

Hatchery Tagging to Estimate *pHOS* Using IPTDS Detections

Kevin See^{1,}*
Mike Ackerman^{1,}*

December 04, 2019

Contents

Introduction	1
Objectives	2
Methods	3
Section 1. Expected precision of <i>pHOS</i> given current PIT tagging of hatchery smolt releases	3
Section 2. Expected precision of <i>pHOS</i> given current PIT tagging of W and HNC adults at LGR .	3
Section 3. Expected precision of <i>pHOS</i> by adding tagging of ad-clipped hatchery (H) adults at LGR	4
Mathematical Reasoning	4
Results	6
Section 1. Expected precision of <i>pHOS</i> given current PIT tagging of hatchery smolt releases	6
Section 2. Expected precision of <i>pHOS</i> given current PIT tagging of W and HNC adults at LGR .	6
Section 3. Expected precision of <i>pHOS</i> by adding tagging of ad-clipped hatchery (H) adults at LGR	6
Mathematical Reasoning	6
Discussion	7
Next Steps	7
Assumptions and Considerations	7
References	8
END WORKING DOCUMENT (SCRAP BELOW)	8
SCRAP	9
Methods Section 3	9
Simulation Description	9

¹ Biomark, Inc.

* Correspondence: Kevin See <Kevin.See@biomark.com>, Mike Ackerman <Mike.Ackerman@biomark.com>

Introduction

The Dam Adult Branch Occupancy Model (DABOM) is currently used to estimate adult escapement of natural origin spring/summer Chinook salmon *Onchorhynchus tshawytscha* and steelhead *O. mykiss* into watersheds and populations of the Snake River Basin (SRB). Escapement estimates are made based on the tagging of a random sample of adults at Lower Granite Dam (LGR) and the subsequent detection of those adults at instream PIT tag detection systems (IPTDS) throughout the SRB. DABOM could also be used to generate estimates of hatchery escapement to various locations as well, which would be useful for estimating hatchery abundance in natural areas and the proportion of hatchery origin spawners (*pHOS*). Doing so, with reasonable precision, may require a random sample (or similar) of adult hatchery fish also be PIT tagged at LGR. Here, we wish to determine expected levels of precision for *pHOS* given current levels of PIT tagging in the SRB, and further, determine the additional tagging, expressed as a tagging rate or number of tags,

required in hatchery adults at LGR to achieve unbiased estimates of hatchery escapement and *pHOS* with reasonable levels of precision (e.g., coefficient of variation [CV] of $\leq 15\%$).

Simultaneously estimating escapement of hatchery origin (HOR) and natural origin (NOR) adults to IPTDS locations throughout the SRB based on PIT tagging at LGR is a trivial extension of the DABOM model. Factually, the previous and existing instance of the DABOM model was developed for the upper Columbia River (upstream of Priest Rapids Dam), which capitalized on the same sampling assumptions applied to natural origin escapement past LGR. However, for multiple reasons (ergonomic limitations at LGR, political considerations, supposed ethical concerns), the random sampling and PIT tagging of HOR adults (specifically ad-clipped adults) at LGR has been precluded. Notably, given trap operations at LGR, HOR adults are trapped, anesthetized, weighed, and measured at the same rate as NOR - a function of "trap rate" constituting a fraction of every hour wherein adults in the ladder are bypassed to the trap. Thus, it is arguable that implantation of a PIT tag constitutes a significantly greater offense to captured adults.

Importantly, NOR and HOR refer to the true origin of a Chinook salmon or steelhead; that is born in the natural or hatchery environment, respectively. However, currently three phenotypic origins of adult Chinook salmon and steelhead are identified at LGR and include wild (W), hatchery ad-intact (HNC), and hatchery ad-clipped (H). NOR only includes W adults; HOR includes both the HNC and H groups despite HOR individuals often being phenotypically W. Currently, all ad-intact fish are systematically PIT tagged at LGR which includes both the W and HNC origins i.e., some HOR adults are currently tagged at LGR, but we need to consider differential tagging rates between HNC and H adults when evaluating the precision of *pHOS* given current levels of PIT tagging.

