CHAPTER 10: FORECASTING

10.1 An age-structured model for probabilistic forecasting of salmon populations

Eric R. Buhle (Northwest Fisheries Science Center, Seattle, WA), Richard W. Zabel (Northwest Fisheries Science Center, Seattle, WA), and Mark D. Scheuerell (Northwest Fisheries Science Center, Seattle, WA)

Introduction

An important component of the Adaptive Management Implementation Plan (AMIP) under the FCRPS BiOp is the capability to determine when evolutionarily significant units (ESUs) reach critically low abundance levels. When ESUs fall to specified levels, the AMIP calls for the Action Agencies, in coordination with NOAA Fisheries, the RIOG, and other regional parties to determine what Rapid Response Actions to implement.

Two metrics have been developed to assess when populations reach dangerously low levels. The abundance metric is based on a four year running average of population abundances. If this metric falls below a pre-determined threshold level, a trigger is tripped and actions will be initiated (see AMIP for details). Similarly, a trend metric measures trend in abundance over 5-year periods, similar to the trend metric adopted by the Biological Review Team (BRT). This metric is used in conjunction with the abundance metric to assess when populations are in trouble.

The AMIP also called for an additional early warning indicator that can provide an indication that a population could possibly fall below critical thresholds in the next two years. In this document, we present a forecasting tool that predicts population abundances in the next two years. The tool uses historical data (smolt counts and age-specific adult returns) to build a predictive model of adult returns rates. The predictive model takes into account how variable ocean conditions affect smolt-to-adult survival rates (SARs). The model also fully accounts for the uncertainty in its predictions. The primary output of the model is the probability that population abundance will fall below certain critical levels over the next two years. Managers
can then use this information to guide decisions on how to begin to undergo actions in the near future.

We note that the objective of this document is to present methods for forecasting adult returns. Further, the methods were designed to produce the type of output that is directly compatible with the decision analysis framework required for AMIP. However, this document does not make any specific recommendations on how this information should be used. In particular, we do not recommend whether the output should be used as an additional trigger mechanism, or whether it should be used as a mechanism to provide early warning for the existing trend and abundance metrics. These are policy choices that require further discussion.

**Overview**

The approach we describe for forecasting future adult returns is empirically and statistically driven, and is designed to take advantage of data on the annual abundance of juvenile outmigrants (smolts) and the abundance and age distribution of returning adults over some time period up to the most recent year. (See *Summary and Future Directions* for a discussion of possible approaches to other types of data sets.) These data provide information on the patterns of temporal variation in SAR and the adult age distribution (i.e., the proportion of total adults from a given cohort that return at age $a$). The task for any forecasting model is to use this retrospective information to make predictions of future population dynamics and to characterize the uncertainty in those predictions.

Although the ultimate test of such a model is its ability to generate accurate and precise forecasts of as-yet-unobserved data, we developed our modeling approach with three additional criteria in mind. First, we sought an approach that efficiently and simultaneously uses all the information available in the data, including smolt-to-adult survival and adult age distribution as well as any environmental covariates found to be useful in predicting these parameters. Second, we attempted to balance parsimony with realism by developing a model that avoids unnecessary complexity while admitting relevant biological details, such as temporal variability in survival and age distributions. Third, we required a formal statistical approach, in particular one that is able to generate predictive probability distributions of future population sizes and associated events (e.g., the probability of the adult population falling below a specified threshold).
Our forecasting approach has at its core a simple age-structured model that projects a cohort of outmigrating juveniles in a given year into surviving adults returning at various ages in future years. In a retrospective context, the parameters of this model (cohort-specific SARs and age distributions) could be estimated simply by fitting the observed time series of data, but predicting the future requires some means of predicting future values of the parameters. To do this, we adopted a hierarchical modeling framework, in which a process model describes the variation in the parameter values over time, as determined by higher-level parameters and predictive covariates, while a data (or observation) model describes the sampling distribution of the data (numbers of adult returns by age) given the parameters, allowing us to estimate the parameters by matching the model predictions to the data.

