (Christensen, 1997) of the numbers in the second stage as a ratio of those in the first stage (e.g. Miller et al., 2012). A typical analysis models the reproductive output from adults (A_t) to the surviving juveniles in the next year (J_{t+1}) as:

$$\ln(J_{t+1}/A_t) \sim N(\alpha + \boldsymbol{\beta} I_t, \sigma^2) \tag{6}$$

or equivalently in a different notation (the former notation is commonly used to describe state-space models and the latter notation commonly used to describe random effect models and can be a more useful description (de Valpine, 2003)).

$$\ln(J_{t+1}/A_t) = \alpha + \boldsymbol{\beta} I_t + \varepsilon_t \tag{7}$$

where $\varepsilon_t \sim N(0, \sigma^2)$, N represents a normal distribution, α and $\boldsymbol{\beta}$ are parameters of the linear model, I_t is a matrix of covariates (forcing functions), and σ^2 is the variance of the error. The observations are often only an index of relative abundance related to the absolute abundance by a constant q, often called catchability in the fisheries literature, such that

$$\ln(q_J J_{t+1}/q_A A_t) = \alpha + \boldsymbol{\beta} I_t + \varepsilon_t \tag{8}$$

so unless $q_J = q_A$, α no longer relates to survival (it also includes reproductive output in our example), but a combination of survival and differences in catchability. However, this does not influence hypothesis tests related to the covariates as long as the q's are constant through time or their temporal variation is random and independent of the covariates.

The parameters can be estimated by maximizing the likelihood based on the assumed error distribution (Eq. (8)). The likelihood function is typically used to represent observation error. However, ε in Eq. (8) includes both process and observation error and ε describes the unexplained variation (process error) in the modeled relationship if J and A are known without error. If J and A are known with error (multiplicative and log-normal):

$$\ln\left(\left(J_{t+1} \exp(\varepsilon_{J,t+1})\right) / \left(A_t \exp(\varepsilon_{A,t})\right) \right) = \alpha + \boldsymbol{\beta} I_t + \varepsilon_t \tag{9}$$

where $\varepsilon_{A,t} \sim N(0, \sigma_{A,t}^2)$, $\varepsilon_{J,t+1} \sim N\left(0, \sigma_{J,t+1}^2\right)$, such that

$$\ln(J_{t+1}/A_t) = \alpha + \boldsymbol{\beta} I_t + \varepsilon_t - \varepsilon_{J,t+1} + \varepsilon_{A,t} \tag{10}$$

illustrating that Eqs. (6) and (7) combine process error and observation error from both measures of abundance into a single error term $\varepsilon_t \sim N(0, \sigma_{J,t}^2 + \sigma_{A,t+1}^2 + \sigma_\varepsilon^2)$.

Often an estimate of the sampling precision of each observation is available (hence the time subscript on the variance terms), which eliminates the need to estimate the observation error variance, but this is generally not the case for the process error. Ignoring observation error may bias the results if the observation error variance differs substantially among observations.

3.1.2. Alternative formulation

The log-linear regression is deterministically equivalent and, depending on assumptions, stochastically equivalent to an exponential growth model. The log-linear model assumes that the unexplained variation in the log of the abundance ratios is normally distributed while the exponential growth model assumes

that the unexplained variation in the abundance in the second stage is log-normally distributed

$$J_{t+1} = \alpha' A_t \exp(\boldsymbol{\beta} I_t + \varepsilon_t) \tag{11}$$

where $\alpha' = \exp(\alpha)$

3.1.3. State-space model

State-space models can be used to include both observation and process error. Non-linear state-space models are flexible in representing process and observation error. Eq. (6) assumes log-normal multiplicative error for both the observation and process error with constant variance. The log-normal assumption as implemented in Eq. (6) will provide an unbiased estimate of α, but the quantity of interest $\alpha' = \exp(\alpha)$ will be biased such that the expected value of $E[\alpha'] = \exp(\alpha + 0.5\sigma^2)$ (Maunder and Deriso, 2011). Eq. (11) could be modified to account for the bias

$$J_{t+1} = \alpha' A_t \exp(\boldsymbol{\beta} I_t + \varepsilon_t - 0.5\sigma^2) \tag{12}$$

Similarly, the likelihood and random effects can be modified to deal with the log-normal bias correction. This may be particularly important when the observations have different variances, resulting in different bias correction factors for each time period. The distribution for the process and observation error need not be normal. For example, the process error may be log-normal, while the observation error might be normal.

3.1.4. Density dependence

Population regulation is controlled by both density-independent and density-dependent factors. The log-linear regression typically includes covariates representing density-independent factors (e.g. the environment). Density dependence can be included in the log-linear regression by adding additional terms related to abundance into the regression. The Ricker model (Ricker, 1954)

$$J_{t+1} = \alpha' A_t \exp(-bA_t + \boldsymbol{\beta} I_t + \varepsilon_t) \tag{13}$$

is often used because it can be linearized by taking the natural logarithm and implemented using multiple linear regression.

$$\ln(J_{t+1}) = \alpha + \ln(A_t) - bA_t + \boldsymbol{\beta} I_t + \varepsilon_t \tag{14}$$

where $\alpha = \ln(\alpha')$. However, the Beverton–Holt model (Beverton and Holt, 1957) may be applicable for some populations, but is non-linear:

$$J_{t+1} = \frac{\alpha' A_t}{1 + bA_t} \exp(\boldsymbol{\beta} I_t + \varepsilon_t) \tag{15}$$

The models are derived based on solving the differential equation for abundance where mortality is a linear function of the cohort abundance and initial abundance for the Beverton–Holt and Ricker models, respectively. The Beverton–Holt model has asymptotic properties, which represent processes such intra-cohort competition, while the Ricker model produces lower abundance from high initial abundance, which represents processes such as cannibalism when used in a stock-recruitment context.