Unbeknownst to the proposal for this evaluation, a condition of the Hatchery and Genetics Management Plan (HGMP) required the operation of IDFG HOR steelhead production to conduct a study on HOR contributions to NOR populations. Thusly, IDFG and NOAA Fisheries endeavored to compile and estimate straying of HOR steelhead upstream of LGR.

This document seeks to begin collaboration between IDFG, NOAA Fisheries, and Biomark, Inc. to:

1. Capitalize on existing progress on stray rates, and
2. Advance statistical modeling to determine whether *pHOS* estimates can be reliably prosecuted by adult HOR PIT tagging at LGR.

In its developed form, the DABOM model includes escapement for NOR and HOR adult salmonids. Demonstrably, the DABOM is fully functional for multiple populations of spring/summer Chinook salmon and steelhead upstream of LGR. The performance of DABOM in the upper Columbia River (above Priest Rapids Dam) strongly suggests that the model can be reliably extended to HOR.

Contents of this document, including findings, will be reported as part of a Final (Annual) Progress Report for Bonneville Power Administration (BPA) project 2019-006-00.

Objectives

This assessment aims to determine the number or tagging rate of HOR adults at LGR necessary to determine *pHOS* across watersheds or populations of steelhead across the SRB estimated by the DABOM model (supported by the operation of IPTDS supported by BPA project 2018-002-00), within "reasonable" statistical parameters (e.g., $CV \leq 15\%$). As a natural extension, we may also explore additional effort necessary to estimate *pHOS* in watersheds or populations of interest for Chinook salmon.

First, we want to know expected CVs for current estimates of *pHOS* that could be generated for populations in the SRB given current levels of PIT tagging. Current tagging could include previously tagged adults that arrive at LGR with a PIT tag (NOR and HOR), and additionally, ad-intact fish that are systematically PIT tagged to evaluate abundance and composition at LGR (e.g., Camacho et al. 2017) and at IPTDS (e.g., Orme and Kinzer 2018) throughout the SRB. Previously PIT tagged hatchery fish would largely be from hatchery smolt releases that occur throughout the SRB.

Second, our goal is to determine levels of additional tagging at LGR of adult HOR fish (H only) that would allow reasonable estimates of *pHOS*. We will define reasonable as an estimate with a CV of 15% or less. Our

evaluation may (or may not) attempt to focus on desired or critical populations/locations.

Methods

Section 1. Expected precision of *pHOS* given current PIT tagging of hatchery smolt releases

This section aims to evaluate the expected precision (CV) of *pHOS* estimates given current PIT tagging fractions of steelhead hatchery smolt releases. We may also perform a simulation that could be naturally extended to make inferences about both steelhead and Chinook salmon. Here, we will leverage the approach of simply estimating the HOR escapement to that DABOM branch using a tag expansion approach:

$$\hat{N}_h = \frac{n_h}{\psi * \phi_h}$$

where \hat{N}_h is the HOR abundance, n_h is the number of observations of HOR fish, ψ is the detection probability, and ϕ_h is the HOR tagging fraction. HOR escapement to a given branch (with uncertainty) can then be combined with estimates of NOR escapement from DABOM to estimate *pHOS* with uncertainty. Importantly, DABOM requires a random sample of tagged individuals (e.g., at the adult trap), and thus, we cannot generate a hatchery abundance estimate to a DABOM model using juvenile tagging. Hence, we will use the a tag expansion approach. To do so, we need to compile data on tagging fractions in hatchery smolt releases by release group, year, and mark type [H vs. HNC]. For steelhead, this has been provided to us by IDFG. To keep the scope of the analysis reasonable, we intend to focus on a limited number of DABOM nodes, perhaps 3-5 per species, that either are representative of other scenarios or are of particular management interest. Particular locations of interest may include South Fork Salmon River Chinook, South Fork Clearwater steelhead, and Big Creek steelhead, but others may be added.