Hierarchical models of this sort have gained popularity in recent years in the environmental sciences (Clark 2005, Royle and Dorazio 2008, Cressie et al. 2009), and well-developed methods exist for fitting and validating them. We used a Bayesian statistical framework to fit our models, generate predictions, and evaluate forecasting performance. While it is possible in principle to analyze hierarchical models within the classical frequentist statistical paradigm (Royle and Dorazio 2008), Bayesian methods enjoy computational advantages (Clark 2005). More fundamentally, only Bayesian analysis can provide results in the currency we need, namely probability distributions of predicted outcomes (Hobbs and Hilborn 2006, Wade 2000), which is why Bayesian methods are a standard tool in risk and decision analysis (Punt and Hilborn 1997).

The modeling framework described below shares some features of other approaches that have been used to forecast salmon returns, but differs in some important ways. Like many previous studies (Scheuerell and Williams 2005, Zabel et al. 2006, Haeseker et al. 2005, Logerwell et al. 2003), our model allows the use of “leading indicator” covariates to predict survival or productivity. Most of these studies, however, have either ignored age structure (e.g., by focusing on species such as pink or coho salmon whose adult returns are dominated by a single age class), or dealt with it in fairly simplistic ways (e.g., by setting the age distribution of future cohorts equal to a recent average) which result in underestimates of forecast uncertainty. On the other hand, sibling regression methods (Peterman 1982, Haeseker et al. 2007) focus primarily on relative age composition of adult returns, using the numbers of younger age classes from a given cohort to predict the numbers of older age classes in subsequent years. Typical applications of sibling regression models, however, assume that the relative proportion of each
age (i.e., the slope of the regression of older siblings on younger ones) is constant through time. Moreover, these regressions do not make the most efficient possible use of the available information, as they use separate, independent relationships to estimate the relative proportions of successive age classes (e.g., 2- vs. 3-ocean, 3- vs. 4-ocean, and so on) when in reality these proportions are inherently correlated because they must sum to one. Our approach combines information on overall marine survival based on covariates with information provided by early returns from a cohort and generates probabilistic forecasts that account for uncertainty in the joint distribution of parameter estimates, the inherent stochasticity of population dynamics, and the sampling variability of the data.

**Methods**

*Observation Model*

The goal of our analysis is to predict adult returns for the next two years and to estimate the probability of a range of returns. To do this, we begin by developing a simple age-structured model that describes the basic population dynamics and makes predictions that can be compared to the available data. In particular, for many salmonid species returning adults in a given year are derived from several different cohorts of outmigrating smolts, and the model reflects this. Further, year-to-year variability in adult returns arises from several sources, and the age-structured model reflects two basic sources of variability: variability in smolt-to-adult return rates (SAR) and age composition. This age-structured model, along with a likelihood function that quantifies the fit between the predicted numbers of returning adults by age and the data, together comprise the observation model component of our hierarchical model. In the subsequent section we describe the process model, which governs the temporal evolution of the parameters in the observation model.
Table 1. Partial data set for aggregate counts of Snake River spring-summer Chinook smolts and returning adults (adjusted for mainstem Columbia River harvest). The full data set includes years 1964-2010. Bold italic entries illustrate two cohorts of outmigrating smolts that can be traced diagonally to recruits by age in subsequent years. The 2000 cohort is fully observed, while the 2008 cohort is only partially observed: 3-ocean adults have not yet been reported. Adults returning in years 2011 and 2012 can be forecast by our hierarchical age-structured model.

<table>
<thead>
<tr>
<th>Year (t)</th>
<th>Smolts (Jt)</th>
<th>Ocean Age (Ra,t)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td><strong>1344382</strong></td>
<td>1601</td>
<td>7415</td>
<td>437</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>490534</td>
<td><strong>1984</strong></td>
<td>45565</td>
<td>1347</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1128582</td>
<td>394</td>
<td><strong>17210</strong></td>
<td>14663</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>1455786</td>
<td>2496</td>
<td>7046</td>
<td><strong>25115</strong></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>1517951</td>
<td>1349</td>
<td>19412</td>
<td>2236</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>1734464</td>
<td>415</td>
<td>6371</td>
<td>2472</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>1227474</td>
<td>297</td>
<td>7167</td>
<td>1598</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>787150</td>
<td>1187</td>
<td>4005</td>
<td>4124</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td><strong>856556</strong></td>
<td>2532</td>
<td>12326</td>
<td>2944</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>929749</td>
<td><strong>3246</strong></td>
<td>10672</td>
<td>3281</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>1219742</td>
<td>1565</td>
<td><strong>28708</strong></td>
<td>1539</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td><strong>R_{1,2011}</strong></td>
<td><strong>R_{2,2011}</strong></td>
<td><strong>R_{3,2011}</strong></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td><strong>R_{2,2012}</strong></td>
<td><strong>R_{3,2012}</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To motivate the development of our observation model, consider a typical data set (Table 1) consisting of the number of juvenile outmigrants \((J_t)\) and returning recruits of ocean age \(a\), adjusted for harvest mortality, in calendar year \(t\) \((Ra,t)\). In order to predict the abundance of a cohort of fish that migrated as juveniles in year \(t\) and returned after \(a\) years in the ocean, we must first predict an SAR \((s_t)\) and an age class proportion \((p_{a,t})\) of adults returning at ocean age \(a\), both indexed by the year of migration:

