MEMORANDUM

Oregon Department of Fish and Wildlife

DATE: October 9, 2014

TO: CCRMP, OASIS, OSCRP

FROM: Matt Falcy, Benjamin Clemens, and Stephanie Gunckel

SUBJECT: **Power analysis: determination of sub-sample size for analysis of fish scales from coastal Chinook salmon from spawning ground surveys**

BACKGROUND: The Fish Life History Analysis Project (FLHAP) processes and reads scale samples from up to several thousand coastal Chinook salmon each year. The vast majority of these samples originate from spawned-out carcasses on spawning ground surveys. The scales are read by FLHAP personnel to generate age estimates, according to established protocols (Borgerson et al. 2014). Most of the scales are received between late December and late February, and fisheries managers require the data by or before mid-March. This is a tight timeline for a project typically staffed by only two people, who process, mount, read the scales, conduct QA/QC, and then send that data on towards OSCRP staff (staff from CCRMP and OSCRP and assisted with mounting scales from time-to-time). The tightness of this timeline was underscored during 2012, when scale samples from ~14,000 Chinook salmon were processed and read by two permanent staff, and one temporary staff in FLHAP in about three months. This work is not sustainable. Therefore a strong impetus existed for FLHAP staff to sub-sample fish scales in a statistically-rigorous fashion.

Sub-sampling for the 2012 and 2013 run years.—Sub-sampling was done for sample sizes > 500 fish from spawning ground surveys conducted by the CCRMP for the 2012 run year (processed and read in 2013) and 2013 run year (processed and read in 2014). For these years, data for each basin was sorted by location, then by date for each sex. Summary statistics were calculated for body length (a proxy for age as the fish had not been aged) for each sex, by basin, for each of three sampling categories: 1) all samples, 2) every 2nd sample, and 3) every 4th sample. For scale samples from 501 – 1,000 fish, every 2nd sample was processed and read, and from ≥ 1,001 fish, every 2nd or 4th sample was processed and read, contingent upon no statistically significant difference between the subsample and the total sample. Finally, sex ratios were compared between the full sample and the sub-sample as another check on representation of each sex. If there was no significant difference between the body size of the entire sample versus the sub-sample for either sex, and there was no significant change in the sex ratio, then sub-sampling commenced.

Sub-sampling for the 2014 run year and subsequent run years.—The authors of this memorandum collaborated to strengthen the rigor and improve the utility of the power analyses to be conducted on coastal Chinook salmon from spawning ground surveys from which every fish encountered was sampled (as per CCRMP). The two goals of this memorandum are:

1) to show the rationale by which this sub-sampling can occur by FLHAP staff *after* they have received all fish scales from the spawning ground surveys conducted by the CCRMP; and

2) to explain these new developments and sub-sampling procedures to be executed by FLHAP staff for the fish scales collected from *all* encountered carcasses on spawning ground surveys 2014 run year and subsequent run years.

When scale samples from > 500 Chinook salmon from spawning ground surveys are collected, *random* sub-sampling will be done by FLHAP.

In view of the on-going interest in sub-sampling, it is intended that this memorandum serve as a reference source for interested parties that can see the scope of inference and rationale of the power analyses used. Those interested in adopting similar protocols will very likely need to change some parameters or assumptions to meet the needs of their particular project.

TERMINOLOGY

***N′* =** Entire spawner population, *N* + an unknown number of fish missed (e.g., due to scavenging, surveyor error and other unknown factors)

N = Statistical population composed of all the fish sampled from the field

n = Subsample from *N*

SUB-SAMPLING RATIONALE AND METHODS

The goal of power analyses is to quantify and convey the certainty in age composition of a population (from a particular basin of interest) that arises from sub-sampling, *n*, given a particular sample size, *N*. Based on available personnel to process and read the fish scales in a high quality and timely manner, this memorandum defines the acceptable magnitude of certainty, and therefore the amount of sub-sampling allowable, given N.

The power analyses can be conducted by one of two methods:

1) closed-form math, via the theory and equations provided in Thompson (2012; section 5.4 on pp. 59-61. PDF available on request). This assumes a “worst-case” age composition that is

the most difficult to accurately estimate. It is biologically unrealistic because it is three age classes with equal proportions, but it is a conservative approach that does not require *a priori* information about the age composition;

or

2) via Monte Carlo simulations with an assumed age composition.

When the "worst-case" age composition from method 1 was used in method 2, the resulting sub-sample size was six fish off from estimate obtained using method 1. This is excellent correspondence between radically different methods, and we know why the correspondence is not perfect.

