State, Federal and Tribal Fishery Agencies Joint Technical Staff

Columbia River Inter-Tribal Fish Commission
Idaho Department of Fish and Game
Oregon Department of Fish and Wildlife
US Fish and Wildlife Service

May 22, 2006

Dr. Richard Zabel
NOAA Fisheries
2725 Montlake Blvd E.
Seattle, WA 98112

Dear Dr. Zabel:

Thank you for the opportunity to review the NOAA Fisheries response to the ISAB comments on NOAA Fisheries' COMPASS model. In a separate May 22, 2006 letter to you we commented on our general concerns regarding the development and structure with COMPASS. In this letter we are providing comments specific to issues raised by the ISAB review of COMPASS.

Overall, we think that the ISAB provided thoughtful comments on the COMPASS model and we share many of their concerns raised. As the model is further developed and individual components of the model are better understood, additional concerns as well as our long-standing concerns are raised. We are providing the following comments for your consideration in NOAA Fisheries' response to the ISAB as well as in continuing development of the model.

While we believe that the ISAB review was premature (FPC memo, February 2006), we share the following preliminary concerns identified by the ISAB:

- **the elimination of data does not appear warranted,**
- **the effects of eliminating data versus weighted versus unweighted model fitting is an important issue that needs further evaluation,**
- **the fish movement component of COMPASS poorly approximates previous observations,**
- **the model has been overfitted to the data and**
- **there has not been enough emphasis on relating model parameters to biological mechanisms.**

We agree with NOAA Fisheries' analysis and justification regarding the use of CJS variance estimates rather than binomial variance estimates mentioned as a first approximation by the ISAB. However, developing a reasonable analytical approach for dealing with differences in precision between the estimates is one of our primary concerns.

We strongly disagree with NOAA Fisheries' analysis and justification regarding their decision to use standard error cutoffs to eliminate data for the following reasons:

> First, we do not agree with NOAA Fisheries' insistence from the onset of this modeling effort that data had to be eliminated. It is inconsistent with the principle that the data should dictate, rather than analysts should dictate, the appropriate models which best describe the data, as stated in earlier COMPASS documents. Elimination of data sets from the "tails" of the migration often represents the weakest stocks and unique life histories that we are specifically trying to protect.
>
> Second, the cutoff approach greatly reduces the contrast in the effects of flow and spill in comparison to an inverse-variance weighted approach using the full data set. Years 1999 and 2001 provided highly-contrasting environmental conditions of high flow and spill, and low flow and spill, respectively. NOAA Fisheries' cutoff approach greatly reduces the influence of these contrasting conditions in comparison to the inverse-variance weighting approach, virtually ensuring that effects of variable levels of flow and spill will be undetected or minimized. The likely result is that COMPASS will be insensitive to changes in flow or spill which have been shown to be important factors affecting survival.
>
> Third, the characterization and effects of the decision to eliminate data is not as simple as was portrayed in NOAA Fisheries' draft response to the ISAB. An inverse-variance weighting approach is not equivalent to equal weighting for data below the cutoff threshold and zero weighting for data above the threshold. Rather, the inverse-variance weighting approach results in different weights being applied to data below and above an analyst-chosen threshold, according to the variance estimates. With the apparent controversy over this weighting issue, it would be reasonable to employ a model selection criterion (e.g., AIC, AICC, BIC, or DIC) to arbitrate between the various estimation approaches to determine which approach results in a model which best explains the variability in the survival data (Burnham and Anderson 2002).
>
> Fourth, we do not agree with NOAA Fisheries' decision to use different cutoffs for different species and different reaches. The cutoffs selected for Chinook are 0.11 and 0.22 for LGR releases and MCN releases. The cutoffs selected for steelhead are 0.15 and 0.30 for LGR releases and MCN releases. Even if there is a variance threshold at which the data become uninformative (and we do not believe that there is one), then why is it not the same across reaches and species? The cutoffs selected result in the elimination of 29-35% of the data, which we believe is excessive. While these data may be imprecise, they may capture important contrast in the environmental conditions (see second comment above).

With this controversy surrounding data elimination and weighting approaches, we are pleased that NOAA Fisheries, in their draft response to the ISAB, has decided to use inverse-variance weighting during the current round of modeling. If combined with a model selection criterion, we believe that this is a reasonable and appropriate approach for evaluating the effects of environmental conditions and hydrosystem operations on in-river survival rates.

