APPENDIX D: Confidence Intervals for 1990 Passage Indices

Introduction

This appendix includes confidence intervals for the 1990 cumulative passage indices of each species at Lower Granite, Rock Island, McNary, John Day, and Bonneville dams. These confidence intervals were computed using a methodology recommended by the FPC contracted Biometrician Group (members are Drs. Lyle Calvin, Cliff Pereira, and Doug Neeley) in their January 1990 report to the FPC, which is presented in this appendix. The recommended methodology was to stratify the migration season into a series of 2-day blocks, and compute variances for each block's passage index total. The passage index totals and associated variances were each summed across the series of 2-day blocks to obtain the cumulative passage index and its respective variance. From this data, 95% confidence interval were then computed for each species and site. In the October 1990 amendment to the January report, the Biometricians noted that this method of using paired-day stratification would lead to conservative confidence intervals. Since the method is also valid for block sizes >2, a comparison to confidence intervals made with blocks of 3-day and 4-day widths is presented in the following tables. The purpose was determine if using 2-day blocks will generally result in narrower confidence intervals than blocks of wider size. Since the methodology already produces conservative confidence intervals, it is preferable to use a blocking size that gives narrowest confidence intervals.

Results and Discussion

Cumulative passage index confidence intervals, based on 2-day blocks, ranged from ±5.1% to ±21.5% of the annual index for all sites and species. The narrowest confidence interval was for yearling chinook at Bonneville Dam and the widest for yearling chinook at Rock Island Dam. About half of the site/species confidence intervals were less than ±10% of the annual passage index. The size of the confidence intervals obtained from the 2-day, 3-day, and 4-day block widths were close. The use of 2-day blocks produced the narrowest confidence intervals for subyearling chinook at each site and for all species at McNary Dam. For spring migrants at the other sites, the 3-day blocks produced slightly narrower confidence intervals compared to the 2-day blocks in two-thirds of the cases, while the 4-day blocks were lower in only one-third of the cases. Overall, the differences between confidence intervals using the three blocking sizes were small, being less than four percentage points different in all cases except yearling chinook at Rock Island Dam. The average size of the confidence intervals across all sites and species was ±11.0% using 2-day blocks, ±11.5% using 3-day blocks, and ±12.1% using 4-day blocks. These results demonstrate that these three levels of blocking will produce similar results, but that the 2-day blocking will, on average, produce slightly narrower confidence intervals.

As stated in the Biometricians report and FPC cover letter on the report, these confidence intervals for cumulative passage indices incorporate day-to-day sampling variability, and provide a measure of precision around the annual fish passage indices. It does not incorporate variability in FGE, nor does it indicate how well the annual indices actually reflect population magnitudes.

Computed Confidence Intervals of 1990 Salmonid Passage Indices at Rock Island Dam

	Passage Index	Confidence Interval Block Size					
		2-Day		3-Day		4-Day	
		CI(±)	%	CI(±)	%	CI(±)	%
Chinook 1s	20,853	4,492	21.5	4,124	19.8	5,592	26.8
Chinook 0s	54,683	3,690	6.8	4,414	8.1	4,747	8.7
Steelhead	18,085	1,236	6.8	1,395	7.7	1,370	7.6
Coho	15,617	2,721	17.4	2,472	15.8	2,572	16.5
Sockeye	4,297	607	14.1	531	12.4	610	14.2

Computed Confidence Intervals of 1990 Salmonid Passage Indices at Lower Granite Dam

	Passage Index	Confidence Interval Block Size					
		2-Day		3-Day		4-Day	
		CI(±)	%	CI(±)	%	CI(±)	%
Chinook 1s	3,199,593	305,079	9.5	300,473	9.4	323,626	10.1
Chinook 0s	N/A						
Steelhead	6,139,888	701,887	11.4	672,907	11.0	694,413	11.3
Coho	N/A						
Sockeye	16,596	2,441	14.7	2,222	13.4	2,532	15.3

Computed Confidence Intervals of 1990 Salmonid Passage Indices at McNary Dam

	Passage Index	Confidence Interval Block Size					
		2-Day		3-Day		4-Day	
		CI(±)	%	CI(±)	%	CI(±)	%
Chinook 1s	2,432,655	143,949	5.9	175,690	7.2	188,557	7.8
Chinook 0s	8,507,935	822,549	9.7	857,362	10.1	868,621	10.2
Steelhead	660,448	48,257	7.3	57,720	8.7	56,205	8.5
Coho	231,034	24,243	10.5	27,264	11.8	28,033	12.1
Sockeye	294,263	18,347	6.2	21,050	7.2	23,739	8.1

