# An Error Identification Approach to Forecasting the Harvest Rate of One Species in a Multispecies Fishery, Illustrated with Two Salmon Fisheries 

Barry D. Smith*<br>Environment Canada--Canadian Wildlife Service, Pacific Wildlife Research Centre, 5421 Robertson Road, Delta, British Columbia V4K 3N2, Canada

Thomas F. Shardlow
South Coast Division, Department of Fisheries and Oceans, 3225 Stephenson Point Road, Nanaimo, British Columbia V9T IW5, Canada

A. W. Argue<br>Program Planning and Economics Branch, Department of Fisheries and Oceans, 555 West Hastings Street, Vancouver, British Columbia V6B 5G3, Canada


#### Abstract

For a multispecies fishery managed to achieve a harvest rate target for a single species, an early and accurate forecast of an in-season or the end-of-season harvest rate index can assist a manager with a decision to control effort. We present a model for rapidly forecasting an inseason or end-of-season harvest rate index for one species, the species of interest, in a multispecies fishery using catch and effort data collected for that fishery. These harvest rate predictions can be statistically evaluated against harvest rate indices calculated independently from demographic and biological data obtained for the fishery. The model structure was defined such that an estimate of both measurement error and process error could be obtained. The data required are the historical time series of catches for the species caught in the fishery, fishing effort, and harvest rate indices for the species of interest calculated independently of the catch and effort data. We illustrate the model with chinook salmon Oncorhynchus tshawytscha as the species of interest in the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery. Other important species caught in these multispecies fisheries are coho salmon $O$. kisutch, sockeye salmon $O$. nerka, and pink salmon O. gorbusha.


Troll and sport fisheries are the primary harvesters of chinook salmon Oncorhynchus tshawytscha along the Pacific coast of Canada. Since 1985, the management of these fisheries has been strongly influenced by conditions set by the Can-ada-United States Pacific Salmon Treaty. One of the conservation goals of the Pacific Salmon Treaty was to double the spawning escapement of Strait of Georgia and west coast of Vancouver Island chinook salmon stocks by 1998 (Anonymous 1994). The mechanisms advanced for rebuilding these chinook salmon stocks included reducing the aggregate exploitation rate to achieve a reduction of the harvest rate on individual stocks. To this end the Pacific Salmon Commission (PSC) has published detailed technical reports on chinook salmon stock status on a stock-by-stock and year-to-year basis. These reports provide composite harvest rate indices (called fishery indices by the

[^0]Received September 5, 1997 Accepted September 2, 1998

PSC) obtained by analyzing the catch, the age structure, and the stock composition of coded-wire-tagged chinook salmon captured by specific fisheries.

To date, the main preseason analytical tool used to design and evaluate regulatory actions aimed at reducing harvest rates, such as area closures, fishing seasons, or size limit changes, has been the troll fishery management model (English et al. 1987). This simulation model was designed around the multispecies nature of the troll fishery because, in addition to chinook salmon, these fisheries catch varying numbers of coho salmon $O$. kisutch, sockeye salmon $O$. nerka, pink salmon $O$. gorbusha, and chum salmon $O$. keta, as well as steelhead $O$. mykiss. Predictions were made about the amount of fishing effort directed toward chinook salmon and the resulting chinook salmon catch. These predictions are based on the amount of effort directed at each species, as well as the predicted abundance and price of each salmon species for the upcoming season. Management decisions to achieve harvest rate targets were based on the performance of the simulations.

In 1993, managers chose to link the conduct of the fishery in terms of catch and effort directly to harvest rate indices developed by the Pacific Salmon Commission. The management objective was to ensure that harvest rates did not exceed those of recent years, based on harvest rate indices provided the Chinook Technical Committee (CTC) of the PSC. Thus, for managers to achieve a harvest rate reduction on an annual basis, they would have to anticipate the postseason harvest rate index for the fishery, based on the recent harvest rate history of the fishery and the anticipated performance of the fishery in the current season. However, formal analyses of stock composition data (i.e., spawning stream, age, brood-year data) for the coastal and Strait of Georgia chinook salmon fisheries are not complete until the summer of the year following a fishery. Thus an official PSC harvest rate index for the previous year's fishery based on these data cannot be produced in time for preseason management discussions.

This frustrating time delay in the calculation of a harvest rate index for the chinook salmon troll and sport fisheries of southern British Columbia stimulated our search for a simple, general method to rapidly and reliably forecast a harvest rate index for one species in any multispecies fishery without having at hand a detailed contemporary knowledge of the stock composition of the species of interest. The method we present in this paper to achieve this goal has at least three benefits. First, for the forecast year, only the catch of each species and the effort expended (e.g., boat-days) are required. Second, fewer assumptions are needed than in a simulation approach like the troll fishery management model mentioned above. That particular simulation model requires, among other things, predictions about prices paid to fishers for each species and grade, assumptions associated with abundance and cohort analysis, and specific predictions about the effect of area or species closures. Third, the model we propose has a statistical basis that allows estimates of error, identification of the source of the error (process error or measurement error), and evaluation of model fit.

We used the approach of Schnute et al. (1989) for solving this forecasting problem because of our desire to forecast a harvest rate based on short time series of historical harvest rates for which more complex time series models (Box and Jenkins 1970; Hipel and McLeod 1994) would not be identifiable. Our purpose was to develop an a priori identifiable model structure that could potentially segregate measurement and process error (Schnute
et al. 1989) and that is general enough to be applied to similar problems and under similar circumstances elsewhere. A particularly useful feature of our model is that distinction of process error and measurement error can improve a manager's understanding of the quality of the information collected about a fishery.

Here we present a model where the only data required are time series of total catch for those species caught by the fishery, of fishing effort, and of the harvest rate (or an index of it). For the time period for which a forecast is to be made, the cumulative in-season catch totals by all species and cumulative fishing effort are required. The model development is typical in that it requires that certain assumptions be made, both to develop a workable formulation and to fulfill the statistical conditions of the model. We show how these assumptions can be challenged retrospectively and that the model can provide useful harvest rate forecasts by using the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery as examples.