3.1.5. State-space population dynamics (life cycle) model

The log-linear regression only models survival from one stage to the next. A sequence of separate log-linear regressions can be used to model the survival between each stage. However this does not link information among stages, which can be useful particularly if there is substantial error in the estimates of abundance or if there are missing abundance estimates. In the case where adults are a
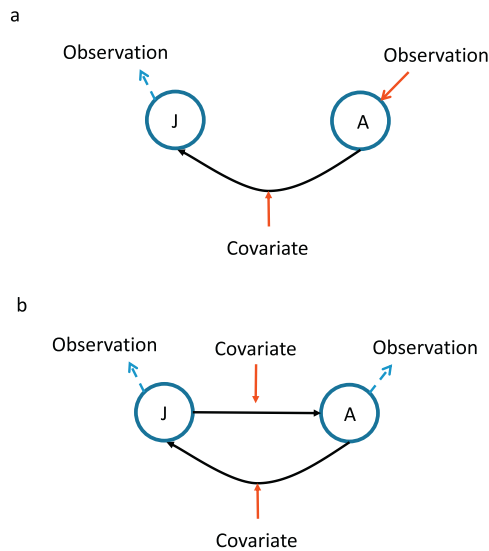
Fig. 1. Conceptual diagram illustrating the differences between (a) the exponential model representation of the log-linear regression and (b) the full state-space population dynamics model. The shaded (red) solid arrows represent forcing functions and the dashed arrows represent predictions of the observations used in the likelihood functions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

year older than juveniles and the juveniles are measured the year after spawning:

$$J_{t+1} = \frac{\alpha'_J A_t}{1 + b_J A_t} \exp(\boldsymbol{\beta_J I_t} + \varepsilon_{J,t}) \qquad (16)$$

$$A_{t+1} = \frac{\alpha'_A J_t}{1 + b_A J_t} \exp(\boldsymbol{\beta_A I_t} + \varepsilon_{A,t}) \qquad (17)$$

where the process errors $\varepsilon_A \sim N(0, \sigma^2_{\epsilon,A})$ and $\varepsilon_J \sim N(0, \sigma^2_{\epsilon,J})$ are treated as random effects and the observation errors $N(\ln(J), \sigma^2_J)$ and $N(\ln(A), \sigma^2_A)$ are implemented using likelihoods.

The initial condition for the population dynamics model, which are the abundances in the first time period for juveniles, J_1, and adults, A_1, have to be estimated as parameters in addition to the parameters of the two Beverton–Holt models, the covariate coefficients, and the standard deviations of the random effects. Fig. 1 illustrates the difference between the exponential model representation of the log-linear regression and the state-space population dynamics model.

3.2. Hypothesis testing and model selection

Various methods can be used for hypothesis testing and evaluating the data-based evidence of support for alternative hypotheses, or, perhaps more accurately, evaluating the measure of evidence from data about alternative models (Hilborn and Mangel, 1997; Hobbs and Hilborn, 2006). The influence of a covariate can be eliminated from the model by fixing its value at zero. This produces a nested model, and model selection can be conducted using likelihood ratio tests. The likelihood ratio test is not appropriate for non-nested models. For example, when comparing between two models that include different covariates or two different density dependence assumptions. In this case, information theory-based methods such as the Akaike information criterion (AIC; Akaike, 1973) are appropriate. They are also appropriate for nested models. We use the AIC adjusted for small sample size (AIC_C) (Burnham and Anderson, 2002)

$$AIC_c = -2\ln L + 2K + \frac{2K(K+1)}{n - K - 1} \qquad (18)$$

where L is the likelihood function evaluated at its maximum, K is the number of estimated parameters, and n is the number of observations. The difference between a given model and the model with the lowest AIC_C value, Δ, is used for comparing models. For model comparison, Burnham and Anderson (1998) recommend: "For any model with $\Delta \leq 2$ there is no credible evidence that the model should be ruled out ... For a model with $2 \leq \Delta \leq 4$ there is weak evidence that the model is not the K–L [Kullback–Leibler] best model. If a model has $4 \leq \Delta \leq 7$ there is definite evidence that the model is not the K–L best model, and if $7 \leq \Delta \leq 10$, there is strong evidence that the model is not the K–L best model. Finally, if $\Delta > 10$, there is very strong evidence that the model is not the K–L best model."

3.3. Application

Data from longfin smelt in the San Francisco Bay-Delta from 1980 to 2009 are used to show the development and advantages of using state-space population dynamics models over simple log-linear regressions for modeling survival. We implement a range of models to determine the difference between the modeling approaches (Table 1). A conceptual model of the San Francisco Bay longfin smelt population (e.g., Rosenfield and Baxter, 2007; Baxter et al., 2008)[1] was used as a basis for identifying potential environmental covariates considered in the models. The covariates reflected various geographic regions of the estuary and seasonal periods based on the life history and seasonality of each lifestage of longfin smelt. A total of 36 potential covariates were identified in the initial selection process (Supplemental Table 1). The covariates included various flow variables (e.g., spring X2 location (a measure of the spatial extent of salinity: position of the 2% isohaline), winter-spring Delta outflow, winter-spring Napa River flow, spring outflow thresholds of 34,500 cfs and 44,500 cfs, spring Sacramento River inflow in addition to various variations of Sacramento and San Joaquin River runoff), zooplankton (prey) densities (e.g., mysid, *Eurytemora*, and *Pseudodiaptomus* densities over various seasonal time periods), predators and competitors (e.g., juvenile Chinook salmon densities in the spring, predators in various regions, and the Asian overbite clam *Potamocorbula*), and a variety of abiotic environmental variables (e.g., Secchi depth as an index of turbidity, water temperature, ammonium loading to various regions of the estuary, and the ratio of ammonium loading to Delta inflow). Based on the conceptual model, the expected sign (positive or negative) in the relationship between each covariate and an expected longfin smelt population response was also assigned to each covariate. All of the environmental covariates were then entered into two formulations of the longfin smelt lifecycle model (a model in which spawners are the adult lifestage (November–March) ages 1 and 2 and an alternative model in which pre-adults (October–March) ages 0 and 1 and adults (November–March) ages 1 and 2 were equally weighted in the model as spawners) and a series of statistical analyses were performed to identify the best model. The models were fit to indices of juvenile and adult longfin smelt abundance created using Bay study otter and mid-water trawl surveys[2]. The covariates that explained the most variation from