We agree with the point that survival estimates greater than one are not necessarily a problem as long as the expected values are within (0, 1). However, we have concerns about how projections of survival with stochasticity, whereby survivals could be greater than one, will be handled. In the draft response to the ISAB, NOAA Fisheries presented five options on how the issue of survival estimates greater than 1.0 could be handled:

> Option 1: Fit models using the current approach but exclude models that have parameter estimates or combinations of parameters that result predicted values > 1.0 within the range of data. We are unclear about how this could be easily implemented, and at this point do not believe it to be a reasonable approach. If the best-fitting model(s) result in predictions greater than one, will poorer-fitting models then be used to generate predictions? What model would be used to generate predictions for the combinations of parameters result in expected values greater than one, or predictions greater than one when stochasticity is imposed?
>
> Option 2: Select the best models without parameter constraints and then set a constraint in COMPASS that predicted survival must be < 1.0. This approach seems reasonable, but the constraint may not be necessary. If it is implemented with stochasticity, there may be a need to capture how residual variability about the expected value changes as a function of the expected value. That is, variability about the mean could differ as a function of the mean (heteroscedasticity), and if so this would be important to capture.
>
> Option 3: Set constraints on regression parameters so that the estimated mortality rates $\lambda = X\beta$ are always positive. This option represents a change from modeling survival to modeling mortality. Given the complications surrounding exactly how it would be implemented and the numerous subtle analytical decisions that would be entailed, it does not appear to be justified at this time. We are also concerned that this approach would not be easily transparent.
>
> Option 4: Truncate data points to have survival < 1.0. We agree that this could bias the survival function and will not necessarily keep predicted values < 1.0. However, it is unclear how "weight[ing] the truncated values in a way such that their influence on the model would mimic that of untruncated values" could be accomplished. This approach seems convoluted and does not appear reasonable to us.
>
> Option 5: Use a transformation of survival that constrains survival to (0,1) space. There are several different options on how one could implement this approach, and there may be an appropriate link function and/or transformation that has desirable statistical properties and can describe the data well. We believe this option deserves additional consideration.

Among this set of options, we believe that option 2 is most reasonable and easily tractable, but option 5 has potential and deserves additional consideration. We therefore disagree with NOAA Fisheries' decision to implement option 1.

As was mentioned in the description of Table 2, the back-transformation of the mean of logit transformed values is not equal to the mean value on the untransformed scale (i.e., back-transformation bias). Given this fact, not properly adjusting for this bias in the presentation of the results in Table 2 gives a misleading impression of the effects of the alternative transformation/estimation approaches evaluated by NOAA Fisheries.

Contrary to the approach used in the COMPASS modeling thus far, we do not agree with modeling survival per-unit-travel-time as the dependent variable. We believe that the dependent variable should simply be survival. This is an important distinction that needs to be made between our view on the most important metric (survival) and that which has been expressed in NOAA Fisheries' COMPASS manual and was mentioned in NOAA Fisheries' draft response to the ISAB (modeling survival per-unit-distance and/or survival per-unit-travel-time as the dependent variable). We disagree with this approach for several reasons. First, in evaluating the relative performance of potential hydrosystem management scenarios on smolts, we are primarily interested in what the expected survival will be, not the survival per-unit-travel-time. Second, previous analyses have shown that survival is highly correlated with water travel time, and to a lesser degree with fish travel time and distance traveled. Therefore it would be inappropriate to divide the dependent variable (survival) by these correlated, and possibly explanatory, factors. Third, the PIT tag survival estimates represent survival between two fixed detection sites, not an instantaneous census of surviving individuals at a point in time. Modeling survival per unit travel time (which varies across individual fish) is inconsistent with how the data are collected (using detections across the season) and what the resulting survival rate estimates represent (survival between release and detection sites). Fourth, because survival rates likely change as smolts grow, develop and migrate through the hydrosystem, even an attempt to standardize the survival rates using per-unit-distance may not be appropriate.

We are also concerned over the inclusion of distance in the model as an explanatory variable. There are several reasons why distance seems problematic as an explanatory variable for modeling survival. First, data from the Snake River show that survival decreased when the dams were built and impoundment occurred, whereas the distance did not change. Second, there is considerable variability in reservoir volumes, which is not captured by their lengths. For example, John Day pool has the approximate volume of all four Snake River reservoirs combined (or all five Upper Columbia Reservoirs combined), well out of proportion to its length. We continue to urge the use of water travel time as an independent variable for explaining smolt survival. Because water travel time accounts for both flow and reservoir volume, and because water travel time is a reasonable surrogate for fish travel time, there is a biological connection with survival that does not exist with distance.

Including distance as an explanatory variable would cause the COMPASS model to be insensitive for evaluating certain management options. Clearly, under breaching or drawdown scenarios, flow and distance would not change whereas water travel time would. Of more immediate importance, the projected benefits of Removable Spillway Weirs (RSWs) will probably depend to a large extent on reductions in forebay delay. Thus, simulations with the COMPASS model would likely underestimate the potential efficacy of RSWs.

