Independent Scientific Advisory Board
for the Northwest Power and Conservation Council,
Columbia River Basin Indian Tribes,
and National Marine Fisheries Service
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204
ISAB@nwcouncil.org

December 2006 Review of the COMPASS Model, Version 1.0

Robert Bilby
Susan Hanna
Nancy Huntly, Chair
Stuart Hurlbert
Roland Lamberson
Colin Levings
William Pearcy
Thomas Poe
Peter Smouse

Richard Alldredge, Ad Hoc Member, ISRP

ISAB 2006-7
December 15, 2006

ISAB Review of the COMPASS Model, Version 1.0

Contents

Background 1
Major Comments.. 2
(1) Specificity of the modeling platform 2
(2) Survival probabilities > 1.0 2
(3) Modeling choices 2
(4) Stochasticity 3
(5) Partitioning the survival variation 3
Minor Comments 5
Appendices.. 7
Post-Bonneville Survival 11

ISAB Review of the COMPASS Model, Version 1.0

Background

At the request of NOAA Fisheries (NOAA), the ISAB provides this report, which assesses the COMPASS Model, Version 1.0. This report is the third in a series of ISAB reports pertaining to the development of this new comprehensive fish passage model, which was created by NOAA Fisheries, along with federal, state, tribal agencies, and the University of Washington for use in developing the new Federal Columbia River Power System Biological Opinion (BiOp). In March 2006, the ISAB completed its first review of the then partially completed COMPASS model specifically addressing several questions regarding the model capabilities, complexity, data usage, statistical protocols, documentation, and graphical interface (ISAB 2006-2[1]). The ISAB concluded that the new COMPASS model should prove to be a welcome addition to the analytical tools available to both scientists and managers alike. The ISAB's critique was explicitly intended to provide a series of strong but constructive suggestions to facilitate the continuing development of what should be a valuable new modeling tool for the region.

The ISAB's second review was a reply to the COMPASS team's responses to the ISAB's initial review (ISAB 2006-6[2]). The points at issue for both the ISAB's report and for NOAA's response were largely confined to statistical usage, over which there remained some differences of opinion that needed further discussion and resolution. The ISAB was encouraged by the efforts of the COMPASS team and provided comments to further the team's discussions and development of the model. This review gives us a chance to see how effectively our earlier comments were incorporated in the model and to again offer constructive suggestions to improve the model.

For this review, in addition to COMPASS Manual and Appendices, we have in hand a letter from Earl Weber (CRITFC) and Rick Kruger and Tim Dalton (ODFW) to Rich Zabel (NOAA), dated 19 October 2006, offering additional commentary. The CRITFC-ODFW commentary makes several good points, to which NOAA can/will respond, but we will take the liberty of flagging just a few of their comments here as well, as they impact ISAB's comments on the COMPASS submission of 13 October 2006.

The COMPASS model has improved markedly since our first examination of it in March 2006. However, we take it as our *operational* charge to report on all aspects of the "in-river" component of the model, and with this review we have not quite accomplished that charge. To complete the review of the in-river component, we need to review the missing appendices and would like to see responses to our five major comments listed in this report and to our comments on appendices 4, 6, and 9. On December 13th, the ISAB also received comments on the COMPASS model from the USFWS. The ISAB did not have time to fully consider these comments for this report but will in our next report. In fact, the ISAB expects that the COMPASS model team will incorporate the USFWS critique and model into their treatment, along with the NOAA and CRITFC models, before our next review. On a related matter, we

[1] www.nwcouncil.org/library/isab/isab2006-2.htm
[2] www.nwcouncil.org/library/isab/isab2006-6.htm

plan on completing our review of the "below-Bonneville" model component by late January 2007, at the earliest. That component is only now beginning to be modeled and is still under active development. The following is our report.

Major Comments

There are a few substantive issues that could use some further resolution.

(1) Specificity of the modeling platform

Page 1, 2nd Para from bottom – The authors indicate that they will "expand the modeling capabilities in the future to other ESUs." We interpret this statement to mean that they will have to run the calibration routine to obtain parameter estimates for the other ESUs. Some indication of what is necessary to extend application to each novel ESU would be in order. Are the authors contemplating recalibration or reconfiguration?

