Expected future performance of abundance forecast models with application to Sacramento River fall Chinook salmon

Arliss J. Winship[a,b,*] Michael R. O'Farrell[b]
William H. Satterthwaite[b,c] Brian K. Wells[b]
Michael S. Mohr[b]

[a] Institute of Marine Sciences, University of California, Santa Cruz

[b] Fisheries Ecology Division, Southwest Fisheries Science Center,
National Marine Fisheries Service, NOAA, 110 Shaffer Road, Santa Cruz, CA 95060, USA

[c] Center for Stock Assessment Research, Applied Mathematics and Statistics,
University of California, Santa Cruz, CA 95064, USA

* Corresponding author
Current address: Biogeography Branch, Center for Coastal Monitoring and Assessment,
National Centers for Coastal Ocean Science, National Ocean Service, NOAA,
1305 East-West Hwy, SSMC-4, 9th Floor, Silver Spring, MD 20910, USA
Tel: +1-301-713-3028,
E-mail: arliss.winship@noaa.gov

September 13, 2013

Abstract

The management of Pacific salmon fisheries relies heavily on abundance forecasts, and there is continued interest in their improvement. Using Sacramento River fall Chinook salmon as a case study, we evaluated the scope for improving the current forecast approach that relates the Sacramento Index (SI; an index of adult age 3–5 ocean abundance) to jack (estimated age 2) spawning escapement from the previous year. Alternative models added effects of density dependence, local environmental conditions, the abundance of the previous cohort, and trends or autocorrelation in the jack-to-SI relationship. Forecast performance was assessed with two cross-validation frameworks allowing evaluation of bias, accuracy, the ability of the models to track trends in the SI, and the potential for forecast errors to be of sufficient magnitude to cause management errors. Several models achieved higher accuracy than the current model, but no single model performed best across all criteria, and substantial forecast error remained across all approaches considered. Models incorporating trends or temporal autocorrelation in the jack-to-SI relationship demonstrated potential for modest forecast improvements with relatively little additional model complexity.

Introduction

Management of Pacific salmon (*Oncorhynchus* spp.) along the west coast of North America relies heavily on annual abundance forecasts. Abundance forecasts for key stocks along the U.S. west coast are used to define annual catch limits and exploitation rate ceilings through the use of harvest control rules (PFMC, 2012a). These stock-specific limits to exploitation rates lead to constrained ocean fisheries through the practice of weak-stock management and are also used to configure terminal-area fisheries. Overforecasting of abundance results in higher allowable exploitation rates and the potential reduction of spawners to suboptimal levels. On the other hand, underforecasting of abundance results in lower allowable exploitation rates and foregone catch (Bocking and Peterman, 1988). Ideally, forecasts should be unbiased in the long term, be able to predict short-term variation in population abundance, and be able to track trends in population abundance. Despite their central role in salmon fishery management, many abundance forecasts have been characterized as relatively inaccurate (Adkison and Peterman, 2000) and a source of management uncertainty (Holt and Peterman, 2006).

In practice, many abundance forecast models currently used for salmon management are variants of the 'sibling model' (Peterman, 1982), where age-specific estimates of freshwater returns are used to forecast ocean abundance or freshwater returns of older members of the same cohort at a later time. Freshwater returns of salmon at early ages (i.e., at age 2) provide an indication of cohort strength as they have survived the high and variable mortality rates associated with early life and have effectively integrated the sources of mortality encountered up to that stage. Yet variation in sibling relationships can be substantial (Noakes et al., 1990), and a variety of factors could contribute to that variation. There has been interest in developing more complex abundance forecast models in an attempt to reduce forecast errors. Some of these models incorporate environmental and biotic variables along with salmon freshwater return data (e.g., Roth et al., 2007; Wang et al., 2009; Rupp et al., 2012; Burke et al., 2013). However, it is not yet clear whether incorporation of environmental or biotic variables into salmon forecast models result in substantial improvements in future forecast performance relative to simpler models (Haeseker et al., 2005). Methods and metrics used to assess future performance of abundance forecast models differ among studies, hence expected future forecast skill can be difficult to assess. Furthermore, it generally remains unclear whether the improvement in forecast performance would be large enough to improve harvest management (Walters, 1989).

The annual, pre-fishery forecast of the abundance index for Sacramento River fall Chinook salmon (SRFC) can have a strong effect on the configuration of ocean fisheries off California and Oregon, USA. SRFC are the largest contributing stock to the mixed-stock commercial and recreational fisheries in this area (O'Farrell et al., 2013), and the stock abundance was historically large enough that it did not constrain ocean fisheries. However, SRFC experienced a steep decline beginning in the early 2000s (Lindley et al., 2009), prompting the development of the Sacramento Index (SI), an index of adult (age 3–5) SRFC ocean abundance. Very low SI forecasts led to complete closures or heavily constrained ocean salmon fisheries off California and Oregon from 2008–2010.

Since its initial development in 2008, the SI has been forecast with a zero-intercept linear model with jack (estimated age-2) spawning escapement as the independent variable (e.g., see Figure II-2 in PFMC, 2013a). In most years, the data used to fit the SI forecast model were SI estimates from 1990 through the year prior to the forecast year and jack spawning escapement data from 1989 through two years prior to the forecast year. The SI was well forecast in 2008 but was overforecast from 2009–2012. In 2012, use of the status-quo forecast model and data range would have predicted a record value of the SI (approximately 2.2 million), but the Pacific Fishery Management Council's (PFMC) Salmon Technical Team did not view this forecast as credible and for this one year limited the SI data used for forecasting to the most recent three years (2009–2011), resulting in a much lower forecast (approximately 819,000 fish). Examined retrospectively, this adjustment to the data range was warranted; the SI in 2012 was approximately 618,000 fish (PFMC, 2013a).

The recent history of overforecasting the SI has led to several hypotheses regarding why the well-defined, linear jack-to-SI relationship for years 1990–2008 had apparently broken down. One leading hypothesis for the recent poor forecast performance was that the jack spawning escapement data, which only confer information about the age-3 cohort, would tend to overforecast the multi-cohort SI in situations when year class strength was increasing from one year to the next. Under such a scenario, the SI would be largely composed of the age-3 cohort, while past data used to fit the jack-to-SI relationship were composed of SI values where older age classes were more fully represented (PFMC, 2012b). Another hypothesis for the change in the jack-to-SI relationship was that a change in the average age-2 maturation rate had occurred, with myriad environmental and genetic explanations proposed for this shift. While it is not currently clear what mechanisms led to the recent history of overforecasting the SI, an investigation into alternative models and data for forecasting the SI is justified.

Using the SI as a case study, we evaluated a variety of salmon abundance forecast models and compared the performance of these models to the status quo approach. The alternative models included intercept-only models, models with jack spawning escapement variables, and models with both jack spawning escapement and environmental variables. The environmental variables considered were previously demonstrated to be related to growth and age at maturity for California Chinook salmon (Wells et al., 2007). The models covered linear and nonlinear functional relationships, temporal trends and autocorrelation in the jack-to-SI relationship, and multivariate statistical methodology. Performance was evaluated using two cross-validation procedures and several measures of forecast error. We judged a forecast model good if it produced unbiased, accurate out-of-sample forecasts that tracked trends in the SI. Translations of forecast errors into management errors, defined as allowing exploitation rates that would result in conservation errors (under-escapement) or fishery errors (foregone fishing opportunity), were used to further evaluate model performance. This multi-faceted approach to forecast evaluation was designed to provide realistic expectations for future performance.

Methods

Data

The SI in year t is calculated as the sum of three estimates: ocean harvest from the period between 1 September in year $t-1$ and 31 August in year t, river harvest in year t, and spawner escapement in year t, each derived from different data sources and methods as described in O'Farrell et al. (2013). We used SI values from 1983–2012 in our analysis (Fig. 1; PFMC, 2013a).

The jack data were estimates of the total number of age-2 spawners returning to natural areas and hatcheries each year. Various methods were used to derive these estimates (O'Farrell et al., 2013; PFMC, 2013b). Jack spawning escapement was estimated, in general, on the basis of a length threshold, and this threshold varied among sites and over time. We used jack data from 1981–2011 in our analysis (Fig. 1; PFMC, 2013b).

Eight local environmental variables were chosen (Table 1) for their potential to explain variability in the survival and maturation rates of fish in the ocean (Wells et al., 2007, 2008). The variables represented sea surface temperature, wind speed and direction, curl, strength and timing of upwelling, and sea level height off the coast of Central California (Fig. 2). The data for these environmental variables were obtained from publicly accessible online sources (Table 1). For each variable, except STI_39N, three average seasonal values were calculated for each year—spring (March–May), summer (June–August), and fall (September–November)—resulting in a total of 22 variables related to the environment. We assumed that the relationship between the SI in year t and the number of jacks in year $t-1$ was influenced by environmental conditions during year $t-1$, the period leading up to and around the time that age-2 fish either return to spawn or remain in the ocean. In other words, the SI in year t was assumed to be partially a function of the environment in year $t-1$. We analyzed environmental data from 1982–2011.

Models

We evaluated 13 models for forecasting the SI (Table 2). Two of these were intercept-only models (Models 1 and 6, the arithmetic and geometric means, respectively, of the SI time series), one was the model currently used to forecast the SI as a function of the number of jacks the previous year (Model 2), and the remaining 11 models were modifications or extensions of Model 2 that covered a range of functional relationships between the SI and previous numbers of jacks and environmental conditions.

Models 2–5 were zero-intercept models that assumed a proportional relationship between the SI and previous numbers of jacks. Models 7–13 were on the log scale with respect to the SI and the number of jacks, allowing for a nonlinear power relationship between these variables on the arithmetic scale that could capture density dependence in the jack-to-SI relationship (Peterman, 1981). Such density dependence could arise from density dependence in the maturation or survival rates, for example.

Models 2 and 7 assumed that the SI was a function of only the number of jacks the previous year. However, the SI indexes the ocean abundance of fish aged 3 and older, thereby representing fish from multiple cohorts. Only the age-3 component of the SI would belong to the same cohort as jacks from the previous year. Models 3 and 9 therefore assumed that the SI was a function of the number of jacks in the previous two years, thus incorporating additional information about the abundance of the next oldest cohort (i.e., the age-4 component of the SI).

Models 4 and 10 were analogous to Models 2 and 7, but in these models the expected ratio of the SI to the number of jacks the previous year was allowed to change over time following a relatively simple, smooth, nonlinear relationship (maximum 3-degree-of-freedom relationship over the 30-year study period). These models were intended to capture gradual changes in average maturation and/or survival rates over time arising from, for example, low-frequency environmental or food web variation. Model 8 also allowed for temporal changes in the expected ratio of the SI to the number of jacks the previous year through autocorrelation in residual errors.

Models 5 and 11–13 allowed environmental variables to modify the jack-to-SI relationship. Model 5 was linear on the arithmetic scale, and environmental effects were incorporated as interactions between the number of jacks the previous year and environmental variables. Model 11 was a multiple linear regression of the log-transformed SI and jack data with main effects of the untransformed environmental variables. To explore nonlinearity in the relationships between environmental variables and the ratio of the SI to the number of jacks we considered Model 12, a generalized additive model (e.g., Rupp et al., 2012) that was similar to Model 11, but whose environmental effects were allowed to be relatively simple smooth functions (≤ 2 degrees of freedom each). Model 13 was a partial least squares regression (PLSR; Geladi and Kowalski, 1986; Mevik and Wehrens, 2007) of log-transformed SI and jack data and untransformed environmental variables. Multivariate statistical techniques such as PLSR or principal component regression allow for the reformulation of explanatory variables as latent variables that are composites of the original explanatory variables, and these models are potentially more efficient for identifying important predictors from a complex set of numerous, often correlated environmental variables (Wells et al., 2008; Burke et al., 2013).