Computed Confidence Intervals of 1990 Salmonid Passage Indices at John Day Dam

	Passage Index	Confidence Interval Block Size					
		2-Day		3-Day		4-Day	
		CI(±)	%	CI(±)	%	CI(±)	%
Chinook 1s	361,968	36,182	10.0	34,588	9.6	36,225	10.0
Chinook 0s	513,687	51,873	10.1	66,870	13.0	70,779	13.8
Steelhead	133,777	19,585	14.6	21,072	15.8	22,278	16.7
Coho	84,342	13,697	16.2	15,105	17.9	13,773	16.3
Sockeye	23,610	2,994	12.7	2,797	11.9	2,774	11.8

Computed Confidence Intervals of 1990 Salmonid Passage Indices at Bonneville Dam

	Passage Index	Confidence Interval Block Size					
		2-Day		3-Day		4-Day	
		CI(±)	%	CI(±)	%	CI(±)	%
Chinook 1s	332,792	16,908	5.1	17,536	5.3	21,093	6.3
Chinook 0s	1,219,778	109,726	9.0	150,762	12.4	139,660	11.5
Steelhead	127,882	9,833	7.7	9,336	7.3	11,009	8.6
Coho	677,413	52,945	7.8	51,477	7.6	63,966	9.4
Sockeye	81,403	14,496	17.8	16,991	20.9	13,808	17.0

COLUMBIA BASIN FISHERY AGENCIES AND TRIBES

FISH PASSAGE CENTER
825 N.E. 20TH AVENUE • SUITE 336 • PORTLAND, OR 97232-2295
PHONE (503) 230-4099

MEMORANDUM

Date: March 27, 1990

To: FPAC

From: Michele DeHart

RE: Report from Biometrician Group regarding confidence intervals for Passage Indices

Last fall NMFS and CRITFC commented on the 1990 Smolt Monitoring Program identifying their need for confidence intervals on the passage indices. This was discussed at FPAC, and subsequently the FPC was requested to have the Biometrician Group, which consists of Drs. Lyle Calvin, Cliff Pereira, and Doug Neeley, address this question. The attached report is the result of their effort to develop a meaningful confidence interval for the cumulative fish passage index. Following discussions with the FPC on applications of the passage index data, they recommended that confidence intervals be developed for the cumulative fish passage index rather than for daily fish passage indices.

The Biometrician Group's report to the FPC provides a methodology for computing a confidence interval. However, it also provides words of caution in how to interpret and use these confidence intervals. It must be emphasized that just as FGE is not factored into the passage index, neither is its variability factored into the confidence intervals defined in this report. A confidence interval around an annual passage index incorporates only a fraction of the variability that would occur around estimates of population magnitude. At most these confidence intervals incorporate day-to-day sampling variability, and provide a measure of precision about the annual fish passage index. They do not necessarily reflect how close the annual index represents the population magnitude.

As stated in our response to NMFS and CRITFC, and concurred by the Biometrician Group report, confidence intervals are desirable, but can be misleading without an explicit definition of what they are intended to measure or represent. The audience who is interested in using the confidence intervals must be cognizant of their limited utility. Their expectation of what the confidence intervals represent will influence what applications they try to make of these confidence interval around cumulative fish passage indices.

199-90.TB

Report of Biometrician Group
on a confidence interval for the
cumulative fish passage index

January 1990

Biometrician Group members:

Cliff Pereira

Lyle Calvin

Doug Neeley

Introduction

In November, 1989 the Fish Passage Center asked the Biometrician Group (Lyle Calvin, Doug Neeley, and Cliff Pereira) to look at the methods used in calculating the daily and annual fish passage indices (FPI's) to determine the feasibility of obtaining confidence interval (CI) estimates. This was the result of discussion between the FPC staff and the Fish Passage Advisory Committee on behalf of the fishery agencies and tribes.

The Biometrician Group agreed to this assignment and received background material and assistance from the FPC in obtaining a full understanding of the problem and the methods presently used. In particular, Michele DeHart and Tom Berggren met with the group on November 21 and Tom Berggren again on December 22. Several documents were developed and exchanged in November and December; these are not included in the report but are being held in the Fish Passage Center. Some of them deal with methods that appeared promising but which were eventually discarded after gaining further understanding of the situation at each of the dams.

A number of issues related to the determination of the daily and annual passage indices as well as to the estimation of confidence intervals arose in our discussions. Some of those seemed important enough to include in our report and do constitute several sections of the report.