(2) Survival probabilities > 1.0

Page 12, Para just below CRITFC model – We remain concerned about using survival probabilities > 1.0. While it is reasonable to use the Monte Carlo runs to evaluate stochastic variation or model uncertainty, one of the things we do know about survival probabilities is that they cannot exceed 1.0. The point here is to predict real survival, allowing for such estimation noise as may exist. Since survival probability from top to bottom of the "in river" phase is not close to 1.0, the partitioning that is to follow should suffice to "smooth out the bumps." Let the estimation routine generate what it will, but then set $S = 1$ if $\hat{S} > 1$. An alternative would be to use a bounded likelihood estimation routine, perhaps with a beta prior, providing a bounded estimate $\hat{S}$. We have discussed this issue for earlier rounds of review, and the issue is that one is trading off precision (inverse variance) against the biases deriving from constrained estimation. For COMPASS modeling, we strongly recommend defining river-specific segments that do not exceed unity.

(3) Modeling choices

Page 9, bottom – Page 13, top – While we earlier accepted the argument that inverse relative variance weighting was better than no weighting at all, we commented that the choice of weights could stand some further exploration. The document needs some justification for using inverse relative variance weighting. Also, even though the negative exponential is a familiar and standard form for survival functions, justification for using this form exclusively should be provided. Is there evidence that the negative exponential is a superior model when compared to other possibilities? We have commented that an appeal to standard survival theory can only take us so far on this one, and we remain concerned that the COMPASS team may omit better modeling considerations with this choice. We are not adamant on the point, but some compelling practical rationale would be in order.

(4) Stochasticity

Page 25, top (et seq.) – One should not use the term "stochasticity" to refer to estimation uncertainty. In Monte Carlo context, *demographic stochasticity* refers to sampling for each demographic decision (live or die, thru the bypass or over the spillway or to the barge, rapid transit or slow meander), fish by fish. *Environmental stochasticity* results from the year-to-year draw of the mean survival rate from an appropriate distribution. If we knew the means and variances of those distributions, we would still have variation from trial to trial of the COMPASS model (luck of the draw on many fish and stages). *Uncertainty* arises when we do not know the means and variances of the distributions, being forced to estimate them.

It does *not* seem reasonable to draw single-fish events from the distribution of *estimation uncertainty*. A better way would be to draw all the parameters for a single run of the COMPASS model randomly from the estimated parameter distributions (for each cohort and year) and then have each fish sample in binary (or multinomial) fashion, project by project, reach by reach, decision by decision, from a single set of cohort-specific parameter values.

There is another possibility. The entire run could be set with a full set of parameters, drawn at random from the estimated parameter distribution, with each run being treated in entirely deterministic fashion, with its own particular parameter set. If each run were deterministic, but starting from a random draw from the distribution of parameter estimation uncertainty, then the delivery would be better couched in terms of "allowing for estimation uncertainty."

It remains unclear, from the delivery, exactly what is being suggested. This is either a minor verbal problem, easily fixable with a little rewriting, or it is a substantive issue with what is being done. In any event, one should not confuse estimation uncertainty with the randomness encountered by individuals. "Life is (still) a crap shoot," even when one knows the odds.

Our preference would be to seed each run with a parameter set, drawn randomly from the estimated parameter distribution, thus allowing for *estimation uncertainty*, with the idea being that each fish (running the gauntlet) would use the same parameter set. However, each fish should draw each decision randomly from the stochastic distribution specified by the chosen parameter set, to evaluate the "crap shoot."

If one were to begin with a million fish, released above Lower Granite Dam, then multiple random runs with the same parameter choices might yield final outcomes of the "crap shoot" that would vary to an insignificant degree, while different starting parameter sets might show substantially larger variation in outcomes. Should that eventuate, one could make a case for treating each COMPASS run in deterministic fashion, but with a randomly chosen parameter set. The interpretation would be that year-to-year *environmental stochasticity*, augmented by *estimation uncertainty*, is more important than the fish-to-fish *variation*, averaged over large numbers of fish. Legitimacy of such a claim would have to be established empirically.