## Harvest Rate Index Forecast Model

## Development

An index of the instantaneous harvest rate,

$$
\begin{equation*}
P_{t}=r \log _{e}\left(1-\frac{C_{t}}{N_{t}}\right) \tag{1}
\end{equation*}
$$

(Ricker 1975), where $C_{t}$ is catch (numbers or weight) and $N_{t}$ is initial population size (numbers or weight) for a discrete time period $t$, includes the scaling parameter $r$ because difficulty in measuring $N_{t}$ usually precludes calculation of an absolute harvest rate. This harvest rate index is related to fishing effort $\left(E_{t}\right)$ by the axiomatic relationship

$$
\begin{equation*}
P_{t}=q E_{t} \tag{2}
\end{equation*}
$$

(Ricker 1975; Gulland 1983), where $q$ is the catchability coefficient.

Consider that competition among fishing vessels (or other factors such as vessel capacity, species value, etc.) can result in $q$ varying among time periods as a function of total effort within the time period $t$ as described by

$$
\begin{equation*}
q=\kappa E_{t}^{\zeta} \tag{3}
\end{equation*}
$$

Typically one would expect $\zeta<0$, with $q$ declining with increasing effort. Brannian (1982) found such a relationship between catchability and effort in
the Togiak Bay, Alaska, sockeye salmon fishery, and Peterman and Steer (1981) and Winters and Wheeler (1985) used this functional form to relate variability in catchability to stock abundance and stock area, respectively. Combining equations (2) and (3) gives

$$
\begin{equation*}
P_{t}=\kappa E_{t}^{1+\zeta} \tag{4}
\end{equation*}
$$

From equation (4) we introduce the ratio model

$$
\begin{equation*}
P_{r}=P, \frac{E_{t}^{1+\zeta}}{E_{t-1}^{1+\zeta}} \tag{5}
\end{equation*}
$$

with a one-time-period lag to relate changes in the instantaneous harvest rate index to changes in effort from period $t-1$ to $t$.

Effort $\left(E_{t}\right)$ in time period $t$ directed at any one species in a multispecies fishery is usually not revealed in the unpartitioned effort data ( $B_{t}$ ) that is typically recorded by fishery management agencies. Therefore we propose the function

$$
\begin{equation*}
E_{t}^{1+\zeta}=B_{t}^{1+\zeta} \exp \sum_{i=1}^{n} \gamma_{j} C_{j, t}^{\prime} \tag{6}
\end{equation*}
$$

with coefficients $\gamma_{j}$ to capture how interperiod variation in the catch ( $C_{j, t}$ ) of the $j=1, \ldots, n$ species caught in the multispecies fishery, including the species of interest, could proportionately affect the amount of effort directed specifically at the species of interest. Fishing effort expended to catch species other than the species of interest will reduce effort directed at that species. Likewise, increased total effort ( $B_{1}$ ) directed at the species of interest can potentially reduce effort directed at that species because of increased handling time required to cope with the quantities of fish caught. In equation (6), $C^{\prime}{ }_{j, t}$ simply represents $C_{j, t}$ scaled to standard deviations ( $\sigma_{j}$ ) from the mean ( $\mu_{j}$ ) of the historical catches over the time periods $t=$ $1,2, \ldots, f-1$ with $f$ being the time period for which a forecast is to be made. This scaling does not affect model predictions but has the useful parameterization that $E_{t}=B_{t}$ when $C_{j, t}=\mu_{j}$ for each species $j$.

The preceding model development leads to the following deterministic expression for predicting the instantaneous harvest rate index for a discrete time period $t$ or a forecast of the index for the forecast time period $t=f$.

$$
\begin{equation*}
P_{t}=P_{t} \frac{B_{i}^{(1+\zeta)} \exp \sum_{j=1}^{n} \gamma_{j} C_{j . t}^{\prime}}{B_{i-1}^{(1+\zeta)} \exp \sum_{j=1}^{n} \gamma_{j} C_{j, t-}^{\prime}} \tag{7}
\end{equation*}
$$

## Error Structure

Deterministic models for biological systems are generally imperfect because they are unable to incorporate sufficient realism, and the measured values for model variables typically include error. Models that include a statistical component to capture error tend to be preferred because they provide an objective means to evaluate the precision, accuracy, and repeatability of model performance. Model error arising from a deficiency in the realism of the model is process error, whereas error arising from an inability to accurately measure the value of a model variable is measurement error.

The model in equation (7) can be made stochastic by adding process error, measurement error, or both. The incorporation of a measurement error term acknowledges that we are probably unable to measure a perfectly predicted harvest rate index $\left(P_{t}\right)$ without random error. If we assume this error enters the model multiplicatively we can define the observed harvest rate indices $\left(I_{t}\right)$ in terms of the true harvest rate indices $\left(P_{t}\right)$ as

$$
\begin{equation*}
I_{t}=P_{t} e^{\theta \delta_{t}} \tag{8}
\end{equation*}
$$

Alternatively, incorporating process error into the model acknowledges that the model is randomly imperfect and that the harvest rate indices ( $P_{t}$ ) are unlikely to be perfectly predicted. Model imperfections can arise from simplified functional relationships, such as the models in equations (4) and (6). By assuming that process error also enters the model multiplicatively, we can define the observed harvest rate indices $\left(I_{t}\right)$ in terms of the predicted harvest rate indices $\left(P_{t}\right)$ as

$$
\begin{equation*}
I_{t}=P_{t} e^{\Theta \epsilon} \tag{9}
\end{equation*}
$$

The model and its possible error structures are now stated. Taking the natural logarithms of equations (7) to (9) leads directly to the linear measurement error model

$$
\begin{aligned}
\log _{e}\left(P_{t}\right)= & \log _{e}\left(P_{1}\right)+\sum_{j=1}^{n} \gamma_{M, j}\left(C_{j, t}^{\prime}-C_{j, 1}^{\prime}\right) \\
& +\left(1+\zeta_{M}\right)\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{1}\right)\right] \\
\text { and } \quad & \log _{e}\left(I_{t}\right)=\log _{e}\left(P_{t}\right)+\Theta \delta_{t} .
\end{aligned}
$$

