Pacific Northwest Hatcheries Smolt-to-Adult Ratio (SAR) Estimation using Coded Wire Tags (CWT) Data

Prepared for:

US Department of Energy
Bonneville Power Administration
P.O. Box 3621
Portland, OR 97232

Prepared by:

John R. Skalski
Richard L. Townsend

Columbia Basin Research
School of Aquatic and Fishery Sciences
University of Washington
1325 Fourth Avenue, Suite 1820
Seattle, Washington 98101-2509

Project No. 1991-051-00
Contract No. 00013690

8 July 2005

Table of Contents

Table of Figures

Table of Tables

1.0 Introduction

This project produces an annual analysis of coded-wire-tag (CWT) returns from a list of hatcheries across the Northwest region (WA, OR and ID) for all available years, prompted by the need to research, monitor, and evaluate smolt-to-adult ratio (SAR) pursuant to the Federal Columbia River Power System (FCRPS) 2000 Biological Opinion (BO) Reasonable and Prudent Alternative (RPA). Smolt-to-adult ratios are based on all available historical CWT data collected since mid-70s. The CWT release and recovery data used in this report were obtained from the Regional Mark Processing Center, managed by the Pacific States Marine Fisheries Commission (PSMFC). Smolt-to-adult ratios from throughout the Columbia Basin, Puget Sound, and outer Washington and Oregon coasts are analyzed to provide a wide geographic representation. The list of hatcheries is provided in Table 1.

1.1 Hatcheries

In total, 91 hatcheries in Washington, Oregon, and Idaho, covering a variety of species, were selected for analysis (Table 1). The color-coded map (Figure 1) shows the geographic distribution of the hatcheries by watershed, based on hydrologic units, in the three states.

Figure 1. Color-coded watershed map, based on hydrologic unit, showing the locations of the hatcheries (some dots overlap at this scale).

