Methods for Postseason Evaluation of the annual Predictions of Run-size and Passage Distributions of Adult Chinook Salmon [*Oncorhynchus tshawytscha*] Returning to the Columbia and Snake Rivers

Executive Summary

The effectiveness of in-season modeling efforts compared to observations of passage and river conditions at the end of the season is analyzed. A pattern matching routine forecasts total run-size and run timing (daily passage) by optimally correlating the shape of the current year's cumulative passage (to date) with truncations of historical cumulative passage data. At the end of the season, for each stock at each observation site, we compute the Mean Absolute Deviation (MAD) for the passage distributions which is a broad measure of the average error in daily passage percentage estimates.

Introduction

Visual counts of returning adult Chinook have been made at Bonneville Dam each year since 1938. The detection of adult Chinook at Bonneville and upstream dams provides a measure of the temporal distribution of the returning adult salmonid populations.

The adult upstream "RealTime" forecaster/passage model was developed to predict the current season's adult salmon run-size at Bonneville Dam and run timing from the Bonneville Dam Tailrace to the upstream dams on the Columbia and Snake Rivers. The forecaster consists of an Escapement Forecaster (EF) that predicts the arrival timing and run-size of adult salmon at Bonneville Dam and an Adult Upstream Model (AUM) that predicts the passage timing of the fish at dams above Bonneville Dam. Each day the predictions are updated on the web.

During the 2001 migration season, Columbia Basin Research launched a prototype run timing system, EF /AUM, to predict run timing with results updated on the World Wide Web. This project was launched in an effort to provide real-time in-season projections of adult salmon migration to managers of the Columbia-Snake River hydrosystem to inform decisions about mitigation efforts such as in-river harvest. The program EF uses an empirical pattern matching routine to predict the arrival distributions for adult Chinook salmon stocks at the first detection point in the migratory route, Bonneville Dam. The AUM model takes the predictions from EF and uses hydrological, fish behavioral and dam geometry information to simulate the movement of the adult salmonid through mainstem Columbia and Snake River dams.

This report is a postseason analysis of the accuracy of the 2016 predictions from EF/AUM. Model results are compared to observations of passage and river conditions at the end of the season. We also compare key results to previous seasons.

Methods

Data

Escapement and travel time data

The fish analyzed here are adult spring and fall Chinook salmon returning to spawn in tributaries (or hatcheries) of the Columbia and Snake Rivers above Bonneville Dam. For the escapement forecasts, the daily visual counts of returning adult Chinook data come from Bonneville Dam. To assess our upstream run timing predictions, the daily visual counts come from additional detection sites at McNary and Lower Granite Dams. Data is retrieved from a link to the Columbia River DART database and provided as a courtesy by U.S. Army Corps of Engineers, NWD (http://www.nwd.usace.army.mil).

Flow and other system operations data

Any forecast of fish movement relies critically on accurate forecasts of flow, and other key system operations. The U.S. Army Corps of Engineers generates operational forecasts at all projects on the Columbia and Snake Rivers where there is fish passage. Water supply forecasts are based on a number of factors: the National Weather Service's Northwest River Forecast Center predictions, flood control requirements from the Army Corps, electrical power demand forecasts, and other criteria. The substantial uncertainty associated with springtime conditions often results in frequent and marked changes in these forecasts during April and May. Moreover, attempts to reduce the biological impacts of dissolved gas generated from high spill levels also results in a shifting of spill between projects within as well as outside the basin. Although the forecasts covered as much as 90 days into the future, it must be recognized that their intended use was in deciding operations for the next week. Forecast accuracy beyond even a few days was itself uncertain. On a monthly basis throughout the season, Bonneville Power Administration provides CBR staff with a long-term system operations forecast.

On a daily basis, forecasts for flow, spill, and elevation are replaced with observations with a query to the Columbia River DART database (CBR 2014a), which includes water quality data from the Army Corps for the majority of monitoring sites in the Columbia Basin. Subsequent fish arrival predictions are therefore based on the forecasted values for flow and spill and the latest available observed data.

Temperature data

The temperature time series is a combination of year-to-date temperature data and forecasted temperatures. The forecasts are based on observed year-to-date temperature and flow data, historical average temperature and flow profiles for 15 locations in the Snake and Columbia rivers, and the flow forecasts. Historic and observed year-to-date data was obtained from the Columbia River DART database. Temperature predictions are made by applying a three-day moving window to fit predicted temperature time series to historical average patterns of temperature change; this method is described in detail in Beer et al. (2004).

Archives of model predictions

The results of EF/AUM runs are stored on the Columbia Basin Research web site. Graphs based on the results are available through web-based query tools at http://www.cbr.washington.edu/crisprt/index_adult.html. Runs are made daily and include daily passage distribution forecasts and run-size forecasts.