We considered different numbers of variables or latent variables in Models 5 and 11–13. For Models 5, 11, and 12 the number of jacks the previous year was always included and the remaining variables were selected based on the model with the lowest second-order Akaike Information Criterion (AIC_c; Burnham and Anderson, 2002) among all possible models with a given number of variables. We explored 1–6 variables for Models 5 and 11 (including jacks the previous year) and 1–5 variables for Model 12. All variables were included when fitting Model 13, but only 1–6 of the resulting latent variables were considered when predicting the SI.

Forecast performance

The forecast performance of the models was evaluated by comparing the observed SI time series to out-of-sample model predictions of the SI (Chatfield, 1996). Two cross-validation frameworks were used to derive out-of-sample predictions: leave-one-out and one-year-ahead. Under the leave-one-out framework, the data for a given year were withheld when fitting the model, then the fitted model was used to predict the SI for that year. Under the one-year-ahead framework, the model was fit to all of the data available prior to a given year, then the fitted model was used to predict the SI for that year. One-year-ahead predictions were made for 1995–2012. The one-year-ahead framework captured how the models would have been used in practice. The performance of Model 8 was only evaluated with the one-year-ahead framework.

For models that had variables selected on the basis of the data (Models 5, 11, and 12), variable selection was conducted using only the data that the model was fitted to, thereby excluding the out-of-sample data and accounting for model selection uncertainty (Chatfield, 1996; Francis, 2006). For comparison we also examined the forecast performance of these models when variables were selected on the basis of the entire dataset (i.e., the same variables were in the model for every prediction). We also examined the forecast performance of Models 5, 11, and 12 when predictions were averaged across all candidate sets of a given number of variables using Akaike weights derived from AIC_c (Burnham and Anderson, 2002).

Out-of-sample predictions from models of log-transformed data (Models 6–13) were adjusted before they were back-transformed to the arithmetic scale so that the predictions represented the expected SI (Beauchamp and Olson, 1973; Sprugel, 1983; Haeseker et al., 2005):

$$\widehat{\mathrm{SI}}_t = e^{\widehat{\log \mathrm{SI}}_t + 0.5\hat{\sigma}^2} \tag{1}$$

where

$$\hat{\sigma}^2 = \frac{\sum \hat{\epsilon}_t^2}{\mathrm{DF}} \quad , \tag{2}$$

$\widehat{\mathrm{SI}}_t$ is the out-of-sample predicted SI for year t, $\widehat{\log \mathrm{SI}}_t$ is the out-of-sample predicted SI on the log scale for year t, $\hat{\sigma}^2$ is the estimated variance of residual errors on the log scale based on the residuals of the model fitted to the sample data ($\hat{\epsilon}_t$), and DF is the number of residual degrees of freedom for the fitted model. For Model 8, the maximum likelihood estimate of the variance of innovation errors ($\hat{v}_t$) was used for $\hat{\sigma}^2$ in Eq. 1.

Six summary performance metrics were considered (Table 3) that reflected different aspects of forecast error (Zhou, 2003; Haeseker et al., 2005; Francis, 2006). Mean error (ME) and mean percent error (MPE) reflected directional bias in raw and relative forecast errors, respectively, with negative values indicating a tendency to overforecast and positive values a tendency to underforecast. Mean absolute error (MAE) and mean absolute percent error (MAPE) reflected overall forecast accuracy (sensu Walther and Moore, 2005) accounting for systematic bias and year-to-year variation. Root mean square error (RMSE) was a second measure of the absolute magnitude of raw errors, but was more sensitive to large errors than was MAE. Percent variance explained (PVE) was an expression of RMSE relative to the naive intercept-only model and reflected forecast skill. All performance metrics were calculated on the arithmetic scale for all models, and ME, MAE, RMSE, and PVE were also calculated on the log scale for models of log-transformed data (Models 6–13).

We focused on two of the summary metrics, ME and RMSE, on the arithmetic scale when comparing the performance of the alternative forecast models. We judged the forecasts from a model better if they were less biased (ME closer to zero) and more accurate (RMSE closer to zero). In addition to these summary performance metrics we assessed how well the forecasts from each model tracked trends in the SI, for example the recent increase since 2009.

Management performance

The management performance of the models was evaluated by comparing the allowable SRFC fishery exploitation rates specified by the observed and predicted SI time series. Each year the forecast SI is used to specify the allowable SRFC fishery exploitation rate via a harvest control rule (Fig. 3). The fishery exploitation rate (F) then corresponds to an expected spawning escapement of SI $\times$ (1 $-$ F). We considered two types of management errors related to conservation and fishing opportunity, each associated with threshold errors in the allowable exploitation rate (Fig. 3). A conservation error was deemed to occur when the SI was overforecast, and the allowable exploitation rate was high enough to result in under-escapement at a level $\leq$75% of the expected escapement given perfect knowledge of the SI. A fishing opportunity error was deemed to occur when the SI was underforecast and the allowable exploitation rate was $<$50% of the rate corresponding to the observed SI. While a fishing opportunity error could theoretically occur at any SI value, a conservation error as we defined it could not occur if the true SI was greater than about 300,000 because of the control rule's maximum allowable exploitation rate. The thresholds used to define management errors were chosen to be reasonable representations of the magnitude of errors that would likely concern fishery managers. The SRFC minimum stock size threshold, the spawner escapement level used to determine overfished status (PFMC, 2012a), is defined as 75% of the maximum sustainable yield spawner escapement. For this reason, we used the 75% escapement threshold to define a conservation error. For a fishery error, the 50% threshold was deemed to be a level that would have a clear impact on the amount of fishing opportunity in California and Oregon, and was chosen based on expert judgement.

Results

Forecast performance

The model with the best forecast performance differed between cross-validation frameworks and summary performance metrics (Fig. 4, Tables S1.1 and S1.2). The current management model (Model 2) had the ME closest to zero (least forecast bias) under leave-one-out cross validation excluding the intercept-only model (Model 1) and Model 5 with 6 variables. Under one-year-ahead cross validation several other models exhibited less forecast bias than did Model 2, with Models 8–10 being the least biased. Overall there was a tendency to overpredict the SI.

Excluding the intercept-only models, RMSE was generally greatest under the current management model (Model 2), but for Models 5 and 11–13 this result depended on the number of variables (Fig. 4, Tables S1.1 and S1.2). Models 4 and 10 had the two lowest RMSE under leave-one-out cross validation, and Model 8 had the lowest RMSE under one-year-ahead cross validation. There was substantial unexplained prediction error across all models with RMSE $\geq$250,000 and PVE $\leq$65%.

The models differed in their ability to track trends in the SI during specific time periods (Fig. 5). Under leave-one-out cross validation, the current management model (Model 2) overpredicted the SI at the start of the time series (1983) and predicted a subsequent decrease in the SI through 1987, opposite to the trend in the observed SI during this period. These predictions were driven by a decrease in the number of jacks from 1982–1986. During the most recent four years (2009–2012) Model 2 predicted a greater increase in the SI than was observed, culminating in a very high overprediction for 2012 under both cross-validation frameworks. These predictions were driven by an increase in the number of jacks from 2008–2011. Models that accounted for the number of jacks two years previous (Models 3 and 9) made qualitatively similar predictions early in the time series, but overpredicted somewhat less during the most recent years. Models 4 and 10 overpredicted less at the start of the time series, and much less at the end, because these models allowed for a smooth change over time in the ratio of the SI to the number of jacks the previous year, which was estimated to first increase and then decrease (Fig. 6). Model 8, with its estimated positive autocorrelation in residual errors, also performed well from 2009–2012. Models with environmental variables (Models 5 and 11–13) had more variable SI predictions at the start of the time series, and some of these models with certain numbers of variables (e.g., Model 13 with 3 latent variables; Fig. 5) performed better from 2009–2012 than Model 2. Some years had a large influence on certain summary performance metrics. For example, RMSE was the most sensitive to large errors, so it was more indicative of performance in years when the SI was greatly overpredicted (e.g., 2005 and 2012).

While models with environmental variables (Models 5 and 11–13) were generally able to achieve a lower RMSE than was the current management model, in many cases simpler models had similar or better performance (Fig. 4, Tables S1.1 and S1.2). There were no consistent trends in performance with increasing numbers of variables in these models. Models 5, 11, and 12 were sometimes able to achieve similar performance to the partial least squares regression model (Model 13). Model averaging across candidate models with a given number of variables did not result in consistently better or worse performance. Forecast performance tended to be better, in some cases much better, when variables were selected on the basis of the entire dataset rather than on the basis of only the training data within the cross-validation framework. Thus, accounting for model selection uncertainty decreased apparent forecast performance. There were similarities and differences in the important variables among Models 5, 11, and 12 (Fig. 7; AIC$_c$-based ‘relative variable importance’ as defined by Burnham and Anderson, 2002). The relative importance of variables varied when individual years were excluded from the analysis under leave-one-out cross validation, but there was some consistency in the important variables across years. For example, in versions of these models with 4 variables, the number of jacks two years previous and sea level height during the previous fall generally had high relative importance. The important variables frequently changed over time under one-year-ahead cross validation, with some exceptions. For example, sea level height the previous fall was consistently important in 4-variable models.

Management performance

Management errors occurred from 2007 onward, with very few exceptions (Fig. 8). Prior to 2007 the SI was always greater than the threshold at which a conservation error could occur (about 300,000; Fig. 3). Also, prior to 2007 the SI was usually much higher than the threshold at which the harvest control rule specifies a reduction in the allowable exploitation rate, so only a large underprediction could have caused a fishing opportunity management error. Between 2007 and 2011 the SI was below these thresholds creating more opportunity for management errors. Conservation errors occurred with multiple models from 2007–2011. The current management model resulted in a conservation error in 2010 under leave-one-out cross validation and in 2011 under both cross-validation frameworks. In fact, a conservation error occurred in 2011 with all models under both cross-validation frameworks. The frequently large overprediction of the SI in 2012 did not result in conservation errors because the observed SI was above the threshold corresponding to the maximum allowable exploitation rate. Fishing opportunity errors occurred with multiple models in 2007 and 2010. Because the number of management errors with each model was low (Table S1.3), it was difficult to draw conclusions about the relative management performance of the models.

Discussion

A suite of salmon abundance forecast models, with a wide range of complexity, were subjected to a set of rigorous tests and evaluations in an attempt to provide realistic expectations for future performance. Model selection, when relevant, was conducted with training data only. Performance was evaluated by (1) examining metrics that summarized forecast bias and accuracy calculated over all years under two cross-validation frameworks (Fig. 4), (2) visually inspecting the ability of the models to track trends in the SI at different points in time (Fig. 5), and (3) quantifying the interaction between forecast performance and a harvest control rule used for annual fishery management (Fig. 8). Our results suggest that the performance of individual models varied across forecast evaluation methods, complicating the selection of a single best model. While these results are derived from the specific case study of the SI, the findings are similar to those of other salmon forecast analyses (e.g., Haeseker et al., 2005) and are germane to other forecast scenarios.