Initially most of our efforts were spent on developing a method of calculating confidence intervals for the daily fish passage index. With the realization that primary interest was really in the cumulative fish passage index (over some number of days within a year) and, speicifically, in the annual fish passage index, our emphasis changed to a procedure for estimating a confidence interval for an annual or other cumulative fish passage index. The method recommended provides a method that can be used at any site where a daily fish passage index can be calculated for each day. It is simple to use and should serve to provide a confidence interval estimate when needed.

The FPI: An index of what?

If the FPI is to be useful, it must be considered to be highly correlated with some parameters of interest. There seem to be two parameters with which the FPI could be correlated and for which other estimates are not readily available. These are (1) the population count (number of fish in the river at the point of collection), and (2) the inherent survival rate. Both are parameters which provide information of interest to fish managers and policy makers.

In this memo we will treat the FPI as being correlated with the population count within a year at a given location. Except for the problem of accounting for the FGE, the FPI is estimated as if it were intended to serve as a population count (i.e. as the population count would be estimated). It reflects the effects of varying inputs of hatchery and wild fish, smolt condition, flow patterns, and project operations in addition to any factors affecting survival of juvenile fish in the river at this location and time. Changes in FGE would affect any actual population estimates as calculated from sample counts but do not affect the FPI since no adjustment for FGE is made. The FPI can, therefore be considered as an estimate of the population count but uncorrected for FGE. As such, it must be considered a finite population index within a year at a given location.

If the FPI were intended to reflect the inherent survival rate, it would have to be adjusted for varying input of fish (particularly hatchery releases) each year and at different locations. Such an input adjustment would be difficult but perhaps no more difficult than interpreting population estimates while recognizing varying inputs.[1]

---

[1]Some attempt is made to account for hatchery releases in the Fish Passage Center annual report by presenting the ratio of FPI to hatchery releases. The Biometrician Group has not studied the use of such a ratio in any detail.

Need or value of confidence intervals (CI) on FPI.

Although the Biometricians Group has been asked to look at the feasibility of calculating confidence intervals for the FPI, there is not a clear recognition of how the CI will be used or how important they are. The FPC uses the FPI in a fairly general way that would probably not change much whether the CI were small or large.

Because only a sample of the fish that enter bypass facilities are counted each day, sampling variation is present. The within-day sampling variation will form the basis for confidence intervals presented in this report. Also present may be non-sampling errors due to such things as varying FGE or the failure of fish passage to be proportional to flow volume. In general, non-sampling errors reduce the value of the FPI as an index and such errors are difficult to assess.

There is validity to the argument that additional information is provided by CI estimates compared to point estimates. Presuming the CI estimates are reliable, an answer is provided to the question of precision of the point estimates. This, in itself, may be sufficient reason for calculating CI when possible. In the case of the FPI, however, there seems to be less use of CI than there would be for the usual statistical estimates because (1) the FPI is an index, not a direct estimate of a parameter, (2) non-sampling errors which cannot be measured may contribute a large part of the total error, and (3) the uses to which the FPI is put are general and do not require high precision.

Comparison of the FPI among projects and among years.

The FPI has been used as a relative measure of the magnitude of the runs of a species at a given location and a given time. The CI estimates that are being recommended apply only to that location and time and are not intended to adequately reflect the variation due to FGE or the changes in the proportion of fish to flow volume.

While these factors vary considerably within any one project and year, the variation among projects and years is even greater and certainly large enough to question the validity of comparing the indices across projects or years.

Looking at the major non-sampling errors, those due to variation in FGE and fish per flow volume, it appears likely that both probably vary more across projects than across years. One might, therefore, expect to have better comparisons among years at the same project, than among projects in the same year. Any such comparisons, however, should recognize the presence of non-sampling errors not included in the CI estimates.

The sampling and non-sampling errors make up the uncontrollable errors constituting the total variation inherent in any comparison of or inference about the FPI. Evaluation of any differences or changes in FPI should also include available information or factors influencing the FPI, e.g. timing and size of hatchery releases, smolt condition, wild stocks, flow patterns and conditions, transportation programs, and project operations.

One may also wish to compare migration timing for different projects. This might be done looking at the consistency of the differences in timing from one project to the next. In addition to the factors listed above, variation in migrational timing may be caused by hatchery releases affecting one dam but not the other.