FLHAP staff have determined that the following parameters for power analyses achieves a good balance between accuracy and economy:

- $\alpha = 0.05$, where α is the Type I error rate
- $d = 0.03$, where d is the critical distance of the estimates from the true age-class proportions in N.

This is a simultaneous error analysis, meaning that we will tolerate a 5% chance that *any* of the age-class estimates are more than 0.03 away from the true proportions in N (Figure 1). The value for "*d*" that we found acceptable was based on an examination of calculated sub-sample sizes across different values for d. At values of $d > 0.03$, our professional opinion was that the sub-sampling regime would be too liberal. By contrast, at values of $d < 0.03$, it was apparent that very little, if any sub-sampling would occur, and FLHAP would experience the same challenges that project staff were trying to avoid, with processing and reading excessive numbers of samples. It is our collective opinion that $\alpha = 0.05$ and $d = 0.03$ represent the right balance between being conservative with sub-sampling, while still allowing it to occur on a basis that will allow project staff to meet deadlines for processing and reading fish scales.

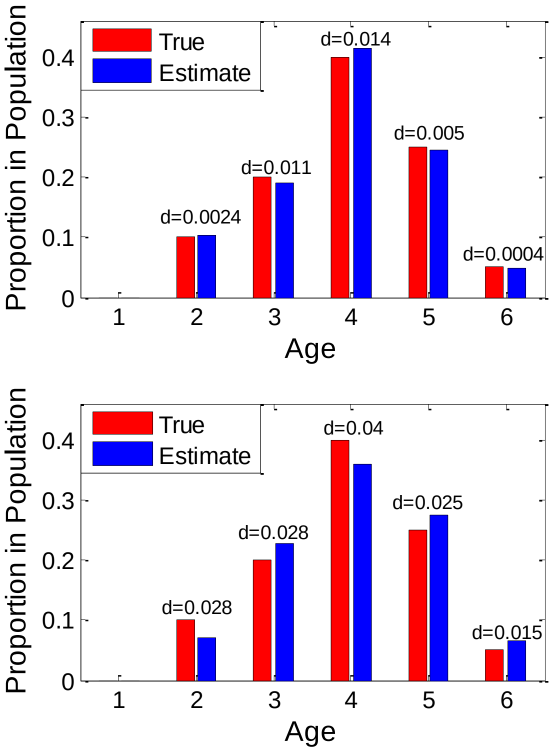

Figure 1. Two examples of estimating age composition. In the top panel, each estimated proportion-at-age (blue) is less than $d = 0.03$ from the true age composition (red). In the bottom panel, the estimated proportion of 4 year old is greater than $d = 0.03$ away from the true proportion. Even though all other estimates are less than $d = 0.03$, the bottom panel is an example of an estimate that is too imprecise. 95% of the time, all age classes will be estimated within $d = 0.03$.

This analysis can be used for any basin. Given a particular sample size of scales from Chinook salmon carcasses from the spawning grounds, the sub-sample size can be found below. Sub-samples are *randomly* selected from the entire collection of scales, so sex and location are preserved within the realm of certainty described above.

CLOSED-FORM CALCULATION

Using values provided in Thompson (2012) for $d^2 n_0$ for a given d and α, n_0 was calculated and used in the equation provided by Thompson for estimating a representative sub-sample:

$n = 1/ [(1/n_0) + (1/N)]$ (**Equation 1**)

Where

n = sub-sample size, after adjusting for finite N, as per above
n_0 = sub-sample sample size tabulated by Thompson (2012) assuming infinite N
N = sample size (from all of the fish scales collected), as per above.

This equation was used to generate Figure 2 and Table 1.