Inference about future forecast performance based on summary metrics alone should be cautious. A comparison of Model 13 (with 3 latent variables) and Model 2 provides an example where a single year of data greatly influenced perceived relative forecast accuracy. Model 13 exhibited a lower RMSE under both cross-validation frameworks when compared to Model 2. However, much of this performance differential was attributable to particularly high overforecasts by Model 2 in a few individual years. For example under one-year-ahead cross validation, Model 13 had a RMSE of 310,797, while Model 2 had a RMSE of 480,748. Omitting 2012 from the analysis resulted in a RMSE of 316,587 for Model 13 and 349,868 for Model 2. RMSE under leave-one-out cross validation also differed little between these models when 2012 was omitted. As another example, RMSE under one-year-ahead cross validation was lower for Model 10 than for Model 2, but when 2012 was excluded the reverse was true. These results illustrate the sensitivity of the summary performance metrics, and RMSE in particular, to a single large deviation between prediction and observation.

Models in which environmental variables modified the relationship between the SI and the number of jacks in previous years were generally able to achieve better performance than the current forecast model. However, the environmental variables that appeared important sometimes changed depending on the model structure and the data that were used to fit the model. As one-year-ahead cross validation indicated, the variables selected in Models 5 and 11–12 would have changed over time had these models been used for forecasting during the study period. A supplementary analysis (unpublished) where these models were fitted over time using a moving data window or down-weighting more distant data resulted in even more frequent changes in the variables that appeared important. The results presented here do not provide compelling evidence for stable, specific functional relationships between the environmental variables that we examined and the jack-to-SI relationship. Sampling error and collinearity among environmental variables almost certainly contributed to some of the apparent changes in variable importance, the former of which would be expected to decrease with a longer time series. It is also likely that the functional relationships were more complex than our models allowed for (e.g., continuous or discrete temporal changes in the strength of the relationships).

Numerous studies have attempted to model and forecast Pacific salmon population abundance, productivity, and vital rates using environmental variables (e.g., Kope and Botsford, 1990; Adkison and Peterman, 2000; Logerwell et al., 2003; Scheuerell and Williams, 2005; Haeseker et al., 2008; Watters and Bessey, 2008; Fujiwara and Mohr, 2009; Wang et al., 2009; Rogers and Schindler, 2011; Wells et al., 2012; Rupp et al., 2012; Burke et al., 2013). Conditions in the freshwater and marine environments certainly affect salmon growth and survival (Quinn, 2005; Wells et al., 2007; Woodson et al., in press), thereby influencing productivity (Beamish and Mahnken, 2001) and transition rates between life stages (Morita and Fukuwaka, 2006; Snover et al., 2006; Satterthwaite et al., 2010). However, identifying and quantifying functional relationships between specific environmental variables and fish population dynamics are very difficult tasks. Any number of environmental variables can be measured and will exhibit variation over a range of spatial and temporal scales, possibly covarying with each other. Statistically significant correlations between population dynamics and some of these variables are an inevitability, regardless of true functional relationships (Walters and Collie, 1988; Megrey et al., 2005). Furthermore, functional relationships between environmental variables and population dynamics are likely to be highly complex given the chains of mechanisms and interactions involved (Deyle et al., 2013). Forecasting future dynamics on the basis of the environment entails further uncertainty about the stability of functional relationships over time. Because of these difficulties, many models relating fish population dynamics to environmental variables have not withstood the test of time (Myers, 1998; Keyl and Wolff, 2008). For example, Rupp et al. (2012) described how select environmental variables used in forecasts for Oregon Coastal Natural coho salmon (*O. kisutch*) had historically accounted for a large amount of variability in abundance, but ultimately became unreliable predictors.

Models 4, 8, and 10 allowed for changes over time in the ratio of the SI to the number of jacks the previous year, but modelled those changes relatively more phenomenologically, rather than functionally relating the changes to jacks two years previous or to environmental variables. The net effect of factors influencing the ratio was implicitly modelled as a gradual change in the expected ratio over time (Models 4 and 10) or as temporally autocorrelated deviations of the SI from the expected SI (Model 8). One would expect age structure and environmental conditions, which are correlated at time scales greater than one year (Stommel, 1963; Haury et al., 1978; Francis and Hare, 1994), to induce serially correlated changes in the jack-to-SI relationship. The forecast performance of these models was often among the best, sometimes better than that of models that incorporated the effects of age structure and environmental conditions explicitly. These models performed particularly well near the end of the study period. They were reasonably flexible yet relatively parsimonious, and were not subject to the difficulties associated with the use of environmental variables described above. Several studies have found that forecast models incorporating temporal and autoregressive changes in abundance, productivity, and recruitment perform relatively well for Pacific salmon populations (Noakes et al., 1990; Peterman et al., 2000; Haeseker et al., 2005, 2008). An advantage of Model 8 over Models 4 and 10 is that one does not have to make a decision about the degree of flexibility to allow for in the estimated temporal change in the jack-to-SI ratio.

For the particular case of the SI, there were substantial forecast errors across all models considered, suggesting that there may be an upper limit to the expected performance of any model. Attributes of the SI certainly contribute to these errors. The SI is a multi-cohort index of abundance, and the use of jacks, both in the previous year and two years prior, is likely insufficient to fully account for variation in cohort strength hidden within the index. There are unknown but likely high levels of measurement error in estimates of the harvest and spawning escapement components that make up the SI, as well as jack spawning escapement estimates. As monitoring programs have changed over time, it is likely that levels of measurement error have changed as well. We note, however, that substantial improvements in tagging and escapement monitoring programs (Bergman et al., 2012) should result in lower measurement error and enable age-specific abundance forecasts in the future. Nevertheless, many of the data quality and quantity issues that currently exist for SRFC apply to other salmon forecast scenarios and the idea that large levels of error in salmon forecasts are unavoidable has been stated before (e.g., Adkison and Peterman, 2000; Mantua and Francis, 2004; Haeseker et al., 2005). If forecasts are inherently inaccurate, Mantua and Francis (2004) suggest that management should de-emphasize preseason forecasts and rely more heavily on inseason monitoring. While this approach may be feasible for terminal salmon fisheries where spawning escapement can be monitored while the fishery is being conducted, it is more difficult to employ such an approach in mixed-stock ocean fisheries where the bulk of the fishery occurs prior to freshwater return, as is the case with SRFC.

Ultimately, the impacts of forecast error on the risk to the harvested population and on fishing opportunity are the bottom line for fisheries management. Simulations by Walters (1989) suggested that the value of improved pre-season forecasts might be limited unless the forecasts were highly accurate. We found little difference in the frequency of management errors (as we defined them) among the models that we explored. The harvest control rule for SRFC includes some precautionary elements that can buffer abundance forecast error. For example, allowable exploitation rates are never specified to be greater than 90 percent of F_{MSY} (PFMC, 2012a). Coupled simulations of the fish population and the management system would be necessary to assess the long-term value of the fishery and the risk to SRFC under different forecast models (e.g., Peterman et al., 2000; Kaje and Huppert, 2007; Rupp et al., 2012).

Our analysis suggests that there is scope to improve the performance of salmon forecast models that rely solely on estimated constant relationships between the abundances of different components of the population. In the case of predicting the SI from the number of jacks the previous year, we found that incorporating local environmental effects, temporal trends, and autocorrelation in the jack-to-SI relationship had the potential to increase forecast accuracy and the ability to track directional changes in abundance. Models that directly incorporated measures of the environment exhibited improved forecast performance in some cases. However, uncertainty about how the strength of particular environmental effects might change in the future and the relative complexity of these models pose challenges for future forecasts. Models that accounted for changes in the jack-to-SI relationship through time in a more phenomenological manner had among the best performance. These models were relatively parsimonious, were able to adequately track changes in population trajectories, and in the case of the model incorporating autocorrelated errors, imposed relatively little structure on changes in the jack-to-SI relationship over time. For these reasons, we believe the model incorporating autocorrelated errors (Model 8) should be given strong consideration for future forecasting of the SI.

Acknowledgements

Thank you to I. Schroeder for providing some of the environmental data. Thank you to R. Kope and N. Mantua for helpful comments on an earlier version of the manuscript. S. Allen-Moran, A. MacCall, S. Munch, and C. Speir also provided helpful suggestions and insights.

References

Adkison, M. D., and Peterman, R. M. 2000. Predictability of Bristol Bay, Alaska, sockeye salmon returns one to four years in the future. N. Am. J. Fish. Manage. **20**: 69–80.

Bakun, A., and Nelson, C. S. 1991. The seasonal cycle of wind-stress curl in subtropical eastern boundary current regions. J. Phys. Oceanogr. **21**: 1815–1834.

Beamish, R. J., and Mahnken, C. 2001. A critical size and period hypothesis to explain natural regulation of salmon abundance and the linkage to climate and climate change. Prog. Oceanogr. **49**: 423–437.

Beauchamp, J. J., and Olson, J. S. 1973. Corrections for bias in regression estimates after logarithmic transformation. Ecology **54**: 1403–1407.

Bergman, J. M., Nielson, R. M., and Low, A. 2012. Central Valley Chinook salmon in-river escapement monitoring plan. Fisheries Branch Administrative Report 2012-1, California Department of Fish and Game, Sacramento, CA, USA. Available from https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=42213 (accessed June 2013).

Bocking, R. C., and Peterman, R. M. 1988. Preseason forecasts of sockeye salmon (*Oncorhynchus nerka*): comparison of methods and economic considerations. Can. J. Fish. Aquat. Sci. **45**: 1346–1354.

Bograd, S. J., Schroeder, I., Sarkar, N., Qiu, X., Sydeman, W. J., and Schwing, F. B. 2009. Phenology of coastal upwelling in the California Current. Geophys. Res. Lett. **36**: L01602.

Burke, B. J., Peterson, W. T., Beckman, B. R., Morgan, C., and Daly, E. A. 2013. Multivariate models of adult Pacific salmon returns. PLoS One **8**: e54134.

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

Chatfield, C. 1996. Model uncertainty and forecast accuracy. J. Forecasting **15**: 495–508.

Deyle, E. R., Fogarty, M., Hsieh, C.-h., Kaufman, L., MacCall, A. D., Munch, S. B., Perretti, C. T., Ye, H., and Sugihara, G. 2013. Predicting climate effects on Pacific sardine. Proc. Nat. Acad. Sci. USA **110**: 6430–6435.

Francis, R. C., and Hare, S. R. 1994. Decadal-scale regime shifts in the large marine ecosystems of the North-east Pacific: a case for historical science. Fish. Oceanogr. **3**: 279–291.

Francis, R. I. C. C. 2006. Measuring the strength of environment-recruitment relationships: the importance of including predictor screening within cross-validations. ICES J. Mar. Sci. **63**: 594–599.

Fujiwara, M., and Mohr, M. S. 2009. Identifying environmental signals from population abundance data using multivariate time-series analysis. Oikos **118**: 1712–1720.

Geladi, P., and Kowalski, B. R. 1986. Partial least-squares regression: a tutorial. Anal. Chim. Acta **185**: 1–17.

Haeseker, S. L., Peterman, R. M., and Su, Z. 2005. Retrospective evaluation of preseason forecasting models for pink salmon. N. Am. J. Fish. Manage. **25**: 897–918.

Haeseker, S. L., Peterman, R. M., and Su, Z. 2008. Retrospective evaluation of preseason forecasting models for sockeye and chum salmon. N. Am. J. Fish. Manage. **28**: 12–29.