where $\log _{e}\left(P_{1}\right)$ is a parameter to be estimated, and
the linear process error model

$$
\begin{align*}
\log _{e}\left(P_{t}\right)= & \log _{e}\left(I_{t-1}\right)+\sum_{j=1}^{n} \gamma_{P, j}\left(C_{j, t}^{\prime}-C_{j, t-1}^{\prime}\right) \\
& +\left(1+\zeta_{P}\right)\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{t-1}\right)\right] \tag{11a}
\end{align*}
$$

and

$$
\begin{equation*}
\log _{e}\left(I_{t}\right)=\log _{e}\left(P_{t}\right)+\Theta \epsilon_{t} . \tag{11b}
\end{equation*}
$$

## Process Error and Measurement Error Model

Both models above (equations 10 and 11) lend themselves to multiple-regression analysis if certain statistical conditions can be met or satisfactorily approximated. One condition is that the model residuals, $\log _{e}\left(I_{t}\right)-\log _{e}\left(P_{t}\right)$, are normal and independently distributed. Another is that the dependent variables of catch $\left(C_{j, t}\right)$ and effort $\left(B_{t}\right)$ ideally must have been measured without error. We proceed under the assumption that both catch and effort are measured with an unimportant amount of error, particularly when compared with the error associated with the predictions $\log _{e}\left(P_{t}\right)$. We justify this by considering that both catch and effort are censuses of information that fishers are obliged to provide to the management agency. They are therefore likely to be measured with much less uncertainty than the harvest rate indices these data are used to predict. This assumption would have to be rejected if a posteriori diagnostics of model fit cast doubt upon it. In that case an errors-in-variables modeling approach might be prescribed (see Schnute et al. 1990; Schnute 1994).

Equations (10) and (11) can be combined to define one model that (1) includes both process error and measurement error, (2) can predict the harvest rate indices for time periods $t=2,3, \ldots, f-1$, and (3) can forecast the harvest rate index for time period $f$. By letting $\omega$ represent the proportion of total variance, $\Theta^{2}=E\left\{\left[\log _{e}\left(I_{t}\right)-\log _{e}\left(P_{t}\right)\right]^{2}\right\}$ contributed by process error and (1- $\omega$ ) the proportion contributed by measurement error, we acknowledge that both error components are likely to contribute to the model error associated with the predictions $\log _{e}\left(P_{t}\right)$. If we combine the models in equations (10) and (11) by (1) using $\omega$ and ( $1-\omega$ ) to represent the proportional contribution of the process error and measurement error models, respectively, (2) using $\beta$ s to represent confounded parameters, and (3) expressing the model in terms of observed harvest rate indices $\left(\log _{e}\left(I_{t}\right)\right.$ ), we get

$$
\begin{align*}
\tau_{M, t}= & \beta_{0}+\omega\left[\tau_{M, t}-\tau_{P, t}+\log _{e}\left(I_{t-1}\right)\right] \\
& +\sum_{j=1}^{n}\left[\beta_{j}\left(C_{j, t}^{\prime}-C_{j, 1}^{\prime}\right)+\beta_{n+j}\left(C_{j, t}^{\prime}-C_{j, t-1}^{\prime}\right)\right] \\
& +\beta_{2 n+1}\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{1}\right)\right] \\
& +\beta_{2 n+2}\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{t-1}\right)\right] \\
& +\beta_{2 n+3} V_{t}+\eta_{t}, \quad t>1 ; \\
\eta_{t}= & \sqrt{(1 \quad \omega)} \Theta\left[(1+\omega) \delta_{t}-\omega \delta_{t-1}\right] \\
& +\sqrt{\omega} \Theta \epsilon_{t}, \quad t>  \tag{12b}\\
\tau_{M, t}= & \log _{e}\left(I_{t}\right)-\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{1}\right)\right] ; \quad \text { and }  \tag{12c}\\
\tau_{P, t}= & \log _{e}\left(I_{t}\right)-\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{t-1}\right)\right] \\
& t>1,
\end{align*}
$$

with variance components
$\operatorname{Var}\left[\eta_{t}\right]=\Theta^{2}\left[(1+\omega)\left(1-\omega^{2}\right)+\omega^{2}(1-\omega)+\omega\right]$,
(13a)
and

$$
\begin{equation*}
\operatorname{cov}\left[\eta_{t}, \eta_{t+1}\right]=\Theta^{2} \omega\left(1-\omega^{2}\right) \tag{13b}
\end{equation*}
$$

As captured in the definitions of $\tau_{M, t}$ (equation 12c) and $\tau_{P, t}$ (equation 12d), the full model contains the following deterministic component:

$$
\begin{align*}
& (1-\omega)\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{1}\right)\right] \\
& \quad+\omega\left[\log _{e}\left(B_{t}\right)-\log _{e}\left(B_{i-1}\right)\right] . \tag{14}
\end{align*}
$$

Model predictions can be partitioned into this deterministic component and the stochastic component represented by equations (4) and (6).

We also added the potential for a step intervention (i.e., the term $\beta_{2 n+3} V_{t}$ ). Such an intervention would change the mean level of the time series of predicted harvest rate indices, $\log _{e}\left(P_{t}\right)$, which might not be otherwise explainable by our proposed model structure (Hipel and McLeod 1994). This could occur if, for example, a management agency suddenly changed the way it collected catch or effort data (or the way it calculated harvest rate indices) between two time periods $t-1$ and $t$. The covariate time series, $V_{t}$, of zeros ( 0 ) and ones (1) is controlled by the analysts according to whether they believe the intervention applies to a time period $t\left(V_{t}=1\right)$ or does not $\left(V_{t}=0\right)$.