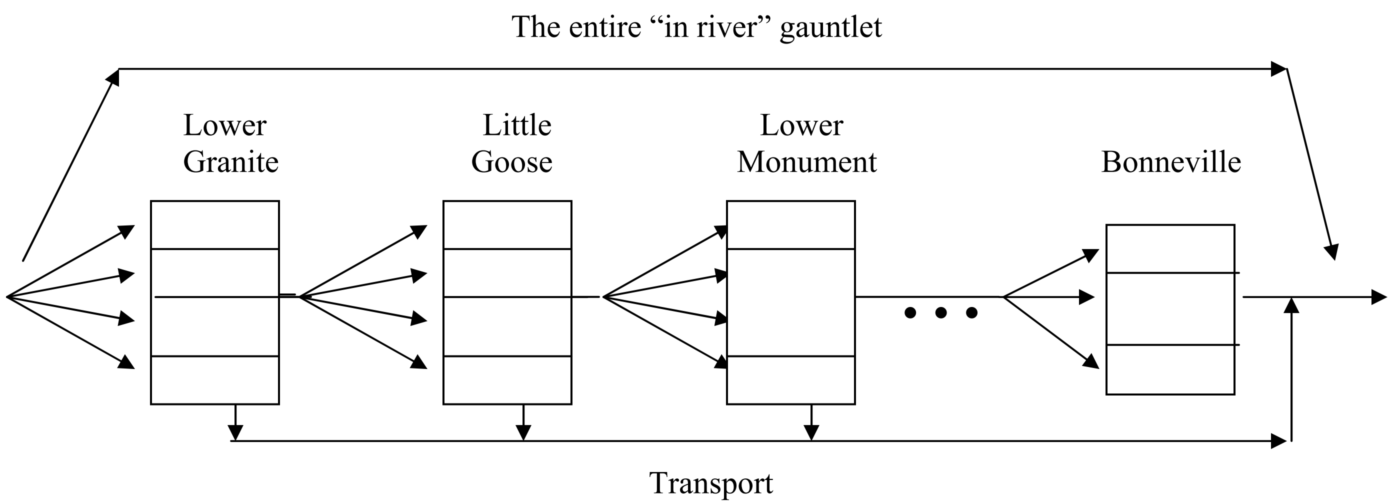
(5) Partitioning the survival variation

There are several assumptions made in the representation of project survival that should be evaluated (or at least rationalized plausibly). Among these is the assumption that weighting by inverse variances is appropriate (commented on above), that error terms are normally distributed (dubious for binary variables), and that variance is linearly related to river segment length.

(a) Having shown in Figure 12 (page 26) that the survival estimates from adjacent projects are negatively correlated ($r \sim -0.38$), as CJS-estimated survival probabilities are guaranteed to be, the authors treat the variance of the log of the survival probability of the total "in river" component as the sum of the separate variances, while ignoring the substantial negative covariances between adjacent projects as "artificial." We are not comfortable with that, and offer the following thoughts on the subject:

(b) If one had a direct estimate of the total survival probability from the "in-river" segment, the probability of successfully running the entire gauntlet, one could partition into segments as described in the document, but the reality is that there are many routes through the "in-river" segment, individually estimated, so the total is estimated as the parts, and the estimates are correlated. It would seem better to proceed in three steps.

(c) First, estimate the probability of surviving the entire "in river" segment, the probability of successfully running the gauntlet. That alone is a major challenge, given all of the alternative routes. One can imagine a large matrix equation, with (say) three or four routes through each project and subsequent survival through the next reach. One could schematize it as below (patterned after Figure 3) for eight segments. They will have to include transportation from four of the dams, a schematic such as the following,



(d) Second, construct the covariance matrix for all of the survival measures of interest, each computed as per Figure 7 on page 19. Our sense is that this is being done now. Next, take the logarithm of the separate passage probabilities (one project and reach at a time). One would have to handle the transported fish separately. The *estimated* total variance for the "in-river" gauntlet (V_T) becomes:

$$V_T = \text{Var}[\log(\text{total in river survival})] = \sum_{j=1}^{8} \text{Var}[\log(S_j)] + 2\sum_{j \neq k}^{8}\sum \text{Cov}[\log(S_j), \log(S_k)] ,$$

remembering that the covariances for adjacent steps (indeed, even those for different routes through each project) are negative. (The CRITFC-ODFW letter raises issues about multiple routes through projects, necessary if one is to evaluate operational changes.) The two-step covariances are probably less correlated. The covariance structure should be discernable from collections of multi-route/reach estimates, as is Figure 12 on Page 26.