Haury, L. R., McGowan, J. A., and Wiebe, P. H. 1978. Patterns and processes in the time-space scales of plankton distributions. *In* Spatial pattern in plankton communities. *Edited by* J. H. Steele, NATO conference series IV: Marine sciences. Plenum Press, New York, NY, USA, pp. 277–327.

Holt, C. A., and Peterman, R. M. 2006. Missing the target: uncertainties in achieving management goals in fisheries on Fraser River, British Columbia, sockeye salmon (*Onchorhynchus nerka*). Can. J. Fish. Aquat. Sci. **63**: 2722–2733.

Kaje, J. H., and Huppert, D. D. 2007. The value of short-run climate forecasts in managing the coastal coho salmon (*Oncorhynchus kisutch*) fishery in Washington State. Nat. Resour. Model. **20**: 321–349.

Keyl, F., and Wolff, M. 2008. Environmental variability and fisheries: what can models do? Rev. Fish Biol. Fish. **18**: 273–299.

Kope, R. G., and Botsford, L. W. 1990. Determination of factors affecting recruitment of Chinook salmon *Oncorhynchus tshawytscha* in Central California. Fish. Bull. **88**: 257–269.

Lindley, S. T., Grimes, C. B., Mohr, M. S., Peterson, W., Stein, J., Anderson, J. T., Botsford, L. W., Bottom, D. L., Busack, C. A., Collier, T. K., Ferguson, J., Garza, J. C., Grover, A. M., Hankin, D. G., Kope, R. G., Lawson, P. W., Low, A., MacFarlane, R. B., Moore, K., Palmer-Zwahlen, M., Schwing, F. B., Smith, J., Tracy, C., Webb, R., Wells, B. K., and Williams, T. H. 2009, July. What caused the Sacramento River fall Chinook stock collapse? U.S.

Department of Commerce, NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-447. Available from http://swfsc.noaa.gov/publications/TM/SWFSC/NOAA-TM-NMFS-SWFSC-447.pdf (accessed September 2012).

Logerwell, E. A., Mantua, N., Lawson, P. W., Francis, R. C., and Agostini, V. N. 2003. Tracking environmental processes in the coastal zone for understanding and predicting Oregon coho (*Oncorhynchus kisutch*) marine survival. Fish. Oceanogr. **12**: 554–568.

Mantua, N., and Francis, R. C. 2004. Natural climate insurance for Pacific Northwest salmon and salmon fisheries: Finding our way through the entangled bank. *In* Sustainable management of North American fisheries. *Edited by* E. E. Knudsen, D. D. MacDonald, and Y. K. Muirhead, Volume 43 of *American Fisheries Society Symposium*. American Fisheries Society, Bethesda, MD, USA, pp. 127–140.

Megrey, B. A., Lee, Y.-W., and Macklin, S. A. 2005. Comparative analysis of statistical tools to identify recruitment-environment relationships and forecast recruitment strength. ICES J. Mar. Sci. **62**: 1256–1269.

Mevik, B.-H., and Wehrens, R. 2007. The pls package: principal component and partial least squares regression in R. J. Stat. Softw. **18**(2): 1–24.

Morita, K., and Fukuwaka, M.-a. 2006. Does size matter most? The effect of growth history on probabilistic reaction norm for salmon maturation. Evolution **60**: 1516–1521.

Myers, R. A. 1998. When do environment-recruitment correlations work? Rev. Fish Biol. Fish. **8**: 285–305.

Nelson, C. S. 1977. Wind stress and wind stress curl over the California Current. NMFS Report SSRF-714, NOAA, Monterey, CA, USA. Available from http://spo.nmfs.noaa.gov/SSRF/SSRF714.pdf (accessed September 2012).

Noakes, D. J., Welch, D. W., Henderson, M., and Mansfield, E. 1990. A comparison of preseason forecasting methods for returns of two British Columbia sockeye salmon stocks. N. Am. J. Fish. Manage. **10**: 46–57.

O'Farrell, M. R., Mohr, M. S., Palmer-Zwahlen, M. L., and Grover, A. M. 2013. The Sacramento Index (*SI*). U.S. Department of Commerce, NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-512. Available from http://swfsc.noaa.gov/publications/TM/SWFSC/NOAA-TM-NMFS-SWFSC-512.pdf (accessed September 2013).

Peterman, R. M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Can. J. Fish. Aquat. Sci. **38**: 1113–1119.

Peterman, R. M. 1982. Model of salmon age structure and its use in preseason forecasting and studies of marine survival. Can. J. Fish. Aquat. Sci. **39**: 1444–1452.

Peterman, R. M., Pyper, B. J., and Grout, J. A. 2000. Comparison of parameter estimation methods for detecting climate-induced changes in productivity of Pacific salmon (*Oncorhynchus* spp.). Can. J. Fish. Aquat. Sci. **57**: 181–191.

PFMC (Pacific Fishery Management Council) 2012a. Pacific Coast Salmon Fishery Management Plan for commercial and recreational salmon fisheries off the coasts of Washington, Oregon, and California as revised through Amendment 16. PFMC, Portland, OR, USA. Available from http://www.pcouncil.org/wp-content/uploads/FMP_through_A-16_Final.pdf (accessed September 2013).

PFMC (Pacific Fishery Management Council) 2012b. Preseason report I: Stock abundance analysis and environmental assessment part 1 for 2012 ocean salmon fishery regulations. Document prepared for the Council and its advisory entities, PFMC, Portland, OR, USA. Available from http://www.pcouncil.org/wp-content/uploads/Preseason_Report_I_2012.pdf (accessed June 2013).

PFMC (Pacific Fishery Management Council) 2013a. Preseason report I: Stock abundance analysis and environmental assessment part 1 for 2013 ocean salmon fishery regulations. Document prepared for the Council and its advisory entities, PFMC, Portland, OR, USA. Available from http://www.pcouncil.org/wp-content/uploads/Preseason_Report_I_2013_FINAL.pdf (accessed June 2013).

PFMC (Pacific Fishery Management Council) 2013b. Review of 2012 ocean salmon fisheries: Stock assessment and fishery evaluation document for the Pacific Coast Salmon Fishery Management Plan. Document prepared for the Council and its advisory entities, PFMC, Portland, OR, USA. Available from http://www.pcouncil.org/wp-content/uploads/salsafe_2012.pdf (accessed June 2013).

Quinn, T. P. 2005. The behavior and ecology of Pacific salmon and trout. American Fisheries Society, Bethesda, MD in association with University of Washington Press, Seattle, WA, USA.

R Core Team 2013. R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Rogers, L. A., and Schindler, D. E. 2011. Scale and the detection of climatic influences on the productivity of salmon populations. Glob. Change Biol. **17**: 2546–2558.

Roth, J. E., Mills, K. L., and Sydeman, W. J. 2007. Chinook salmon (*Oncorhynchus tshawytscha*) – seabird covariation off central California and possible forecasting applications. Can. J. Fish. Aquat. Sci. **64**: 1080–1090.

Rupp, D. E., Wainwright, T. C., and Lawson, P. W. 2012. Effect of forecast skill on management of the Oregon coast coho salmon (*Oncorhynchus kisutch*) fishery. Can. J. Fish. Aquat. Sci. **69**: 1016–1032.

Rupp, D. E., Wainwright, T. C., Lawson, P. W., and Peterson, W. T. 2012. Marine environment-based forecasting of coho salmon (*Oncorhynchus kisutch*) adult recruitment. Fish. Oceanogr. **21**: 1–19.

Satterthwaite, W. H., Beakes, M. P., Collins, E. M., Swank, D. R., Merz, J. E., Titus, R. G., Sogard, S. M., and Mangel, M. 2010. State-dependent life history models in a changing (and regulated) environment: steelhead in the California Central Valley. Evol. Appl. **3**: 221–243.

Scheuerell, M. D., and Williams, J. G. 2005. Forecasting climate-induced changes in the survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*). Fish. Oceanogr. **14**: 448–457.

Snover, M. L., Watters, G. M., and Mangel, M. 2006. Top-down and bottom-up control of life-history strategies in coho salmon (*Oncorhynchus kisutch*). Am. Nat. **167**: E140–E157.

Sprugel, D. G. 1983. Correcting for bias in log-transformed allometric equations. Ecology **64**: 209–210.

Stommel, H. 1963. Varieties of oceanographic experience. Science **139**: 572–576.

Walters, C. J. 1989. Value of short-term forecasts of recruitment variation for harvest management. Can. J. Fish. Aquat. Sci. **46**: 1969–1976.

Walters, C. J., and Collie, J. S. 1988. Is research on environmental factors useful to fisheries management? Can. J. Fish. Aquat. Sci. **45**: 1848–1854.

Walther, B., and Moore, J. L. 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography **28**: 815–829.

Wang, S., Morishima, G., Sharma, R., and Gilbertson, L. 2009. The use of generalized additive models for forecasting the abundance of Queets River coho salmon. N. Am. J. Fish. Manage. **29**: 423–433.

Watters, G. M., and Bessey, C. 2008. Variation in the probability that male coho salmon will mature early: inferences from hierarchical models. Trans. Am. Fish. Soc. **137**: 70–95.

Wells, B. K., Field, J. C., Thayer, J. A., Grimes, C. B., Bograd, S. J., Sydeman, W. J., Schwing, F. B., and Hewitt, R. 2008. Untangling the relationships among climate, prey and top predators in an ocean ecosystem. Mar. Ecol.-Prog. Ser. **364**: 15–29.

Wells, B. K., Grimes, C. B., and Waldvogel, J. B. 2007. Quantifying the effects of wind, upwelling, curl, sea surface temperature and sea level height on growth and maturation of a California Chinook salmon (*Oncorhynchus tshawytscha*) population. Fish. Oceanogr. **16**: 363–382.

Wells, B. K., Santora, J. A., Field, J. C., MacFarlane, R. B., Marinovic, B. B., and Sydeman, W. J. 2012. Population dynamics of Chinook salmon *Oncorhynchus tshawytscha* relative to prey availability in the central California coastal region. Mar. Ecol.-Prog. Ser. **457**: 125–137.

Wood, S. N. 2006. Generalized additive models: An introduction with R. Chapman and Hall/CRC, Boca Raton, FL, USA.

Woodson, L. E., Wells, B. K., Weber, P. K., MacFarlane, R. B., Whitman, G. E., and Johnson, R. C. In press. Size, growth, and origin-dependent mortality of juvenile Chinook salmon *Oncorhynchus tshawytscha* during early ocean residence. Mar. Ecol.-Prog. Ser. **000**: 000–000.

Worley, S. J., Woodruff, S. D., Reynolds, R. W., Lubker, S. J., and Lott, N. 2005. ICOADS release 2.1 data and products. Int. J. Climatol. **25**: 823–842.

Zhou, S. 2003. Application of artificial neural networks for forecasting salmon escapement. N. Am. J. Fish. Manage. **23**: 48–59.

Table 1. Environmental variables.

Variable	Description	Units	Resolution	Location	Source
SST	sea surface temperature	°C	monthly; 1° lon/lat	5 cells centered on 235.5°E, 40.5°N 235.5°E, 39.5°N 236.5°E, 38.5°N 236.5°E, 37.5°N 237.5°E, 37.5°N	ICOADS[a]
SCALAR	non-directional (scalar) surface wind speed	m/s	*see SST*	*see SST*	ICOADS[a]
NPSEUDO	northerly pseudo-wind stress; product of scalar and vector surface wind	m^2/s^2	*see SST*	*see SST*	ICOADS[a]
EPSEUDO	easterly pseudo-wind stress; product of scalar and vector surface wind	m^2/s^2	*see SST*	*see SST*	ICOADS[a]
CURL	surface wind-stress curl; influences upwelling and offshore transport[b]	dynes/cm^2/cm × 1E10	monthly, 6-hourly; 15 point locations few ° apart	125°W, 39°N	PFEL[c]
UPWELLING	upwelling index "based on estimates of	t/s/100 m of coast	*see CURL*	*see CURL*	PFEL[c]

Table 1. continued.