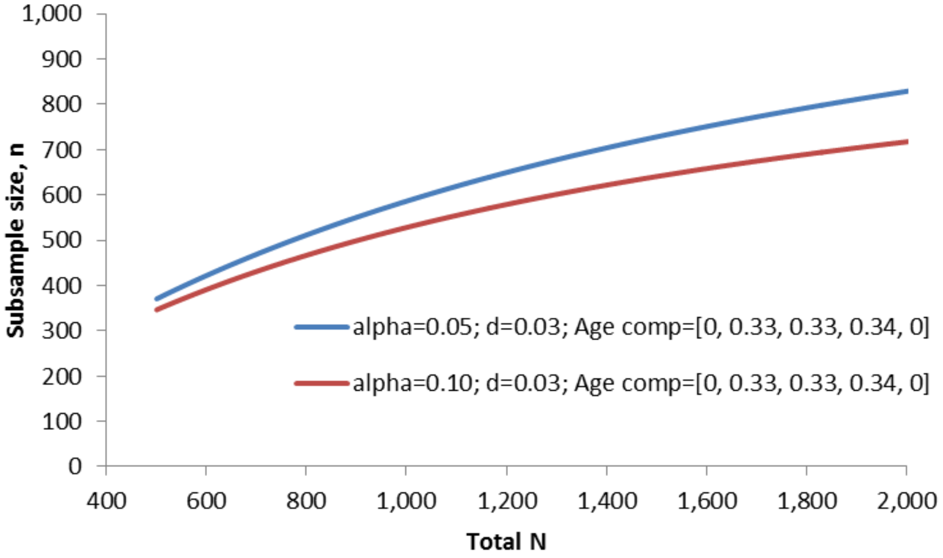

Figure 2. Subsample size required to achieve *d* and α for different *N*. The "worst-case" age-composition is assumed here. FLHAP has decided that an alpha of 0.05 is desirable, but an alpha of 0.10 is shown for comparison.

Table 1. Results from equation 1, given $d^2 n_0$ values from Thompson (2012), $d = 0.03$, and $\alpha = 0.05$.

N	***n***
501	371
750	491
1,000	586
1,250	664
1,500	728
1,750	783
2,000	829

MONTE CARLO SIMULATIONS

Prior to discovering Thompson's work, we developed a simulation technique to conduct the power analysis. Although the simulation is less elegant than Thompson's mathematics, it has two advantages. First, any age composition can be supplied — it does not presume the "worst-case." Second, it does not assume normality or use Bonferroni corrections; rather, sampling occurs from the actual mulitnomial distribution of ages, and simultaneous error can be directly verified. A drawback of the technique is that it is inherently stochastic, so slightly different estimates can be obtained with identical inputs. However, simply increasing the number of simulation replicates will result in a stable answer.

The code was initially developed to compute α given the age composition, *N, d,* and *n*. However, we wish to know *n*, given α. Rather than sweeping through various input values of *n* and then trying to map them onto the resulting α, we just turned the code into a function that could be numerically optimized to find the *n* that minimizes the difference between computed α and the target α. This is an efficient solution. As noted previously, the simulation produced results nearly identical to the Thompson's mathematics when we forced the simulation to evaluate the "worst-case" age composition.

The simulation code is provided below using the "R" computational language. We provide this so that others can conduct the power analysis if they wish to explore the effects of more realistic age-compositions on the subsample size needed to achieve α and *d.* Users should change numerical values in the first three lines to obtain a desired scenario. The code will return two numbers. The first number, listed under "$minimum," is the subsample size. The second number pertains to the performance of the optimization. Using extreme values will cause errors. Heed any warning messages.

#####RCode#####

```
opt.n<-function(n,N=1000,d=0.05,alpha=0.05){
a<-c(0, 0.33, 0.33, 0.34, 0, 0)#worst case age comp. Must sum to 1.
reps<-50000 #increase this to obtain better precision. Repeat and compare results.

sample.p<-n/N
a.pop<-a*N
A<-matrix(NA,length(a),2)
out<-matrix(NA,reps,length(a))
out2<-vector()
for (rep in 1:reps){
  a.samp<-vector()
  Pop<-c(rep(1,a.pop[1]),rep(2,a.pop[2]),rep(3,a.pop[3]),rep(4,a.pop[4]),rep(5,a.pop[5]),rep(6,a.pop[6]))
  samp<-Pop[sample(N,N*sample.p,replace=F)]
  a.samp[1]<-sum(samp==1)
  a.samp[2]<-sum(samp==2)
  a.samp[3]<-sum(samp==3)
  a.samp[4]<-sum(samp==4)
  a.samp[5]<-sum(samp==5)
  a.samp[6]<-sum(samp==6)
  A[,1]<-a.pop/N
  A[,2]<-a.samp/(N*sample.p)
  out[rep,]<-A[,2]
  out2[rep]<-as.numeric(sum(as.numeric(abs(a-out[rep,])>d))>0)
}

return(abs(alpha-sum(out2)/reps))
}
optimize(opt.n,lower=10,upper=0.9*N)
```

REFERENCES

Borgerson, L., B. Clemens, K. Bowden, & S. Gunckel. 2014. Fish life history analysis project: methods for scale analysis. Information Report 2014-10. Oregon Department of Fish and Wildlife.

Thompson, S. K. 2012. Sampling, 3rd edition. John Wiley & Sons, Inc., Hoboken, New Jersey.