Umpqua River Spring Chinook Sliding Scale

Guidance for pre-season inland harvest management

January 27, 2022

Peter M. Stevens
Conservation and Recovery Program
Oregon Department of Fish and Wildlife

Chris Lorion
Conservation and Recovery Program
Oregon Department of Fish and Wildlife

Greg Huchko and Evan Leonetti
Umpqua Watershed District
Oregon Department of Fish and Wildlife

Contents

Summary

- The sliding scale relates harvest to cycles of wild spring Chinook salmon abundance, allowing more harvest when abundance is high and minimizing harvest impacts when abundance is low.
- Sliding scale bag limits for wild spring Chinook salmon are determined annually based on abundance of the North Umpqua population (Table 1); abundance is calculated as the average of the current year's pre-season forecast and the previous year's observed spawner abundance.
- Abundance thresholds for the sliding scale were selected so that each category (Table 1) occurs with an expected frequency (25% high, 60% medium, 15% low) based on a benchmark spawner abundance time series from 1986 to 2015.
- A harvest closure (i.e., "conservation closure") for wild spring Chinook salmon will be implemented <u>on the mainstem Umpqua River and North Umpqua River</u> if abundance (average of pre-season forecast and previous year's spawner abundance) of the North Umpqua population falls below the critical abundance threshold of 2,000 fish identified in the Coastal Multi-Species Conservation and Management Plan (CMP).
- In addition, a conservation closure will be implemented on the <u>mainstem Umpqua River</u> if abundance (average of pre-season forecast and previous year's spawner abundance) of the South Umpqua population falls below a critical abundance threshold of 150 spawners.
- Bag limits for <u>hatchery spring Chinook salmon</u> are not affected by the sliding scale.
- Notable changes from the CMP are: 1) identification of a critical abundance threshold for the South Umpqua population; and 2) more conservative bag limits due to increased conservation concern about these unique populations.

Table 1. Umpqua River wild spring Chinook salmon harvest sliding scale.

Sliding Scale Abundance Categories **abundance** = average of current year's forecast and previous year's spawner abundance	**Wild Spring Chinook Bag Limits**** (daily/seasonal) **February 1 – June 30**	
	Mainstem Umpqua	**North Umpqua**
If North Umpqua abundance **> 4,500** spawners, **Abundance = High**	**2/5**	**2/10**
If North Umpqua abundance **2,000 – 4,500** spawners, **Abundance = Medium**	**1/5**	**1/10**
If North Umpqua abundance **< 2,000** spawners, **Abundance = Low**	**No wild Chinook harvest allowed**	**No wild Chinook harvest allowed**
<u>In addition, for the Mainstem Umpqua only:</u> If **South Umpqua** abundance < **150** spawners, **Abundance = Low**	**No wild Chinook harvest allowed**	**Wild Chinook harvest <u>may be</u> allowed based on North Umpqua abundance**

** **Aggregate Bag Limit:** No more than 10 wild adult Chinook per year may be harvested in aggregate from the Mainstem Umpqua River and the North Umpqua River from February 1 – June 30; daily aggregate bag limit is 1 fish per day at medium abundance and 2 fish per day at high abundance.

Background

The CMP called for a "sliding scale" of harvest as a function of abundances at the population level. The sliding scale is intended to determine the allowable bag limits (both daily and seasonally) for fisheries on wild Umpqua River spring Chinook. A goal of the sliding scale is to protect individual populations from harvest when abundances are significantly depressed, and to allow greater harvest impacts when abundances are good, or significantly better than average. Another goal of the sliding scale is to eliminate unnecessary volatility in the regulations so that the angling community can more easily predict and understand the regulations.

This sliding scale builds on work done and feedback provided for the coastal fall Chinook sliding scale implemented from 2019 through present as part of the CMP. Feedback from that process indicated there was concern that the proposed sliding scale did not provide sufficient ability to react to deteriorating conditions. That concern was addressed through two mechanism which we have incorporated into the Umpqua River wild spring Chinook sliding scale presented here: (1) bag limits can change on a yearly basis, and (2) the decision criteria for setting bag limits is the average of a pre-season forecast and the previous year's escapement. Unlike the fall Chinook sliding scale, the criterion for conservation closure of wild Chinook harvest on either the mainstem or North Umpqua occurs when the pre-season forecast and the previous year's escapement for the North Umpqua River spring Chinook population falls below 2,000 fish as specified in the CMP. Also, this sliding scale incorporates an additional conservation closure criterion for the mainstem Umpqua River fishery based solely on the South Umpqua River spring Chinook population to assure sufficient escapement for that small population. South Umpqua spring Chinook returns have been very low in recent years, resulting in conservation closures for the mainstem Umpqua River and the need for a critical abundance threshold to guide decisions about the fishery. A precautionary approach is warranted given that this population is below the CMP's Minimum Equilibrium Threshold (MET = 500). The critical abundance threshold of 150 spawners identified here is 25% below median abundance for the period assessed in the CMP and well above historical lows for this population.