Multiple regression would not be prescribed for this model if $0<\omega<1$ because the model errors would be autocorrelated, though somewhat weakly with a absolute value maximum of $r=-0.22$ when $\omega=0.58$. We therefore rely on the solution of

Table 1.-Official Canadian Department of Fisheries and Oceans statistics for effort (in boat-days) and catch (in pieces) of chinook salmon, coho salmon, sockeye salmon, and pink salmon; and the Pacific Salmon Commission harvest rate index for the west coast of Vancouver Island troll fishery.

| Year | Boat-days | Chinook <br> salmon | Coho <br> salmon | Sockeye <br> salmon | Pink <br> salmon | Index |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |

Schnute et al. (1989). Our parameter $\omega$ is related to the parameter $\lambda$ of Schnute et al. (1989) by the following quadratic relationship, $\lambda$ being the smaller root of

$$
\begin{align*}
0= & \omega\left(1-\omega^{2}\right) \lambda^{2}+\left(2 \omega^{3}-2 \omega-1\right) \lambda \\
& +\omega\left(1-\omega^{2}\right) . \tag{15}
\end{align*}
$$

To convert model forecasts, $\log _{e}\left[P_{f}\right]$, to arithmetic harvest rate predictions, note that the expected value of an observed harvest rate index ( $I_{f}$ ) is

$$
\begin{equation*}
E\left[I_{f}\right]=P_{f} \exp \left(\frac{\operatorname{Var}\left[\eta_{f}\right]}{2}\right) \tag{16a}
\end{equation*}
$$

with variance

$$
\begin{equation*}
V\left[I_{f}\right]=P_{f}^{2} \exp \left(2 \operatorname{Var}\left[\eta_{f}\right]\right)-P_{f}^{2} \exp \left(\operatorname{Var}\left[\eta_{f}\right]\right) \tag{16b}
\end{equation*}
$$

Log-normal confidence intervals for $I_{f}$ can be constructed using equation (16) where $\operatorname{Var}\left[\eta_{f}\right]$ can be calculated using equation (20.34) of Zar (1984).

When interpreting the results of a particular analysis, consider that if $\omega$ in equation (12a) is estimated to lie between zero and one $(0 \leq \omega \leq$ 1), then this can be interpreted as an estimate of the proportion of model error attributable to process error as defined by equation (12). If the confidence limits for $\omega$ include zero ( 0 ) or one (1), then the analyst should consider parsimony and reduce the model to a measurement error model ( $\omega=0$ ) or a process error ( $\omega=1$ ) model. An analyst is most likely to face this decision when analyzing short time series or when a proposed model has difficulty explaining the observed data. When the value for $\omega$ falls outside the range 0 to

1 the analysts might consider that the model is inadequately posed, especially if they insist that model error be defined strictly in terms of process error and measurement error. At this point, a wisely chosen intervention might result in $0 \leq \omega \leq 1$. Likewise, a parsimonious model with $\omega=0$ (all measurement error) or $\omega=1$ (all process error) might include the independent variables from the alternate error model (equation 10 or 11 ). In such a case, one should think of the variables from the alternate model as interventions.

## Two Examples of Model Performance

To illustrate the model, we chose two salmon fisheries for which the appropriate data are available. One fishery is the west coast of Vancouver Island commercial troll fishery for which we had data available data on catch, effort, and harvest rate indices for 1979-1992 (Table 1; Figure 1). Because of its high commercial value, chinook salmon can be considered the target species in this multispecies fishery. Our other example is the Strait of Georgia sport fishery. Historically chinook salmon or coho salmon have been the target species for the sport fishery (Argue et al. 1983), but in recent years, catches of sockeye salmon and pink salmon have increased (Collicutt and Shardlow 1990, 1992). For this fishery, we have data on catch, effort, and harvest rate indices for 19801991 (Table 2; Figure 2). Our analyses for both examples considered only catches of chinook salmon, coho salmon, sockeye salmon, and pink salmon taken by both these fisheries. Chum salmon and steelhead are excluded because of the small numbers of fish caught.


Year

Year

Year
Figure 1.-Time series of (a) effort for those species retained in the models for the west coast of Vancouver Island troll fishery and catches of (b) chinook salmon and (c) pink salmon.

## Catch and Effort Data

As a condition of the license to fish, one copy of each sales slip produced by the sale (delivery) of salmon caught by a member of the troll fleet is forwarded to the Canadian Department of Fisheries and Oceans Catch Statistics Unit. Each sales slip documents the effort (i.e., the number of days fished or boat-days) and the weight of each species of salmon caught since the last delivery. Average weight per fish factors are then used to convert the weight of fish delivered to number of pieces
caught. Catch for all five species of salmon and steelhead appear on these sales slips. This formal collection of catch and effort data provides a reliable and timely census of catch and effort for this troll fishery. These same species are caught by the Strait of Georgia sport fishery. Catch and effort (boat-trip) statistics for this fishery have been obtained since 1980 by creel survey. The estimation errors for catch and effort calculated by this methodology tend to be about $5-10 \%$ of the estimate (Collicutt and Shardlow 1990, 1992).

Table 2.-Official Canadian Department of Fisheries and Oceans statistics for effort (in boat-trips) and catch (in pieces) of chinook salmon, coho salmon, sockeye salmon, and pink salmon; the analyst's choice for an intervention $\left(V_{f}\right)$ : and a harvest rate index calculated using cohort analysis for the Strait of Georgia sport fishery.

| Year | Bual-days | Chinook <br> salmon | Coho <br> valmon | Sockeye <br> salmon | Pink <br> salmon | $V_{t}$ | Index |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |

## Harvest Rate Indices

The PSC publishes harvest rate indices for major chinook salmon fisheries in British Columbia, Washington, Oregon, and Alaska (e.g., Anonymous 1994). The PSC formally refers to these harvest rate indices as fishery indices. A fishery index for a particular fishery is calculated by using the exploitation rate estimates of individual chinook salmon stocks that are exploited by a particular fishery in a particular year. To calculate this index. each fishery is declared to have a number of chinook salmon indicator stocks identified by their spawning stream. Mark-recovery and cohort analysis are used to estimate an exploitation rate for those components of each indicator stock defined by their age ( $3-5$ years) and brood year. An unscaled chinook salmon exploitation rate for an entire fishery (e.g., the west coast of Vancouver Island troll fishery) is then calculated by weighting the contributions to the fishery of each indicator stock by age-class and brood year. These rates are then scaled to obtain an average aggregate PSC harvest rate index of close to 1.0 over the four base years 1979-1982.