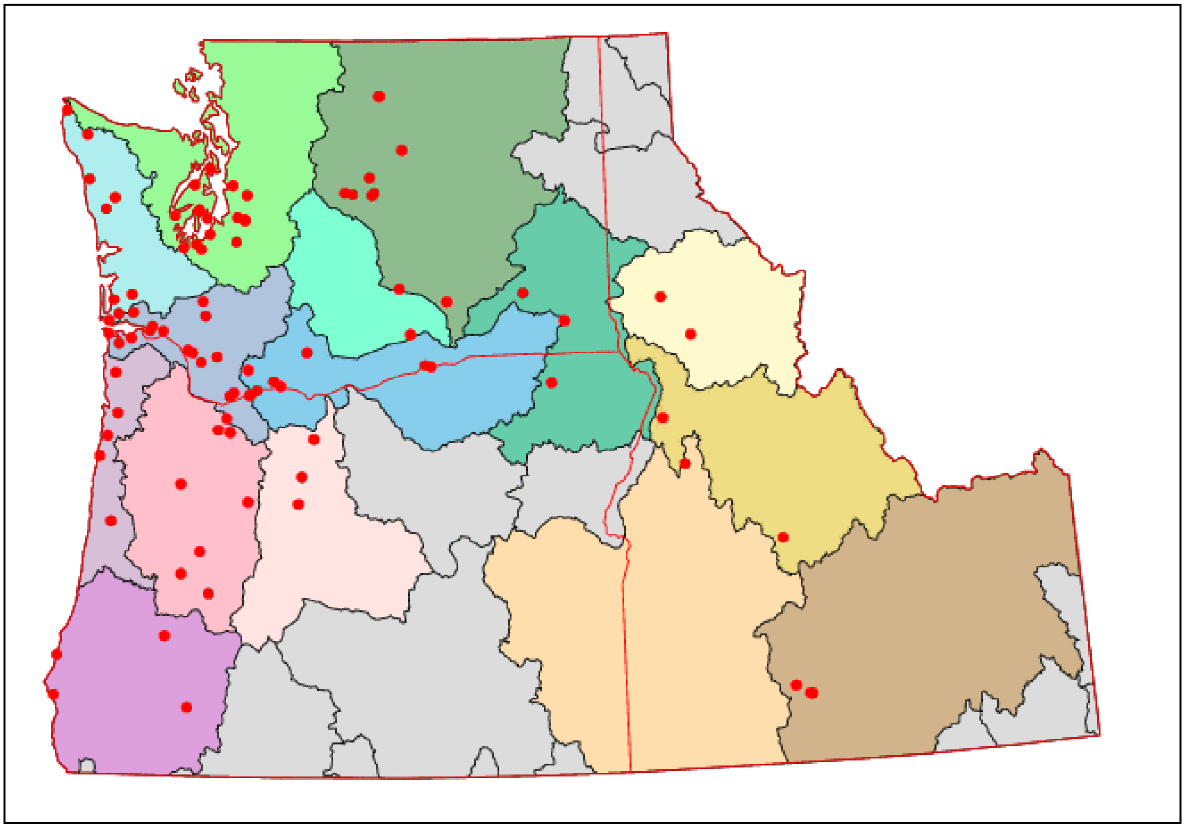

Table 1. A list of the hatchery release sites and species used in the CWT SAR analyses. Listed under each species is the run type. “X” for Coho and Sockeye. An “ND*” means no designation for run type—i.e. there wasn’t a run type recorded in the CWT database.

Hatchery	Chinook	Coho	Sockeye	Steelhead
Abernathy SCDC Hatchery	spring, fall			
Bandon Hatchery	spring, fall, late fall	X		
Beaver Creek Hatchery	fall	X		winter
Big Beef Creek Hatchery	fall			
Big Creek Hatchery	fall, late fall, none			
Bonneville Hatchery	spring, fall, late fall		X	
Capitol Lake Rearing Pond	spring, fall, hybrid			
Carson National Fish Hatchery	spring	X		
Cascade Hatchery		X		
Cedar Creek Hatchery	spring, fall	X		
Clackamas Hatchery	spring			
Cole Rivers Hatchery	spring, fall, late fall	X		
Cowlitz Salmon Hatchery	spring, fall	X		
Crisp Creek Rearing Pond	fall	X		
Dexter Pond	spring, late fall			
Dryden Ponds	summer			
Dworshak National Fish Hatchery	spring	X		summer, ND*
Eagle Creek National Fish Hatchery	spring	X		winter, ND*
Eastbank Hatchery	summer			
Elk River Hatchery	fall			
Elochoman Hatchery	fall	X		
Entiat National Fish Hatchery	spring			
Fall Creek Hatchery	fall	X		
Fallert Creek Hatchery	spring, fall	X		
Forks Creek Hatchery	fall	X		
Fox Island Net Pens	fall	X		
Garrison Hatchery	fall	X		
Grays River Hatchery	spring, fall	X		
Grovers Creek Hatchery	fall			winter
Hagerman National Fish Hatchery	spring, fall			summer
Hood Canal Marina Net Pens	fall			
Hupp Springs Rearing Pond	spring, fall			

Table 1: (continued)

Hatchery	Chinook	Coho	Sockeye	Steelhead
Irrigon Hatchery	spring, fall, late fall	X		summer, late fall
Issaquah Hatchery	fall, hybrid			
Kalama Creek Hatchery	fall	X		
Kalama Falls Hatchery	spring, summer, fall	X		
Keta Creek Hatchery	fall	X		
Klaskanine Hatchery	fall, late fall	X		
Klickitat Hatchery	spring, summer, fall	X		
Kooskia National Fish Hatchery	spring			
Leavenworth Hatchery	spring			summer
Lewis River Hatchery	spring, fall	X		
Lookingglass Hatchery	spring			
Lyons Ferry Hatchery	spring, fall, URB late fall			summer, ND*
Magic Valley Hatchery				spring, summer, ND*
Makah National Fish Hatchery	fall	X		winter
Marion Forks Hatchery	spring, late fall, ND*			
McAllister Hatchery	fall			
McCall Hatchery	summer, ND*			
McKenzie Hatchery	spring, late fall	X		
Methow Hatchery	spring, summer			
Minter Hatchery	spring, fall	X		
Naselle Hatchery	fall	X		
Nehalem Hatchery	fall	X		
Nemah Hatchery	fall	X		
Niagara Springs Hatchery				spring, summer
North Toutle Hatchery	spring, fall	X		
Oak Springs Hatchery				summer, winter, late fall
Oxbow Hatchery	spring, fall, late fall	X		
Portage Bay Hatchery	fall	X		
Priest Rapids Hatchery	spring, fall		X	
Prosser Hatchery	late fall			
Quinault Hatchery/Net Pens	summer, fall	X	X	winter
Quinault National Fish Hatchery	fall	X		winter
Rapid River Hatchery	spring, ND*			
Ringold Springs Hatchery	spring, fall			summer
Rock Creek Hatchery	spring, fall, late fall	X		

Table 1: (continued)