Variable	Description	Units	Resolution	Location	Source
	offshore Ekman transport driven by geostrophic wind stress"; "geostrophic winds are derived from . . . surface atmospheric pressure fields"[d]				
SLH	sea level height	mm	monthly, hourly, daily; multiple stations around the world	122°27.9'W, 37°48.4'	UHSLC[e]
STI_39N	spring transition index; date on which the cumulative coastal upwelling index (beginning 1 January) begins to increase from its minimum value[f]	day of year	*see CURL for spatial resolution*	*see CURL*	PFEL[c]

[a] International Comprehensive Ocean-Atmosphere Data Set (Worley et al., 2005), available from http://coastwatch.pfeg.noaa.gov/erddap/griddap/esrlIcoads1ge.html

[b] Nelson (1977), Bakun and Nelson (1991)

[c] Pacific Fisheries Environmental Laboratory, available from http://www.pfeg.noaa.gov/products/pfel/modeled/indices/transports/transports.html

[d] http://www.pfeg.noaa.gov/products/PFEL/modeled/indices/upwelling/NA/how_computed.html

[e] University of Hawaii Sea Level Center, available from http://ilikai.soest.hawaii.edu/uhslc/htmld/d0551W.html. Station 551, San Francisco, California.

[f] Bograd et al. (2009)

Table 2. Alternative models for forecasting the Sacramento Index (SI) as a function of the number of jacks the previous two years and the environment the previous year. Model 2 is currently used to forecast the SI. Model variables, parameters, and terms are defined as follows: J_t - jacks in year t, $\mathrm{E}_{i,t}$ - environmental variable i in year t, β_i - model intercept (β_0) and coefficients, $f_{i(n)}(\mathrm{X}_i)$ - smooth function of variable X_i with cubic spline basis and maximum n degrees of freedom, ϵ_t - SI residual for year t, ρ - first-order temporal autocorrelation in SI residuals, v_t - 'innovation' for year t, and σ^2 - error variance.

Model	Formula	Error structure	Model selection	Selected terms (X_i)
1	$\mathrm{SI}_t = \beta_0 + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
2	$\mathrm{SI}_t = \beta_1 \mathrm{J}_{t-1} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2 \mathrm{J}_{t-1}\right)$	none	
3	$\mathrm{SI}_t = \beta_1 \mathrm{J}_{t-1} + \beta_2 \mathrm{J}_{t-2} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2 \mathrm{J}_{t-1}\right)$	none	
4[a,b]	$\mathrm{SI}_t = f_{1(3)}(t)\, \mathrm{J}_{t-1} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2 \mathrm{J}_{t-1}\right)$	none	
5	$\mathrm{SI}_t = \beta_1 \mathrm{J}_{t-1} + \sum_i \beta_i \mathrm{X}_i + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2 \mathrm{J}_{t-1}\right)$	AIC_c	J_{t-2}, $\mathrm{J}_{t-1} \times \mathrm{E}_{j,t-1}$
6	$\log \mathrm{SI}_t = \beta_0 + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
7	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
8[c]	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \epsilon_t$	$\epsilon_t = \rho \epsilon_{t-1} + v_t,\ v_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
9	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \beta_2 \log \mathrm{J}_{t-2} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
10[a]	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + f_{1(2)}(t) + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	
11	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \sum_i \beta_i \mathrm{X}_i + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	AIC_c	$\log \mathrm{J}_{t-2}$, $\mathrm{E}_{j,t-1}$
12[a]	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \sum_i \mathrm{X}_i + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	AIC_c	$\beta_2 \log \mathrm{J}_{t-2}$, $f_{j(2)}(\mathrm{E}_{j,t-1})$
13[d]	$\log \mathrm{SI}_t = \beta_0 + \beta_1 \log \mathrm{J}_{t-1} + \beta_2 \log \mathrm{J}_{t-2} + \sum_i \beta_i \mathrm{E}_{i,t-1} + \epsilon_t$	$\epsilon_t \sim \mathrm{N}\left(0, \sigma^2\right)$	none	

[a] Generalized additive model fit with 'mgcv' package (Wood, 2006) for R (R Core Team, 2013)

[b] Varying coefficient model (Wood, 2006)

[c] First-order autoregressive error structure fit with 'arima' function in R (R Core Team, 2013)

[d] Partial least squares regression model fit with 'pls' package (Mevik and Wehrens, 2007) for R (R Core Team, 2013); data were centered and scaled

Table 3. Forecast performance metrics. Symbols are defined as follows: SI_t - observed Sacramento Index in year t, $\widehat{\mathrm{SI}}_t$ - predicted Sacramento Index in year t, and n - sample size. ME, MAE, RMSE, and PVE were also calculated on the log scale for models of log-transformed data (Models 6–13). For calculating PVE, RMSE_i was the RMSE of the appropriate intercept-only model (Model 1 for PVE on the arithmetic scale, and Model 6 for PVE on the log scale).

Metric	Acronym	Formula
mean error	ME	$\frac{\sum_t \left(\mathrm{SI}_t - \widehat{\mathrm{SI}}_t\right)}{n}$
mean absolute error	MAE	$\frac{\sum_t \left\lvert\mathrm{SI}_t - \widehat{\mathrm{SI}}_t\right\rvert}{n}$
mean percent error	MPE	$100 \times \frac{\sum_t \left[\left(\mathrm{SI}_t - \widehat{\mathrm{SI}}_t\right)/\mathrm{SI}_t\right]}{n}$
mean absolute percent error	MAPE	$100 \times \frac{\sum_t \left\lvert\left(\mathrm{SI}_t - \widehat{\mathrm{SI}}_t\right)/\mathrm{SI}_t\right\rvert}{n}$
root mean square error	RMSE	$\sqrt{\frac{\sum_t \left(\mathrm{SI}_t - \widehat{\mathrm{SI}}_t\right)^2}{n}}$
percent variance explained[a]	PVE	$100 \times \left(1 - \frac{\mathrm{RMSE}^2}{\mathrm{RMSE}_\mathrm{i}^2}\right)$

[a] Francis (2006); forecast skill relative to the naive intercept-only model (RMSE_i)

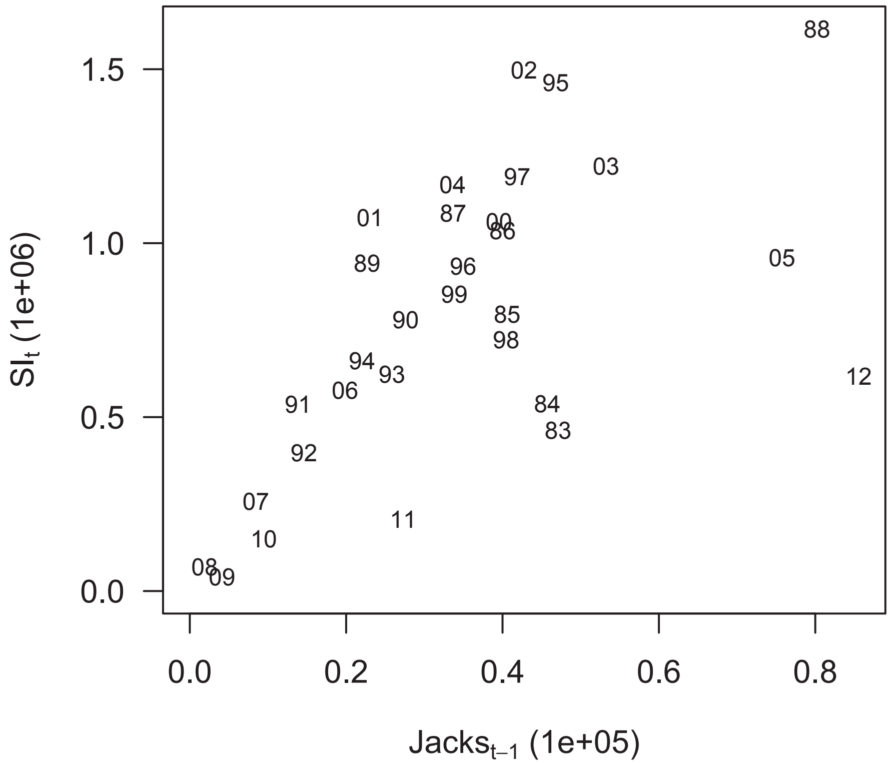

Figure 1. The Sacramento Index (SI) for year t (1983–2012) versus the number of jacks the previous year.

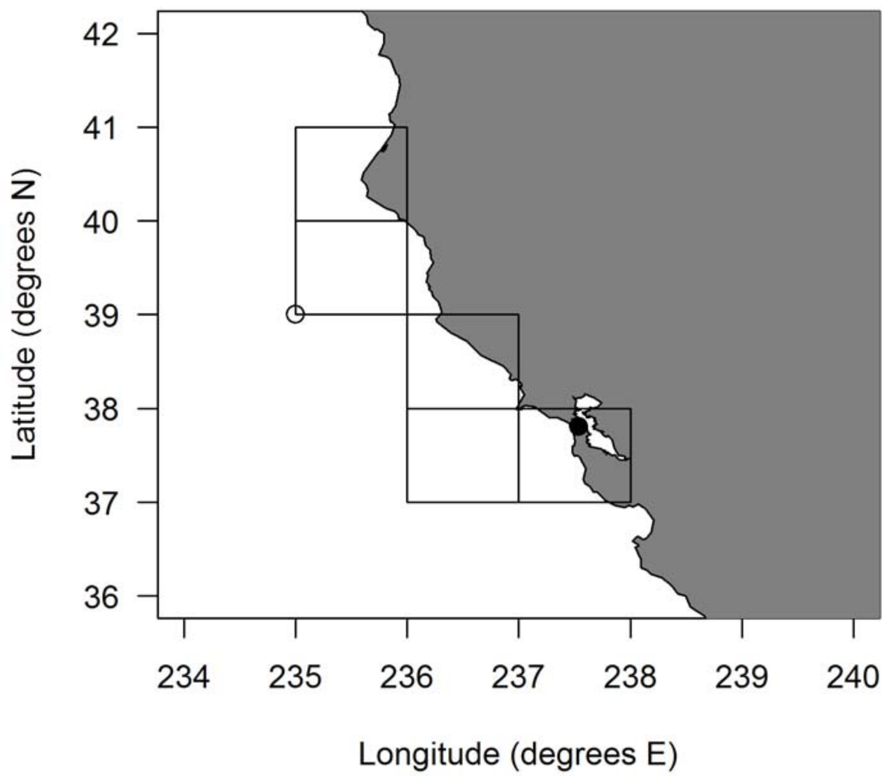

Figure 2. Geographic area from which environmental data were derived. UPWELLING and CURL were derived from the location marked by the open point, SLH was derived from the location marked by the solid point, and SST, SCALAR, NPSEUDO, and EPSEUDO were derived from the indicated 1-degree cells.