Harvest rate indices for Strait of Georgia chinook salmon fisheries were generated by cobort analysis without regard to the stock composition or the explicit use of mark-recovery information (A. W. Argue, unpublished data). Age-2 and older chinook salmon from streams contributing to fisheries in the Strait of Georgia are primarily resident in the strait until they mature and return to these coastal streams to spawn (Argue et al. 1983). Therefore, given estimates of catch at age and escapement for age- 5 fish, natural mortality, net emigration from the Strait, and maturation rates for age-3-4 fish, initial population sizes and fishing
moralities can be estimated. These data were used in cohort analysis (Pope 1972) to calculate harvest rates for age-3-5 chinook salmon caught in the Strait of Georgia sport fishery. The harvest rates calculated by this method were also scaled to obtain an average harvest rate index of 1.0 over the four PSC base years 1979-1982. Strait of Georgia harvest rate indices were analyzed only for the years 1980 and later because 1980 was the first year sport effort data were collected by creel survey.

The harvest rate index series for both the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery are calculated using arithmetic harvest rates. Though we have previously defined $I_{i}$ and $P_{t}$ as instantaneous indices and have developed the model on that basis, we justify using these arithmetic indices by relying on the similarity of the values for arithmetic and instantaneous indices when arithmetic harvest rates are relatively low. The annual arithmetic harvest rates for both of these example fisheries are generally lower than $25 \%$ (Anonymous 1994: Appendix D). justifying this assumption.

## Results

As Zar (1984) states for multiple regression, there is no general rule for judging the "best" fit of a model. Therefore, we used two criteria to define the best fit. The best fit for our so-called open model yielded the smallest unbiased mean squared error among all possible candidate models (i.e.. including all four salmon species) without consideration for parsimony. The best fit for our parsimonious model required that all coefficients (excluding the intercept $\beta_{0}$ ) be significant at $\alpha=$ 0.05 , irrespective of the value for the mean squared


Figure 2.-Time series of (a) effort for those species retained in the models for the Strait of Georgia sport fishery and catches of (b) chinook salmon and (c) pink salmon.
error. All accepted models using either criterion for identifying the "best" fit passed a posteriori diagnostics that tested the assumptions of normal and independently distributed residual error and no skewness.

Satisfying models were obtained using both criteria for both the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery when the complete data sets were analyzed (Tables 3, 4; Figures 3a, and 4a). For both fisheries, there was statistical evidence only of measurement error ( $\omega=0$ ), although terms from both equations (10)
and (11) were retained in both model fits. The SEs reported in Tables 3 and 4 are less than those that would be reported for an equivalent multipleregression model (applicable only when $\omega=0$ or $\omega=1$ ) because Schnute et al. (1989) do not account for small sample bias in the estimate of variance $\left(\Theta^{2}\right)$. That measurement error dominates the model error is consistent with the PSC harvest rate indices and those derived using cohort analysis having been calculated using uncertain data. For both our examples, the accepted results depended only on the time series of chinook salmon and pink

Table 3. -Model statistics for the west coast of Vancouver Island troll fishery data of Table 1 for 1979-1992

| Source or covariale | Sum of squares | dt | Mean cepuare | $r^{2}$ | $P$ | Comefficient | $\begin{aligned} & \text { Appowi- } \\ & \text { mate Sl: } \end{aligned}$ | $P$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Open mudel: -2 loge (likelihood) $=-37.02$ |  |  |  |  |  |  |  |  |
| Source of variation |  |  |  |  |  |  |  |  |
| Model | 0.7685 |  | 0.2562 | 0.93 | $<0.001$ |  |  |  |
| Residual | 0.0580 | 9 | 0.0064 |  |  |  |  |  |
| Total | 0.8265 | 12 |  |  |  |  |  |  |
| Covariates |  |  |  |  |  |  |  |  |
| Intercept |  |  |  |  |  | -0.976 | 0.12 | $<0.001$ |
| Chinook salmon ( $\mathrm{C}_{1, t}-\mathrm{C}_{1, \ell-1}$ ) |  |  |  |  |  | 0.0730 | 0.03 | $>0.05$ |
| Pink salmon ( $\mathrm{C}_{4, t}-\mathrm{C}_{4,1}$ ) |  |  |  |  |  | -0.600 | 0.06 | $<0.001$ |
| Pink salmon ( $\mathrm{C}_{4, t}-\mathrm{C}_{4, t-1}$ ) |  |  |  |  |  | 0.279 | 0.03 | <0.001 |
| $\omega$ |  |  |  |  |  | 0 |  |  |
|  | Parsimonious model: $2 \log _{e}($ likelihood $)=\mathbf{- 3 1 . 2 2}$ |  |  |  |  |  |  |  |
| Source |  |  |  |  |  |  |  |  |
| Model | 0.7360 | 2 | 0.3679 | 0.89 | $<0.001$ |  |  |  |
| Residual | 0.0906 | 10 | 0.0091 |  |  |  |  |  |
| Total | 0.8265 | 12 |  |  |  |  |  |  |
| Covariates |  |  |  |  |  |  |  |  |
| Intercept |  |  |  |  |  | -1.098 | 0.13 | <0.001 |
| Pink salmon ( $\mathrm{C}_{4, t}-\mathrm{C}_{4,1}$ ) |  |  |  |  |  | -0.651 | 0.06 | <0.001 |
| Pink salmon ( $\mathrm{C}_{4, t}-\mathrm{C}_{4,1-1}$ ) |  |  |  |  |  | 0.265 | 0.03 | $<0.001$ |
| $\omega$ |  |  |  |  |  | 0 |  |  |

salmon catches. The analyses of the Strait of Georgia sport fishery (Table 4) required an intervention time series ( $V_{i}$ ) starting in 1985 (see Table 2) with $\beta_{2 n+3}=0.389$ for the open model and $\beta_{2 n+3}=$ 0.484 for the parsimonious model.