---

[1] Rosenfield, J.A. 2010. Life History Conceptual Model and Sub-Models for Longfin Smelt, San Francisco Estuary Population. Unpublished Report. Available at: http://www.dfg.ca.gov/erp/cm_list.asp. Hanson, C. H. 2014. Covariates for Consideration in Developing a Lifecycle Model for the San Francisco Bay-Delta Population of Longfin Smelt. Hanson Environmental, Inc. Unpublished contract report. 93pp. http://new.baydeltalive.com/projects/7012

[2] Maunder, M.N. and Deriso, R.B. 2013. Empirical estimates of abundance indices and standard deviation for longfin smelt from the bay study otter and mid-water trawl surveys. Unpublished QRA contract report. 13pp. http://new.baydeltalive.com/projects/7012

Table 1
Description of modeling scenarios. The symbol under the "Analysis type" column is based on the entries in the other columns with symbols: juvenile = "J", adult = "A", juvenile divided by adult = "J/A", both juvenile and adult = "J+A" None = "-", likelihood = "L", random effects = "re", Beverton–Holt = "BH", Ricker = "R".

Name	Analysis type	Dependent variable	Adult observation error	Juvenile observation error	Process error	Density dependence	Equation
Log-linear	J/A--L-	Juvenile divided by Adult	None	None	Likelihood	None	7
Exponential	J--L-	Juvenile	None	None	Likelihood	None	11
Log-linear with observation error	J/ArereL-	Juvenile divided by Adult	Random effect	Random effect	Likelihood	None	7
Exponential with juvenile observation error only	J-L--	Juvenile	None	Likelihood	None	None	11
Exponential with juvenile observation error and process error	J-Lre-	Juvenile	None	Likelihood	Random effect	None	11
Exponential with observation an process error	JreLre-	Juvenile	Random effect	Likelihood	Random effect	None	11
Ricker	JreLreR	Juvenile	Random effect	Likelihood	Random effect	Ricker	13
Beverton–Holt	JreLreBH	Juvenile	Random effect	Likelihood	Random effect	Beverton–Holt	15
Population dynamics (Life cycle)	J+ALLreBH	Juvenile and Adult	Likelihood	Likelihood	Random effect for both A and J	Beverton–Holt	16 and 17

each category of covariate[3] (e.g. flow, prey, predators, environmental conditions) were then used in the application below that illustrates the benefits of state-space models.

AIC_C was used to conduct forward stepwise covariate selection. The covariates were normalized (mean subtracted and divided by the standard deviation) to improve model performance. Several covariates were chosen as candidates for the model selection procedure (Table 2 and Supplemental 2). These covariates were chosen based on initial analysis of the wider range of factors in supplemental Table 1. Many of the factors in the larger set were highly correlated and so were eliminated. We kept two flow variables that were highly correlated to illustrate some of the difficulties in hypothesis testing. The model is fit to relative abundance indices for each stage (Supplemental Table 3), as appropriate. The models were implemented using AD Model builder and the Laplace approximation was used for random effects. The observation error in Eq. (10) was implemented by treating the true population abundance as a random effect and using the sampling distribution as the likelihood for abundance. The true abundance was then used in the calculation of the regression model and the likelihoods for the observations were combined with the likelihood for the regression equation. The lognormal bias correction is not used since α is not of interest and the temporal variation in the observation error is low.

4. Results

In general, all scenarios support the two flow-related covariates (Sacramento and Napa river runoff) when a single covariate is tested (Fig. 2) followed closely by the prey species *Eurytemora*. However, after including a flow covariate, support for *Eurytemora* is lost and it is not selected in any of the final models. In all models, ammonia is the second covariate selected and temperature is the third covariate selected (Table 3). Adding density dependence (models JreLreR and JreLreBH) results in more support for Sacramento River runoff over Napa River runoff, and over the other covariates in general, when comparing single covariate models. Using observation error only for juveniles and no process error (model J-L--; Table 1) creates greater differences in the

likelihood between covariates and gives increased relative support to temperature and ammonia.

The likelihood values from the log-linear model (model J/A--L-) and the exponential model (model J--L-) are identical as expected (Table 3). The results from the log-linear model with observation error (model J/ArereL-), which implies both observation and process error, and the exponential model with both observation and process error (model JreLre-) are identical despite the likelihood and random effects representing different error components.