Models

Initial Run size

The year's initial run-size is determined from a linear regression of each year's adult return vs. the previous year's jack return for fall Chinook and from NMFS methods at https://www.nwfsc.noaa.gov/research/divisions/fe/estuarine/oeip/g-forecast.cfm for spring Chinook. The timeframe for the spring run at Bonneville is March 15 to June 15 and the timeframe for the fall run is August 1 to November 15.

Escapement Forecaster

The Escapement Forecaster predicts the arrival timing and run-size of adult salmon at Bonneville Dam. It consists of an expected distribution based on the previous year's jack counts in the early season, and switches to a pattern matching algorithm as the season progresses. There is also a blending routine to switch smoothly between the jack-based and pattern match methods.

The arrival distribution is taken as the historic daily mean scaled to produce the correct total run-size. The pattern matching routine forecasts total run-size and run timing (daily passage) by optimally correlating the shape of the current year's cumulative passage (to date) with truncations of historical cumulative passage data. This returns the fraction of the run complete, *f*. Total run-size is then predicted by $\tilde{r} = P_c / f$ where P_c is the total passage (current year) to date.

To compare the current year's passage to that of historic runs, the cumulative current passage data is partitioned into N time intervals. The pattern matching optimization is performed as least-squares minimization; comparing slopes S_i^c over each subinterval *i* of the current run with slopes $S_i^h(f)$ of subintervals of each historic year run truncated after *f* fraction of the historic run has passed. The optimization to determine *f* is then performed as:

$$\underset{f \in (0,1)}{\text{minimize}} \sum_{h \in H} \sum_{i=1}^{N} \left(S_i^c - S_i^h(f) \right)^2$$

where *H* is the set of historical data years being used.

After the pattern matching method determines the completed fraction *f* of the current run, the passage forecast for each remaining day of the season is produced by appending the historic daily mean passage for each day of the final $1 - f$ fraction of the season, scaled to produce the correct total run-size. In this way, the forecast may be a forward or backward shift in time as compared to the historic average, thereby forecasting not just run-size, but also run timing.

Adult Upstream Model

The Adult Upstream Model (AUM) describes in detail fish movement through reaches and dams and the effects of various river operations on their migration. Essential elements are described in: https://www.monitoringresources.org/Document/Protocol/DownloadDescription/2174

For in-season forecasts, we use the projected escapement at Bonneville as input to AUM and predict the arrival timing at the upstream dams. The model contains a temperature and flow based submodel for upstream passage. River flow and temperature are modeled using portions of the COMPASS smolt passage model. Fish travel time has been calibrated using PIT-tag data of adult Chinook detected at multiple dams. The temperature and flows encountered by upstream migrating salmon are the main factors determining reach migration speed and a submodel controls this process. The flow encountered should subtract directly from the swimming speed in order to compute net up-river velocity. In each reach, the travel time distribution is determined by the migration velocity and environmental variables.

$$FishVel = FishVel_{mean} + Vel_{H_20} + Vel_{H_20}{}^2 + Temp + Temp^2 + Vel \bullet Temp$$

Example travel time parameters are shown in Table 1.

Table 1 Calibration parameters used for AUM runs in 2016. Updates are made as new data becomes available.

Stock	MedianVelUp	MeanVelUp	b0up	bVelup	bVel2up	bTempup	bTemp2up	bVelTempup	vVarup
SnakeSpr	31.575	31.432	31.432	0.347	-0.008	-1.713	0.043	0.025	160.772
UColSpr	28.699	28.341	28.341	-0.148	---	0.018	-0.003	0.013	96.411
SnakeFall	31.413	31.351	31.351	-2.381	0.015	6.451	-0.292	0.072	85.126
UColFall	26.866	26.560	26.560	1.447	0.010	-2.390	0.121	-0.084	78.143
SnakeStlhd	21.175	20.584	20.584	0.148	0.008	-0.489	0.037	-0.031	83.491
UColStlhd	22.523	22.874	22.874	-0.495	0.004	1.580	-0.069	0.011	62.013
SlowSnakeStlhd	24.088	23.918	23.918	-3.065	0.013	4.599	-0.205	0.117	64.320
SlowUColStlhd	21.759	22.418	22.418	0.109	0.008	-2.290	0.135	-0.032	60.868

Schematic of data and modeling

The relationship of the data and models is depicted in Figure 1.