The key interpretation of the fits obtained for
both models is that only the time series of chinook salmon and pink salmon catches correlate with the time series of harvest rate indices. Both the longterm variation in chinook salmon or pink salmon catch and the change since the previous year can be correlated with year-to-year changes in the har-

Table 4.-Model statistics for the Strait of Georgia sport fishery data of Table 2 for 1980-I991

| Source or covariate | Sum of squares | df | Mean square |  | $P$ | Coefficients | Approximate SE | $P$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Open model: $-2 \log _{e}($ likelihood $)=-42.52$ |  |  |  |  |  |  |  |  |
| Source of variation |  |  |  |  |  |  |  |  |
| Model | 0.4751 | 5 | 0.0950 | 0.96 | 0.002 |  |  |  |
| Residual | 0.0210 | 5 | 0.0042 |  |  |  |  |  |
| Total | 0.4960 | 10 |  |  |  |  |  |  |
| Covariates |  |  |  |  |  |  |  |  |
| Intercept |  |  |  |  |  | 0.450 | 0.04 | $<0.001$ |
| Chinook salmon ( $\mathrm{C}_{1, t}-\mathrm{C}_{1,1}$ ) |  |  |  |  |  | 0.178 | 0.03 | $<0.01$ |
| Chinook salmon ( $\mathrm{C}_{1, t}-\mathrm{C}_{1, i-1}$ ) |  |  |  |  |  | 0.082 | 0.03 | $>0.05$ |
| Pink salmon $\left(\mathrm{C}_{4, t}-\mathrm{C}_{4,1}\right)$ |  |  |  |  |  | 0.156 | 0.08 | $>0.05$ |
| Pink salmon ( $\mathrm{C}_{4, t}-\mathrm{C}_{4, t-1}$ ) |  |  |  |  |  | -0.120 | 0.05 | $>0.05$ |
| $V_{t}$ |  |  |  |  |  | $0.389$ | 0.06 | $<0.01$ |
| $\omega$ |  |  |  |  |  |  |  |  |
| Parsimonious model: $\mathbf{- 2} \log _{e}($ likelihood $)=-36.24$ |  |  |  |  |  |  |  |  |
| Source of variation |  |  |  |  |  |  |  |  |
| Model | 0.4589 |  | 0.1530 | 0.93 | $<0.001$ |  |  |  |
| Residual | 0.3071 |  | 0.0053 |  |  |  |  |  |
| Total | 0.4960 |  |  |  |  |  |  |  |
| Covariates |  |  |  |  |  |  |  |  |
| Intercept |  |  |  |  |  |  | 0.05 | <0.001 |
| Chinook salmon ( $\mathrm{C}_{1, t}-\mathrm{C}_{1,1}$ ) |  |  |  |  |  | 0.175 | 0.03 | $<0.01$ |
| Chinook salmon ( $\mathrm{C}_{1, t}-\mathrm{C}_{1, f-1}$ ) |  |  |  |  |  | 0.116 | 0.02 | $<0.01$ |
| $V_{t}$ |  |  |  |  |  | 0.483 | 0.05 | $<0.001$ |
| $\omega$ |  |  |  |  |  | 0 |  |  |

a) Contemporary fit for 1980-1992

b) Retrospective forecasts for 1983-1992

c) Correlation of forecasted and observed values 1983-1992


## Forecasted

Figure 3.-West coast of Vancouver Island troll fishery 1980 to 1992 . (a) The values of $\log _{e}\left[P_{f}\right]$ predicted by the model in Table 3 is compared with the observed values calculated by the Pacific Salmon Commission ( $\log _{e}\left[I_{t}\right]$ ). (b) Observed value ( $\log _{e}\left[I_{t}\right]$ ) is superimposed upon the $95 \%$ confidence interval (CI, error bars) of the forecasted value, $\log _{e}\left[P_{f}\right]$, for that same year. (c) Correlation between the forecasted value, $\log _{e}\left[P_{f}\right]$, and the observed value, $\log _{e}\left[I_{f}\right]$, for a given year, 1983-1992. The $1: 1$ line indicates a perfect forecast (i.e., $\log _{e}\left[I_{f}\right]=\log _{e}\left[P_{f}\right]$ ). Open circles indicate forecasts for the early years (1983 and 1984) when the degrees of freedom were smallest.
a) Contemporary fit for 1981-1991

b) Retrospective forecasts for 1984-1991

c) Correlation of forecasted and observed values 1984-1991


Forecasted
Figure 4.-Strait of Georgia sport fishery 1981-1991. (a) The values of $\log _{e}\left[P_{f}\right]$ predicted by the model in Table 4 is compared with the observed values calculated using cohort analysis $\left(\log _{e}\left[I_{t}\right]\right)$. (b) Observed value $\left(\log _{e}\left[I_{t}\right]\right)$ is superimposed upon the $95 \%$ confidence interval (CI, error bars) of the forecasted value, $\log _{e}\left[P_{f}\right]$, for that same year. (c) Correlation between the forecasted value, $\log _{e}\left[P_{f}\right]$, and the observed value, $\log _{e}\left[I_{f}\right]$, for a given year, $1984-$ 1991. The $1: 1$ line indicates a perfect forecast (i.e., $\log _{e}\left[I_{f}\right]=\log _{e}\left[P_{f}\right]$. Open circles indicate forecasts for the early years (1984-1986) when the degrees of freedom were smallest.
vest rate indices. The 1-year lag coefficients disconnect time series data more historic than 1 year from the harvest rate index for year $t$. This implies that some characteristic of these fisheries is changing over time in a manner that prevents the longterm trends in chinook salmon and pink salmon catches from explaining the pattern in the harvest rate indices over the same time period.