\[
R_{a,t+\alpha} = \hat{s}_t \cdot \hat{p}_{a,t} \cdot J_t,
\]

where the “hat” notation signifies that these are unknown parameters to be estimated.

Given annual estimates of smolt-to-adult survival and the age distribution of recruits, we could predict the number of recruits by age in each calendar year and obtain a prediction of the
total return by summing across age classes. We might also be interested in subsets of the total return; for example, in the Snake River spring/summer Chinook case study below, we report forecasts as the total number of 2- and 3-ocean fish, ignoring 1-ocean fish (jacks):

\[ R_{T,t} = R_{2,t} + R_{3,t} . \]

In this example, predicting returning adults over a 2-year period requires predicting two age-classes from each of two cohorts (e.g., those representing 2- and 3-ocean adults in 2011 and 2012 in Table 1).

In order to estimate the model’s parameters we must specify a likelihood function, which defines the probability of observing the data \((R_{a,t} \text{ for all adult ages } a \text{ and return years } t)\) given some parameter values \((s_t \text{ and } p_{a,t} \text{ for all } a \text{ and } t)\). We assumed the numbers of recruits of each age from a given cohort follow a multinomial distribution where the sample size is the number of outmigrants, \(J_t\):

\[
[R_{1,t+1}, \ldots, R_{\text{maxAge},t+\text{maxAge}}, U_t] \sim \text{Multinomial}(J_t, [s_t, p_{1,t}, \ldots, s_t, p_{\text{maxAge},t}, 1 - s_t]). \tag{1}
\]

Here \(U_t\) is shorthand for the number of presumed mortalities at sea; that is, the number of fish that outmigrated in year \(t\) and never returned:

\[
U_t = J_t - \sum_{a=1}^{\text{maxAge}} R_{a,t+a} .
\]

By including \(U_t\) in the vector of observations, we have a likelihood function that depends on both SAR and the age distribution, allowing us to estimate these parameters jointly. Thus the distribution of “fates” of fish from cohort \(t\) includes those that survived and returned to freshwater at a given age, with probabilities given by the survival rate times the age proportion, and those that did not survive.

For partially observed cohorts (e.g., outmigration years 2008-2010 in Table 1), we pooled the unobserved “fate” categories in the data and the corresponding multinomial cell probabilities.
For example, if 1- and 2-ocean recruits from cohort \( t \) have been observed but 3-ocean recruits have not yet returned, the likelihood is

\[ [R_{1,t+1}, R_{2,t+2}, U_t] \sim \text{Multinomial}(J_t, [s_t, p_{1,t}, s_t, p_{2,t}, 1 - s_t(p_{1,t} + p_{2,t})]), \quad (2) \]

where \( U_t = J_t - R_{1,t+1} - R_{2,t+2} \). It might appear that the SAR and age proportions for partially observed cohorts are not uniquely identifiable, and indeed this would be the case if the parameters for each cohort were treated as separate and independent from those in all other cohorts. This is one of our main motivations for adopting a hierarchical approach: it allows earlier, fully observed cohorts in the data set to provide information (via the process model) about the likely distribution of parameters that are not well-defined by the data, while avoiding the often unrealistic assumption that the parameters simply remain constant over time. The next section describes the process model for SAR and the age proportions in detail.