Adding observation error (e.g. compare model J--L- with model JreLre-) makes little difference in relative likelihoods (Table 3), but changes the variables selected (Table 3). Sacramento River runoff is selected in the first stage of the stepwise regression in place of Napa River runoff when allowance is made for observation error. This is in part because Napa River runoff and Sacramento River runoff are highly correlated. The stepwise procedure also selects Napa River runoff as a fourth covariate. However, if Sacramento River runoff is dropped from the final model (that is the model chosen by the stepwise procedure that includes both flow variables) the AICc drops by 2.58 units. The AIC_C for the model which only includes Napa River runoff as the flow variable is 5.39 units lower than the model which only includes Sacramento River run off as the flow variable (Fig. 3) providing "definite" evidence of Napa River runoff over Sacramento River runoff in models that do not include density-dependence; evidence favors Napa River runoff over Sacramento River runoff in all the various model configurations, but not as definitive as the ones above (Table 3).

Ignoring process error and including observation error only for the juvenile abundance (model J-L--) leads to much greater changes in the likelihood causing all covariates to be selected except for those that are rejected because the coefficient has the wrong sign.

The Ricker (model JreLreR) and Beverton–Holt (model JreLreBH) forms of density dependence lead to different results, with the Beverton–Holt model including Napa River runoff as a fourth covariate resulting in a better AIC_C, but it is only 1.65 units lower than the Ricker model providing "no credible" evidence to differentiate between the two forms of density dependence. The AIC_C for the Beverton–Holt model is 4.19 units less than the exponential model with observation error providing "definite" evidence for density dependence. If the Sacramento outflow is discarded from the Beverton–Holt model, the AICc is only 0.25 units less than the final model, and is only 1.21 units lower than if Napa River runoff is not included and Sacramento runoff is

[3] Maunder, M.N. and Deriso, R.B. 2013. Evaluation of factors impacting longfin smelt – summary analysis. Unpublished QRA contract report. 9 pp. http://new.baydeltalive.com/projects/7012

Table 2
Covariates used in the longfin smelt application (Hanson, C.H. 2014. Selection of Environmental Covariates for Consideration in Developing a Lifecycle Model for the San Francisco Bay-Delta Population of Longfin Smelt. Hanson Environmental, Inc. Unpublished contract report. 93pp. http://new.baydeltalive.com/projects/7012).

| Factor	Time	Stage	Sign of coefficient
Mysid	May–June	Adult to Juveniles	+
Secchi depth	April–June	Adult to Juveniles	−
Eurytemera	April–May	Adult to Juveniles	+
Napa River flow	January–March	Adult to Juveniles	+
Predators central + San Pablo	Annual	Adult to Juveniles	−
Average temperature	April–June	Adult to Juveniles	−
San Pablo ammonium	April–June	Adult to Juveniles	−
Sacramento River runoff	Previous October–July	Adult to Juveniles	+
Overbite clam presence	Year round	Adult to Juveniles	−
Mysid	July–September	Juveniles to pre-adult	+
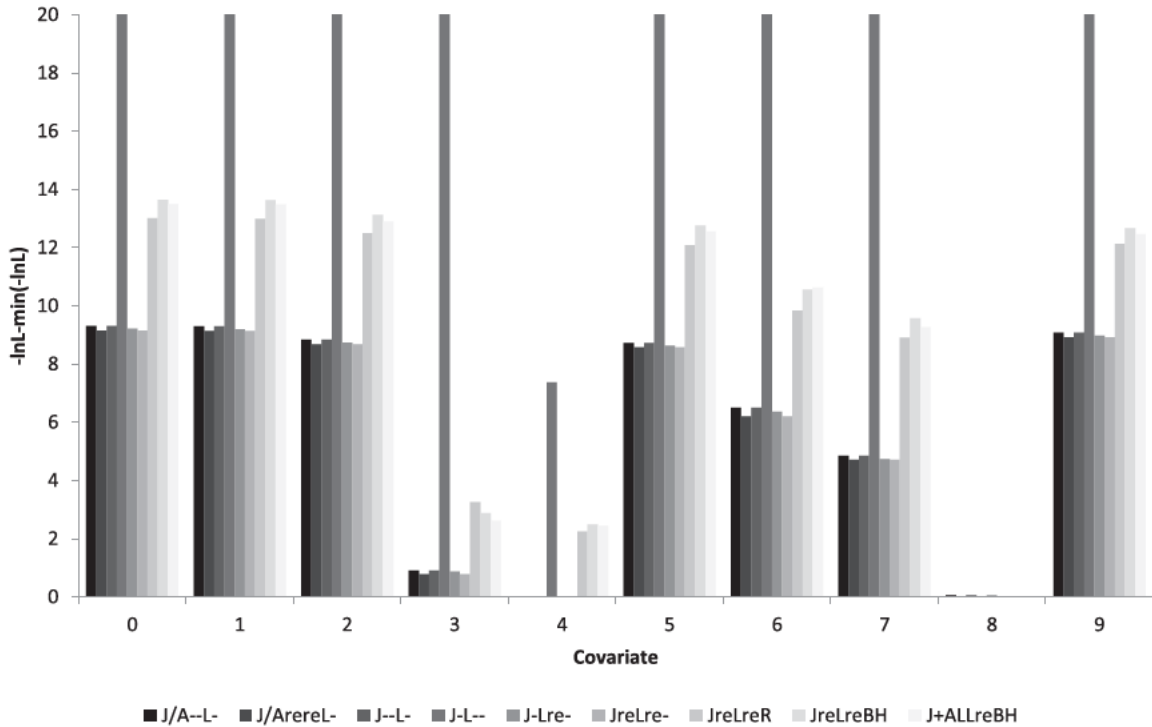
Fig. 2. Difference in negative log-likelihood from the model with the covariate minus the lowest negative log-likelihood for a scenario with any covariate $[-\ln L - \min(-\ln L)]$. A smaller value represents more support for that covariate compared to the other covariates in that scenario. The value for Model J-L- - is truncated. The covariates are presented in the same order as they are defined in Table 2, with the exception that "0" represents no covariates.