One interpretation of the fits to the west coast of Vancouver Island troll fishery is that the implementation of the Pacific Salmon Treaty in 1985 changed the dynamics of the fishery. This change might have resulted in, for example, the troll fishery directing more effort at pink salmon because chinook salmon were made less attractive by the treaty. Note that despite huge biannual variability, landings of pink salmon appeared to be declining until about 1984, at which point they seemed to level off (Table 1; Figure 1c). Effort also tended to stabilize after 1985 (Table 1; Figure 1a). For the Strait of Georgia sport fishery (Table 4; Figure 4a), inserting a step intervention (Hipel and McLeod 1994) between 1984 and 1985 dramatically improved the model. This implies that an abrupt event occurred between 1984 and 1985 that was not captured in the independent time series candidates for the model. Although there is no need to explain the underlying cause that leads to this intervention in order to validate the model in a statistical sense, it would be more satisfying to propose a plausible explanation. Again, the implementation of the Pacific Salmon Treaty in 1985 is one such proposal, although the catch trends portrayed in Figure 2 do not suggest to us a specific mechanism.

## Validation

The results we obtained for the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery (Tables 3, 4; Figures 3, 4a) demonstrate an ability to identify statistical correlations between catch, effort, and harvest rate. However, this result alone does not constitute a test of the model's utility for forecasting harvest rate indices. We validated the model's forecasting ability by retrospectively and blindly obtaining "best" fits, as defined by our two previously stated criteria, to the historical data for the two examples we presented. Because we were primarily interested in evaluating how the multispecies nature of these fisheries relates to chinook salmon harvest rate, we took advantage of hindsight and included the intervention time series, $V_{t}$, in the analyses for the Strait of Georgia sport fishery. For the open
model we proceeded by blindly generating a harvest rate index forecast for each past year $(t)$ where there were sufficient data for years up to $f$, the forecast year. For the parsimonious model we accepted the covariate series suggested by the analyses of all available data presented in Tables 3 and 4 and used only those series (pink salmon catches for the west coast of Vancouver Island troll fishery and chinook salmon catches for the Strait of Georgia sport fishery) to generate retrospective forecasts. Then to judge model performance we formally compared those values for $\log _{e}\left(I_{f}\right)$ forecasted in year $t=f-1$ to those values obtained in year $t=f$, for all years 1980 and later where the degrees of freedom were greater than zero (1983 and 1984 for the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery, respectively).

Our results for both examples demonstrate that we have identified useful forecast models and that both the deterministic and stochastic components of the model contribute to model success. Not surprisingly, the reliability of a forecast increased rapidly as the degrees of freedom rose. Figures 3b and 4 b indicate that the forecasts of $\log _{e}\left(I_{f}\right)$ tended to fall within the $95 \%$ confidence interval for the forecasts starting in 1985 for the west coast of Vancouver Island fishery and starting in 1987 for the Strait of Georgia sport fishery. This indicates that as the lengths of the time series increase, the model shows statistical improvement (i.e., a narrowing of forecast confidence intervals). Another indicator of model performance is the correlation between forecasted and observed harvest rate indices. We obtained the best correlation between forecasted and observed values of $\log _{e}\left(I_{f}\right)(0.94$ for the west coast of Vancouver Island troll fishery, 0.90 for the Strait of Georgia sport fishery, Table 5) for the later years and when using the most parsimonious model identified using all available data (Tables 3, 4; Figures 3c, 4c). Regression of the observed values on forecasted values for these time periods detected neither absolute nor relative bias in the forecasts (Schnute et al. 1990). A posteriori diagnostics detected neither autocorrelation nor skewness among the model residuals.

One contrast between the model for the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery is that the success of the former model is due mainly to identification of the stochastic component of the model, whereas the success of the latter model is due mainly to the deterministic component of the model (i.e., that component related to changes in total effort, equa-

Table 5.-Correlations between forecasted, $\log _{e}\left(P_{t}\right)$, and observed, $\log _{e}\left(I_{i}\right)$, values of the natural logarithms of the harvest rate index for the west coast of Vancouver Island troll fishery and the Strait of Georgia sport fishery. For any forecast year, the open model is that chosen based on the minimum mean squared error of competing models without consideration for parsimony. The parsimonious model is that model fit to the complete data set for each fishery for which all coefficients are significant when $\alpha=$ 0.05 (Tables 3 and 4). The "stochastic" component of the model is that due to the processes represented by equations (4) and (6); the "full" model includes the deterministic component (equation 14).

| Model and forecast years | Stochastic | Full |
| :---: | :---: | :---: |
| Open model |  |  |
| 1982-1992 | 0.62 | 0.78 |
| 1985-1992 | 0.90 | 0.96 |
| Parsimonious model |  |  |
| 1983-1992 | 0.71 | 0.90 |
| 1985-1992 | 0.89 | 0.94 |
| Strait of Georgia |  |  |
| Open model |  |  |
| 1983-1991 | 0.34 | 0.63 |
| 1987-1991 | -0.10 | 0.90 |
| Parsimonious model |  |  |
| 1984-1991 | 0.26 | 0.59 |
| 1987-1991 | 0.06 | 0.90 |

tion 14). The parsimonious model of the Strait of Georgia fishery during 1987-1991 could attribute only $6 \%$ of the variance of the $\log _{e}\left(I_{t}\right)$ s to identification of the stochastic component of the model. The full model accounted for $90 \%$ (Table 5). This occurs despite the parsimonious analysis of Table 4 indicating that the time series of chinook salmon catches relates significantly to the $\log _{e}\left(I_{t}\right)$ s. In contrast, the parsimonious model of the west coast of Vancouver Island troll fishery attributed $89 \%$ of the variance of the $\log _{e}\left(I_{t}\right)$ s during 1985-1992 to the stochastic component of the model (Table 5).

