CODED-WIRE TAG REPLICATION STUDIES

by

Frank de Libero

Pacific Marine Fisheries Commission/
Washington Department of Fisheries

CODED-WIRE TAG REPLICATION STUDIES
PRELIMINARY RESULTS

Abstract

The necessary groundwork has been completed to estimate the internal variability of CWT data via replication. A preliminary analysis was done on replicate releases from Abernathy. The analysis and results are discussed.

Tag codes have been identified by most agencies along the coast that can either be used as replicates, i.e., there are no significant differences between tagged groups except the code, or as reasonable approximations to replicates. Latter groups include pond effects or genetic differences. The criteria for selecting replicates were that they should represent normal hatchery practices and that there were no known practical biological differences between them.

Among the replicates identified is a fall chinook family study conducted at the Abernathy Hatchery, Fish and Wildlife Service. Because of the way the replication was carried out, it is possible to approximate the variance associated with the CWT within a female.

For the broods 1974 through 1977, 50 females were sequestered; the eggs from each being divided into two replicates. For the most part, the sperm from a common male was used to fertilize five successive egg lots. Occasionally, two different males were used, one for each replicate from a single female. A unique tag code was applied to the fingerlings that developed from each replicate. Part of the 1975 data is tabled as an example:

Female number	Egg lot	Tag code	Male number
1	A	14-10-1	1
	B	14- 4-3	
2	A	14- 5-3	
	B	14- 6-3	
3	A	14- 7-3	
	B	14- 8-3	2
4	A	14-10-3	
	B	14- 9-3	

etc.

This process was completed each year for all fish in the study. If there had been no missing values, there would have been 100 tag codes per year, two replicates per female.

Three of the four years did not have 100 tagged groups. In order to avoid possible problems in interpreting the analytical results because of unbalanced data, and since there was ample data from the study, females were randomly selected out until there were 40 females, or 80 tag codes left for each brood year.

Males were not considered during this process or the ensuing analysis. There was no evidence that including them would contribute to the results.

A nested random effects analysis of variance was used to estimate the components of variability[1]. The dependent variable was recovery rate which was calculated as the sum of the total estimated harvest plus escapement to Abernathy as a percent of total number of tagged fish released. Female salmon were nested within brood years and the determinations were the recovery rates within each female. The program used was SPSS MANOVA[2] on the CYBER computer at the University of Washington.

The data is not normally distributed, but sharply skewed to the right. The analysis was redone using the arcsin, square root and log transformations with little appreciable effect on the final result. Because it is easier to interpret, the analysis is presented here in the original units.

The model used:

$$Y_{ijk} = m + a_i + b_{j(i)} + e_{ijk}$$

where a_i represents the ith brood year ($i = 74,\ldots,77$)

---

[1]Snedecor, George W., William G. Cochran. 1980. Statistical Methods. 7th Edition.

[2]Hull, C. Hadlai, Norman H. Nie. 1981. SPSS Update 7-9. New Procedures and Facilities for Releases 7-9.

$b_{j(i)}$ represents the jth fish within the ith year (j = 1,...,40) and e_{ijk} is the estimate of error (k = 1, 2).

The analysis of variance results are:

Source of variation	Degrees of Freedom	Mean Square	Parameters Estimated
Between brood years	3	5.804	$S^2 + nS_b^2 + bnS_a^2$
Fish within years	156	0.674	$S^2 + nS_b^2$
Tag codes within fish	160	0.549	S^2

$s^2 = 0.549$ estimates S^2

$s_b^2 = (0.674 - 0.549)/2 = 0.0625$ estimates S_b^2

$s_a^2 = (5.804 - 0.674)/80 = 0.0641$ estimates S_a^2

The within-fish error accounts for 80% of the variability in the data. This is interpreted to be the result of small numbers of observed recoveries, two on the average, for each tag code.

Ideally, the replication analysis will yield an estimated relationship between statistical error and information. This in turn will help determine the extent of tagging and sampling necessary to realize different levels of precision for the estimates. The Abernathy data was used to make a first cut at that relationship.

The Abernathy family study began with the 1973 brood. That year, however, 100 females were used with one tag code for the progeny of each female. Unlike the other years, no estimate of within-female variability is available, but otherwise the study protocol was the same. Because the analysis of variance results indicates that the between-female variability was small compared to the estimate of pure error, the difference between the 1973 protocol and the other years was not considered important for the regrouping that follows. There was no statistical difference between the 1973 error mean square and the pooled variance estimate for 1974 through 1977.

In order to include as many years as possible, the 1973 data was incorporated in the following analysis after randomly selecting out successive pairs of females (as if they were a single female with two groups of eggs) until there were 80 tag codes left.

The coefficient of variation was calculated across all tag codes. The number of observed recoveries was estimated by dividing the estimated number recovered by the average expansion factor of 5 for Abernathy salmon. Then the data from successive pairs of codes was grouped to increase the number of recoveries within a replicate, now comprised of two tag codes. Again, the relative error was estimated. This grouping of the data and reestimating the coefficient of variation was continued until there were over 50 observed recoveries per group. (Over half the tag codes in the CWT data base have fewer than 50 observed recoveries.)

The following table summarizes the results:

Number of tag codes per group	Average coefficient of variation	Average number of observed recoveries	Estimated Minimum and maximum recovery rates
1	110%	2	0 - 5.17%
2	80%	5	0 - 3.20%
4	60%	11	0 - 2.69%
10	40%	27	0.15 - 1.71%
20	25%	54	0.24 - 1.16%

These results indicate that for those stocks that can be represented by the Abernathy experiment, 1 to 5 observed recoveries is not enough for stable estimates. There simply is insufficient power. It takes approximately 15 observed recoveries to achieve a relative error of 50%.

It is anticipated that an analysis of the other replicated groups will yield relationships that are at least qualitatively similar to the one tabulated above. A small sample of WDF coho planned replicates tends to confirm this.

The anticipated benefits of the CWT replication analysis, once complete, are:

* As previously noted, as an aide in determining tagging and sampling levels.

* To give minimum error estimates which can be used as bench marks.

* To help detect outliers. For example, during the above analysis, one replicate within a fishery had 58 estimated recoveries while the other replicate showed zero captures.

* To help determine probability distributions to model the CWT estimation process.

* To futher document the value of replication for CWT experiments.