Whether V_T is larger or smaller than the sum of the variances depends on the collective extent to which survival estimates from multi-step/multi-route reaches and projects are positively/negatively correlated, that is, whether the second term is positive or negative. If one estimates survival through the "in river" gauntlet by stitching the pieces together, then V_T must allow for both the variances and covariances of the piecewise estimates.

(e) <u>Third</u>, partition V_T into pieces that are additive and proportional to relative distances of the respective reaches, as described in the current document. We are not arguing that survival through subsequent reaches/projects is biologically correlated, though it might be in practice, or even that we should build the covariances into the partition. Take the estimated logarithm of the total "in river" survival probability, and partition it into single-reach additive pieces, but sample from V_T, not from the sum of piecewise variances.

Minor Comments

There are a number of stylistic or informational lapses that could profitably be fixed before release of this report. In reading frame:

(6) <u>Page 2, bottom</u> – How does the USFWS model differ from the NOAA-CRITFC models? Are the differences in the modeling equations or something requiring more elaborate changes?

(7) <u>Page 4, Table 1</u> – SARs are not defined before being presented in this table. It is not until later in the document that SARs are formally defined. One could do it right here in the Table, and then remind the reader once or twice in subsequent narrative.

(8) <u>Page 6, bottom</u> – The authors note that the "current routine does not produce standard errors for weighted regression", indicating that this will be rectified in the near future. We agree that measures of statistical variation are important, and the current lack of those measures represents an incompleteness of the current product that needs attention soon.

(9) <u>Page 13</u> – The first use of w_i is as an inverse relative variance weighting, while the second is for AIC-weights. This is confusing. Use different notation for the AIC-weights.

(10) <u>Pages 14 and 15, Tables 2 and 3</u> – For comparative purposes, these tables are not quite working. The information could easily be reorganized to facilitate the critical comparisons, which are <u>*not*</u> cross-species. For that reorganization, we suggest the following:

(a) Use Table 2 (see next page) exclusively for Sp/Su Chinook and Table 3 exclusively for Steelhead. The top few rows of each table should look about as follows:

(b) They will need a couple of extra rows to accommodate the CRITFC parameters. It would also be good to have a final row for the AIC criteria, so that one can see how well these two models (or any other pair) compare, relative to predictive efficacy. The USFWS model (or any other specifications) could provide entries for this table as well, once those become available. The CRITFC-ODFW memo queries the absence of the USFWS model, and the ISAB wondered about that also. The report describes it as having been only recently submitted for inclusion, but we suggest that it be added to the mix at an early opportunity. It might also be useful to provide some of the partial models and those on page 12 for the NOAA set. The tables could be presented broadside, as necessary.

Table 2: Spring/Summer Chinook Model Comparisons

Model	Modeled	Snake River		Columbia River	
Parameter	Covariate	NOAA Full	CRITFC	NOAA Full	CRITFC
γ	Grand intercept				
α_0	distance	0.00223		0.0105	
α_1	Flow-distance	–0.05937			
. . .	. . .	. . .		. . .	
AIC	Criterion				

(11) Page 16, Figure 5 - Some of the plots of log (predicted survival) against log (observed survival) suggest a bias, but it is a little difficult to tell. Plots of residuals versus predicted values would be more useful for identifying adequacy of model fitting.

(12) Page 20, Delay in Dam Passage - This is for the future, and should probably be relegated to a section at the end of the report, indicating intended future additions to (elaborations of) the COMPASS Model. Don't clutter up the delivery of what is currently available with what will eventually become available. It confuses the reader.

(13) Page 21, top – Reference to the Gurarie model should be relegated to the "future intentions" section described just above.

(14) Page 21, Figure 9 – The fact that the empiric peak is sharper than the model peak is probably not a biological or statistical artifact. A comment is in order.