**Process Model**

Conceptually, our approach to the problem of estimating the parameters (SAR and the age proportions) for the partially observed or unobserved cohorts in the most recent years involves a combination of (1) using predictive covariates such as ocean climate indices, which are typically hypothesized to act primarily on survival during the first year at sea and thus can be measured before any adults from the cohort return (i.e., “leading indicators”), and (2) using information gained from previous, fully observed cohorts to help inform the likely parameter values. The cohort-specific parameters are thus constrained on the one hand by the process model, and on the other hand by the observed data, which act via the likelihood to give more support to some parameter values than to others. The process model in turn has its own higher-level parameters, which are themselves unknown quantities to be estimated.

Specifically, the process model consists of two components, corresponding to smolt-to-adult survival and the recruit age distribution. We assumed that smolt-to-adult survival for each cohort \( t \) is drawn from a beta distribution,

\[ s_t \sim \text{Beta}(\phi, m_t), \]
where $\phi_t$ is the mean or expected value of $s_t$ and $m_s$ is a precision parameter (the larger $m_s$ is, the less survival is likely to differ from its expected value). The beta distribution is a flexible, natural choice that is commonly used for parameters such as survival that are bounded on the interval (0,1).

We modeled the cohort-specific mean survival as a function of covariates $X_1, \ldots, X_k$, using a logit transformation to linearize the relationship:

$$\log\left(\frac{\phi_t}{1-\phi_t}\right) = \beta_0 + \sum_{i=1}^{k} \beta_i X_{i,t},$$

(3)

where the coefficients include an intercept ($\beta_0$) and possibly one or more slopes ($\beta_i$). This model of covariate effects on survival is analogous to a beta regression model (e.g., Cribari-Neto and Zeileis 2010). Some versions of the model allowed temporal autocorrelation in the sequence of $s_t$ values, distinct from the effects of exogenous covariates, by including a first-order autoregressive [AR(1)] term in the residual errors of the logit-linear regression for mean survival:

$$\log\left(\frac{\phi_t}{1-\phi_t}\right) = \beta_0 + \sum_{i=1}^{k} \beta_i X_{i,t} + \rho_s \varepsilon_{t-1},$$

(4)

where $\rho_s$ is the autocorrelation coefficient (-1 ≤ $\rho_s$ ≤ 1) and the residual at the previous time step, $\varepsilon_{t-1}$, is defined as the difference between $s_{t-1}$ and $\phi_{t-1}$ on the logit scale. In this case a year in which survival is higher than expected, given the covariates, is likely to be followed by another higher-than-expected year ($\rho_s > 0$), or by a lower-than-expected year ($\rho_s < 0$). If $\rho_s = 0$ then Eq. 4 reduces to Eq. 3, in which case consecutive values of $s_t$ are independent after accounting for any covariate effects.

The second component of the process model assumes that the age distribution of recruits from each cohort is drawn from a Dirichlet distribution,

$$p_t \sim \text{Dirichlet}(\alpha, m_p),$$
where \( \boldsymbol{p}_t = [p_{1,t}, ..., p_{\text{maxAge},t}] \) is the vector of proportions of each ocean age up to the maximum in the data, \( \boldsymbol{\alpha} = [\alpha_1, ..., \alpha_{\text{maxAge}}] \) is the corresponding vector of mean (expected) proportions, and \( m_p \) is a precision parameter (the larger \( m_p \) is, the less the proportions in each age class will vary among cohorts). The Dirichlet is the multivariate generalization of the beta distribution, and is a standard choice for modeling vectors of probabilities, proportions, or relative compositions that are constrained to sum to one (Aitchison 2003). Because survival and the age distribution jointly determine the likelihood of the data, their precision parameters cannot be estimated separately, so we defined \( m = m_s + m_p \) as the overall precision. In some versions of the model, we allowed temporal autocorrelation in the sequence of \( \boldsymbol{p}_t \) vectors. To facilitate this, we transformed \( \boldsymbol{\alpha} \) using the additive log ratio transformation (Aitchison 2003), which is the multivariate generalization of the logit transformation, and included first-order autoregressive [AR(1)] terms in the residual errors of the transformed parameters:

\[
\log \left( \frac{\alpha_i}{\alpha_{\text{maxAge}}} \right) = \mu_i + \rho \varepsilon_i,t-1, \tag{5}
\]

where \( \rho_p \) is the autocorrelation coefficient \((-1 \leq \rho_p \leq 1)\) and the residual at the previous time step \((\varepsilon_i,t-1)\) is defined as the difference between the log ratio of \( \alpha_i \) and its mean, \( \mu_i \).

