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

Review of *draft* Viability Criteria for Application to Interior Columbia Basin Salmonid ESUs

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

John Epifanio, Ad Hoc Member, ISRP
Chris Wood, Ad Hoc Member

ISAB 2007-4
July 18, 2007

Review of *draft* Viability Criteria for Application to Interior Columbia Basin Salmonid ESUs

Contents

Review of *draft* Viability Criteria for Application to Interior Columbia Basin Salmonid ESUs

Introduction

In a March 16, 2007 letter from NOAA Fisheries, Dr. Usha Varanasi requested that the ISAB review the March, 2007 technical review draft of the *Viability Criteria for Application to Interior Columbia Basin Salmonid ESUs*.

The review request included three specific questions:

1) Are the specific approaches and methods used by the Interior Columbia Technical Recovery Team (ICTRT) to develop viability criteria scientifically sound? Are there any significant conceptual flaws in the ICTRT's approach to developing the criteria?

2) Can the criteria be assessed using the types of data that are, or could reasonably be, collected for the Interior Columbia River Salmonid evolutionary significant units (ESUs)?

3) Is the report clearly written? Are the methods described in sufficient detail for a reader to understand and replicate what was done? Are the assumptions and uncertainties about the analyses clearly described?

The answer to each of the questions is at least a qualified yes. The weakest part of the report may be the clarity of writing (undefined acronyms, typos, jargon, missing references, incorrectly numbered figures, etc.).

Abundance and Productivity

The centerpiece of a science-based recovery plan should be a rigorous approach that has as its foundation a model or models by which one can judge the strategies for recovery, progress toward recovery, and determine needs for further research, data collection, and analyses leading to adaptive management decisions. The model proposed in this plan is simple in terms of variables and parameters, but it is applied in a rather sophisticated manner. The authors develop viability curves based on a "hockey-stick" model using combinations of abundance and productivity. More specifically, they developed the viability curves using a simulation procedure that incorporates a quasi-extinction threshold. They then modified the resulting viability curves by superimposing minimum abundance thresholds. Populations are categorized by how their geometric mean abundance and productivity compare with the particular levels of vulnerability provided by the viability curves (e.g., does it fall above or below the 5% probability of extinction in the next 100 years). Variability and uncertainty are addressed by requiring that the productivity/abundance point be approximately one standard error above the viability curve claimed as the risk level for the population.

The great uncertainty regarding salmon and steelhead dynamics in the ocean precludes the development of a full life history model at the present time, so some type of simple stock recruit model is the only reasonable choice for the underlying model in the recovery plans pending an increased understanding of salmon dynamics in the ocean. The choice of the hockey-stick model will likely serve as well as a Ricker or Beverton-Holt model. Though the model appears to be applied appropriately, the accompanying text is often not completely clear on the relationship between the variables and the parameters in the implementation of this model.

Spatial Structure and Diversity

To address spatial structure and the diversity within the ESUs of concern, the authors have broken down these criteria to reflect the extent of demographic connection between and within the various levels of subgroups: constituent major population groups (MPGs) and populations. At the lowest level, populations are categorized by the spatial complexity of their spawning areas: linear, dendritic, trellis, and major spawning areas with separated minor areas (core-satellite). A categorical assessment is applied at several levels to generate a final evaluation of spatial structure and diversity. This categorical assessment applies sets of quantitative or qualitative rules to characterize various aspects of spatial structure and diversity, and a final framework is provided which guides the integration of these metrics and provides a final spatial structure and diversity rating.

In a 2003 document (ISAB 2003) the ISAB reviewed earlier work on the identification of independent salmonid populations within ESUs. In their comments the ISAB was critical of the work's lack of transparency. General guidelines had been given, but few concrete rules for their application were provided. The new document (the subject of this review) is much improved in that regard. It lays out the basis for the breakdown of the ESUs into their component parts and the assessment of these components with regard to both spatial structure and diversity – thus addressing the earlier concern about transparency. It also provides a clearer framework for aggregating the several metrics to determine a final spatial structure and diversity rating.