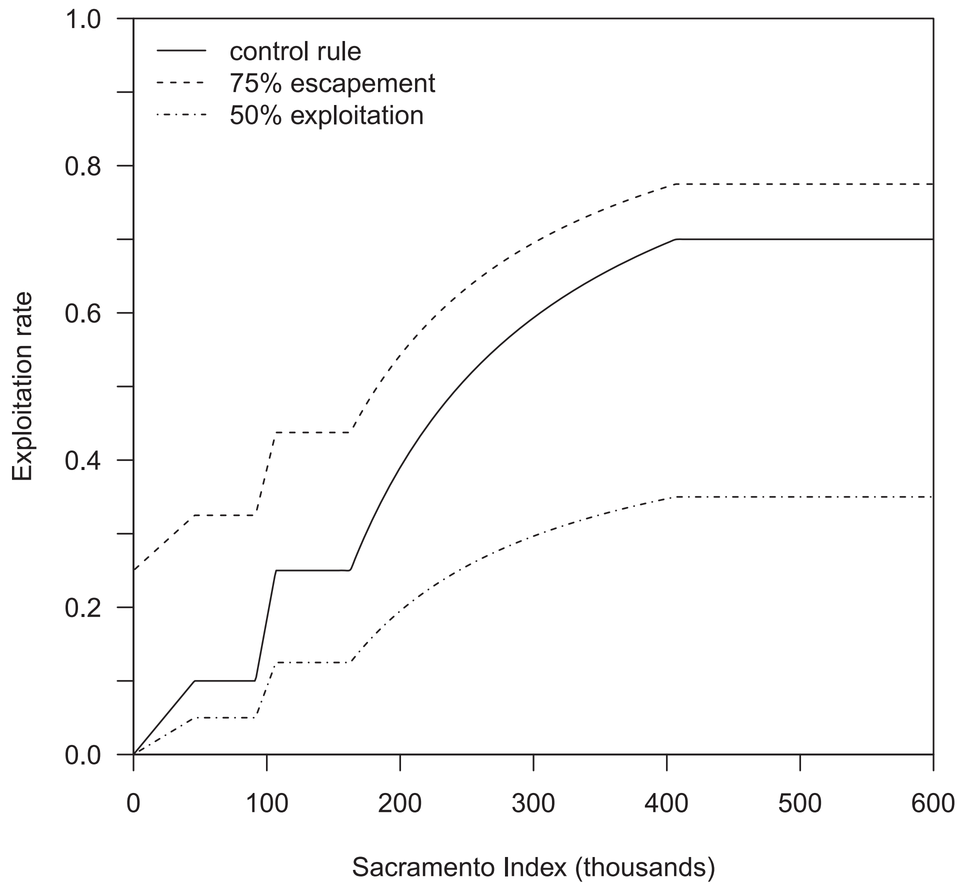

Figure 3. Harvest control rule for Sacramento River Fall Chinook salmon. The solid line indicates the allowable exploitation rate as a function of the Sacramento Index. The upper dashed line indicates the exploitation rate that would result in 75% of the spawning escapement specified by the control rule (threshold for conservation error). The lower dashed-dotted line indicates 50% of the exploitation rate specified by the control rule (threshold for fishing opportunity error).

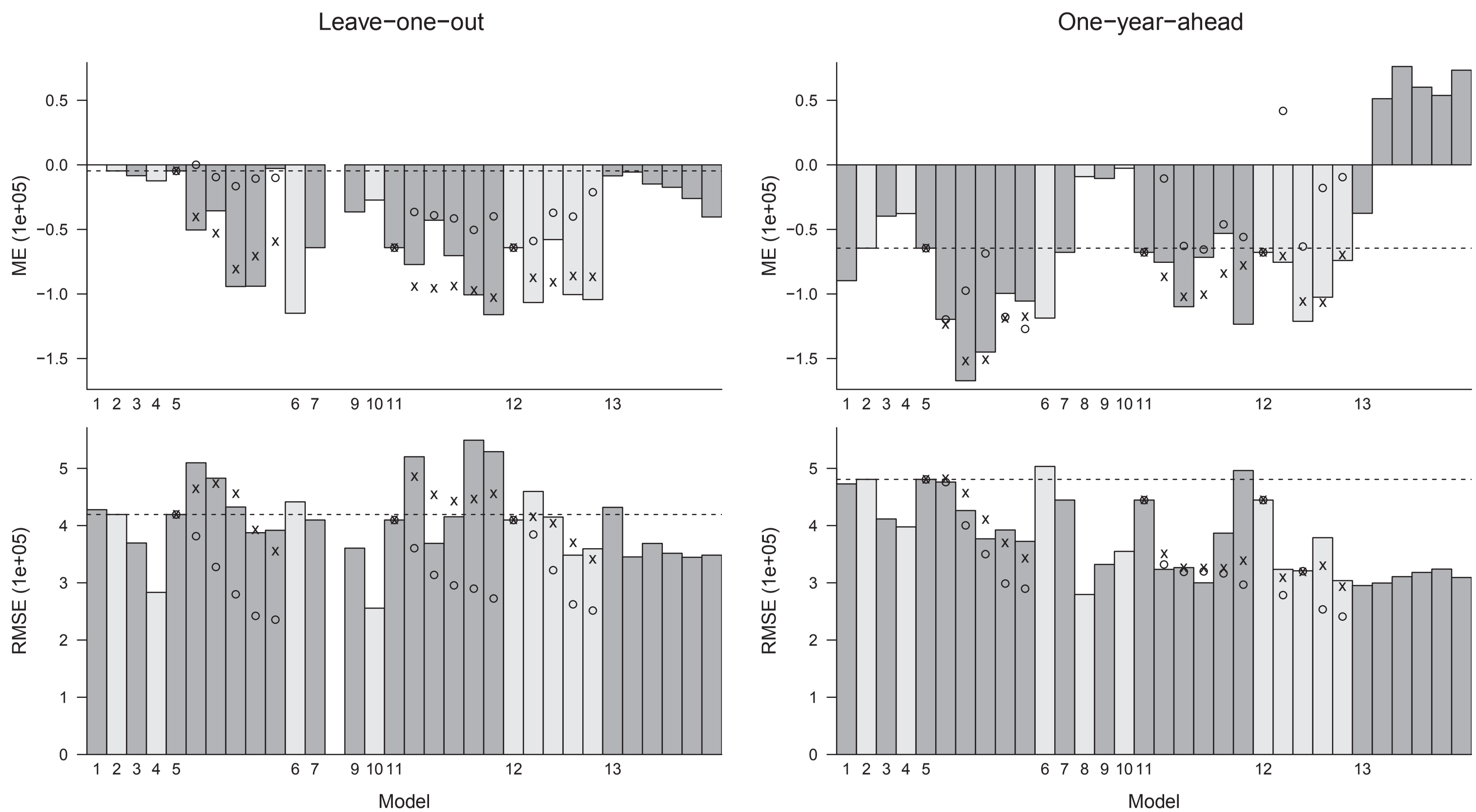

Figure 4. Model forecast performance in leave-one-out and one-year-ahead cross validation. Performance metrics are defined in Table 3. Crosses represent model averaging results and circles represent results when model selection was not included in cross validation (i.e., the model was selected on the basis of the entire dataset). The multiple bars for Models 5 and 11–13 represent the results for sequential numbers of (latent) variables in these models, beginning with one variable. The number of jacks the previous year was always included, so the one-variable versions of Models 5, 11, and 12 had only this variable. The dashed lines reference the performance of Model 2. Leave-one-out cross validation was not conducted for Model 8.

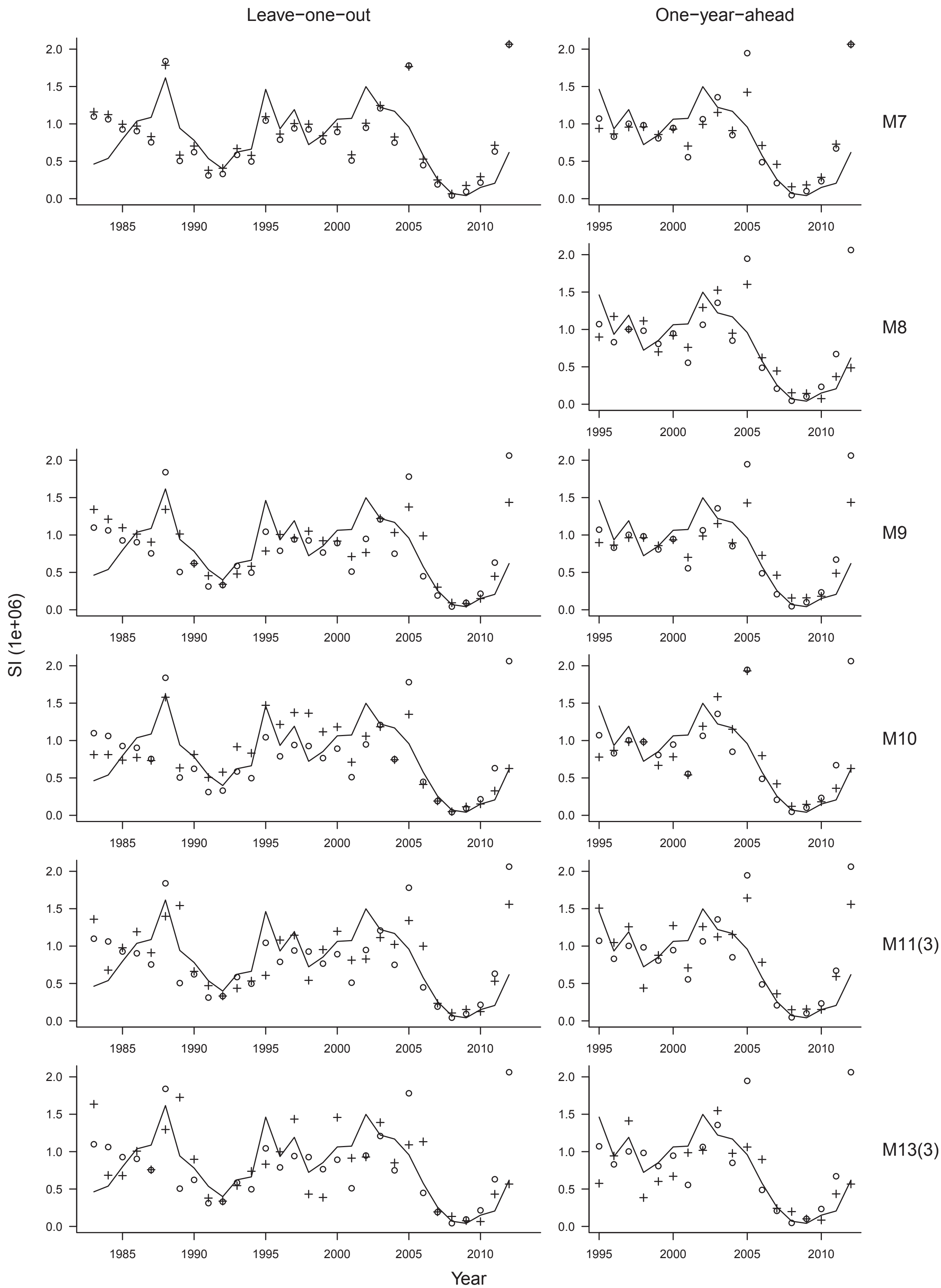

Figure 5. Observed and predicted Sacramento Index (SI) under leave-one-out and one-year-ahead cross validation for Models 2, 7–11, and 13. The latter two models had three variables each. Solid lines represent observations, circles represent predictions under Model 2 (current management model), and crosses represent predictions under the model indicated in the right margin.