**Bayesian Model-Fitting and Forecasting**

We used a Bayesian framework to fit the model to data, where the data consist of time series of smolts, adult recruits by age, and any covariate values used in predicting mean SAR. As discussed above (see Overview), we chose the Bayesian framework partly for its computational tractability (relative to frequentist inference) in dealing with nonlinear hierarchical models such as ours, but more importantly because it provides intuitive probability statements about model-derived quantities, such as the number of recruits in some future year, and for this reason is well suited to forecasting and decision-making problems in general (Clark 2005, Hobbs and Hilborn 2006).
We placed noninformative proper uniform priors on all the parameters (the $\beta$ coefficients, the mean log ratio $\mu_i$ for each age proportion $i$, $\rho_s$, $\rho_p$ and $m$). We drew samples from the joint posterior distribution of the parameters using Markov Chain Monte Carlo (MCMC) simulation (Gelman et al. 2003) implemented in the JAGS software, used standard diagnostics to assess convergence, and visually compared prior and posterior distributions to check for prior sensitivity.

To make a prediction for a given cohort we drew a sample from the multinomial data distribution (Eq. 1), using a vector of parameters $s_t$ and $p_t$ drawn from the posterior distribution. Repeating this for every set of parameters in the posterior sample generated by MCMC then gives a posterior predictive distribution of recruits of each age from the cohort. This posterior predictive distribution integrates three levels of variability: (1) parameter estimation uncertainty due to fitting the model to a finite set of data, (2) temporal stochasticity in the underlying population dynamics (described by the process model), and (3) sampling variability via the data-generating model. Predicted returns-at-age can easily be aligned by calendar year, and summing across age classes in a given calendar year gives a prediction of total annual returns. Such predictions may be retrospective (i.e., hindcasts), such as for years 2000-2010 in Table 1, or prospective (i.e., forecasts), as for years 2011-2012 in Table 1. In either case, the posterior predictive distribution naturally lends itself to probability statements such as, “There is a 95% probability that total returns in year $t$ will lie in the interval $(R_{lo}, R_{hi})$” or “The probability that total returns in year $t$ will fall below some specified management threshold $R^*$ is $P$."

It is worth pointing out that in the case of partially observed cohorts, the posterior predictive distribution of forecasts—that is, numbers of recruits in the as-yet-unobserved older age classes—automatically incorporates the information provided by the younger age classes whose numbers are already known. This information enters through the likelihood (Eq. 2) and adjusts the cohort-specific parameters to achieve a balance between fitting the observations in the younger age classes and conforming to the process model, which describes patterns common to all cohorts. Thus our prospective forecasts of older fish are conditioned on returns of their younger siblings, albeit in a more subtle way than in the traditional sibling regression models.
Application to Snake River Spring-Summer Chinook

As a case study to illustrate the potential applications of our modeling framework, we used aggregate run size data for the Snake River spring-summer Chinook ESU, a subset of which is shown in Table 1 (see Scheuerell and Williams 2005 for details). We chose this ESU because long time series of smolt and age-structured adult counts are available and because previous studies have identified possible environmental predictors of marine survival. The full data set includes estimates of the total number of smolts passing Lower Granite Dam in calendar years 1964-2010 and the total number of adults at each ocean age (1, 2, or 3, ignoring 4-ocean adults as explained above) returning in calendar years 1964-2010, which gives full return data for the 1964-2007 smolt cohorts and partial data for the 2008-2009 cohorts. As covariates of marine survival, we considered (1) the annual average of the monthly Pacific Decadal Oscillation index (PDO, Mantua et al. 1997), (2) the annual average of the multivariate El Niño Southern Oscillation Index (ENSO1), (3) three monthly coastal upwelling indices (April, September, and October CUI; see Scheuerell and Williams 2005), and (4) monthly sea surface temperatures (April-December SST) just offshore of the Columbia River mouth2, all in the year of ocean entry. This is by no means an exhaustive list of potentially useful covariates, and is intended merely as an illustrative example.