Table 3
AIC_C values for the steps (step order given in parenthesis) in the forward stepwise selection procedure and for models with no covariates and with different combinations of flow variables (temperature and ammonia included). AICc scores cannot be compared among some models because the data used to fit the model differs. Models with observation error in both abundance time series fit to both abundance time series are comparable (indicated by "b") but cannot be compared to models that fit to only the juvenile abundance time series (indicated by "a"). The two stage model (J+ALLreBH) includes two random effects and due to the method used to model random effects cannot be compared to the other models. The row labeled "Delta AICc" is the absolute difference in AICc from the selected model compared to the model without covariates for each scenario.

| Covariates	Analysis type								
	J/A–L- (a)	J/ArereL- (b)	J–L- (a)	J-L- (a)	J-Lre- (a)	JreLre- (b)	JreLreR (b)	JreLreBH (b)	J+ALLreBH (c)
Mysid May–June									
Secchi depth									
Eurytemera									
Napa River flow	105.14 (1)	−19.46 (4)	105.14 (1)	359.95 (4)	50.32 (1)	−19.46 (4)		−23.64 (4)	41.33 (5)
Predators central +San Pablo				351.12 (5)					
Average temperature	88.53 (3)	−16.65 (3)	88.53 (3)	405.73 (3)	33.44 (3)	−16.65 (3)	−21.99 (3)	−22.68 (3)	44.52 (3)
San Pablo ammonium	95.23 (2)	−13.10 (2)	95.23 (2)	650.94 (2)	40.28 (2)	−13.10 (2)	−17.39 (2)	−18.98 (2)	47.62 (2)
Sacramento River runoff		−4.72 (1)		1006.95 (1)		−4.72 (1)	−10.84 (1)	−12.17 (1)	55.11 (1)
Overbite clam presence									
Mysid July–Sept									42.16 (4)
Delta AICc	32.76	30.57	32.76	1222.22	32.83	27.77	34.48	36.09	38.02
No covariates	121.29	11.11	121.29	1573.34	66.27	11.11	12.49	12.45	79.35
Napa River runoff	88.53	−22.04	88.53	365.05	33.44	−22.04	−23.59	−23.89	43.76
Sacramento River runoff	94.07	−16.65	94.07	405.73	38.89	−16.65	−21.99	−22.68	44.52
Both flow variables	91.23	−19.46	91.23	359.95	36.10	−19.46	−23.01	−23.64	43.55
Best forward stepwise	88.53	−19.46	88.53	351.12	33.44	−19.46	−21.99	−23.64	41.33
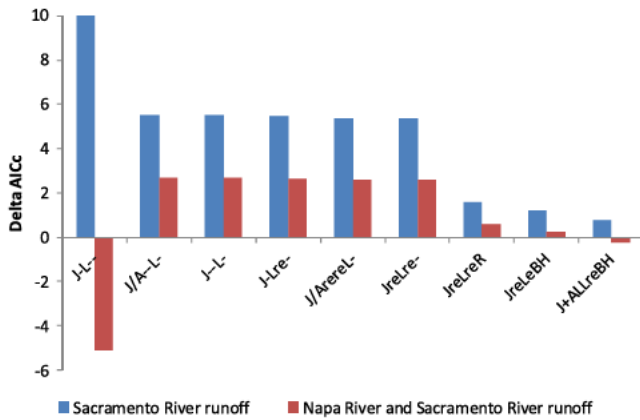
Fig. 3. Difference in AICC between the models with different flow variables. The blue histogram includes only Sacramento River runoff and the red histogram includes both Napa River and Sacramento River runoff. The ΔAICc values are the AICc values for these models minus the AICc values for the model with only Napa River runoff. The Sacramento River runoff value for model J-L-- is truncated. The models are ordered by ΔAIC$_C$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

included (Fig. 3). Consequently, there is "no credible" evidence supporting one runoff covariate over the other in the presence of density dependence. This differs from the result without density dependence, which shows "definite" evidence of Napa River runoff over Sacramento River runoff.

Using a population dynamics model by linking both stages using a Beverton–Holt relationship (model J+ALLreBH) produces nearly identical support for the covariates compared to the Beverton–Holt model when evaluating single covariate models. The final selected model adds the additional prey covariate for survival from juveniles to adults.

5. Discussion

We have illustrated the progression from traditional log-linear models for estimating the factors influencing survival to state-space population dynamics life-cycle models. State-space models accommodate both observation and process error, which can be vital to avoid bias in parameter estimates, confidence intervals, and hypothesis tests (de Valpine and Hastings, 2002; Maunder and Watters, 2003; Deriso et al., 2007). Our model that ignored process error selected prey as an additional covariate, which was not selected by any other model, and would have selected additional covariates if they had not been discarded because the coefficient was the wrong sign. In our application, ignoring observation error did not have a large impact on the relative support for the various covariates. However, it did change which covariates were selected because the two flow covariates were highly correlated. In other applications, the influence of including observation error is likely to be greater where observation error is larger and particularly if it varies among data points. Explicitly modeling process error and separating it from observation error is also important in estimating the probability of future events such as extinction (Maunder, 2004) and evaluating the uncertainty in the relationships between survival and covariates so this uncertainty can be included in management advice (Maunder and Deriso, 2011).

5.1. Observation error

The observation error standard deviations used in our application, calculated from bootstrap analysis of the survey data, were assumed known and were used to represent the random sampling error. They do not include variation due to other factors such

as annual changes in survey catchability. This additional observation error may influence hypothesis testing. The standard deviation representing additional variation in the observation process could be estimated analytically (Maunder and Starr, 2003; Deriso et al., 2007) or covariates could be added to the observation model, perhaps using finer scale data (e.g. Maunder, 2001; Besbeas and Freeman, 2006). Estimating the additional observation error variance adds one more parameter, which will increase the variance of parameter estimates and will probably reduce the statistical significance of covariates.

