FISH PASSAGE CENTER

1827 NE 44th Ave., Suite 240, Portland, OR 97213
Phone: (503) 230-4099 Fax: (503) 230-7559
http://www.fpc.org/
e-mail us at fpcstaff@fpc.org

MEMORANDUM

TO: Agnes Lut, Oregon Department of Environmental Quality
Andrew Kolosseus, Washington Department of Ecology

FROM: Fish Passage Center Staff

DATE: June 2, 2008

RE: Response to comments on FPC's Presentation "Importance of spill in Juvenile Hydro-system Survivals and SARs"

Thank you for the opportunity to respond to the comments provided on behalf of BPA by John Skalski and Jim Anderson. We realize that BPA requested the reviewer's to comment on a power point presentation and they were not present at the AMT meeting when the presentation was given. Consequently, some of their reviewer's comments show a misinterpretation of our analysis based on the difficulty of interpreting the presentation. Other comments have provided useful to further improving the ongoing analysis.

Response to Comments from Jim Anderson

"Prior to any claims for the impacts of spill on fish McCann needs to include causative factors which are currently ignored in the analysis."

We included a seasonally progressive variable called "Date Group" in our analysis. Many environmental factors related to juvenile reach survival are highly correlated. We considered this variable to be a surrogate for temperature, day length and perhaps other seasonally progressive variables. As such Date Group was important in explaining the variability in reach survivals—see slides numbered 6 and 7. And for adult returns (especially for steelhead) we found a high relative variable importance for Date Group of 0.98 compared 1.000 for average spill proportion

(Slide 32). This would indicate that temperature and other seasonally progressive variables are important. We were asked to focus on the effects of spill in this presentation. The predictive model we used included an estimated coefficient for spill effects in the presence of other variables considered including the flow variable WTT and the seasonal variable Date Group. As such the bivariate predictive plots are not simply based upon a bivariate regression analysis, but rather use a multivariate approach and isolate the effects of one environmental variable, by holding other variables constant.

"McCann's hypothesis is just the opposite. He claims that SAR is independent of fish conditions at very low passage survivals and SAR is highly sensitive to passage conditions at high passage survivals. Analyses relating passage conditions to ocean survival are complex and it is inappropriate to make statements on the coupling from simplistic correlations."

It must be understood that the impetus for presenting juvenile survival plots versus SAR's was in response to several presentations and reports by NOAA Fisheries who used this approach to imply that juvenile reach survival was not correlated to adult returns. And therefore ocean productivity was the only important factor in explaining adult return rates. From the lack of correlation they found between juvenile reach survival and SAR's they have discounted the role of hydro-system survival in affecting adult returns. My analysis began with an exploration of that method of analyzing the role of reach survival in explaining adult returns. That lead FPC to develop ocean index criteria and finally to the multivariate approach to predicting SAR's based on both freshwater and ocean productivity variables.

As frustrating as fisheries science can be due to lack of data, the so-called "simplistic correlations" we employ have been widely used to explore the relationships between in-river survival and adult returns rates. Anderson seems to argue for more complex models that require complex parameterizations and large numbers of assumptions, which may provide interesting insights into potential mechanisms of survival, but in this case would lack solid grounding in empirically measured data.

"The claim that that spill is wholly responsible for the relationship is unfounded."

No where did we claim that spill is wholly responsible for the relationship between reach survival and spill proportion. The plot shown was simply a step used to check the fit of the model predictions with the observed data. Perhaps this is an example of reading too much into a figure. The figure should be interpreted as the predicted reach survival with varying average spill proportion with the influence of other variables. In other words, the other variables are affecting survival, it is simply an attempt to measure the predictive capability of spill versus observed data. As can be seen from the plot the model shows less variability than the empirical data.

Response to Overall Comments from John Skalski

1. In the analysis of inriver juvenile survival, only three covariates were considered:
a. Average spill proportion
b. Water particle travel time
c. Inseason group release order
Other variables such as water temperature, turbidity, etc., also known to affect survival were not considered. A thorough analysis should have used a wider range of potential covariates.