Coastal Multi-Species Plan (CMP)

The CMP was approved by the Oregon Fish and Wildlife Commission on June 6, 2014. Page 61 of the Plan describes "Sliding Scale Management." The intent is to acknowledge cycling abundances, anticipate them with a forecast, and allow greater harvest when expected returns are high.

Sliding scale bag limits identified in the CMP are modified here for two reasons. First, the expected frequency of North Umpqua abundance falling below the critical abundance threshold of 2,000 fish is similar to the expected frequency of the low abundance category. Therefore, at low abundance a conservation closure will be implemented instead of the 1/1 bag limit identified in the plan. The second reason for changes in bag limits is increased conservation concern for these unique populations. Reduced bag limits for the medium and high abundance categories relative to the plan are intended to minimize risk from harvest while maintaining opportunity and simplifying fishing regulations.

Abundance Summaries

Abundance estimates for North and South Umpqua River spring Chinook populations were generated from adjusted passage counts at Winchester Dam and resting pool counts in the South Umpqua River, respectively. The North Umpqua counts were adjusted for wild fish retained for hatchery broodstock and harvest above Winchester Dam. Appropriate adjustments were also made for changes in counting methods (e.g. - shift from in-person to video review). An estimate of escapement is generated based on these adjusted numbers. This estimated escapement is used in generating the sliding scale and forecast models. The South Umpqua population counts are adjusted for origin (i.e. – hatchery v. "natural") and then resting pool counts are expanded to reflect estimated abundance throughout the entire basin. The methods for these estimates were standardized as part of the CMP process and the same numbers are used here. As noted in the CMP, estimates may be refined in the future as sampling methods or protocols are improved over time. If estimation methods change, ODFW will make appropriate adjustments to maintain valid comparisons with the historical baseline and will describe the changes in annually updated CMP Wild Fish Monitoring Summaries.

Both populations of Umpqua River spring Chinook exhibit mild autocorrelation (Figure 1). A premise of the sliding scale is that natural "smoothness" in abundance time series can be used to ensure that bag limits do not vary drastically year to year. Autocorrelation can be leveraged in forecasts.