(15) Page 22, Second line below equation – It should read "passing through the spillway" not "passing the spillway".

(16) Page 22, Table 4 – This needs repacking into a pair of tables, one for Sp/Su Chinook, one for Steelhead. The authors go to considerable trouble to present competing predictive

models of travel time, one involving distance, the other without it (page 12). It would be useful to have a table for each species, laid out in the pattern shown above, where the authors could present both the full model (top of page 22) and some of the sub-models attempted. It would help the reader determine how stable the model parameters are to alternative specifications. Again, a line for AIC Criteria at the bottom of the table would be good.

(17) Page 22, Migration Rate Model - Also, the equation has a superfluous coefficient β_5. Some of the estimated regression coefficients do not appear to make sense. For example, negative coefficients for velocity for Columbia River and spill for Snake River Chinook. Some interpretation/explanation would be useful. Also, plots of residuals versus predicted values would be useful.

Appendices

The following comments are in order.

Appendix 1**:** *Hydrological Processes* **–** We have worked our way through most of this section, and it appears to be correct, but they need to indicate what Hd is in the second figure and what (Hd + Hu/2) is in the third.

Appendix 2**:** *Calibration Routine* – This one seems well done, but the panels need to be larger, perhaps by being presented in broadside mode.

Appendix 3**:** *Sensitivity Analysis* - This appendix was provided late in our review process, and we plan to review it in our next review.

Appendix 4**:** *COMPASS User's Manual* – The large I/O Table is so large (and the entries are so small) that it cannot be read, which means the reader cannot follow along. Either break it up on separate pages, or use the current Figure to show the overall structure, and then use smaller sections of the table on separate pages (perhaps three of them) to illustrate with actual entries that might be readable and useful. It might also be useful to introduce the table for a particular example (which always helps), and then walk the user through the example.

In the 19 October letter from CRITFC and ODFW to NOAA, the Manual comes in for a number of criticisms, and there are several suggestions for inclusions. Our own view is that a manual that is large and all-encompassing becomes more of an impediment to usage than a help. The challenge is to strike the right balance between easy usage and exhaustiveness. Having said that, our sense is that all of the features mentioned in the CRITFC-ODFW letter need to be considered. A workable compromise might be to have the Manual describing the overall structure of the model, showing the reader/user how it all fits together, but directing the reader/user to specific modules for the nitty-gritty details. A useful analogy would be a website homepage, with modular components accessible as "go to" buttons. The trick is to keep an eye on the overall model, while being able to pursue nitty-gritty of all of the pieces.

The Manual will need continuous "tweaking" as user experience accumulates. It should be treated as an evolving document. In that spirit, the latent mortality section is still under active development, as of this writing, and the modeling history should be viewed as backdrop and prelude, a separate section. Both should be available in the form of "go to" buttons.

Irrespective of how much detail is presented in the Manual, in the process of rolling the COMPASS model out, the COMPASS team will have to offer a series of workshops for the many people who will want to use the instrument. Appendix 4 will not be sufficient for the new user. Teach them to use it by using it, and it will then be *used*. It might even be useful to have a series of workshops, so that the users can progressively master programmatic features, while simultaneously contributing precious feedback. The strengths and weaknesses of the product will become apparent, once it is in usage, and user feedback becomes priceless.

Appendix 5**:** *Prospective Flow Modeling* - This appendix is not yet available, but presumably dealing with some of the issues raised by the CRITFC-ODFW critique.

Appendix 6**:** *Prospective Temperature Modeling* – The object of this Appendix is to estimate daily temperatures at each project for the 1975 – 2005 period, using data on flow rates for those same days. There are substantial amounts of missing temperature data, and those that are available for the purpose are evidently of variable quality. The flow rate data are thought to be of better quality and more reliable. The correlation between them is said to be generally high ($r_{TF} > 0.90$). The following thoughts surfaced, in reading frame:

(a) The authors claim that they have not used either scroll case temperatures or flow data from Bonneville, because there are no reservoir survival estimates for Bonneville. There are about 10 years of survival estimates for Chinook yearlings. Are they referring to subyearlings? Bonneville Flow data are used (as surrogates) for the Dalles, so we conclude that flow data (and probably temperature data as well) are available for Bonneville. Snake River projects were seeded with surrogate flow data from Lower Granite and McNary.