Final Population-level Risk Assessment

The authors develop a simple matrix or categorical approach to rating the overall population-level risk. Populations are ranked as high, moderate, low, or very low risk, with respect to both spatial structure/diversity and abundance/productivity. For example, the combination of *low* spatial structure/diversity risk and *very low* abundance/productivity risk might provide an overall *highly viable* population-level rating while *low* risk abundance and productivity with *very low* spatial structure and diversity risk may only result in an overall *viable* population-level rating.

Again, structurally this seems to be a reasonable approach to combining the two parallel assessments (spatial structure/diversity and abundance/productivity) into one overall risk status assessment.

Overall

The proposed criteria are based on a well-reasoned and well-supported set of scientific principles and are laid out in a relatively transparent fashion with clear guidelines and examples for their application. The definition of viability includes self-sustainable and persistent "without [ongoing] input of hatchery-produced fish" and is tied to an articulated general policy decision (i.e., 5% risk of extinction probability in a 100 year period), which has a basis in the NRC (1995) report. The establishment of metrics specific to Interior Columbia Basin salmon and steelhead implies objectives that are quantitatively based and establish an inherent need for monitoring and evaluation within an adaptive recovery framework. The criteria establish approximate thresholds but are effectively "silent" on the specific approach to reach a level of risk, viability, or recovery. Decisions on the specific approaches are left to the managers responsible for achieving such objectives.

The ICTRT encourages the use of multiple approaches (models) as independent lines of evidence to address uncertainties and evaluate risks because there is no single best, unbiased, or universally accepted predictor. The viability criteria appear to be developed independent of specific ESU designation and criteria (isolation and importance to species' diversity), which permits more general application beyond the focus of the specific case examined by the ICTRT. In cases where ESUs have a single MPG, criteria are more stringent in recognition of the special risks in these situations. All of these are very positive features.

It is important to emphasize the adaptive nature of the recovery plans and their implementation since the recovery of these species will not be taking place in a static environment. Global climate change is already responsible for earlier peak flows on many of the rivers and streams in the Columbia River Basin and more substantial changes are predicted for the future (ISAB 2007a). At the same time the Pacific Northwest is experiencing a significant population increase. The accompanying development will likely have a negative impact on water quality in the rivers and streams in the region along with increased demands for hydropower and water for human consumption (ISAB 2007b). All of these factors make it imperative that recovery plans incorporate rigorous monitoring and evaluation and be adaptive as the environmental changes evolve.

The Questions

1. Conceptual Issues

Are the specific approaches and methods used by the ICTRT to develop viability criteria scientifically sound? Are there any significant conceptual flaws in the ICTRT's approach to developing the criteria?

The general approach employed by the ICTRT and the underlying understanding of conservation biology and salmon biology are scientifically defensible with no apparent conceptual gaps that might negate the defensibility of the approach. The approach used to establish recovery criteria:

1) is quantifiable and objective-based;
2) recognizes ecological and demographic elements as well as evolutionary and genetic ones;
3) recognizes the interaction of compositional attributes (i.e., abundance and productivity) with structural attributes (i.e., spatial and temporal structure) and functional attributes (i.e., life-history diversity) of the ESUs in question;
4) takes a hierarchical approach toward population, meta-population, and ESU complexity;
5) provides a high-level of transparency in most areas, especially in that it articulates analytical assumptions and has built-in decision maps to guide application;
6) recognizes (and articulates) key uncertainties in available information and understanding along with a need for effectiveness monitoring and evaluation;
7) assumes that estimates may carry error, have variances around means, and may change through time in reaction to natural and human-caused environmental changes;
8) uses real-world examples to test models and assumptions;
9) identifies where science informs policy, but recognizes that specific kinds of decisions or judgments are policy rather than scientific; and
10) makes full use of VSP, RSRP, ISAB, NRC, and published peer-reviewed literature as the basis for developing criteria.