HOR smolt tagging fractions ϕ can differ by release group g , year y , and mark type m (HNC or H). For any given return (i.e., spawn) year, we perhaps need to account for several (perhaps 10 - 50) different tagging fractions. On October 22, 2019, IDFG provided us tagging fraction information for hatchery smolt steelhead releases from brood years 2013-2018. In many cases, these tagging fractions ϕ can be small, but in particular and perhaps important locations, tagging fractions can be high (e.g., South Fork Clearwater steelhead supplementation releases). Although for a given DABOM branch i one likely only needs to account for a handful of tagging fractions $\phi_{g,y,m}$, it will still be helpful to have tagging fractions for all smolt releases.

$$\hat{N}_{h,i} = \sum_{g,y,m} \frac{n_{g,i}}{\phi_{g,i} * \psi_i}$$

Ugh, my notation is atrocious here. Essentially, I want to expand the number of observations for each group by the tagging fraction for that group and the detection probability to get an estimate of abundance for that group. And then sum across release groups to get a total hatchery abundance for that location.

Note that in this scenario we don't need to account for differential survival among groups; we are only interested in the tagging fractions within groups as long as we are willing to assume equal survival between tagged and untagged smolts within each group.

Section 2. Expected precision of *pHOS* given current PIT tagging of W and HNC adults at LGR

Our aim here is to evaluate the expected precision (CV) of *pHOS* estimates given current PIT tagging rates of ad-intact adults (W and HNC) at LGR. This leverages the fact that all ad-intact (W and HNC) adults trapped at LGR are already implanted with a PIT tag. Here, we will re-run DABOM for HNC individuals to get HNC abundance estimates for watersheds and populations where we already report W abundance using DABOM. The HNC and W abundance estimates, at the same spatial scales, can then be easily combined to estimate *pHOS* with uncertainty. This assessment would effectively ignore ad-clipped H fish, but could be of

interest for particular supplementation locations (e.g., South Fork Clearwater steelhead). Unfortunately, it seems there is no easy way to combine juvenile smolt release tagging information (tag expansion) and adult tagging information at LGR (DABOM model) and so this section and the previous need to be completed separately.

Section 3. Expected precision of *pHOS* by adding tagging of ad-clipped hatchery (H) adults at LGR

Finally, we will evaluate the expected precision (CV) of *pHOS* estimates if additional (theoretical) tagging of ad-clipped hatchery (H) adults was performed at LGR. Again, we must ignore PIT tagged hatchery smolt releases. We'll perform a simulation and evaluate a range of tagging numbers (or rates) at LGR for H adults, estimate abundance of W, H, and HNC adults to a given DABOM branch, and then combine H and HNC abundance with W abundance to estimate *pHOS* with uncertainty.

Again, we need to account for the fact that adult HOR fish at LGR come in two flavors: ad-intact (HNC) and ad-clipped (H). Currently, all ad-intact fish that are trapped (W and HNC) are PIT-tagged, assuming they weren't previously tagged. So here we will focus on ad-clipped (H) adults only. In other words, we may need to account for differential tagging rates between ad-clipped (H) and ad-intact (HNC) when estimating *pHOS* here.

Mathematical Reasoning

First, we demonstrate how an estimate of *pHOS* can be generated under the DABOM framework. The estimate of *pHOS* at any branch i in the DABOM model can be written as

$$p\hat{H}OS_i = \frac{\mu_{h,i}}{\mu_{h,i} + \mu_{w,i}}$$

where $\mu_{h,i}$ and $\mu_{w,i}$ are the estimates of hatchery and wild fish who have escaped to spawn in branch i . Ignoring the subscript i , let σ_h^2 and σ_w^2 represent the variance of those estimates. Then we can re-write our equation for *pHOS* as a function of hatchery and wild escapement estimates, assuming they are independent of one another (and assuming multivariate normality):

$$\begin{aligned} p\hat{H}OS &= \frac{h}{h+w} = g\left(\begin{matrix} h \\ w \end{matrix}\right) \\ g\left(\begin{matrix} h \\ w \end{matrix}\right) &\sim N\left(\begin{matrix} \mu_h \\ \mu_w \end{matrix}, \begin{pmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_w^2 \end{pmatrix}\right) \\ &\equiv N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \end{aligned}$$