(b) There are some worrisome points here, among them:

(1) If the object is to evaluate the connection between temperature and flow, the 1st order of business should be to fit those pairs of dates and places that have both a flow and a temperature reading, without plugging in *surrogate* data from elsewhere/elsewhen to fill the sampling holes. There should be enough complete data to do this more than adequately, particularly if one includes Bonneville in the available data set.

(2) Having estimated the regression coefficients from actual data, a 2nd order of business would be to predict the temperatures for those same data points, evaluating how well one is doing (R^2 and AIC for a model cascade). Having chosen a model that makes some general sense (fits well and is robust over time and space, relative to which variables are used to predict), one could then judiciously plug some surrogate flow data into sampling holes, and predict those temperatures for *modeling* purposes.

(3) If one has flow and temperature data from Bonneville, one should include them in the regression analysis to estimate the parameters and choose predictive variables. We are <u>*not*</u> predicting survival here, just temperature. Using terrific Bonneville data on flow and temperature to establish the pattern makes a lot more sense than using flow data from Bonneville and temperature data from the Dalles to evaluate that pattern.

(c) The authors make the point that this is a very crude, first approximation model, and we understand that, but one should strive to do better at the outset. We have commented above on data usage, which if altered, should probably raise the R^2-values, at least where actual paired data (temp, flow) do exist. In fitting the eight models of the Table that will follow; the authors convey no sense of just how the model improves as one adds variables to the predictive equation (see below). That lack needs to be rectified.

(1) It is hard to imagine how the average daily flow rate or its squared value, averaged over a 12-month period is going to be even remotely predictive of the temperature on a particular Julian day for that same year.

(2) Similar skepticism is appropriate for the average flow/average temperature on that Julian day, averaged over 30 years. They're correlated, of course, but why bother?

(3) In this situation, it would make sense to create a model cascade, moving from the simplest model to the more complex models, thusly

$$T_{j,n} = \beta_0$$

Temperature is a constant, the null hypothesis

$$T_{j,n} = \beta_0 + \beta_1 \bullet F_{j,n}$$

Temperature is linear in flow on that same day

$$T_{j,n} = \beta_0 + \beta_1 \bullet F_{j,n} + \beta_2 \bullet \overline{F}_j$$

Next comes a model that is linear in flow and the average temperature for that same day, averaged over 30 years, etc. Compute and table the R^2 and AIC values for the cascade, project by project. The object is to come up with a relatively simple model that works for all the projects (same set of parameters), but with different estimates of those parameters, project by project. As long as one is content to be "crude", there is no reason to fit exquisite improvements by addition of more and more predictors. If one were to construct a generic model, say the third equation above (just to illustrate), but with different estimates of α, β_1 and β_2 for the different projects, and which can account for no more than R^2 = 0.90 of the variation in temperature, project after project, year after year, Julian day after Julian day, that would probably suffice, pending the availability of some directly measurable (and credible) temperature data.

(d) The authors point out that the available temperature data are going to need some serious QA/QC treatment, before much more effort is invested in modeling from the estimated relationships, and we concur - inevitable and necessary. But don't overfit shaky data.

(e) Tables and Figures: All the figures are too small, and the captions cannot be deciphered.

Figures A6.2A & B and A6.3A & B. One of these pairs may be sufficient to make the point. Turn them sideways, with increasing Julian day for the *X*-axis, and frequency of fish on the *Y*-axis. The bars above the plot don't seem to add anything – suppress. There is nothing wrong with listing the quartiles, but they don't really help the visualization. One can put Chinook and Steelhead plots on one page, perhaps both sets on one page.

Figure A6.4. Use predicted temperature for the *Y*-axis and observed temperature for the *X*-Axis, do it broadside, and make it larger. Suppress the box. One also needs an R^2 and an indication of the model used for the prediction (in the figure caption).