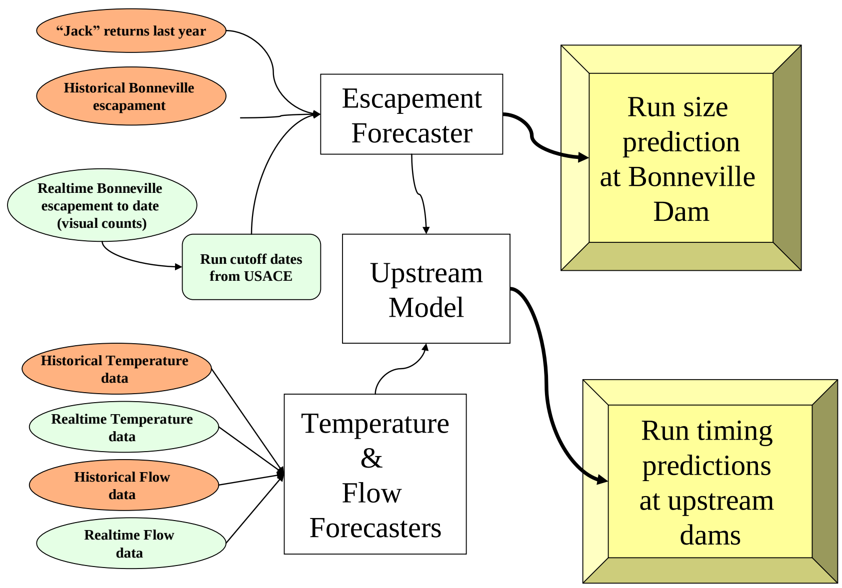

Figure 1 Schematic of data, models and products. Brown is used for historical data, green is real time up-to-date information, white boxes are modeling processes and the yellow frames are final products.

Postseason Assessment of Predictions

Mean Absolute Deviation

To assess the performance of run-size predictions, we compute the first day when the run-size estimate was within 10, 20 and 30% of the true run-size, and we determine what percent of the run had been completed on that day. Run size predictions are important for catch allocations, and compliance with federal and state regulations on fishery management. There is no established standard by which these predictions are evaluated.

To assess the performance of passage timing predictions, we apply the same measure used to assess RealTime/COMPASS predictions (Beer et al. 2008). For each stock at each observation site, we compute the Mean Absolute Deviation (MAD) for the day (*j*) on which the prediction was made. This measure is based on the average deviation between predicted and observed cumulative passage on prediction dates during the season. MAD is computed as:

$$MAD_j = \frac{1}{N}\sum_{t=1}^{N}\left|F_{Day_t} - \hat{F}_{Day_{tj}}\right| \times 100$$

where:
j = forecast day on which MAD_j is calculated;
t = index of prediction day (from 1 to N);
N = number of days on which a prediction and observation were made for the stock at the site during the season;
Day = vector of length N which identifies the days of the year from first observation of the stock at the site until two weeks past last observation (this is fixed for each site and each stock);
F = observed cumulative passage on Day_t; and
$\hat{F}$ = predicted cumulative passage on Day_t.

The MAD summation is performed over each of the dates on which model predictions were implemented – approximately every day during the season. This provides a snapshot of how well the model performs as the season progresses based on the final, "true" data. Ideally, there would be general decrease in MAD as t goes from 1 to N because the true distribution of the run should be better known and the true state of the flow and spill profiles should be known.

A second measure for run timing is the Maximum Absolute Daily Deviation (MADD)

$$MADD = \max\left\{\left|F_{Day_t} - \hat{F}_{Day_t}\right|\right. \times 100$$

All estimates of the run passage percentage are as good or better than this estimate.

Predictions of passage at Bonneville shape the forecasts of passage at other dams, so all the predictions are sensitive to these important first observations. Any errors end up affecting upstream passage predictions, and run-size predictions interact with the passage percentage predictions.

References

Beer, W.N., S. Iltis, C. Van Holmes, and J.J. Anderson. 2004. Evaluation of the 2003 Predictions of the Run-Timing of Wild Migrant Yearling Chinook and Water Quality at Multiple Locations on the Snake and Columbia Rivers using CRiSP/RealTime. Columbia Basin Research, School of Aquatic and Fishery Sciences. University of Washington, Box 358218 Seattle, WA 98195.

Beer, W.N. 2008. Stock Separation of Chinook salmon at Bonneville Dam for Adult Upstream Migration Model (2008). White Paper. Columbia Basin Research, School of Aquatic and Fishery Sciences. University of Washington, Box 358218 Seattle, WA 98195. Available March 31 at: http://www.cbr.washington.edu/papers/StockSeparationforAUM2008.html

CBR. 2013.(Columbia Basin Research). Data Access in Real Time (DART) Available on-line at http://www.cbr.washington.edu/dart/.

NMFS. 2013. (National Marine Fisheries Service) Forecast of Adult Returns for coho and Chinook salmon http://www.nwfsc.noaa.gov/research/divisions/fed/oeip/g-forecast.cfm