5.2. Process error

The estimated observation (sampling) error variance often incorporates the process error in models such as the log-linear and simple exponential models. They do not explicitly model the process error, but accommodate it by ignoring the observation error variances in the likelihood and estimating the variance of the likelihood function. However, it is important to understand that the variance estimates from these models represent a combination of process error and observation error. In more complex population dynamics models, such as those used in fisheries stock assessment (Maunder and Punt, 2013; Punt et al., 2013, Methot and Wetzel, 2013), which model many processes, only one type of process error is typically modeled (e.g. annual recruitment variability) and estimation of the observation error variance for a variety of data types or the modeled process error is implicitly assumed to accommodate the unmodeled process error.

Contemporary fisheries stock assessment models are often too complicated to model in a state-space framework, although some success has been achieved (McAllister and Ianelli, 1997; Maunder and Deriso, 2003; Nielsen and Berg, 2014), particularly in a Bayesian context (Punt and Hilborn, 1997). The standard approach is to use penalized likelihood, with the variance of the process error for annual recruitment fixed at a pre-determined value (Maunder and Deriso, 2003). Misspecified process error variance will bias confidence intervals and hypothesis tests. Adding covariates to explain process error will reduce the process error variance, and the variance needs to be adjusted for this. Hopefully, fisheries stock assessment models can be implemented in the state-space framework as computers and estimation algorithms get more efficient, so the process error variance can be estimated. In the meantime, it might be prudent to estimate the parameters and conduct hypothesis tests under different assumptions about the process error variance to ensure that results are consistent.

We found that modeling either process error or observation error as random effects or likelihood functions gave the same results. This was an interesting result and it is not clear if this is a general phenomenon or if it is a consequence of comparing linear Gaussian models. Further research is needed.

5.3. Model selection

Our results corroborate other studies that have found that evaluating factors in isolation can lead to different results than evaluating them in combination (e.g. Deriso et al., 2008; Maunder and Deriso, 2011). Similarly, our results parallel those of Maunder and Deriso (2011) who found that some final models had a coefficient with confidence intervals that cover zero, and removing that covariate improved the AIC$_c$. As with Maunder and Deriso's (2011) study, the covariate in question (Sacramento River flow) was highly correlated with another covariate (Napa River flow) included in the model.

Maunder and Deriso (2011) recommend that all possible combinations of covariates and density dependent factors should be evaluated because some factors may only be detected in

combination with other factors or in the presence of density dependence. However, conducting analyses of all possible combinations can be computationally demanding. To reduce the computational time, Maunder and Deriso (2011) applied a strategy that evaluates two covariates at a time and uses AIC_c summed over all possible one and two covariate combinations to select a covariate that has general support. In contrast, Anderson et al. (2000) warn against testing all possible combinations unless model averaging is used. Practical advice is to ensure that covariates included in the model have *a priori* support and that the framework of Maunder and Deriso (2011) is followed to identify the life stage and the relationship to density dependence before conducting an all combinations analysis. Results should be used to rank models and provide an idea of the data based evidence for alternative hypotheses rather than strict acceptance–rejection hypothesis testing (Maunder and Deriso, 2011).

5.4. Integrated analysis

We illustrated how multiple life stages of a species, each with their own data sets, can be integrated into a population dynamics model. This is an elementary form of the contemporary integrated analysis (also known as data assimilation), which attempts to include all relevant data into a single analysis (e.g. Maunder, 2003; Buckland et al., 2007; Schaub and Abadi, 2010; Maunder and Punt, 2013). Integrated analysis facilitates the propagation of information and uncertainty, particularly when states are linked from one time period to the next in a population dynamics model. For example, one life stage in the analysis of Maunder and Deriso (2011) did not have an abundance index until partway through the modeling time frame and the processes related to this stage were informed by the indices of abundance for other stages. However, the years that the index was available for were enough to help determine which stages the covariates influenced. Similarly, Tenan et al. (2012) showed how integrating different types of data allowed for the estimation of population processes not directly measured in the field. We found that adding data and a covariate for survival from juveniles to adults did not influence the support for the covariates of survival from adults to juveniles. This is somewhat reassuring since the application had good data for all time periods and therefore it would not be desirable for the results of one stage to influence those of another. If process error was not modeled, the added data may have inadvertently influenced the covariate selection. If the data were poor or missing for some time periods, then it would be reasonable and desirable for data for one stage to influence the other stages.

5.5. Model structure

The models we used to illustrate state-space models were simple compared to those used in many real applications. Alternative functions could be used to model the transition among stages. For example, Maunder and Deriso (2011) used the three-parameter Deriso–Schnute stock-recruitment model (Deriso, 1980; Schnute, 1985) and also allowed the flexibility to implement covariates before or after density dependence. The covariates were included as simple log linear terms and there may be more appropriate relationships between survival and covariates. For example there may be a dome shaped relationship between survival and temperature, with lower survival at lower and higher temperature or temperature may interact with prey availability.

5.6. Longfin smelt application

We found that multiple factors and density dependence influenced the survival of longfin smelt. The AIC_c was over four units higher for the Beverton–Holt model compared to the exponential model suggesting there is "definite" evidence for density dependence. The level of evidence is less if the models with Napa River flow are used. We also found that flow, ammonia, and temperature were consistently supported by the data for longfin smelt. Thomson et al. (2010) found that X2, which is related to flow, and water clarity explained longfin abundance. Mac Nally et al. (2010) also found that X2 explained longfin abundance, but in addition found a correlation with prey species. Among candidate flow variables, we did not find X2, OMR flow, or the two outflow threshold variables in supplemental Table 1 to be important covariates in our initial screening after the inclusion of flow variables that had higher support in the data.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fishres.2014.10.017.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Proceedings of the 2nd International Symposium on Information Theory. Publishing house of the Hungarian Academy of Sciences, Budapest, pp. 268–281.