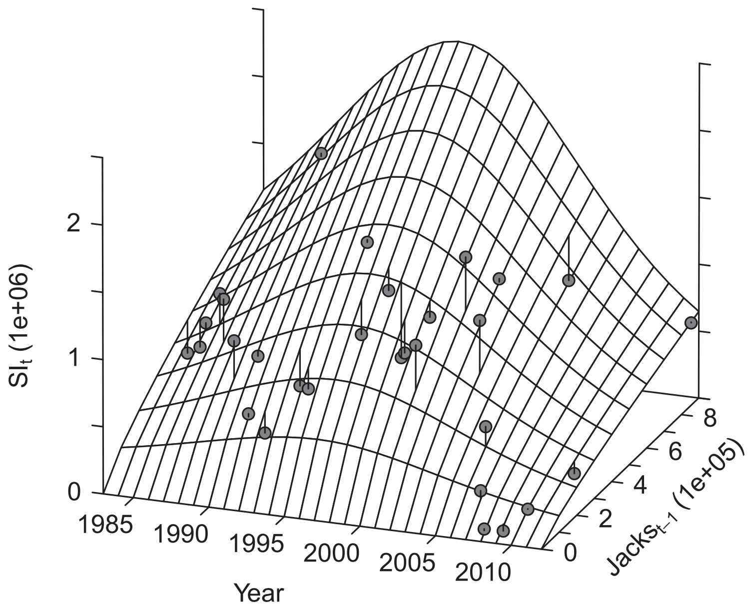

Figure 6. Predicted Sacramento Index (SI) as a function of year and the number of jacks the previous year (Model 10). Grey dots represent the data.

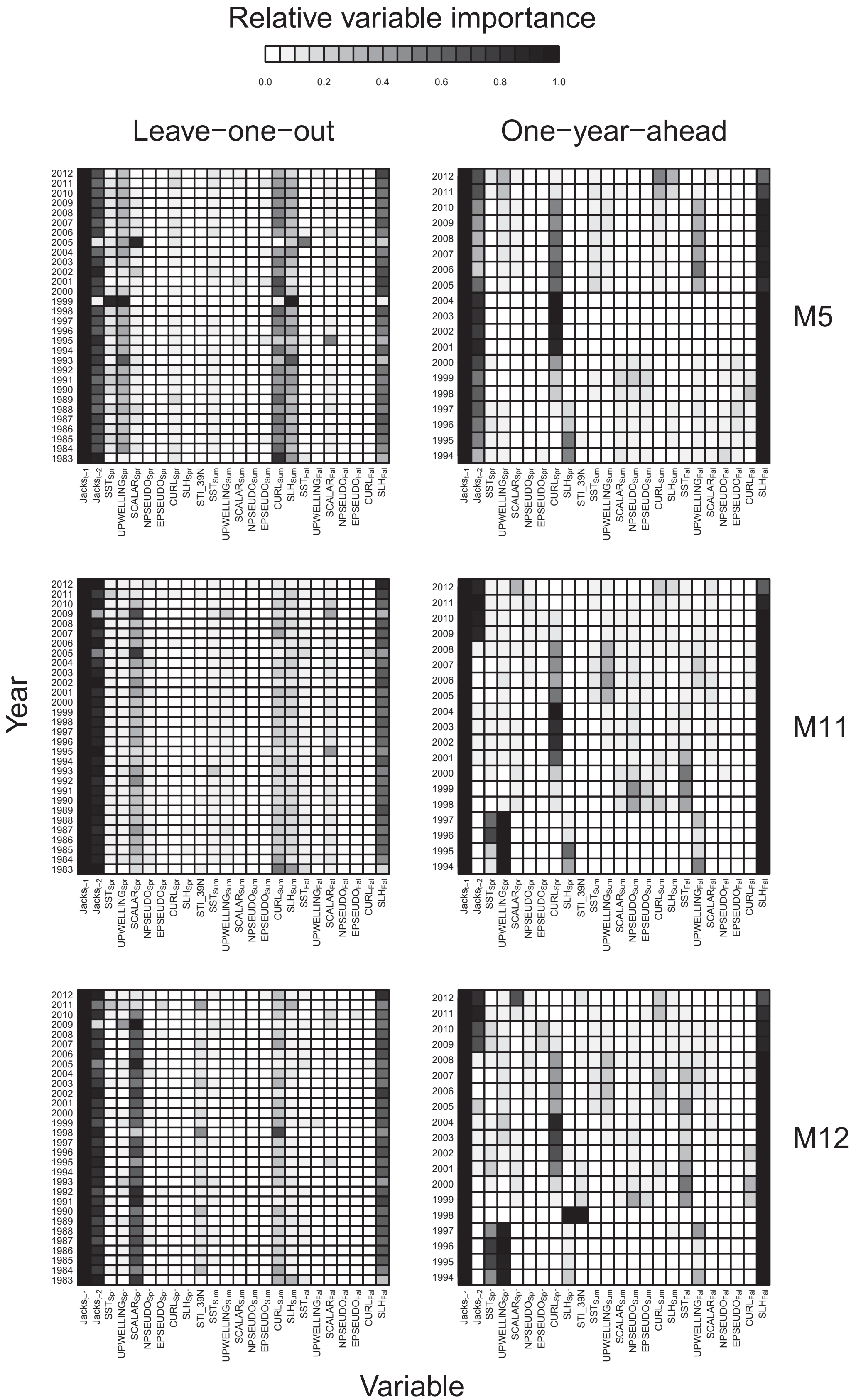

Figure 7. AIC$_c$-based ‘relative variable importance’ (Burnham and Anderson, 2002) in Models 5, 11, and 12 with 4 variables under leave-one-out and one-year-ahead cross validation. For leave-one-out cross validation, each year represents relative variable importance when data from that year were excluded. For one-year-ahead cross validation, each year represents relative variable importance based on the data available up to and including that year. The number of jacks the previous year (Jacks$_{t-1}$) was always included so its relative importance was always 1. The subscript ‘t - 1’ has been suppressed for environmental variables, which are defined in Table 1.

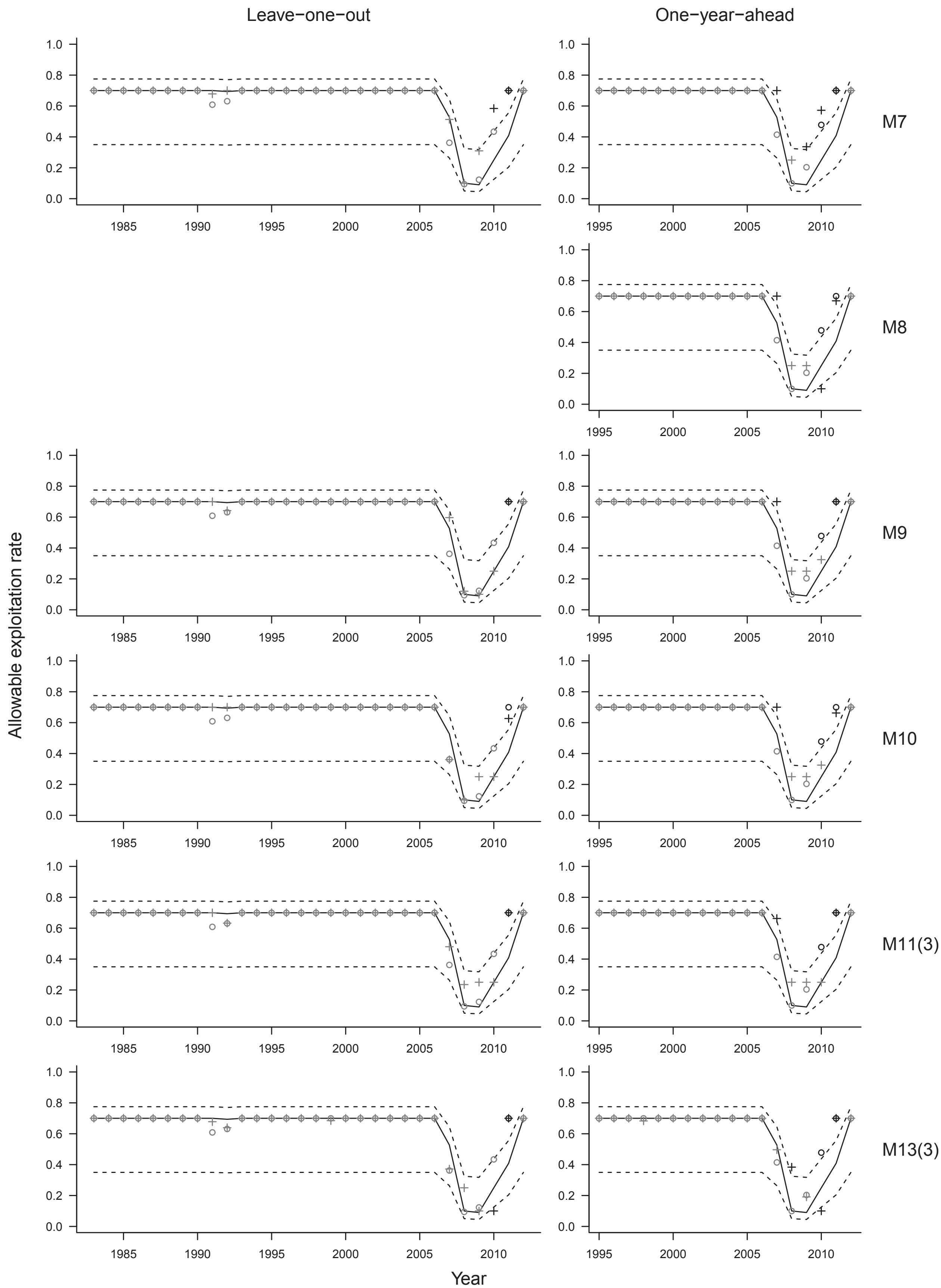

Figure 8. Allowable exploitation rate over time under leave-one-out and one-year-ahead cross-validation for Models 2, 7–11, and 13. The latter two models had three variables each. Solid lines represent the rate corresponding to the observed Sacramento Index (SI), upper and lower dashed lines represent the threshold rates corresponding to conservation and fishing opportunity errors, respectively, circles represent the rate corresponding to the predicted SI under Model 2 (current management model), and crosses represent the rate corresponding to the predicted SI under the model indicated in the right margin. Black symbols indicate management errors.

Supplementary material S1

Table S1.1. Leave-one-out cross-validation performance of alternative models for forecasting the Sacramento Index. Models are defined in Table 2, and performance metrics are defined in Table 3.

Model	#	Arithmetic scale						Log scale			
	variables	ME	MAE	MPE	MAPE	RMSE	PVE	ME	MAE	RMSE	PVE
1	0	0	353717	-128	156	427787	0				
2	1	-4668	297309	-16	50	419377	4				
3	2	-8389	265585	-18	42	369555	25				
4	2	-12400	211469	-4	31	283256	56				
5	1	-4668	297309	-16	50	419377	4				
	2	-50419	337914	-24	56	509726	-42				
	3	-35594	297372	-21	47	482794	-27				
	4	-94236	281188	-20	43	432431	-2				
	5	-94018	262022	-20	44	387572	18				
	6	-2790	303255	-11	49	391801	16				
5[a]	1	-4668	297309	-16	50	419377	4				
	2	-40480	292596	-20	48	464182	-18				
	3	-53137	286016	-22	46	473220	-22				
	4	-80935	296105	-22	47	455480	-13				
	5	-70851	255057	-18	42	392273	16				
	6	-59498	246702	-16	42	355056	31				
5[b]	1	-4668	297309	-16	50	419377	4				
	2	131	237141	-13	43	381608	20				
	3	-9472	216207	-16	37	327782	41				
	4	-16450	197063	-12	33	280042	57				
	5	-10610	172802	-12	34	242328	68				
	6	-10011	162592	-12	32	235719	70				
6	0	-114941	364435	-156	175	441469	-6	0	0.641	0.888	0
7	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65

Table S1.1. continued.