## In-Season Forecasting

Equation (6) guarantees that an increase in total effort in the forecast year ( $B_{f}$ ) will cause a monotonic increase in directed effort ( $E_{f}$ ) and, therefore, also in the forecasted harvest rate index. Figure 5 demonstrates how weekly harvest rate index forecasts for the west coast of Vancouver Island troll fishery will monotonically increase with increasing total effort within the forecast season. Thus, up-to-date information on total effort expended and cumulative in-season catch will facilitate a forecast of the harvest rate index to that point in the season, notwithstanding delays in the collec-
tion and reporting of catch and effort information. An analyst could also informally use adjunct historical information on in-season effort patterns or, for our examples, salmon run-timing to anticipate total effort to some point in the season, then use that information to generate a corresponding harvest rate forecast. Decisions to control effort could be made in light of harvest policy, confidence in the forecast model and its forecasts, and the timeliness of the catch and effort information.

## Conclusions

We have shown that harvest rate indices calculated independently of knowledge of the catch and effort statistics of a fishery could have been reliably predicted and forecasted from those catch and effort statistics. Further, the potential to identify the nature of model error can contribute to a fishery manager's understanding of the fishery's dynamics and also the quality of the data used to manage the fishery. If this model structure had been used to forecast harvest rate indices for the west coast of Vancouver Island troll fishery from 1985 until 1992 and the Strait of Georgia sport fishery from 1987 to 1992, it would have performed well. The model has thus withstood the scrutiny of retrospection. Four requirements for identifying a usable model have therefore been met: a model structure was identified, the values for coefficients were estimated, diagnostics to validate assumptions were passed, and retrospective analysis to assess model utility were satisfying. These retrospective analyses did not cast doubt upon our assumption that the time series of $B_{t}$ and $C_{j, r}$ are adequate measurements of effort and catch.

The good performance of the model for our two examples would justify its continued use by a fishery manager if they were confident that the characteristics of a particular fishery and related data collection had not changed. This cannot always be guaranteed, however; indeed, the examples we use seem to indicate that implementation of the Pacific Salmon Treaty in 1985 altered certain characteristics of these fisheries. There can be other subtle influences. For example, using the model to justify controls on effort in order to achieve a certain harvest rate could result in a change in the characteristics of a fishery as fishers attempt to counter the management measures. Nevertheless, we think that the model structure introduced in this paper can be useful to managers requiring a tool to assist them in achieving a harvest rate target for one species in a multispecies fishery. Model covariates and their coefficients would have to be chosen ob-


Figure 5.-Weekly in-season harvest rate index forecasts for the west coast of Vancouver Island troll fishery in 1992. Also shown are the weekly cumulative boat-days and chinook salmon and pink salmon catches upon which the forecasts are based.
jectively and a retrospective analysis performed. A risk-averse approach to model use would see it applied to avoid unacceptably high harvest rates, given the explicit uncertainty in model forecasts. If a manager were to use a forecast to justify a control on effort, the manager would have to evaluate the future benefit of a conservative harvest rate against the consequences of an inaccurate forecast that forgoes catch.

## Acknowledgments

We thank members of the Pacific Stock Assessment Review Committee, the Pacific Salmon Commission, and anonymous reviewers for constructive criticisms of earlier versions of this model and manuscript.

## References

Anonymous. 1994. Pacific salmon commission joint chinook technical committee 1993 annual report. Pacific Salmon Commission Report Tcchinook, Vancouver.
Argue, A. W., R. Hilborn, R. M. Peterman, M. J. Staley, and C. J. Walters. 1983. The Strait of Georgia chi-
nook and coho fishery. Canadian Bulletin of Fisheries and Aquatic Sciences 211.
Box, G. E. P., and G. M. Jenkins. 1970. Time series analysis: forecasting and control. Holden Day, San Francisco.
Brannian, L. K. 1982. The estimation of daily escapement and total abundance from catch per unit effort of the sockeye salmon fishery in Togiak Bay, Alaska. Alaska Department of Fish and Game, Information Leaflet 206, Juneau.
Collicutt, L. D., and T. F. Shardlow. 1990. Strait of Georgia sport fishery creel survey. Statistics for salmon and groundfish, 1989. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2087.
Collicutt, L. D., and T. F. Shardlow. 1992. Strait of Georgia sport fishery creel survey. Statistics for salmon and groundfish, 1990. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2109.
English, K. K., W. J. Gazey, T. F. Shardlow, and M. Labelle. 1987. Development of troll fishery management models for southern British Columbia. Canadian Technical Report of Fisheries and Aquatic Sciences 1526.
Gulland, J. A. 1983. Fish stock assessment. A manual of basis methods. Wiley, Toronto.
Hipel, K. W., and A. I. McLeod. 1994. Time series modelling of water resources and environmental sys-
tems. Elsevier, Developments in Water Science 45, New York.
Peterman, R. M., and G. S. Steer. 1981. Relation between sport-fishing catchability coefficients and salmon abundance. Transactions of the American Fisheries Society 110:585-593.
Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. International Commission for the Northwest Atlantic Fisheries Research Bulletin 9:65-74.
Ricker, W. E. 1975. Computation and interpretation of biological statistics of fish populations. Fisheries Research Board of Canada Bulletin 191.
Schnute, J. T. 1994. A general framework for developing sequential fisheries models. Canadian Jour-
nal of Fisheries and Aquatic Sciences 51:16761688.

Schnute, J. T., L. J. Richards, and A. J. Cass. 1989. Fish survival and recruitment: investigations based on a size-structured model. Canadian Journal of Fisheries and Aquatic Sciences 46:743-769.
Schnute, J., T. J. Mulligan, and B. R. Kuhn. 1990. An errors-in-variables bias model with application to salmon hatchery data. Canadian Journal of Fisheries and Aquatic Sciences 47:1453-1467.
Winters, G. H., and J. P. Wheeler. 1985. Interaction between stock area, stock abundance, and catchability coefficient. Canadian Journal of Fisheries and Aquatic Sciences 42:989-998.
Zar, Z. H. 1984. Biostatistical analysis, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.


[^0]:    * Corresponding author: barry.smith@ec.gc.ca