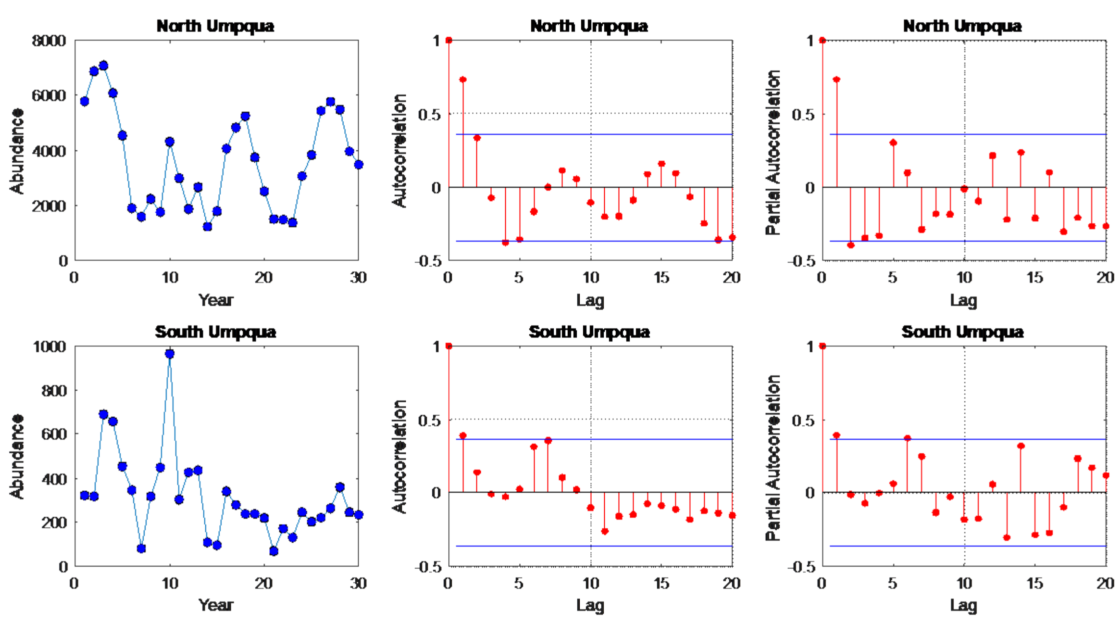

Figure 1. Time series of adult spawners of Umpqua River spring Chinook exhibit minimal autocorrelation. The middle panels show significant positive autocorrelation at lag 1, 2 and 3 (above blue line) and lag 1 and 2 for North and South Umpqua spring Chinook, respectively. Both occur due to mild cycling. The panel on the right shows residual autocorrelation after removing the expected autocorrelation for shorter lags.

Methods

1. The North Umpqua population's time series (Figure 2a) of abundance is used to compute the mean and standard deviation of a lognormal distribution (Figure 2b).
2. Each estimate of abundance can then be re-expressed as a percentile of the lognormal distribution (Figure 2c).
3. A two-year running average is taken for the North Umpqua spring Chinook population (Figure 3, *). The two-year running average represents a pseudo decision metric in this example. In practice, a forecast for the focal year will be substituted for the empirical abundance used here.
4. The North Umpqua has a distribution of decision metrics (*). These metrics are bounded between zero and one, and thus modeled with a beta distribution. Beta distributions are fitted to the series of decision metrics (Figure 4). Using the parameterized beta distribution, it is possible to back calculate threshold values of the decision metric that produce the probabilities of occurrence (Figure 4, colors) specified in the CMP.

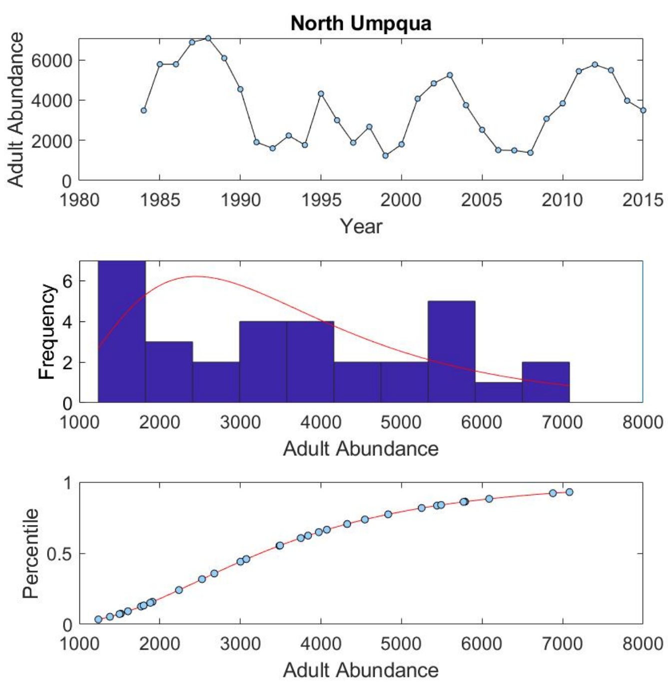

Figure 2. A population's time series of abundance (a), can be replotted into a histogram (b, grey). A lognormal distribution can be fitted these data (b, red), and all data points mapped into percentiles of the lognormal distribution (c).

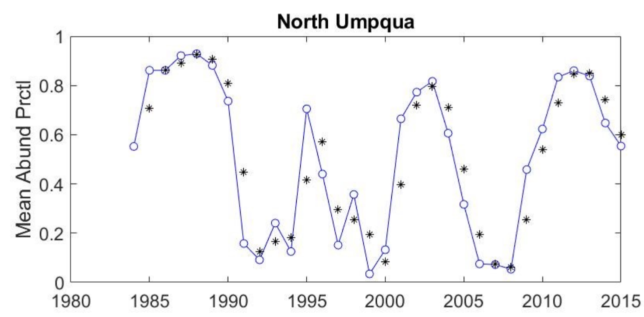

Figure 3. The percentiles of North Umpqua River spring Chinook populations (blue). A two-year running average yields a pseudo decision metric (*). In practice, the actual decision metric will use a pre-season forecast instead of the empirical estimates used here.