We developed a set of candidate models (Table 2) by initially fitting a set of beta regressions (Cribari-Neto and Zeileis 2010) to the full time series of observed SAR. These models are simple GLMs fitted by maximum likelihood, and were used only to identify promising sets of covariates for the SAR process model (Eq. 3). In this initial screening we used Akaike’s information criterion, corrected for sample size (AICc, Burnham and Anderson 2002) to rank all subsets of the global model, which included all covariates. Because it would be computationally prohibitive in terms of CPU time to explore this full set of models in a Bayesian forecasting context, we retained just the subset of models whose AICc values were within 2 units of the best model (ΔAICc ≤ 2) for the full analysis described above. In addition, we included two models with no covariates and only an intercept for survival, either with autocorrelated errors in survival and the age distribution (Eq. 3) or without (Eq. 2). We fit each of these models to the full dataset using

1 Available from http://www.esrl.noaa.gov/psd/enso/mei/
2 Available from http://www.esrl.noaa.gov/psd/data/timeseries/
Bayesian MCMC and calculated the deviance information criterion (DIC, Spiegelhalter et al. 2002), a Bayesian measure of model performance that, like AIC in a frequentist context, balances goodness of fit to the data against parsimony (Burnham and Anderson 2002). Like AIC, DIC is derived from a theory based on optimizing out-of-sample predictive ability, but its actual calculation is based only on the retrospective fit to the data.

To directly assess forecasting performance, we iteratively fitted each model to subsets of the data, beginning with years 1964-1973 and sequentially adding one year at a time through 2010. For each subset of years 1, 2, ..., \( t \), we generated forecasts of adult returns (the sum of ocean ages 2 and 3, ignoring jacks) in years \( t+1 \) and \( t+2 \). These forecasts can be compared to the observed adult returns in years 1974-2010 (for the one-year-ahead forecasts) or 1975-2010 (for the two-year-ahead forecasts), and the comparisons can be formalized by calculating quantitative metrics of forecast performance. We used two such metrics. The first, root mean squared error (RMSE) is defined as the squared deviation between the observed \( (R_{t,t}) \) and predicted \( (\hat{R}_{t,t}) \) values, averaged over the \( M \) samples from the posterior distribution and then averaged across the \( Y \) forecast years:

\[
RMSE = \frac{1}{Y} \sum_{t=1}^{Y} \left[ \frac{1}{M} \sum_{m=1}^{M} \left( R_{t,t} - \hat{R}_{t,t} \right)^{2} \right].
\]

The second, posterior predictive deviance (PPD, Gelman et al. 2003) is based on the multinomial likelihood of the data under the model (Eq. 1) and is defined as two times the negative log-likelihood of a future observation (here, the numbers in each age class in year \( t+1 \) or \( t+2 \) ), averaged over the posterior distribution. These values are then summed over all years to give an overall PPD for one- or two-year-ahead forecasts:

\[
PPD = \sum_{t=1}^{Y} \left\{ \frac{1}{M} \sum_{m=1}^{M} \sum_{a=1}^{3} \left[ R_{a,t} \log \left( s_{t-a} p_{a,t-a} \right) + \left( J_{t-a} - R_{a,t} \right) \log \left( 1 - s_{t-a} p_{a,t-a} \right) \right] \right\}.
\]

Lower values of PPD imply that the observed data were more likely to occur, given the forecasts.
Results

The models we fit to the Snake River spring-summer Chinook salmon data illustrate the benefits of incorporating complexity in the form of predictors of marine survival and serial autocorrelation in survival and the adult age distribution. The simplest model we considered, with no covariates and serially independent errors in survival and the age distribution, performed relatively poorly at fitting the full data set (reflected in a high DIC) and at predicting the near future (reflected in high RMSE for one-year-ahead forecasts and high RMSE and PPD for two-year-ahead forecasts; Table 1). Allowing serial autocorrelation in survival and the age proportions did not substantially change the fit to the data (DIC), but did improve forecasting ability. Including covariates of survival resulted in improvements in DIC that ranged from negligible to substantial, but all models with covariates showed major improvements in forecast performance metrics. Despite large differences in DIC among the four models with covariates, each gave very similar forecasts, underscoring the fact that the ability of a model to fit retrospective data well does not necessarily translate into success at predicting the future.

The one-year-ahead forecasts of total returns (ocean ages 2 and 3) based on the more complex models captured the patterns of the observed data fairly well, as we illustrate (Fig. 1A) for a model with CUI, PDO, and SST covariates (model 3 in Table 2). As expected, two-year-ahead forecasts are slightly less accurate than one-year-ahead forecasts (Fig. 1B), and also more clearly reveal the relative performance of different candidate models, especially if PPD is used as the criterion (Table 1).