Then, via the delta method (or Taylor approximation) [Doob1935], we can derive the variance of our estimates of *pHOS*:

$$\begin{aligned}
Var(g) &= \left(\frac{\partial g}{\partial h}, \frac{\partial g}{\partial w} \right) \Sigma \begin{pmatrix} \frac{\partial g}{\partial h} \\ \frac{\partial g}{\partial w} \end{pmatrix} \\
&= \left(\frac{w}{(h+w)^2}, \frac{h}{(h+w)^2} \right) \begin{pmatrix} \sigma_h^2 & 0 \\ 0 & \sigma_w^2 \end{pmatrix} \begin{pmatrix} \frac{w}{(h+w)^2} \\ \frac{h}{(h+w)^2} \end{pmatrix} \\
&= \left(\frac{w\sigma_h^2}{(h+w)^2}, \frac{h\sigma_w^2}{(h+w)^2} \right) \begin{pmatrix} \frac{w}{(h+w)^2} \\ \frac{h}{(h+w)^2} \end{pmatrix} \\
&= \frac{w^2\sigma_h^2 + h^2\sigma_w^2}{(h+w)^4},
\end{aligned}$$

so the standard error of $pHOS$ is

$$SE(g) = \frac{w\sigma_h + h\sigma_w}{(h+w)^2},$$

which implies that the CV of $pHOS$ can be derived as

$$\begin{aligned}
CV(g) &= \frac{SE(g)}{g} \\
&= \frac{w\sigma_h + h\sigma_w}{(h+w)^2} * \frac{(h+w)}{h} \\
&= \frac{w\sigma_h + h\sigma_w}{h(h+w)} \\
&= \frac{h\sigma_w}{h(h+w)} + \frac{w\sigma_h}{h(h+w)} \\
&= \frac{\sigma_w}{(h+w)} + \frac{w}{(h+w)} * \frac{\sigma_h}{h} \\
&= \frac{w}{(h+w)} \left(\frac{\sigma_w}{w} + \frac{\sigma_h}{h} \right),
\end{aligned}$$

which is a combination of the proportion of wild fish, the CV of estimate of wild fish and the CV of the estimate of hatchery fish.

In other words, we can predict the CV of $pHOS$ for a variety of situations for a given branch i by varying the percent wild fish in a stream, and the precision (or CV) of our escapement estimates of wild and hatchery fish in branch i . We selected several levels of precision in our estimates of wild fish, with CVs from 0.01 to 0.5, and a range of precision in our hatchery estimates, with CVs from 0 to 1.5, and then calculated the predicted CV of $pHOS$ on those values.