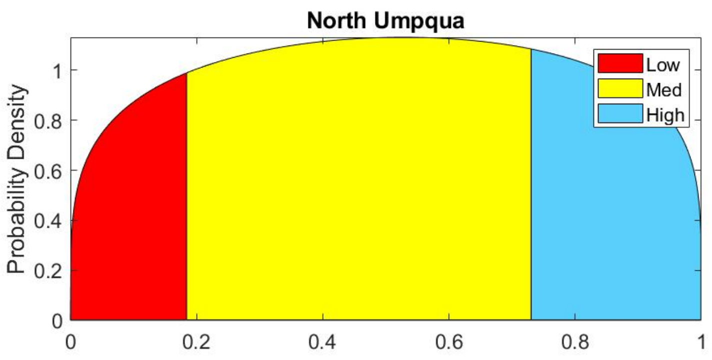

Figure 4. Beta distributions fitted to each stratum's pseudo decision metrics (see Figure 5). Colors indicate bag limit categories. Breaks between colors/categories are intentionally chosen to reflect probabilities of occurrence. In the North Umpqua and Umpqua River, low, medium and high bag limits occur with probabilities 0.15, 0.65, and 0.25, respectively.

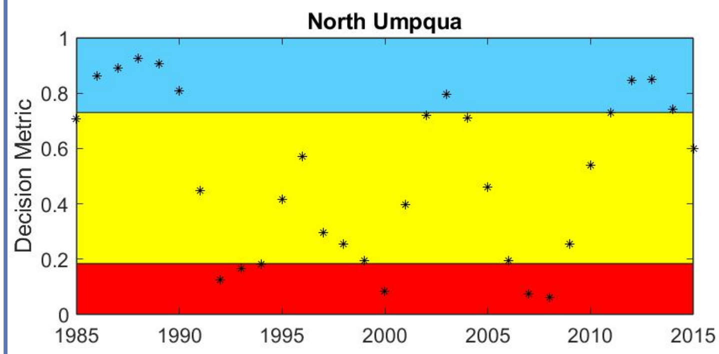

Figure 5. Time series of pseudo decision metrics (*) relative to the bag limit categories (colors). These are pseudo decision metrics because they average the previous and current abundances. In practice, decision metrics are the average of the previous abundance and a preseason forecast. Autocorrelation in the decision metrics prevents large annual changes in bag limits.

Forecasting

Forecasting is a critical element of the sliding scale. All forecasting techniques used leverage the explanatory power of the previous year's escapement (Figure 1). Thus including the previous year's escapement in the decision metric (mean of forecast and the previous year's escapement) is not intended to improve the 'accuracy' of the decision metric. Rather, it is intended to reduce the year-to-year variability in the bag limit increasing predictability and stability for anglers. The best estimate of the current year's escapement is the forecast itself.

Forecasting Techniques

Four "classes" of forecasting models have been included in the present ensemble model:

Autoregressive Integrated Moving Average with Covariates (ARIMAX): A classic forecasting technique for time-series that leverages the inherent autocorrelation in the data series and external covariates. AIC_c was used to define optimal lags, trends, and window sizes for each population separately. These lags, trends and window sizes were then held constant across the suite of models for each population.

Sibling Regression: A well-established salmon forecasting technique which uses regression relationships to predict abundance-at-age ($N_{t,a}$) based on the previous year's abundance of back-shifted ages ($N_{t-1,a-1}$), which are earlier returning siblings thereby capturing cohort effects. The earliest age class must be estimated using either a naïve model or a time-series model (i.e. – ARIMA or exponential smoothing). AIC is used to select which model best represents the data. Also, this model requires age compositions which, in this case, had to be imputed for each population based on a limited number of sample years. Proportions for each age-class at each year were drawn from a normal approximation to the binomial distribution of the mean frequency of each age-class across all sampled years. These proportions were then standardized to sum to one within a given year. Standardized proportions were then multiplied by the total escapement for a given year to estimate the number of returners in that age class that year.

Generalized linear models (GLM): These simple extensions of classical regression are easy to use, easy to understand and leverage the power of likelihood and information theory to prevent overfitting. However, autocorrelation or cross-correlation must be incorporated as a covariate to be included in the model. See the discussion of "quasi-covariates" in the Covariates section below for details.

Nonlinear Autoregressive Artificial Neural Network with Covariates (NARX) A machine learning technique inspired by the human brain designed to capture extremely complex relationships within and between the abundance time series and predictors. However, this technique requires a lot of observations for accuracy. Also, users must pre-specify network architecture and rules for preventing overfitting. Finally, as with all neural networks, the fitting process itself is considered a "black box."