Hatchery	Chinook	Coho	Sockeye	Steelhead
Round Butte Hatchery	spring, fall, late fall			
Salmon River Hatchery	spring, fall	X		
Sandy Hatchery	spring	X		
Sawtooth Hatchery	spring, ND*			
Sea Resources Hatchery	fall			
Shale Creek Hatchery		X		
Skamania Hatchery				summer, winter
Solduc Hatchery	spring, summer, fall, hybrid	X		
Soos Creek Hatchery	spring, fall, hybrid	X		
Speelyai Hatchery	spring, fall	X		
Spring Creek National Fish Hatchery	fall, hybrid			
Stayton Pond	fall, late fall			
Trask River Hatchery	spring, fall, late fall, winter	X		
Tucannon Hatchery	spring			summer, ND*
Turtle Rock Hatchery	summer, fall	X		summer
Umatilla Hatchery	spring, fall, late fall			summer, late fall
Vanderveldt Ponds	fall	X		
Voights Creek Hatchery	fall	X		
Warm Springs National Fish Hatchery	spring			summer
Washougal Hatchery	fall	X		
Wells Hatchery	summer			summer
Willamette Hatchery	spring	X		
Willard National Fish Hatchery	spring, fall, late fall	X		
Winthrop National Fish Hatchery	spring, summer	X		summer

2.0 Statistical Methods

The following section describes the statistical methods used to estimate the smolt-to-adult ratios and their associated variances.

2.1 SAR Estimators

Define the estimator of the smolt-to-adult ratio (SAR) as follows:

$$\widehat{SAR} = \frac{\frac{1}{\left(1-\hat{L}\right)} \sum_{i=1}^{Y} \sum_{t=1}^{T} \sum_{a=1}^{A} \sum_{f=1}^{F} \frac{x_{tafi}}{r_{tafi}}}{R} \quad (1)$$

where

$\hat{L}$ = estimated fractional tag-loss rate post-release;

R = release size;

x_{tafi} = number of CWT recovered in the *a*th area ($a = 1, \ldots, A$) during the *t*th time period ($t = 1, \ldots, T$) and *f*th fishery ($f = 1, \ldots, F$) in the *i*th year ($i = 1, \ldots, Y$) of returns;

r_{tafi} = sampling fraction in the *a*th area ($a = 1, \ldots, A$) during the *t*th time period ($t = 1, \ldots, T$) and *f*th fishery ($f = 1, \ldots, F$) in the *i*th year ($i = 1, \ldots, Y$) of returns.

Note $\sum_{i=1}^{Y} \sum_{t=1}^{T} \sum_{a=1}^{A} \sum_{f=1}^{F} \frac{x_{tafi}}{r_{tafi}} = \hat{N}$ is an estimate of the total CWTs recovered for a release group, taking into account the sub-sampling process of examining fish for tags.

Hence, the estimate of SAR can be rewritten as

$$\widehat{SAR} = \frac{1}{\left(1-\hat{L}\right)} \left(\frac{\hat{N}}{R} \right). \quad (2)$$

In practice, data from extended-holding studies to estimate tag loss are unavailable. In which case, it must be assumed post-release tag loss is negligible (i.e., $L = 0$) or constant across all released compared, such that

$$\widehat{SAR} \propto \frac{\hat{N}}{R}. \tag{3}$$

The reporting format for RMIS where the CWT-data are stored presents each individual recovered fish as an expanded count. Designations concerning area, time, or method of capture are omitted. Consistent with this reporting format in RMIS, the $\widehat{SAR}$ for a tag-release group can be reexpressed as

$$\widehat{SAR} = \frac{\sum_{j=1}^{C} \frac{1}{r_j}}{R} \tag{4}$$

where

C = total number of fish recovered from the release of size R,

r_j = sampling fraction associated with the recovery of the *j*th fish caught ($j = 1, \ldots, C$).

2.2 Calculating the Variance of $\widehat{SAR}$

The variance of $\widehat{SAR}$ can be derived, taking the variance in stages, where

$$Var\left(\frac{\hat{N}}{R}\right) = Var_2\left[E_1\left(\frac{\hat{N}}{R}\middle|2\right)\right] + E_2\left[Var_1\left(\frac{\hat{N}}{R}\middle|2\right)\right] \tag{5}$$

where 2 = binomial sampling of N of R,

1 = estimation of $\hat{N}$. Then

$$Var\left(\frac{\hat{N}}{R}\right) = Var_2\left[\frac{N}{R}\right] + E_2\left[\frac{1}{R^2}Var\left(\hat{N}\right)\right]$$

$$Var\left(\frac{\hat{N}}{R}\right) = \frac{\frac{N}{R}\left(1 - \frac{N}{R}\right)}{R} + \frac{Var\left(\hat{N}\right)}{R^2}.$$