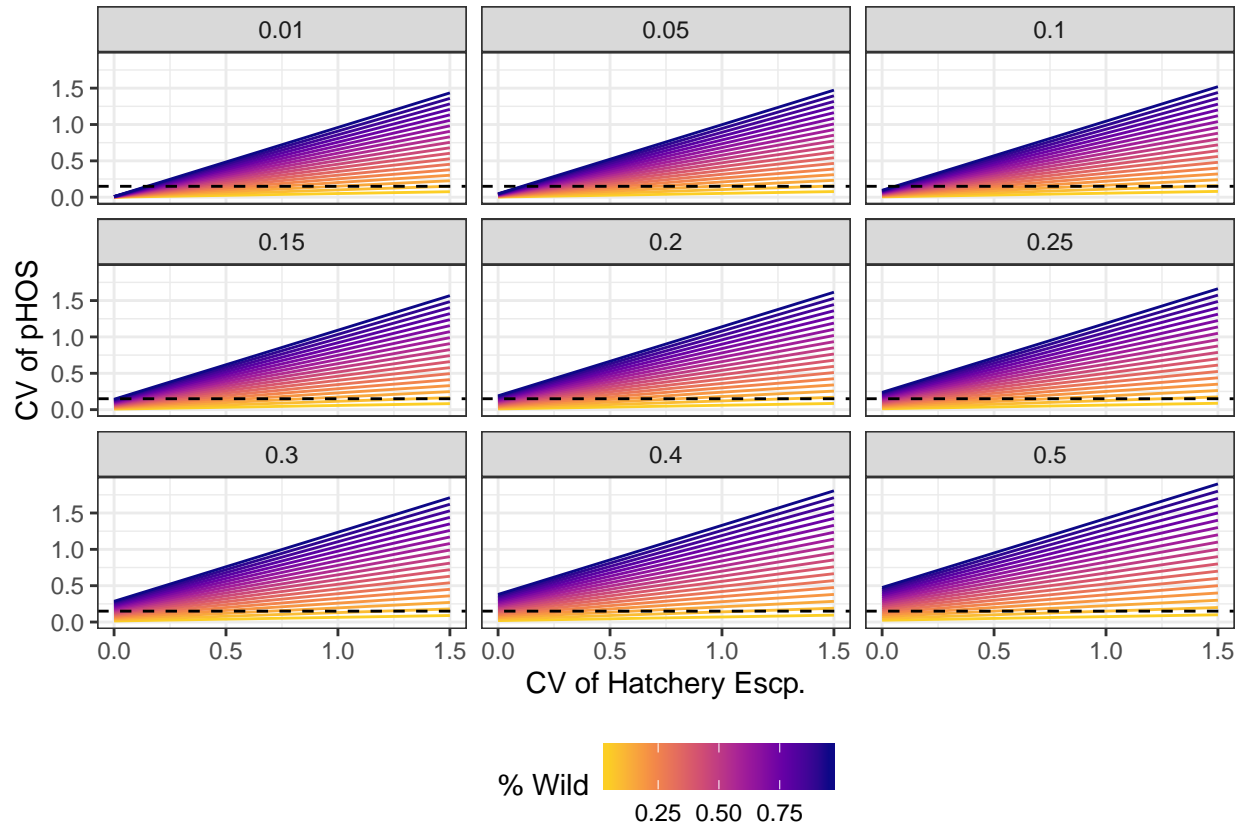


Figure 1: Predictions of the CV of $pHOS$ across ranges of the CV hatchery and wild escapement estimates and the proportion wild (inverse of $pHOS$). The horizontal dotted line depicts a CV of 15%.

Results

Section 1. Expected precision of $pHOS$ given current PIT tagging of hatchery smolt releases

Section 2. Expected precision of $pHOS$ given current PIT tagging of W and HNC adults at LGR

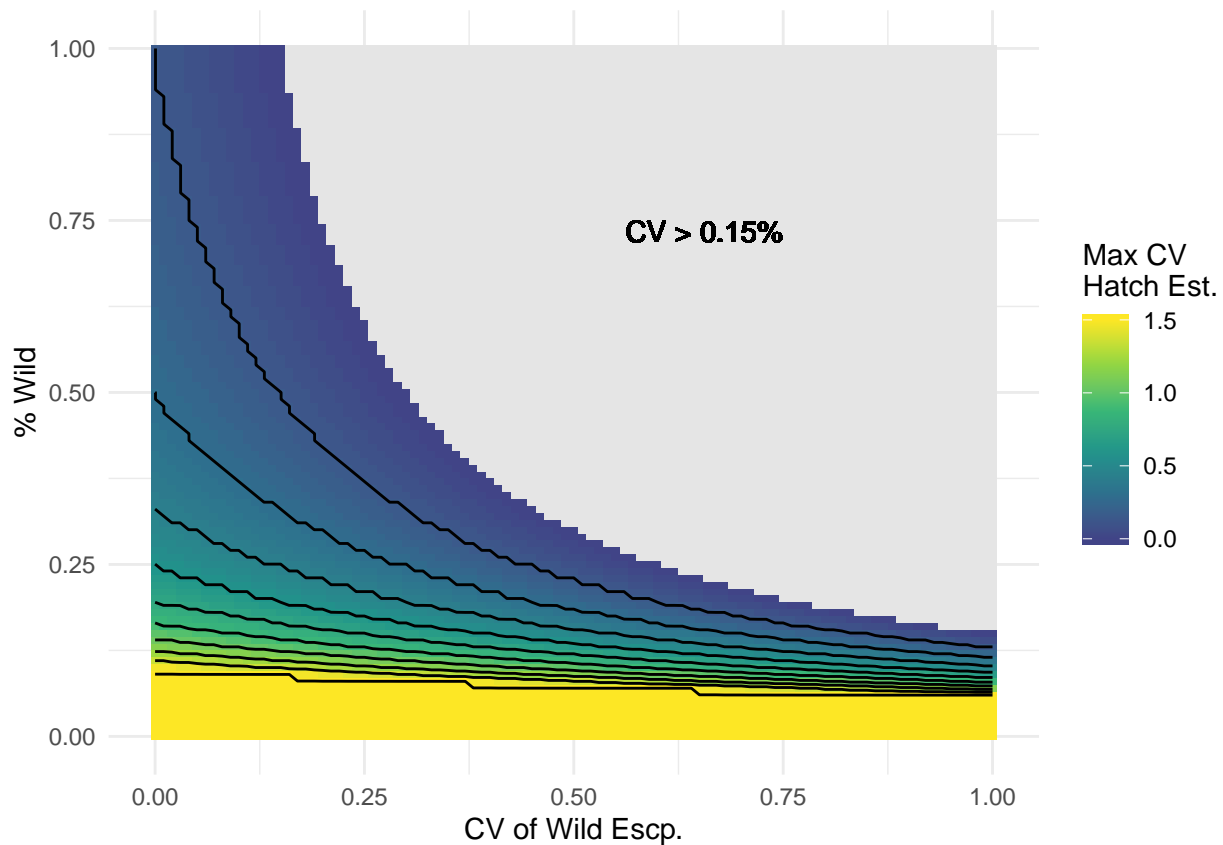
Section 3. Expected precision of $pHOS$ by adding tagging of ad-clipped hatchery (H) adults at LGR