We have significant concerns with how the COMPASS model is structured with respect to separating reservoir and dam mortality from the reach survival estimates derived from PIT-tag data. Our primary objection is to the practice of calculating reservoir survival values when no dam survival data was collected in the year/project being evaluated. Our objections are exacerbated by the practice of implementing this approach within-season, when almost no study in the entire system has calculated within-season changes in dam survival. Because of these objections, it is difficult for us to comment on the NOAA Fisheries response to the ISAB on model complexity in the reservoir survival functions. We believe that COMPASS has major problems with over-parameterization and complexity, but these problems do not solely reside within the reservoir functions of the model. The ability of the model to predict the behavior of fish is being portrayed at a level of detail that goes well beyond the precision of the underlying fisheries data. Even the ability to separate dam mortality from reservoir mortality on an annual basis is dubious.

We share the ISAB concern over the lack of fit in the travel time model presented in Zabel and Anderson (1997). The NOAA Fisheries draft response to the ISAB includes a plot of yearling Chinook arrival distribution at McNary Dam in 2000 from Zabel (2002). We need to know more about the differences in model form and predictive capacity between the work presented in Zabel and Anderson (1997) versus that presented in Zabel (2002) to evaluate this component of the COMPASS model.

The ISAB suggests (page 8) that the COMPASS documentation's claim that years of very low flow are associated with poor survivals is overstated. The ISAB's apparent belief is that the sole evidence for this relationship is "one anomalous observation." In fact, evidence supporting poor expected survival in years with poor migration conditions (low flows, little or no spill, high water travel time) is found in survival rate estimates for multiple species in multiple years, as well as in studies of smolt behavior in the FCRPS.

In historically low flow years other than 2001, reach survival estimates for steelhead and yearling migrant Chinook have consistently been low. Williams et al. (2001) used reported survival rate from the late 1960s onwards to estimate what they called "overall system survival"—i.e., survival from the upper reservoir on the Snake River to the tailrace of Bonneville Dam. The two lowest flow years for which estimates were available for yearling Chinook and steelhead at the time of that paper were 1973 and 1977. The overall system survival estimates for these years were by far the lowest in the time series for both Chinook and steelhead. Similarly, USFWS (2003) estimated survival rate per fixed distance for PIT-tag data for 1994-2001, and found that the 2001 survival estimate was by far the lowest for steelhead and the second lowest for Chinook. In the time series, spill and flow during the spring migration season were by far the lowest in 2001. Survival estimates for 2001 were from Zabel et al. (2002).

In 2003, the ISAB reviewed the evidence for relationships between flow and smolt survival through the hydrosystem. From the executive summary of ISAB (2003), pg. 3:

First, for the lower Snake River studied between 1996 and 2001, there is a range of lower flows over which survival of PIT-tagged smolts increases with increasing flow and a range of high flows in which fish survival appears to be independent of incremental changes in flow. There is an apparent breakpoint in the relationship between survival and flow (i.e., a rather sharp change in the slope of models fitted to the data or a sharp corner on a “broken-stick” model) near 100 kcfs for yearling Chinook salmon and steelhead and fan appearance of a break near 50 kcfs for underyearling fall chinook salmon. The pattern is evident despite a large amount of scatter in the data. Whether these apparent breakpoints represent two distinctly different mechanisms at different flows or one continuous mechanism with a general curve is a subject we discuss. Survival data within years do not consistently show a pattern because the range of flows within a given year is relatively small.

Second, radiotelemetry of yearling chinook salmon and steelhead smolts in Lower Granite Reservoir (1996-2001) also showed an apparent breakpoint in the relationship of smolt behavior and flow near 100 kcfs. Above that level, smolts moved through the reservoir at decreasing velocities as flow decreased and the dam was approached, but passage was generally not hindered. At flows below 100 kcfs, smolt migrations were slowed markedly at the dam, often to zero, and they began extensive upstream movements. These upstream wanderings often exceeded the length of Lower Granite pool and took many days, potentially exposing smolts to more predators, excessive use of stored energy, or other factors affecting survival.

We thank you for this opportunity to provide comments on NOAA Fisheries’ draft response to the ISAB comments on the COMPASS model. It has been our intent to provide constructive criticism to help increase the likelihood that the model developed will be acceptable to all fishery management agencies in the Region.

Sincerely,

Raymond R. Boyce

Ron Boyce, ODFW

Earl C Weber

Earl Weber, CRITFC

Charles Petrosky, IDFG

Steve Haeseker, PhD, USFWS