Note:

$$
\begin{aligned}
Var\left(\hat{N}\right) &= Var\left(\sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\frac{x_{tafi}}{r_{tafi}}\right) \\
&= \sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F} Var\left(\frac{X_{tafi}}{r_{tafi}}\right) \\
&= \sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{X_{tafi}r_{tafi}\left(1-r_{tafi}\right)}{\left(r_{tafi}\right)^2}\right] \\
Var\left(\hat{N}\right) &= \sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{X_{tafi}\left(1-r_{tafi}\right)}{r_{tafi}}\right]
\end{aligned}
$$

where X_{tafi} = actual number of fish returning in t, a, f, i. This variance can be estimated by

$$
\widehat{Var}\left(\hat{N}\right) = \sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{x_{tafi}\left(1-r_{tafi}\right)}{r_{tafi}^2}\right].
$$

Therefore,

$$
\widehat{Var}\left(\widehat{SAR}\middle| SAR\right) = \frac{\frac{\hat{N}}{R}\left(1-\frac{\hat{N}}{R}\right)}{R} + \frac{\sum_{i=1}^{Y}\sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{x_{tafi}\left(1-r_{tafi}\right)}{r_{tafi}^2}\right]}{R^2}. \tag{6}
$$

Variance estimator (6) can be reexpressed in terms of the total catch C as follows

$$
\widehat{Var}\left(\widehat{SAR}\middle| SAR\right) = \frac{\frac{\hat{N}}{R}\left(1-\frac{\hat{N}}{R}\right)}{R} + \frac{\sum_{j=1}^{C}\frac{\left(1-r_j\right)}{r_j^2}}{R^2} \tag{7}
$$

when analyzing data from RMIS.

2.3 Results Across Replicate CWT Release Groups

In any one year at a hatchery, multiple CWT groups may be released. Using these replicate release groups, an overall estimate of SAR can be calculated. The general form of the overall estimate is a weighted average of the form

$$\widehat{\widehat{SAR}} = \frac{\sum_{k=1}^{K} W_k \cdot \widehat{SAR}_k}{\sum_{k=1}^{K} W_k} \tag{8}$$

with variance estimator

$$\widehat{VAR}\left(\widehat{\widehat{SAR}}\right) = \frac{\sum_{k=1}^{K} W_k \left(\widehat{SAR}_k - \widehat{\widehat{SAR}}\right)^2}{(K-1)\sum_{k=1}^{K} W_k} \tag{9}$$

and where K = number of replicate CWT releases.

There are several choices for the values of the weights (W_k; $k = 1,\ldots,K$). If the weights are set equal to the release sizes (i.e., $W_k = R_k$; $k = 1,\ldots,K$), then estimator (8) is simply a pooled estimate of the form

$$\widehat{\widehat{SAR}} = \frac{\sum_{k=1}^{K} \hat{N}_k}{\sum_{k=1}^{K} R_k} \tag{10}$$

with variance estimator

$$\widehat{VAR}\left(\widehat{\widehat{SAR}}\right) = \frac{\sum_{k=1}^{K} R_k \left(\widehat{SAR}_k - \widehat{\widehat{SAR}}\right)^2}{(K-1)\sum_{k=1}^{K} R_k}. \tag{11}$$

The pooled estimator (10) would be a reasonable estimator if all replicate release groups are experiencing a common SAR. Furthermore,

$$\frac{1}{Var\left(\hat{N} \middle| R_k\right)} \propto R_k ,$$

making R_k a good candidate for a weight.

An alternative weight is to weight inversely proportional to the $Var\left(\widehat{SAR}\right)$ where

$$Var\left(\widehat{SAR}\right) = \sigma_{SAR}^2 + Var\left(\widehat{SAR}_k \middle| SAR_k\right) \quad (12)$$

where

σ_{SAR}^2 = random variation in SAR between replicate releases.

From the replicate CWT releases, an estimate of σ_{SAR}^2 can be calculated as

$$\hat{\sigma}_{SAR}^2 = s_{\widehat{SAR}}^2 - \frac{\sum_{k=1}^{K} \widehat{Var}\left(\widehat{SAR} \middle| SAR\right)}{K} \quad (13)$$

for $\hat{\sigma}_{SAR}^2 \geq 0$. Should Equation (13) estimate a negative variance components for σ_{SAR}^2, then $\hat{\sigma}_{\widehat{SAR}}^2$ should be set to zero.

In this report the more intuitive estimator for the $\widehat{\widehat{SAR}}$ will be used, the pooled estimate using Equations (10) and (11).