All models were implemented in R. Sibling Regression was implemented using the ForecastR package. Other models are not currently available in ForecastR and were implemented using purpose-built scripts.

Covariates.

Covariates can improve fit and forecasting accuracy of a model. However, too many covariates or covariates with spurious correlations can lead to overfitting and ultimately reduce forecasting accuracy. Moreover, examining multiple potential relationships from many possible covariates (i.e. – "data dredging") also creates the possibility of developing spurious correlations. As such, only a selected suite of covariates with established relationships to run size and/or with an *a priori* biological rationale for why they might impact population abundance were considered for inclusion in the forecasting models. These included Pacific Decadal Oscillation (PDO), three candidate indices of upwelling (Logerwell's upwelling index, CUTI and BEUTI), minimum summer flow and maximum winter flow. However, correlation between nearly all these covariates and population abundance tended to be insignificant and/or highly unstable across all plausible temporal back-shifts (i.e. – 1 - 5 years). Only minimum summer flow demonstrated a significant, stable relationship across both populations and was therefore included in forecasting models.

In the GLM and NARX models only, a one year offset of population abundance (ylag) and a mean z-score of population abundance across both populations, also offset by one year, were included as quasi-covariates in the model structure. Neither model inherently incorporates autocorrelation or cross-correlation so these quasi-covariates were included to incorporate the auto- and cross-correlation between and among populations. Also, the mean z-score allowed the populations to "share" data with the intent of improving overall quality and reliability in years where one abundance estimate or the other may have been more reliable. This is particularly important in the case of the South Umpqua population where low and highly variable abundances can create uncertainty.

Performance.

Leave-one-out cross-validation (LOOC) was used to assess the performance and fit of the models to data while reducing the likelihood of overfitting. Overfitting occurs when a model fits the available data too well. This produces impressive diagnostics but reduces predictive accuracy as the model does not generalize. Modeling techniques that produce likelihoods provide a method to control for over-fitting. However, not all models considered here produce likelihoods. LOOC presents an effective, alternative method for controlling for overfitting, comparing models and combining multiple models (see Ensemble Averaging below).

LOOC is implemented by withholding the most recent empirical observation from the model fitting process. The fitted model is then used to predict the withheld observation and the absolute percent error, $\left|\frac{x-\hat{x}}{x}\right|$, is calculated. The second most recent empirical observation is then withheld and all preceding observations are used to predict it. This was process was repeated for each of the five most recent empirical observations (i.e. – 2017 - 2021). The median of these five absolute percent errors (median APE) could then be calculated to assess model performance and weight model outputs for inclusion in the ensemble (See tables below for performance diagnostics and ensemble weighting). Model performance was also assessed against a simple historical average using the same techniques. This demonstrates any improvement of more sophisticated techniques over a naïve approach.

Ensemble Averaging

Each model "class" has strengths and weakness relative to other models and the relative performance therefore varies across years. Rather than selecting a "best" model, I used an ensemble modeling approach where each model's estimate was weighted by the inverse of the median APE of that model calculated as described above.

Performance metrics for each forecast technique used in the ensemble forecast are presented in the tables below. As stated above, these values are computed from a leave-one-out cross-validation (LOOC) trials. They reflect the accuracy of the forecast techniques as compared to actual estimated returns from 2017 through 2021.

Table 1. Proposed models used in the ensemble forecast

ID	MODEL
1	N. Umpqua – ARIMA (2,0,2); S. Umpqua – ARIMA (0,1,0)
2	N. Umpqua – ARIMAX (2,0,2), MinFlow; S. Umpqua – ARIMAX (0,1,0), MinFlow
3	Sibling Regression
4	GLM Poisson, y ~ pop + ylag + Zscore
5	GLM Poisson, y ~ pop + ylag + Zscore + MinFlow
6	Artificial Neural Network, 8Neurons, Bayesian regularization, Autoregressive, covs= Zscore, Age
7	Artificial Neural Network, 8Neurons, Bayesian regularization, Autoregressive, covs= ylag, Zscore, MinFlow

Table B2. 2022 Forecasts for individual models (see Table B1), associated weights, and weighted ensemble prediction.