Mathematical Reasoning

We predicted the CV of $pHOS$ across a range of values for both hatchery escapement CV (0 - 1.5) and proportion wild (i.e., inverse of $pHOS$). The results are shown in Figure 1.

The general patterns show that as the proportion of wild fish in the stream decreases, the precision of $pHOS$ improves for a given precision level of wild and hatchery escapement estimates. Also, as the precision of the wild estimates improves (i.e. smaller CV), there are more scenarios under which the precision of the $pHOS$ estimate is reasonable. As a reference, the CVs of the estimates for spring/summer Chinook escapement to the Lemhi range from 0.1 to 0.25 for years 2010 - 2018.

Another way to look at it would be to examine the maximum CV of hatchery estimates that would provide a reasonable estimate of $pHOS$ for a given proportion of wild fish and precision of wild escapement. One way to visualize these results in in the plot below.



From this plot, it is clear that as long as the proportion of wild fish in a stream is below a certain threshold (around 10%), then regardless of the CV of wild or hatchery escapement, reasonable precision of $pHOS$ will be obtained.

Discussion

Next Steps

- Incorporate results and implications from Craig Busack's script located in ../R/idahopit.R
 - from Lance Hebdon, August 5, 2019: I would contact Busack directly, we never saw any specific analysis only Craig's description of how low a confidence we would have given the low tag rates and detection rates at a single array in a single year basis. Also some of the issues we had with Craig's analysis was his assumption that straying was equal across the landscape, as opposed to straying being more likely near the release points, the effect of transportation on straying, the ability to aggregate information from multiple years to further understand patterns in straying, and a few other things that I don't have on the top of my head.
- Identify 2-3 locations of interest for each Chinook salmon and steelhead to perform evaluations
- Do we need to perform a PTAGIS query for particular locations to identify HOR observations; and can we track down release 'groups' for each observation?

Assumptions and Considerations

- What are the appropriate transition probabilities for hatchery fish along model branches? Need to consider harvest of hatchery fish. Presumably transition probabilities are lower for hatchery fish (relative to natural origin fish) due to harvest.
- Do hatchery fish mimic movement of wild fish? Probably not, due to the location of origin hatcheries and releases.

- Do some branches contain 0 hatchery fish?
- Should branches further upstream of LGR automatically include lower transition probabilities due to fishing on the mainstem?
- Do we need to accomodate time-varying movement probabilities in the simulation?
- Appropriate detection probabilities for each detection node in the simulation model.
- We must assume a fraction of hatchery fish will be harvested prior to reaching tributary habitats.
- How many total hatchery fish escape past LGR? We can run STADEM to determine range of values over the last several years.
- Is 15% CV for PHOS an appropriate, or even achievable, goal?