FGE adjustment for the FPI

The calculation of the FPI includes no adjustment for the FGE of screens, as one would want to do for a population estimate. This is not necessarily a great concern, however, for an index that only needs to be highly correlated with the population count. At least it is of little concern unless the FGE varies considerably among days. If it does, and there is evidence that it does, this will decrease the correlation between the

FPI and the population count and make the FPI less useful.

If the FGE could be measured well each day, an adjustment could be made which would increase the correlation between FPI and the population count. Unfortunately the FGE cannot be measured with any reasonable degree of precision, at reasonable cost, each day and has therefore been excluded from consideration in the FPI calculation. An average FGE value could be used for adjustment but this would not increase the correlation and hence the value of the FPI.

Two additional factors that can reduce the usefulness of the FPI

If hand counts are available only every 24 hours and the sampling rate varies within a day, then variation in the within-day species composition can have the effect of decreasing the correlation between the FPI and the population count. This happens because the data is not available to calculate species-specific 24-hour average sampling rates. The single overall average sampling rate that is used in calculating the daily FPI may be too high for some species and too low for others. Some insight into the effect of within-day variation in species composition may be obtained at projects where hourly hand counts are made. Note that there will not be a problem of this type as long as there is a constant within-day sampling rate or there are separate holding areas for fish sampled at different rates (as at McNary Dam).

Another factor which can vary within a day is the percent of the flow sent through the sampled unit or powerhouse. The flow adjustment used in calculating the daily FPI is based on a single average percent of flow for the day. Depending on the relationships within each day between 1) flow percent, 2) number of fish passing the facility and 3) species composition, the currently used daily flow adjustment can result in decreased correlation between FPI and population count. This problem may potentially be studied at facilities with hourly hand counts, however, a long lag time between entry into the bypass and entry into the sampling device could make it difficult or impossible to adequately relate the hand counts to within-day flows for this purpose.

Possible confidence intervals for the FPI

We discussed three approaches for generating confidence intervals for the FPI. After briefly describing each, the rationale for rejecting the first two will be summarized.

Approach 1. The first is to consider the day to be stratified into L periods, within which the hourly counts could be assumed to be random counts from a common population (but possibly different for each period). Confidence intervals are then calculated for the total daily count as for a stratified random sample. Consideration needs to be given to the finite population correction needed when subsampling is used. Confidence intervals for cumulative FPI can be obtained in the usual way by summing variances over days. Approaches 1 and 2 both require hourly counts of fish, either by Smith Root counters or hand counts.

Approach 2. The second approach would be to fit a common pattern of hourly counts within the day to be use for all days within some time period and use the deviations from the expected count at each hour to estimate the error for the daily FPI. This should work well if the pattern is fairly consistent across days within the period. The average pattern might be obtained from 7 or 14 days and deviations taken from it. Using the same data to establish the pattern and to calculate the error causes the estimated error to be a little too small, but this is counterbalanced by the failure of the pattern to be common over 7 or 14 days.

Approach 3. This approach does not require hourly counts, although, like the first, it is a stratified approach. The difference is to define strata as periods of several days and use the variation among daily FPI within periods as an estimate of error. For example, if strata of two days are established, the estimated error variance for each stratum would have one degree of freedom. The error variance for the annual (or other cumulative) FPI is obtained by summing the variances over all strata and the confidence interval is calculated from this.

The third approach was taken because the first two assume that the variance for the daily FPI could be estimated using variation in hourly counts. This would be possible if appropriate within-day stratification of hourly counts were possible. However, there are many daily patterns and activities that could not easily lend themselves to stratification. These include:

1. within-day changes in flow patterns at a project,

2. within day surges of fish caused by hatchery releases, and

3. the long lag-time at some facilities between entry of fish into the bypass and their subsequent entry into the sampling device.

There were other problems associated with the use of variation in hourly counts as a basis of estimating the variance. At some facilities there are no hourly counts of any kind, so that neither approaches 1 nor 2 would be possible. A number of assumptions would be required to use approaches 1 and 2 including that:

1. species composition can be treated as constant over a day,

2. percent flow through the sampled unit or powerhouse can be treated as constant over a day, and

3. (if SR counts are used) the relationship between hand-counts to SR counts can be treated as constant (and known) over a day.

It should be mentioned that approach 3 is not without problems. True day to day variations in such things as hatchery fish inputs, flows, and FGE will tend to bias the confidence interval estimates toward being too large. However, since the variance estimate is based on variation in FPI's between adjacent days only, the bias should be minimized.