It is straightforward to combine these forecast distributions to predict quantities of interest for management, such as running sums (Fig. 1C) or averages of recruitment over a specified time horizon. One quantity of particular interest is the probability that the population size will fall below some specified threshold in the near future, which can be read off from the posterior cumulative distribution function (Fig. 2). From the perspective of forecasting the risk of decline or quasi-extinction, it is arguably more critical for the quantiles of the forecast distribution to accurately match the quantiles of future observations than for the mean or median forecast to match the observed values. In other words, events that the model predicts to be rare (e.g., next year’s return falling at or below the 5th percentile of the forecast distribution) should in fact occur rarely (in this case, approximately 5% of the time). Although both the RMSE and PPD metrics
reflect this distributional forecasting accuracy to some degree, we also present an intuitive graphical assessment. Each observation of the total adult return is expressed as a quantile of the same year’s forecast probability distribution, and the cumulative frequency distribution of these forecast quantiles is shown. This relationship should follow the 1:1 line on average if the forecast is distributionally accurate. These plots generally show close agreement between the predicted and observed distributions of future adult returns (Fig. 3), indicating that the model performs well at predicting the probability of population growth or decline. There is a slight excess of observations that fell in the lowest quartile of the corresponding forecast distribution (e.g., 12/37 or 32% of the observations fell at or below the 20th percentile of the one-year-ahead forecasts), suggesting that this particular model slightly underestimates the probability of population decline. Many of these overly optimistic forecasts were made in roughly the first seven years of forecasting (1974-1980; see Fig. 1A), when <20 years of data were available to fit the model and when SAR was declining precipitously. We discuss possible further refinements to model performance in the next section.

Summary and Future Directions

We believe the modeling framework presented here will be broadly applicable to short-term forecasting of adult population size in salmonids. The strengths of our approach are that it efficiently uses the available information (the size of outmigrating smolt cohorts, relationships between marine survival and “leading indicator” covariates, and patterns of covariance among survival and age distribution parameters in historical data), and that it provides a realistic accounting of multiple sources of uncertainty (parameter estimation error, temporal variability in population dynamics, and sampling error) in the forecasts. In contrast to other commonly used forecasting models, we avoid making unrealistic assumptions about the temporal constancy of parameters such as the adult age distribution. Like all forecasting approaches, our model requires up-to-date data in order to make forecasts of the future. Moreover, better data allow better predictions. All data, particularly when aggregated to the scale of a salmon population or ESU, contain observation error. One possible future modification of our approach would involve explicitly modeling the observation error in smolt and adult estimates within a state-space framework (Newman et al. 2006).
The application to Snake River spring-summer Chinook illustrates the benefits of using covariates to help predict marine survival and including process variability, e.g. in the adult age distribution. The simple example presented here is not intended to be a final, definitive model for this ESU, and clearly there is room for improvement. One avenue will be to explore other possible covariates of marine survival, including physical oceanographic variables such as characteristics of the Columbia River plume, as well as biological variables such as the abundance and species composition of prey (Peterson et al. 2010). In addition, we will examine other covariates, such as river flow, that have also been demonstrated to influence marine survival. Another avenue of model development will be to incorporate additional structure into the process model, such as environmental covariates to help predict the distribution of age-at-return probabilities. As a comparison, we will also examine several simpler models, such as standard sibling regression models and naive (that is, not population dynamics-based) time series models that relate abundance in the next year or two to current abundance. Finally, it is often the case that different models perform best under different conditions, and predictive ability is often improved by averaging forecasts across multiple models. Bayesian averaging of multiple model forecasts (Gelman et al. 2004) would be a fairly straightforward extension of our approach.