2.4 Interval Estimation

An asymptotic (1 - α) 100% confidence interval for the weighted average of the SAR estimates was computed by

$$\widehat{\widehat{SAR}} \pm Z_{1-\frac{\alpha}{2}} \sqrt{Var\left(\widehat{\widehat{SAR}}\right)}. \quad (14)$$

For a 95% confidence interval, $Z_{0.975}$ = 1.96.

2.5 Variance for Annual Return Numbers

Define the estimate of total adult returns for the *i*th year of returns as

$$\hat{N}_i = \sum_{t=1}^{T} \sum_{a=1}^{A} \sum_{f=1}^{F} \frac{x_{tafi}}{r_{tafi}} \quad (15)$$

where the SAR is correspondingly defined as

$$\widehat{SAR} = \frac{\sum_{i=1}^{Y} \hat{N}_i}{R}.$$

The variance of $\hat{N}_i$ can be expressed by

$$Var\left(\hat{N}_i\right) = \sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{X_{tafi} r_{tafi}\left(1-r_{tafi}\right)}{r_{tafi}^2}\right]$$
$$= \sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{X_{taf}\, i\left(1-r_{tafi}\right)}{r_{tafi}}\right]$$

and estimated by

$$\widehat{Var}\left(\hat{N}_i\right) = \sum_{t=1}^{T}\sum_{a=1}^{A}\sum_{f=1}^{F}\left[\frac{x_{tafi}\left(1-r_{tafi}\right)}{r_{tafi}^2}\right]$$

or equivalently

$$Var\left(\hat{N}_i\right) = \sum_{j=1}^{C_i}\frac{\left(1-r_{ij}\right)}{r_{ij}^2} \qquad (16)$$

where

C_i = total catch return for the *i*th year from a release of size R,

r_{ij} = sampling fraction for the *j*th fish caught ($j=1,\ldots,C_i$) in the *i*th year of returns from a release of size R.

2.6 Exceptions to the statistical methodology

Occasionally, some calculations resulted in an obviously erroneous SAR estimate for a tag group. This was due to one of two circumstances:

1. The estimated number of recovered tags from a particular tag group was greater than the number released.
2. The expanded count reported by RMIS for a recovered tag was less than 1.

The first error occurred when the release size for a tag group was very small (i.e. 1 or 2 fish), and 100% of the release was recovered. Because recovery numbers are expanded to compensate for the fact that only a fraction of a particular catch is sampled, these recovered tags can be expanded to be greater than the number actually released. For example, 1 fish is released under a tag code and recovered in a survey that looked at 25% of a total catch. This 1 fish is expanded to represent a total of 4 fish that might have been present and recovered, had the sampling process looked at 100% of the fish in the catch. Taking the numbers at face-value, this would give a SAR of 400%, along with an erroneous negative variance estimate. To correct this, the total number recovered from a particular release is limited to the total number released, resulting in a SAR equal to 100%. This is a rare occurance, and due to the low release size, would not contribute much to a weighted average for a particular release year and release site.

The second error is likely due to a typo. Individual recovered fish are presented as an expanded count to adjust for the fraction sampled in a particular survey. Valid sampling fractions *(s.f.)* range from zero to 100%, and the recovery numbers are expanded by multiplying the actual catch by $\frac{1}{s.f.}$. Valid ranges of expanded recoveries therefore range from 1 fish to the total number released (previous paragraph). A reported expanded recovery count of less than 1 suggests that over 100% of a catch was sampled, which is impossible. As the true value of the expanded count cannot be discerned without going back to the source of the recovery report, and appear to be quite rare, these fish are discarded from the analysis.

3.0 Discussion

The historical CWT data used in this project are provided online by the Regional Mark Information System (RMIS) on their website (www.rmis.org). However, data summaries readily provided by RMIS do not provide the level of detailed information necessary for formally estimating the SARs and associated variances. A SAR estimate can be calculated for a tag code from the total recovered tags and release size, but not the associated variance. For proper variance calculation, the recovery data needs to be reported separately for each tag return along with its specific sampling fraction.

Our preliminary analysis had found large heterogeneity in adult return rates between replicate tag releases in the same year from the same hatchery. It is therefore important to include that source of variability into the overall precision of an annual SAR estimate for a hatchery. To this end, tag return data were not pooled across replicate tag-code release groups. Instead, a weighted average across replicates was reported and the information by tag-code preserved. This project reports annual estimates of SARs by hatchery and provide electronic summaries of the data by tag-code.

The SAR estimates are updated annually as additional return information becomes available after the beginning of the calendar year.