In summary, because the primary interest in confidence intervals is for cumulative FPI's (and not daily FPI's), approach 3 above is the one recommended for use by the Biometrician Group. It is also simpler to compute and incorporates the appropriate components of variances better than the other methods and with fewer assumptions. The method is illustrated below.

Recommended confidence interval for FPI

Let y_i = sample count of a species on day i

f_i = sampling fraction = fraction of time used in collecting sample fish[2]

p_i = proportion of flow from which sample fish taken = index flow

If the sampling fraction changes during the day, an average sampling fraction is calculated by weighting the known sampling fraction in each period (partial day) by the estimate of total fish collected in that period. The average sampling fraction is then used as the sampling fraction for that day.

The FPI for day i is given by

$$(FPI)_i = \frac{y_i}{f_i p_i} = I_i$$

[2]The sampling fraction is a weighted average on days when the sampling fraction varies.

Establishing L strata, with n_h days in each stratum, the estimated variance of the cumulative FPI over the n_h days in the hth stratum, is

$$\hat{V}(I_h) = n_h s_h^2$$

where s_h^2 is the sample variance of the daily FPI in the hth stratum calculated in the usual way.

The annual FPI (API) is obtained by summing the I_h over all strata, i.e.

$$API = \sum I_h$$

and the estimated variance of the API from

$$\hat{V}(API) = \sum \hat{V}(I_h)$$

The (1-α) CI for the API is given by

$$API \pm Z_\alpha \sqrt{\hat{V}(API)}$$

where Z_α is the normal deviate at $P = \alpha$.

The above confidence interval should work well when the sampling fraction within each day is fairly small. If the sampling fraction is large on a number of days each year, the Fish Passage Center may want to again consult with the Biometrician's Group to have them consider ways to reduce the positive bias of the recommended method. If the sampling fraction on any particular day reaches 100%, some adjustment should be made to account for the fact that the sampling variation is zero for the FPI on that day. One approach that could be considered by the Biometrician Group is to break up the stratum in question, placing the day with less than 100% sampling into an adjacent stratum and placing the 100% sampling day into its own stratum with a known total and, hence, zero variance.

Amendment to

January 1990
Report of Biometrician Group
on a confidence interval for the
cumulative fish passage index

October 1990

Biometrician Group members

Cliff Pereira

Lyle Calvin

Doug Neeley

The Biometrician Group on pages 10 and 11 of the report suggested that, in-addition to the recommended paired-day stratification system, a stratification procedure might be considered that structured strata so that, on days when 100% sampling were realized, a stratum be created to contain those days. This stratum would then have a zero variance associated with the number of fish passing through the bypass.

The Group, based on discussions in a meeting held on 10 October 1990, recommends that the paired-day stratification still be used but that the suggestion for considering a stratum of 100% sampling days be ignored for the time being. The rationale for this recommendation stem from legitiment observations made by Tom Berggren: 100% sampling would only account for the fish in the bypass system; it does not account for FGE. Bypass sampling represents a conditional sampling in that it depends on the number of fish that entered the bypass system. There is another sampling source associated with probabilities of fish entering the bypass. Since this source of sampling is outside the control of the sampling effort, we do not know the variance associated with this source.

According to Tom Berggren, there is a 100% bypass sampling at Rock Island Dam, but the FGE at that project is assumed to be about 5%, meaning that only the 5% of the population entering the gatewell is actually 100% sampled. Apparently, a true 100% sampling rate is rarely approached.

The Group now suggests that no adjustment for finite population be made at the present time. The recommended cumulative-count variance estimate based on paired-day strata estimates of daily count variances is expected to be positively biased. This bias will lead to conservative confidence intervals. The reasons for the positive bias are as follows:

1. There would probably be negative covariances among daily counts because, for a given contribution of fish, an increase in number passing through on one day would tend to reduce the number passing through on subsequent days. These negative covariances are not taken into account. The failure to estimate and add these negative covariances in obtaining the estimated variance of the cumulative count would result in a positive bias in the cumulative count variance estimate. (Failure to account for finite population sampling is included in this source of bias.)

2. The between paired-day variance estimates for the daily count variances is biased. Their expectations include the variances between expected paired-day counts as well as the variances associated with the daily counts. The variance between expected paired-day counts is a source of positive bias.

3. There is an additional source of bias in the estimates of the daily-count variances. The estimates include the subtraction of negative paired-day covariances (mentioned under 1). Subtraction of negatives is an additional source of positive bias.

If less biased estimates of variance of the cumulative count are desired, the Fish Passage Center may want to contact the Biometrics Group to explore and, if necessary and possible, to adjust for these biases.