See comments above for Jim Anderson. We should mention that reliable turbidity data was not available or all sites and years included in this analysis. Further, we focused on environmental variables that could potentially be affected by managers such as water transit time, spill and date of migrations. While water transit times and spill can be readily manipulated in the hydro-system, perhaps date of passage is considered more a function of annual conditions. However, the latter might be more difficult to manipulate except that the start of transportation or date of hatchery releases could affect population timing distribution to some extent. It is unclear how turbidity or temperature could be manipulated in the spring, since both are inexorably linked to flows and runoff.

2. The so-called "relative weight of evidence" for the individual variables is a term formulated by the presenter. Burnham and Anderson (2002), developers of the concept, call it "relative variable importance" (RVI). It expresses the relative strength of the covariates to each other and has no absolute meaning. The value of RVI depends on the models considered and is very sensitive to the set of models analyzed.

We stand corrected. The term we used "relative weight of evidence" is not technically the correct term. It should be called "relative variable importance". However, we should point out that in their text Burnham and Anderson use the terms "relative weight of evidence and "relative variable importance" interchangeably. It does not change how we would interpret the analysis.

3. In the first section on inriver survivals, the calculation of the RVI appears to be correct. A balanced set of models with each variable represented an equal number of times was used. In the last section on SARs, however, the calculation of RVI values seems to be incorrect. It appears from the presentation that all or most of the models compared included average spill proportion as a variable. If this is the case, the RVI for that variable will have a value at or near 1 regardless of its importance, simply because of the nature of the models compared. In this situation, the RVI values are not reasonable.

This comment is fair given the data presented and represents the problem with reviewing a PPT presentation without input from the presenter. The models presented were only the top several models that had weights of > 0.01 as seen in slides 28 to 31. All models estimated for each model selection would have numbered between 40 and 60 for each relationship. For brevity only the top models were included in the presentation.

Response to Detailed comments from John Skalski

A. Juvenile salmon reach survival analyses

1. Overall, the analyses found average spill proportions to be a useful predictor of spring/summer Chinook salmon smolt inriver survival. However, for steelhead, average spill proportions were no more useful than the other two covariates (water particle travel time and date of group release).

We agree. Our intent was not to downplay the importance of other variables, rather to explore the importance of spill as a management tool for improving reach survival.

2. The model selection procedure was appropriate with a balanced design of alternative models for the three covariates considered. Of greater concern, however, is that only three covariates were considered in the analyses. Other variables such as water temperature and turbidity, known to affect inriver survival, were not considered. If the goal is to investigate spill− survival relationships, the best way to find it is to exclude alternative explanations.

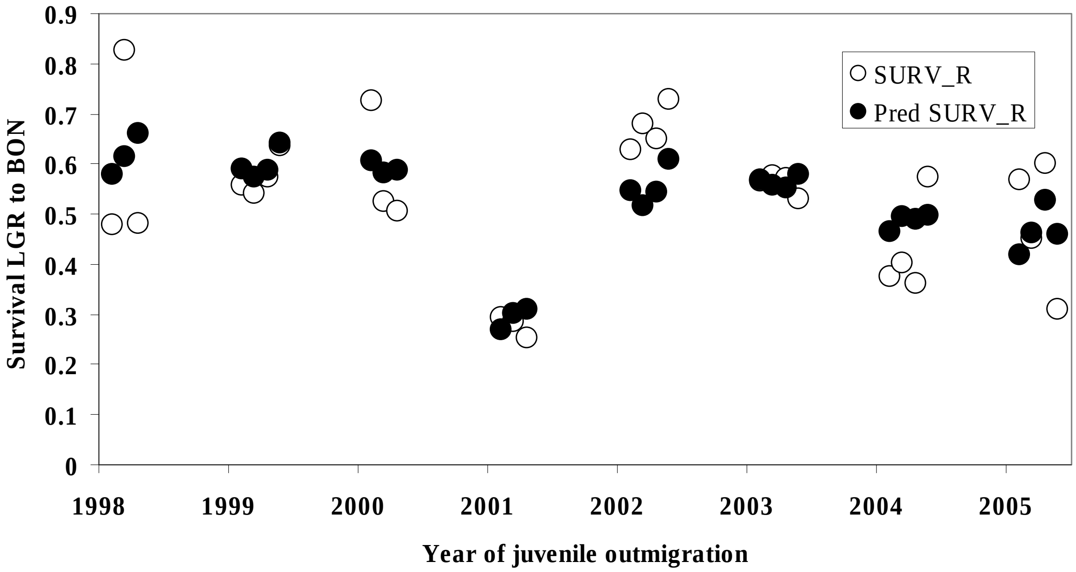
See response to Jim Anderson's first comment.