Table A6.1. Show the cascade, with R^2 and AIC criteria. Provide the estimates and SE's to no more than three decimal places. The authors point out that all of the effects are significant, but that says no more than that they have lots of data. Significance is not the issue. What is at issue is how many loosely correlated variables we have to use, in order to raise R^2 from $R^2 = 0.90$ to $R^2 = 0.97$. What one really needs is a generic predictive model that is simple, sensible, credible, and robust. It doesn't have to be exquisitely precise and accurate, particularly since the data on which it is based are not.

Appendix 7**:** *Prospective Modeling Results* - This appendix is not yet available.

Appendix 8**:** *Dam Passage Parameters* - This appendix is not yet available, but presumably dealing with some of the issues raised by the CRITFC-ODFW critique.

Appendix 9**:** *Passage Efficiency Relationships* – The dam passage efficiency relationship models logit(P) versus logit(F), where P is the proportion of fish passing through a passage route and F is the proportion of the flow through the passage route requires more justification. As part of the justification, Appendix 9 claims that there is no guarantee that the relationship between logit (Spill Passage Efficiency) and spill is linear. One could also argue that there is no guarantee that the relationship between logit (SPE) and logit (%Spill) is linear. Is there evidence that the logit-logit model is better than the logit model?

The form of the variance estimator for SPE_{new} in Appendix 9 should be justified with additional details or a citation. In the results section of Appendix 9, many of the plots show a problem with the adequacy of the prediction. For example, it appears that the models for Bonneville spillway for Spring Chinook and Steelhead overestimate for lower Spill % and underestimate for larger %Spill. It is confusing to have the estimates of the regression coefficients for logit (Spill) while the plots are in terms of %Spill. Plots of residuals versus predicted values would be useful for evaluating the adequacy of model fitting.

In deriving the basic model, the first two equations on the first page are too small to be read. One could write them in a single line, as shown below (for example):

$$\text{SPE} = \exp\{f(\%\text{Spill})\} / \ [1 + \exp\{f(\%\text{Spill})\}] \ .$$

It seems reasonable to skip this one, and go directly for the more general form,

$$\text{SPE} = \exp\{\beta_0 + \beta_1 \bullet \%\text{Spill}\} / [1 + \exp\{\beta_0 + \beta_1 \bullet \%\text{Spill}\}],$$

which translates immediately into Eq. (1) and from there into Eq. (2). The equations are too small (at the moment) for routine readability in a Manual.

Careful examination of the Figures and estimates also yields some intriguing observations.

(a) For any given project, the intercept differs between Chinook and Steelhead, but the slope coefficient for logit (Spill) are virtually identical - very close to 1.00. Most of the curves contain very few points, and a simplified model, such as that below, could be used to fit larger amounts of data from several projects and ESUs credibly at the same time,

$$\text{logit (SPE for Species } j \text{ and project } k) = \beta_{0,jk} + \beta_1 \bullet \text{ logit (Spill)} \ ,$$

where it seems that $0.99 < \beta_1 < 1.01$, virtually 1.00. The $\beta_{0,jk}$ coefficients do vary among species and projects, and substantially so, but β_1 appears to be virtually a constant. In logit form, the lines should be virtually parallel, but starting from different intercepts.

(b) Given a preference for logit models, the plots should be presented as logit (SPE) vs logit (% Spill), not as SPE vs % Spill. The data points should be larger and more visible. The adequacy of the fits should be assessed by comparing logit residuals vs logit predictions.

Appendix 10**:** *Alternative Reservoir Survival Models* - This appendix is not yet available, but presumably dealing with some of the issues raised by the CRITFC-ODFW critique.

We note that Appendices 5, 7, 8 and 10 are still missing (at time of writing). In addition, we received Appendix 3 late in this review and have not reviewed it. This report must be viewed as incomplete, pending review of those appendices.

Post-Bonneville Survival

There were a few preliminary comments offered, and we have made a decision to evaluate the below Bonneville or “Latent Mortality” component of the model in a separate report. At the time of this writing, voluminous briefing materials had just arrived for this last component, and the below Bonneville phase of the modeling has just begun. It seems clear to us that the below Bonneville component will require substantial work by the modeling team and that it is going take some time. For us to rush even a preliminary review to completion does not seem profitable. We will review Below Bonneville next, but will not delay this report for it.

w:\em\ww\isab projects and reports\1 isab final reports\isab 2006-7 compass dec06.doc