References

Camacho, C.A., K.K. Wright, J. Powell, W.C. Schrader, T. Copeland, M.W. Ackerman, M. Dobos, M.P. Corsi, and M.R. Campbell. 2017. Wild adult steelehead and Chinook salmon abundance and composition at Lower Granite Dam, spawn years 2009-2016. Cumulative Report 2009 through 2016. Idaho Department of Fish and Game Report 17-06.

Doob, J.L. 1935. The limiting distributions of certain statistics. The Annals of Mathematical Statistics 6:160-169.

Orme, R. and R. Kinzer. 2018. Integrated In-stream PIT Tag Detection System Operations and Maintenance; PIT Tag Based Adult Escapement Estimates for Spawn Years 2016 and 2017. Nez Perce Tribe Department of Fisheries Resource Management. Prepared for: Quantitative Consultants, Inc.

END WORKING DOCUMENT (SCRAP BELOW)

SCRAP

But then what's the next step here? To do an empirical evaluation we would then need to do a PTAGIS query of HOR observations across SRB IPTDS and be able to 'assign' each observation back to a release 'group'. Not a trivial task. And going down this path seems to require HOR runs of DABOM.

Maybe it makes sense to include Craig Busack's demonstration here to show that, even given high detection probabilities, it may be difficult to observe HOR fish at a given location if tagging rates and number of fish are low, which could lead to biased estimates of $pHOS$ at given locations if only juvenile observations are used.

Methods Section 3

Now we need to figure out, depending on results from Section #2, what level of additional tagging (number or rate) is needed for HOR adults at LGR to achieve reasonable CVs for estimates of $pHOS$. This additional tagging would likely be targeted at just ad-clipped (H) fish. In this case, we could also easily calculate the tagging fraction by species x release group x brood year for adults except in this case we can ignore both g and y because those should be equal across groups! i.e., because we are tagging at LGR. However, we do need to account for differential tagging fractions between H and HNC, however, HNC could be set equal to the W tagging fraction (i.e., all ad-intact fish are tagged at the same rate). The final consideration to make is whether we need a time varying component in tagging fractions for juveniles and adults. I'm thinking we make a simplifying assumption here and ignore a time-varying component for this exercise.

Here's where I need help. How do we then calculate the 'overall or global' tagging fraction for a group i.e., the tagging fraction after considering juvenile and adult tagging for a given release group? And then use that to calculate abundance for that release group? And then sum across groups? Are tagging fractions for a group a summation across life stage tagging fractions? Need to clean up some of my notation and thoughts here?

In the end, for Methods section #3, perhaps it makes most sense to ignore juvenile PIT tagging fractions and simply focus on adult tagging at LGR due to 1) low juvenile tag rates and 2) complexity of problem.

Alternatively, to determine the appropriate level of tagging, we may examine the last several years of DABOM results and focus on the number of tags observed within a few different tributaries and the level of precision of the accompanying escapement estimates, accounting for detection probability. We could then build a relationship to predict the precision based on the number of tags detected and the detection probability. After making some assumptions about the transition probability of hatchery fish to a particular tributary, we can use the estimate of detection probability to predict the precision of the escapement estimate under a variety of tagging rates.

Simulation Description

The $p\hat{H}OS$ to a given location j can simply be estimated as the ratio of all hatchery fish returning to a location, divided by total (i.e., hatchery and wild) fish returning. And in the case of hatchery fish, we need to account for broodstock removal and harvest (if any).

$$p\hat{H}OS = \frac{\hat{N}_{j,H} - (Broodstock + Harvest)}{\hat{N}_{j,W} + (\hat{N}_{j,H} - (Broodstock + Harvest))}$$

Typically, in STADEM and DABOM we estimate weekly escapement past LGR (STADEM) and weekly movement probabilities (DABOM). However, for the purposes of this simulation, perhaps we can assume similar run timings between hatchery and wild fish and among branches. And thus, wild and hatchery abundance to a given location j could be simply estimated as

$$\begin{aligned}\hat{N}_{j,H} &= X_H \psi_{j,H} \\ \hat{N}_{j,W} &= X_W \psi_{j,W}\end{aligned}$$

where X is the escapement past LGR and ψ is the movement probability for hatchery (H) and wild (W) fish, respectively.