3. The model averaging approach of Burnham and Anderson (2002) should be based on only "reasonable" models. This means biologically and quantitatively reasonable models. Unfortunately, "reasonable" is in the eye of the beholder. The author appears to have averaged all models tested. The other approach would consider only models with delta AIC_c of 10 or less from the best model. The consummate model would then change.

We chose a list of variables that have been demonstrated to be important in explaining the variability in adult return rates of salmon, as well as reach survivals. We used relatively simple combinations of those variables to avoid overly complex and unrealistic models given the strength of the dataset. As such we did not consider any of our models "unreasonable" nor does Dr. Skalski provide examples of models he considered unrealistic. Instead he cautions us against considering too few models in an earlier comment. Burnham and Andersons' method of using weighted coefficients to develop an overall weighted model essentially selects the top several models because the value of the coefficient in the overall model is a function of the weighted values of the coefficients in the models. Thus the relative contribution to the value of any coefficient in the weighted predictive model from a model with a weight less than 0.05 for example would be trivial or approaching none.

4. The fitted model for inriver survival of yearling Chinook salmon seems to possess systematic bias. For the majority of years, predictions exceed observed values.

We have recalculated the values in the predictive model and those can be seen in the corrected figure shown below. In this figure there appears to be no bias.



5. There is a contradiction in the regression results. The multivariate final models have smaller R_2 values than the univariate models with only average spill proportions (Chinook salmon: Model average, R_2 = 0.5198, univariate = 0.5347; steelhead: Model average, R_2 = 0.534, univariate = 0.5467). The R_2 expresses the proportion of variability explained by the regression model. In classical multiple regression, R_2 values increase monotonically as the number of variables increases.
6. In the univariate regression of in survival against average spill proportion, the author also included a regression of the predicted values (black dots) against spill and obtained high R_2 values. For example, for yearling Chinook salmon, R_2 = 0.9103; steelhead, R_2 = 0.8301). In short, the author used spill to predict survival and then used the predictions to refit the spill model. The resulting conclusion that spill is a good predictor is circular in logic, and thus cannot be supported.

We agree that the use of the R2 value reported on the figure mentioned in his comments is not appropriate. That figure was meant as simply a representation of the way in which the model predicted reach survival using spill as the variable. It shows a fairly representative fit, with similar slope to the curve. However the model underestimates variability when a single variable such as spill is allowed to vary and other variables are held at constant values. That was the point of the figure as mentioned above in responses to Jim Andersons' comments.

B. Juvenile salmon reach survival and relationship to adult returns

1. The presentation uses a creative way of ranking years by oceanographic

conditions. However, transforming the four continuous oceanographic variables into a single categorical variable is fraught with problems. Bishop et al. (1975:301) suggests such approaches may be useful as a preliminary step for detecting interactions but the consummate analysis should use the continuous values to avoid loss of information. There are other concerns as well.
a. The choice to subdivide the ocean data into thirds is arbitrary.
b. The ranking weights the four oceanographic variables equally, which may or may not be correct if they are correlated—in which case, methods such as principal component analysis would be more appropriate.
c. The classification of years into good, moderate, and poor is again arbitrary but also inconsistent. Using the values good (3), moderate (2), and poor (1), as specified in the presentation, the various classifications can be given an overall score. The overall scores indicate contradictions in classification. For example, there are two oceanographic conditions with a score of 9, one classified as a good year, the other as moderate. There are two oceanographic conditions with a score of 7, one classified as moderate, the other as poor.