We are currently compiling datasets for other populations/ESUs (Snake River steelhead, Chiwawa River spring Chinook, Umatilla River steelhead) to expand the modeling. We will also continue to search for other appropriate datasets and hope to inform monitoring efforts on the most useful types of data to collect to build datasets for future analyses. Conversely, we also plan to explore modeling approaches that might be used in more data-limited situations where, for example, smolt counts are not available. One possibility in this case would be to use a spawner-recruit model to predict the total number of returning adults from a given brood year, thus replacing the smolt counts and SAR values in our model with spawner counts and productivity and capacity values. An alternative approach would be to ignore smolt-to-adult survival or adult-to-adult productivity altogether and focus solely on the relative abundance of different age classes, as in sibling regression models but within a more unified framework allowing for interannual variation in the adult age distribution. Another topic that will merit further discussion is how to scale model results up or down to correspond to the appropriate conservation units.
Literature Cited


Table 2. Forecasting models applied to Snake River spring-summer Chinook salmon. Models included monthly coastal upwelling indices (CUI), monthly sea surface temperatures (SST), mean annual Pacific decadal oscillation (PDO), and mean annual El Niño Southern Oscillation Index (ENSO) as predictors of smolt-to-adult survival, or no covariates. Some models included autocorrelated errors [AR(1) terms] in cohort-specific mean survival and age distribution parameters. The deviance information criterion (DIC, scaled so the best model has a value of zero), root mean squared error (RMSE), and posterior predictive deviance (PPD) for one- and two-year-ahead forecasts are shown for each model (see Methods for details).

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariates</th>
<th>Survival AR(1)</th>
<th>Age AR(1)</th>
<th>ΔDIC</th>
<th>One year ahead</th>
<th>Two years ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>PPD</td>
<td>RMSE</td>
<td>PPD</td>
<td>RMSE</td>
</tr>
<tr>
<td>1</td>
<td>None</td>
<td>No</td>
<td>No</td>
<td>14.99</td>
<td>23087.66</td>
<td>22111.46</td>
</tr>
<tr>
<td>2</td>
<td>None</td>
<td>Yes</td>
<td>Yes</td>
<td>15.43</td>
<td>19660.66</td>
<td>21093.41</td>
</tr>
<tr>
<td>3</td>
<td>CUI_{Apr} + CUI_{Oct} + PDO + SST_{Jun} + SST_{Nov}</td>
<td>Yes</td>
<td>Yes</td>
<td>16.45</td>
<td>11968.43</td>
<td>14151.59</td>
</tr>
<tr>
<td>4</td>
<td>CUI_{Apr} + CUI_{Oct} + PDO + SST_{Apr} + SST_{Jun} + SST_{Nov}</td>
<td>Yes</td>
<td>Yes</td>
<td>4.29</td>
<td>10857.14</td>
<td>14091.69</td>
</tr>
<tr>
<td>5</td>
<td>CUI_{Apr} + CUI_{Sep} + CUI_{Oct} + PDO + SST_{Jun} + SST_{Nov}</td>
<td>Yes</td>
<td>Yes</td>
<td>0.00</td>
<td>11623.82</td>
<td>13391.42</td>
</tr>
<tr>
<td>6</td>
<td>CUI_{Apr} + CUI_{Oct} + ENSO + PDO + SST_{Jun} + SST_{Nov}</td>
<td>Yes</td>
<td>Yes</td>
<td>16.09</td>
<td>14447.08</td>
<td>18630.05</td>
</tr>
</tbody>
</table>
Figure 1. Forecasts of total adult returns (ocean age 2 and 3, adjusted for mainstem Columbia River harvest) of Snake River spring-summer Chinook salmon. Forecasts were generated by fitting model 3 (see Table 2) to subsets of the data, beginning with the first 10 years and sequentially adding one year at a time, and predicting returns (A) one or (B) two years in the future, as well as (C) the two-year running sum of returns. Medians (thick line) and 95% credible intervals (thin lines) are shown for all forecasts. The vertical gray bar indicates 2010, the last year of adult return estimates in our data set.

Figure 2. Cumulative posterior distribution of the adult (ocean age 2 and 3, adjusted for mainstem Columbia River harvest) return forecast of Snake River spring-summer Chinook salmon for 2011, based on model 3 (see Table 2) fitted to data from 1964-2010. The dotted lines indicate a hypothetical conservation or management threshold (the x-axis value) and the predicted probability that the population will fall below that threshold (the y-axis value).

Figure 3. Cumulative frequency distributions of observed adult Snake River spring-summer Chinook salmon returns (ocean age 2 and 3, adjusted for mainstem Columbia River harvest), expressed as quantiles of the forecasted returns in (A) one-year-ahead or (B) two-year-ahead forecasts based on model 3 (see Table 2). The 1:1 line indicates the agreement between predicted and observed quantiles that would occur on average if the forecast perfectly captured the distributional properties of the data.
Figure 1.
Figure 2.
Figure 3.