Model	# variables	Arithmetic scale ME	MAE	MPE	MAPE	RMSE	PVE	Log scale ME	MAE	RMSE	PVE
9	2	-36447	256136	-22	41	360594	29	-0.016	0.343	0.429	77
10	2	-27266	198608	-14	33	255723	64	-0.006	0.294	0.361	83
11	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-77192	341102	-33	58	520302	-48	-0.033	0.44	0.556	61
	3	-42914	261421	-26	48	368960	26	-0.041	0.37	0.485	70
	4	-70307	299517	-28	53	415412	6	-0.046	0.423	0.523	65
	5	-100675	392736	-31	63	549191	-65	-0.047	0.497	0.58	57
	6	-116031	323332	-28	55	529188	-53	-0.033	0.443	0.578	58
11[a]	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-94429	298228	-34	53	485259	-29	-0.052	0.396	0.51	67
	3	-95790	281245	-30	49	453521	-12	-0.054	0.371	0.479	71
	4	-93973	278968	-29	49	442337	-7	-0.049	0.375	0.478	71
	5	-97397	285365	-29	50	446205	-9	-0.043	0.385	0.483	70
	6	-102918	295873	-29	51	455384	-13	-0.036	0.4	0.498	69
11[b]	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-36447	256136	-22	41	360594	29	-0.016	0.343	0.429	77
	3	-38982	230836	-17	37	313906	46	-0.02	0.316	0.384	81
	4	-41426	218403	-17	36	295645	52	-0.024	0.313	0.38	82
	5	-50334	219194	-16	36	289909	54	-0.031	0.315	0.375	82
	6	-39841	202403	-14	33	272681	59	-0.025	0.295	0.352	84
12	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-106608	308835	-34	55	459631	-15	-0.073	0.395	0.526	65
	3	-57860	294744	-22	51	414902	6	-0.006	0.417	0.505	68
	4	-100491	258124	-28	48	348342	34	-0.096	0.362	0.461	73
	5	-104328	266112	-26	44	359470	29	-0.116	0.338	0.428	77
12[a]	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-87596	275161	-30	49	415308	6	-0.047	0.367	0.47	72
	3	-91150	269044	-29	49	403674	11	-0.048	0.346	0.471	72

Table S1.1. continued.

Model	# variables	Arithmetic scale ME	MAE	MPE	MAPE	RMSE	PVE	Log scale ME	MAE	RMSE	PVE
	4	-86222	259554	-27	48	369793	25	-0.043	0.361	0.462	73
	5	-86814	261164	-24	44	340888	37	-0.06	0.352	0.422	77
12[b]	1	-64154	272098	-34	55	409768	8	-0.009	0.412	0.528	65
	2	-58932	231828	-19	39	384563	19	-0.008	0.31	0.419	78
	3	-37064	237002	-17	38	322236	43	-0.017	0.323	0.393	80
	4	-40009	202329	-13	33	262522	62	-0.018	0.294	0.347	85
	5	-21133	194040	-9	27	251642	65	-0.008	0.251	0.3	89
13	1	-8521	313130	-28	55	431908	-2	-0.007	0.435	0.546	62
	2	-5572	270125	-17	45	345258	35	0.006	0.386	0.454	74
	3	-14834	263435	-14	43	368833	26	0.022	0.38	0.476	71
	4	-17385	264838	-14	43	351591	32	0.014	0.39	0.472	72
	5	-26060	248363	-15	43	344582	35	0.001	0.378	0.478	71
	6	-40366	257405	-16	43	348359	34	-0.01	0.386	0.486	70

[a] Model averaging.

[b] Model selection not included in cross validation.

Table S1.2. One-year-ahead cross-validation performance of alternative models for forecasting the Sacramento Index. Models are defined in Table 2, and performance metrics are defined in Table 3.

Model	#	Arithmetic scale						Log scale			
	variables	ME	MAE	MPE	MAPE	RMSE	PVE	ME	MAE	RMSE	PVE
1	0	-89697	388747	-228	252	472821	0				
2	1	-64532	317535	-32	59	480748	-3				
3	2	-39731	290653	-29	49	411566	24				
4	2	-37724	261274	-10	39	397725	29				
5	1	-64532	317535	-32	59	480748	-3				
	2	-119659	281818	-39	59	476077	-1				
	3	-167229	262874	-45	57	426301	19				
	4	-144996	253728	-47	61	376898	36				
	5	-99556	275879	-41	59	392318	31				
	6	-105506	270174	-39	58	372433	38				
5[a]	1	-64532	317535	-32	59	480748	-3				
	2	-123772	284232	-40	60	481784	-4				
	3	-152007	277560	-44	58	456339	7				
	4	-151096	262627	-44	58	410355	25				
	5	-119056	245194	-41	55	369343	39				
	6	-117654	229452	-41	53	342570	48				
5[b]	1	-64532	317535	-32	59	480748	-3				
	2	-119659	281818	-39	59	476077	-1				
	3	-97390	256358	-37	52	400452	28				
	4	-68593	243914	-27	45	350233	45				
	5	-117758	199451	-33	44	298842	60				
	6	-127042	190707	-34	43	289831	62				
6	0	-118684	412604	-247	270	503396	-13	-0.36	0.724	1.084	0
7	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
8	1	-9105	230730	-27	51	279596	65	-0.075	0.405	0.491	80
9	2	-10538	256871	-40	59	332135	51	-0.133	0.424	0.527	76
10	2	-2633	256182	-28	49	354910	44	-0.088	0.384	0.48	80

Table S1.2. continued.

Model	# variables	Arithmetic scale						Log scale			
		ME	MAE	MPE	MAPE	RMSE	PVE	ME	MAE	RMSE	PVE
11	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	-75457	229655	-49	63	323488	53	-0.221	0.394	0.559	73
	3	-109867	220946	-46	57	326589	52	-0.234	0.374	0.533	76
	4	-71626	232302	-41	59	300023	60	-0.183	0.396	0.536	76
	5	-53100	274071	-43	66	386791	33	-0.154	0.446	0.596	70
	6	-123440	353400	-43	71	496209	-10	-0.154	0.522	0.626	67
11[a]	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	-86780	240228	-51	65	350280	45	-0.229	0.402	0.567	73
	3	-102330	223352	-46	58	325923	52	-0.228	0.372	0.531	76
	4	-100702	231407	-46	58	325689	53	-0.22	0.382	0.535	76
	5	-84263	231923	-43	58	325148	53	-0.192	0.386	0.542	75
	6	-78030	241585	-42	59	338595	49	-0.17	0.407	0.55	74
11[b]	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	-10538	256871	-40	59	332135	51	-0.133	0.424	0.527	76
	3	-62778	218924	-37	51	319178	54	-0.175	0.354	0.494	79
	4	-65492	237683	-36	51	319986	54	-0.175	0.365	0.481	80
	5	-46041	229831	-34	50	316578	55	-0.161	0.357	0.479	81
	6	-55839	223728	-26	44	296954	61	-0.123	0.349	0.431	84
12	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	-75457	229655	-49	63	323488	53	-0.221	0.394	0.559	73
	3	-121207	223523	-48	59	320990	54	-0.252	0.379	0.538	75
	4	-102529	301181	-47	69	378881	36	-0.203	0.481	0.606	69
	5	-74070	244995	-32	52	303838	59	-0.149	0.397	0.501	79
12[a]	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	-70840	224260	-49	63	308659	57	-0.212	0.385	0.545	75
	3	-106017	222219	-49	60	318852	55	-0.237	0.375	0.539	75
	4	-106809	243531	-49	62	329676	51	-0.228	0.394	0.552	74
	5	-69946	233906	-41	58	293000	62	-0.173	0.397	0.533	76

Table S1.2. continued.

Model	# variables	Arithmetic scale						Log scale			
		ME	MAE	MPE	MAPE	RMSE	PVE	ME	MAE	RMSE	PVE
12[b]	1	-67765	307624	-59	78	444607	12	-0.21	0.489	0.622	67
	2	41829	232987	-36	59	278630	65	-0.105	0.425	0.527	76
	3	-63209	219355	-37	51	320449	54	-0.176	0.355	0.495	79
	4	-17854	183956	-32	46	253662	71	-0.14	0.327	0.468	81
	5	-9469	182740	-20	36	241125	74	-0.095	0.29	0.376	88
13	1	-37524	216476	-83	98	295208	61	-0.234	0.455	0.726	55
	2	51251	215086	-22	47	299597	60	-0.043	0.38	0.503	79
	3	76190	230717	-15	47	310797	57	0.018	0.412	0.517	77
	4	60140	238742	-17	50	318147	55	0.001	0.432	0.542	75
	5	53755	238571	-19	53	323993	53	0.005	0.454	0.578	72
	6	73339	226056	-13	48	309322	57	0.04	0.429	0.553	74

[a] Model averaging.

[b] Model selection not included in cross validation.

Table S1.3. Management performance of alternative models for forecasting the Sacramento Index under leave-one-out and one-year-ahead cross validation. The number of conservation (C) and fishing opportunity (F) errors are shown. Models are defined in Table 2.

Model	# variables	Leave-one-out		One-year-ahead	
		C	F	C	F
1	0	5	0	5	0
2	1	1	0	2	0
3	2	1	0	1	0
4	2	1	1	1	0
5	1	1	0	2	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	1
	5	1	1	1	0
	6	1	3	1	0
5[a]	1	1	0	2	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	0
	5	1	1	1	0
	6	1	2	1	0
5[b]	1	1	0	2	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	0
	5	1	1	1	0
	6	1	1	1	0
6	0	5	0	5	0
7	1	2	0	4	0
8	1			2	1
9	2	1	0	2	0

Table S1.3. continued.

Model	# variables	Leave-one-out C	Leave-one-out F	One-year-ahead C	One-year-ahead F
10	2	1	0	2	0
11	1	2	0	4	0
	2	1	0	2	0
	3	1	0	2	0
	4	1	0	1	0
	5	1	2	1	0
	6	1	2	1	1
11[a]	1	2	0	4	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	0
	5	1	0	1	0
	6	1	2	1	0
11[b]	1	2	0	4	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	0
	5	1	1	1	0
	6	1	1	1	0
12	1	2	0	4	0
	2	2	0	2	0
	3	1	1	2	0
	4	1	0	2	0
	5	1	1	1	0
12[a]	1	2	0	4	0
	2	1	0	2	0
	3	1	0	1	0
	4	1	0	1	0

Table S1.3. continued.

Model	# variables	Leave-one-out C	Leave-one-out F	One-year-ahead C	One-year-ahead F
	5	1	1	1	0
12[b]	1	2	0	4	0
	2	2	0	4	0
	3	1	0	1	0
	4	1	0	1	0
	5	1	1	1	0
13	1	2	0	4	0
	2	2	0	2	0
	3	1	1	2	1
	4	1	1	2	1
	5	1	2	2	1
	6	1	2	2	1

[a] Model averaging.

[b] Model selection not included in cross validation.