The division of ocean data into thirds was arbitrary and that was stated in the presentation. Using a division our groupings of poor, average and good ocean productivity designations agreed well with those that Scheurell and Williams 2005 identified as good productivity years in their publication. The so-called contradictions identified within overall ocean productivity scores is a reality of the ocean data. Most years do not show all indices lining up as good or bad, typically two or three having indices associated with good returns are the best years. This is not a contradiction but the nature of observational data sets.

The “overall scores” column in his comments that he reported and refers to in his comments are not scores we reported. It appears that the numbers in parentheses that accompany each category confused him. Those are not values assigned to each ranked number. We apologize if this was confusing but they did not affect how data were assigned to ocean productivity categories. His comment about a “more rigorous use of the data” anticipates the analysis that is reported on in the multivariate analysis section where we used raw values of ocean productivity in regression analysis.

2. The analyses use an inappropriate scale for the observations. Because oceanographic conditions are measured at an annual scale, use of the biweekly release groups is pseudo-replication. Accordingly, the author has four years of poor oceanographic conditions, two years moderate and five years of good on which to base these analyses. All releases within the same year share the same oceanographic conditions, so they are not independent observations but repeated measures. Thus, analyses should have been based on the 11 yearly observations or adjusted for subsampling. Note in the above figure, SARs within a year are clustered, illustrating their lack of independence within a year, particularly for yearling Chinook salmon.

The scale of observations is not inappropriate nor is it pseudo-replication nor repeated measures for that matter. This type of data set is called hierarchical or multilevel data set. Pseudo-replication is a term that was used widely to criticize multiple observations that may be correlated in space and time. However, in the past 15 years such observational and even experimentally controlled data sets have been routinely analyzed in the ecological literature using mixed regression models. Repeated measures are repeated observations of a single group or individual. Our analysis uses separate time blocks and measures survival on those fish. We routinely see differences in both reach survival and adult return rates based on passage timing through the hydrosystem as a result of changing conditions either in the fish, the hydrosystem. The terms pseudo-replication and repeated measures are incorrect in the case of this analysis.

We are in the process of analyzing the data using a mixed regression model. Our preliminary review suggests that the variable importance will not change greatly in the final version.

3. Note that while SAR curves were generated for "good ocean" and "moderate ocean" data, they were not produced for poor ocean conditions. For both yearling Chinook salmon and steelhead, their data indicate SARs are unaffected by inriver survival during poor ocean years. Under any inriver conditions, SARs are at or near zero under poor ocean conditions.

We found no relation between poor ocean conditions as we defined them and SAR's. Because we found no significant relation, we did not fit a curve to the data.

Relationship between SAR (and SOA) and both inriver and ocean variables

1. The relative variable importance (RVI) results in this section share the same interpretation problem mentioned in the previous section (see comment A7). Again, high values do not imply high absolute importance. The higher the value, the more important this variable is relative to the other variables considered.

We pointed out in the presentation that these were measures of the importance of each variable given the data analyzed and the models considered to describe the relationships.

...This time, however, there is also a fundamental analysis flaw. Burnham and Anderson (2002:169) state when assessing the relative importance of variables, it is necessary to achieve a balance in the number of models that contain each variable. It appears that all or at least perhaps most of the models analyzed included the variable for average proportion spill.

...By design, then, this action will always result in that variable having an RVI value at or near 1, as reported. As such, the results reported are of little value...

The statement is incorrect and the flaw he identifies does not exist. The comments are likely due to the attempt by John Skalski to review a power point presentation without having heard the presentation. We only presented the top several models (those with weights greater than 0.01) on each slide for sake of brevity. We considered and included in the weighting process many more models which had much lower ranks. We understand the importance of balancing the number of variables contained in the models considered in the weighting and we did exactly that. Consequently, the results are not devalued as he suggests, rather the comment is of little value because he did not understand that he was reviewing an abbreviated table.

2. In plotting percent fish passing in spill vs. percent spill, the spill efficiency curves that are usually expressed in terms of the ratio may not be familiar to most people. Note, therefore, an odds ratio of 4:1 does not mean a spill efficiency of 4:1.

The odds ratio is not equivalent with a spill efficiency of 4:1. We used the odds based approach to fit a curve to the data. We clarified that point.