Anderson, D.R., Burnham, K.P., Thompson, W.L., 2000. Null hypothesis testing: problems, prevalence, and an alternative. J. Wildl. Manag. 64 (4), 912–923, http://dx.doi.org/10.2307/3803199.

Baxter, R., Breuer, R., Brown, L., Chotkowski, M., Feyrer, F., Gingras, M., Herbold, B., Mueller-Solger, A., Nobriga, M., Sommer, T., Souza, K., 2008. Pelagic Organism Decline Progress Report: 2007 Synthesis of Results. California Department of Water Resources, Sacramento, USA (Interagency Ecological Program Technical Report 227).

Bennett, W.A., 2005. Critical assessment of the delta smelt population in the San Francisco estuary, California. San Franc. Estuary Watershed Sci. 3 (2), 1–71.

Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A., 2002. Integrating mark—recapture—recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58 (3), 540–547.

Besbeas, P., Lebreton, J.D., Morgan, B.J.T., 2003. The efficient integration of abundance and demographic data. Appl. Stat. 52, 95–102.

Besbeas, P., Freeman, S.N., 2006. Methods for joint inference from panel survey and demographic data. Ecology 87, 1138–1145.

Beverton, R.J.H., Holt, S.J., 1957. On the Dynamics of Exploited Fish Populations, Fisheries Investigations Series 2, 19. Ministry of Agriculture, London, U.K.

Bolker, B.M., Gardner, B., Maunder, M., Berg, C.W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., deValpine, P., Ford, J., Gimenez, O., Kery, M., Kim, E.J., Lennert-Cody, C., Magnusson, A., Martell, S., Nash, J., Nielsen, A., Regetz, J., Skaug, H., Zipkin, E., 2013. Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. Methods Ecol. Evol. 4, 501–512.

Brook, B.W., Bradshaw, C.J.A., 2006. Strength of evidence for density dependence in abundance time series of 1198 species. Ecology 87 (6), 1445–1451.

Buckland, S.T., Newman, K.B., Thomas, L., Koesters, N.B., 2004. State-space models for the dynamics of wild animal populations. Ecol. Model. 171 (1–2), 157–175.

Buckland, S.T., Newman, K.B., Fernandez, C., Thomas, L., Harwood, J., 2007. Embedding population dynamics models in inference. Stat. Sci. 22 (1), 44–58.

Burnham, K.P., Anderson, D.R., 1998. Model Selection and Inference: A Practical Information-theoretical Approach. Springer Verlag, New York.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. Springer, New York.

Christensen, R., 1997. Log-Linear Models and Logistic Regression, 2nd ed. Springer.

Ciannelli, L., Chan, K.-S., Bailey, K.M., Stenseth, N.C., 2004. Nonadditive effects of the environment on the survival of a large marine fish population. Ecology 85 (12), 3418–3427.

Deriso, R.B., 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37 (2), 268–282.

Deriso, R.B., Maunder, M.N., Skalski, J.R., 2007. Variance estimation in integrated assessment models and its importance for hypothesis testing. Can. J. Fish. Aquat. Sci. 64, 187–197.

Deriso, R.B., Maunder, M.N., Pearson, W.H., 2008. Incorporating covariates into fisheries stock assessment models with application to Pacific herring. Ecol. Appl. 18 (5), 1270–1286.

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

de Valpine, P., 2003. Better inferences from population dynamics experiments using Monte Carlo state-space likelihood methods. Ecology 84, 3064–3077.

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

Dupont, W.D., 1983. A stochastic catch-effort method for estimating animal abundance. Biometrics 39 (4), 1021–1033.

Edgington, E.S., 1987. Randomization Tests. Marcel Dekker, New York, New York, USA.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., Sibert, J., 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Opt. Met. SOFT 27, 233–249.

Hilborn, R., Mangel, M., 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton, N.J.

Hobbs, N.T., Hilborn, R., 2006. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. Ecol. Appl. 16 (1), 5–19.

Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends Ecol. Evolut. 19 (2), 101–108.

Lele, S.R., Dennis, B., Lutscher, F., 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. Ecol. Lett. 10, 551–563.

Lunn, D., Spiegelhalter, D., Thomas, A., Best, N., 2009. The BUGS project: evolution, critique and future directions. Stat. Med. 28 (25), 3049–3067.

Mac Nally, R., Thomson, J.R., Kimmerer, W., Feyrer, F., Newman, K.B., Sih, A., Bennett, W., Brown, L., Fleishman, E., Culberson, S.D., Castillo, G., 2010. Analysis of pelagic species decline in the upper San Francisco Estuary using multivariate autoregressive modeling (MAR). Ecol. Appl. 20 (5), 1417–1430.

Maunder, M.N., 2001. A general framework for integrating the standardization of catch-per-unit-of-effort into stock assessment models. Can. J. Fish. Aquat. Sci. 58, 795–803.

Maunder, M.N., 2003. Paradigm shifts in fisheries stock assessment: from integrated analysis to Bayesian analysis and back again. Nat. Resour. Model. 16 (4), 465–475.

Maunder, M.N., 2004. Population viability analysis, based on combining integrated, Bayesian, and hierarchical analyses. Acta Oecol. 26, 85–94.

Maunder, M.N., Deriso, R.B., 2003. Estimation of recruitment in catch-at-age models. Can. J. Fish. Aquat. Sci. 60, 1204–1216.