	1	2	3	4	5	6	7	Weighted Ensemble
North Umpqua	3156	3256	1667	2444	2192	2491	2264	2437
South Umpqua	146	112	49	282	246	188	126	122
Median % error - North	0.29	0.80	0.53	0.14	0.14	0.33	0.30	
Median % error - South	0.44	0.05	0.62	1.81	1.48	0.91	0.84	
Weight - North	0.130	0.046	0.070	0.259	0.261	0.113	0.122	
Weight - South	0.085	0.721	0.061	0.021	0.025	0.041	0.045	

Model 1: North Umpqua – ARIMA (2,0,2) South Umpqua – ARIMA (0,1,0)

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN (2,0,2)	0.1763	1.46	0.16	0.29	0.76	0.2859
UmpS (0,1,0)	0.1435	8.00	0.44	0.60	0.67	0.5981
	median=	0.52	mean=	1.27		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	0	0	0.6
UmpS	1	1	1	1	1	1
			mean=	0.80		

Model 2: North Umpqua – ARIMAX (2,0,2), MinFlow South Umpqua – ARIMAX (0,1,0), MinFlow

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN (2,0,2)	0.16	1.58	0.11	0.29	0.80	0.2935
UmpS (0,1,0)	0.15	8.05	0.45	0.59	0.72	0.5929
	median=	0.52	mean=	1.29		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	0	0	0.6
UmpS	1	1	1	1	1	1
			mean=	0.80		

Model 3: Sibling Regression

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN	0.77	0.36	0.56	0.53	0.22	0.532
UmpS	0.09	3.86	0.62	0.89	0.59	0.615
	median=	0.58	mean=	0.85		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	0	1	0	0	1	0.4
UmpS	1	1	1	1	1	1
			mean=	0.70		

Model 4: GLM Poisson, y ~ pop + ylag + z-score

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN	0.01	0.46	0.15	0.03	0.14	0.1434
UmpS	0.48	11.51	6.30	1.81	3.42	3.4166
	median	0.47	mean=	2.43		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	1	1	1
UmpS	0	1	0	0	1	0.4
			mean=	0.70		

Model 5: GLM Poisson, y ~ pop + ylag + z-score + minSumFlow

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN	0.03	0.38	0.01	0.14	0.17	0.1422
UmpS	0.43	10.33	5.21	1.48	3.38	3.3759
	median=	0.40	mean=	2.15		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	0	1	0.8
UmpS	1	1	1	1	1	1
			mean=	0.90		

Model 6: Artificial Neural Network, Autoregressive, 8Neurons, Bayesian Regularization, z-score, age

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN	0.12	0.95	0.17	0.33	0.68	0.3281
UmpS	0.26	8.47	2.07	0.35	2.47	2.0705
	median=	0.52	mean=	1.59		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	0	0	0.6
UmpS	1	1	1	1	1	1
			mean=	0.80		

Model 7: Artificial Neural Network, Autoregressive, 8Neurons, Bayesian Regularization, z-score, ylag, minSumFlow

Forecast Percent Error

	2017	2018	2019	2020	2021	Median
UmpN	0.20	0.90	0.05	0.30	0.39	0.3041
UmpS	0.32	10.87	3.63	0.82	2.32	2.3193
	median=	0.61	mean=	1.98		

Closer than historical average? 1=yes, 0=no.

	2017	2018	2019	2020	2021	Mean
UmpN	1	1	1	0	1	0.8
UmpS	1	1	1	1	1	1
			mean=	0.90		

Appendix A – Raw Data

Table A1. Abundance estimates for North Umpqua and South Umpqua spring Chinook. Estimates for the North Umpqua are based on adjusted counts at Winchester Dam. Estimates for South Umpqua are based on adjusted and expanded resting pool counts from the South Umpqua River.

Year	North Umpqua	South Umpqua
1986	5786	322
1987	6879	318
1988	7084	690
1989	6087	658
1990	4545	454
1991	1908	346
1992	1604	82
1993	2237	318
1994	1766	449
1995	4325	965
1996	3000	305
1997	1881	428
1998	2676	435
1999	1237	109
2000	1799	96
2001	4073	340
2002	4834	279
2003	5249	239
2004	3749	239
2005	2524	220
2006	1514	69
2007	1499	172
2008	1381	132
2009	3074	246
2010	3839	202
2011	5440	221
2012	5769	264
2013	5491	360
2014	3973	245
2015	3492	234
2016	2570	247
2017	2942	216
2018	1374	24
2019	2298	43
2020	3563	107
2021	2100*	116

* Based on preliminary harvest estimates