Maunder, M.N., Deriso, R.B., 2011. A state-space multistage life cycle model to evaluate population impacts in the presence of density dependence: illustrated with application to delta smelt (*Hyposmesus transpacificus*). Can. J. Fish. Aquat. Sci. 68, 1285–1306.

Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61–74.

Maunder, M.N., Starr, P.J., 2003. Fitting fisheries models to standardised CPUE abundance indices. Fish. Res. 63, 43–50.

Maunder, M.N., Watters, G.M., 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and examples. Fish. Bull. 101, 89–99.

McAllister, M.K., Ianelli, J.N., 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54, 284–300.

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

Miller, W.J., Manly, B.F.J., Murphy, D.D., Fullerton, D., Ramey, R.R., 2012. An investigation of factors affecting the decline of delta smelt (*Hypomesus transpacificus*) in the Sacraento-San Joaquin Estuary. Rev. Fish. Sci. 20, 1–19.

Nielsen, A., Berg, C.W., 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158, 96–101.

Newman, K.B., 1998. State-space modeling of animal movement and mortality with application to salmon. Biometrics 54 (4), 1290–1314.

Newman, K.B., Lindley, S.T., 2006. Accounting for demographic and environmental stochasticity, observation error and parameter uncertainty in fish population dynamics models. N. Am. J. Fish. Manag. 26 (3), 685–701.

Newman, K.B., Fernandez, C., Thomas, L., Buckland, S.T., 2009. Monte Carlo inference for state-space models of wild animal populations. Biometrics 65 (2), 572–583.

Pawitan, Y., 2003. In All Likelihood: Statistical Modeling and Inference using Likelihood. Oxford University Press, Oxford, UK.

Polacheck, T., Hilborn, R., Punt, A.E., 1993. Fitting surplus production models: comparing methods and measuring uncertainty. Can. J. Fish. Aquat. Sci. 50, 2597–2607.

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

Punt, A.E., Huang, T.-C., Maunder, M.N., 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. ICES J. Mar. Sci. 70 (1), 16–33.

Quinn II, T.J., Deriso, R.B., 1999. Quantitative Fish Dynamics. Oxford University Press, New York.

Ricker, W.E., 1954. Stock and recuitment. J. Fish. Res. Board Can. 11, 559–623.

Rivot, E., Prevost, E., Parent, E., Bagliniere, J.L., 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. Ecol. Model. 179 (4), 463–485.

Rose, K.A., Cowan Jr., J.H., Winemiller, K.O., Myers, R.A., Hilborn, R., 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish Fish. 2, 293–327.

Rosenfield, J.A., Baxter, R.D., 2007. Population dynamics and distribution patterns of longfin smelt in the San Francisco Estuary. Trans Am. Fish. Soc. 136, 1577–1592.

Sæther, B.E., 1997. Environmental stochasticity and population dynamics of large herbivores: a search for mechanisms. Trends Ecol. Evol. 12 (4), 143–149.

Schaub, M., Abadi, F., 2010. Integrated population models: a novel analysis framework for deeper insights into population dynamics. J. Ornithol. 152, 227–237.

Schnute, J., 1985. A general theory for the analysis of catch and effort data. Can. J. Fish. Aquat. Sci. 42 (3), 414–429.

Schnute, J.T., 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51, 1676–1688.

Seber, G.A., Wild, C.J., 1989. Nonlinear Regression. John Wiley and Sons, New York.

Skalski, J.R., 1996. Regression of abundance estimates from mark—recapture surveys against environmental covariates. Can. J. Fish. Aquat. Sci. 53, 196–204.

Skalski, J.R., Townsend, R.L., Steig, T.W., Hemstrom, S., 2010. Comparison of two alternative approaches for estimating dam passage survival using acoustic-tagged sockeye salmon smolts. N. Am. J. Fish. Manag. 30, 831–839.

Skaug, H.J., 2002. Automatic differentiation to facilitate maximum likelihood estimation in nonlinear random effects models. J. Comput. Graph. Stat. 11 (2), 458–470.

Skaug, H., Fournier, D., 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Comput. Stat. Data Anal. 51 (2), 699–709.

Sommer, T., Armor, C., Baxter, R., Breuer, R., Brown, L., Chotkowski, M., Culberson, S., Feyrer, F., Gingras, M., Herbold, B., Kimmerer, W., Mueller-Solger, A., Nobriga, M., Souz, K., 2007. The collapse of pelagic fishes in the upper San Francisco Estuary: El colapso de los peces pelagicos en la cabecera del Estuario San Francisco. Fisheries 32, 270–277.

Tenan, S., Adrover, J., Navarro, A.M., Sergio, F., Tavecchia, G., 2012. Demographic consequences of poison-related mortality in a threatened bird of prey. PLOS One 7 (11), 1–11.

Thomson, J.R., Kimmerer, W.J., Brown, L.R., Newman, K.B., Mac Nally, R., Bennett, W.A., Feyrer, F., Fleishman, E., 2010. Bayesian change point analysis of abundance trends for pelagic fishes in the upper San Francisco Estuary. Ecol. Appl. 20 (5), 1431–1448.

Walters, C.J., Ludwig, D., 1981. Effect of measurement errors on the assessment of stock-recruitment relationships. Can. J. Fish. Aquat. Sci. 38, 704–710.

Wang, S.-P., Maunder, M.N., Aires-da-Silva, A., Bayliff, W.H., 2009. Evaluating fishery impacts: application to bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. Fish. Res. 99 (2), 106–111.

White, G.C., Garrott, R.A., 1990. Analysis of Wildlife Radio-Tracking Data. Academic Press, San Diego.