Potential concerns:

- The hierarchical approach (populations, MPGs, ESUs) seems well reasoned and based on decades of scientific findings concerning the complex structure of salmonid meta-populations. It is logical that a viable ESU will have viable components, but what are the limits for assigning viability at the higher levels based on cumulative effects at a next lower level? To what extent can a viable ESU have nonviable parts (populations)? Does this lead to saying that a few strong sub-components might lead to ESU viability while letting other sub-components effectively go extinct? The ISAB's previous review of TRT documents addressed this issue. See www.nwcouncil.org/library/isab/isab2003-4.htm). Also see www.nwcouncil.org/library/isab/isab2005-2.htm: "Evaluation of ESU viability should rest not only on the numbers of component populations or on the abundance and productivity of those individual populations, but also should be based on the population dynamics within the ecosystem as a whole."

- Historical information is extremely difficult to assemble and interpret. Therefore, baseline is necessarily a "best judgment." Especially problematic is how best to "recreate" or recover extirpated populations or MPGs because we don't know the most complementary source or because such, if known, is not accessible.

- Historic capacity and population size was modeled, as we understand it, from estimates of spawning and rearing habitat based on stream morphologies, sediment, erodibility, and flow velocities. Temperatures are mentioned, but water quality, to our knowledge, is not. Are contemporary conditions assumed? Contemporary and historic conditions could be very different.

- Page 8: The ESU viability criterion is "*All extant MPGs and any extirpated MPGs critical for proper functioning of the ESU should be at low risk.*" The paragraph below the criterion defines MPGs, stating that they are "*geographically and genetically cohesive groups of populations within an ESU and are thus critical components of ESU-level spatial structure and diversity*."

  This definition seems at odds with the ESU viability criterion, implying that an ESU with even a single extirpated MPG is non-viable. The implication is probably not intended, judging from later text about criteria to determine whether or not an extirpated MPG is critical (pages 14 and 15). However, there seems not to have been any consideration of whether some extant MPGs might not be critical. Presumably, the same criteria could be applied. Also, how can extirpated MPGs be at low risk?

- Page 9: The biological or scientific basis for the five MPG viability criteria (specifically #1 and #2) is not transparent. The rationale would add to the discussion.

- MPG viability criterion #3 – population size: It is difficult to follow what is intended by "size" here, and to judge whether the logic or just the writing is muddled. The bullet on page 9 states that size categories are based on "*historical intrinsic potential*" and generally reflect the historical proportions present in the MPG. The corresponding paragraph on page 11 does not define historical intrinsic potential or historical proportions; instead, it refers to "*size distribution of populations*" and states that large populations are more likely to have served historically as "*source*" areas, and that they are at lower risk of total loss from localized catastrophic risks. It seems that three concepts have been conflated: abundance, productivity, and spatial extent. Later in the document, it seems that the metric developed to address this criterion refers only to historical spatial extent. If this is true, the justification on page 11 seems inadequate to link the metric to rescue potential (most related to productivity, as suggested by the terms "historical intrinsic potential" and "source area").

- MPG viability criterion #4 - is it correct to assume that this holds only for life history variants (e.g., runs) that are not separate ESUs?

- MPG viability criterion #5 – maintained populations: The statement on page 11 that "*The primary intent is to avoid situations where one or more of these populations serve as an overall 'sink' for production across an MPG*" seems at odds with the definition of populations as units that are demographically independent within a 100-year period (McElhany et al. 2000). The phrase "*a sink for production*" implies demographic linkage in the metapopulation sense, which would preclude delineation as an independent population. The logic would be more coherent if the intention was merely to provide a longer term evolutionary connection by maintaining gene flow at levels too low to link populations demographically. Perhaps this is what was really intended.

- There needs to be some clarification of the full method used for adapting the viability curves in regard to spawning and rearing habitat.

- The biological or scientific basis for the four population size categories is not transparent (page 16, last paragraph). Also, is this the same as Weighted Area Category in Table 2a?

- The sensitivity analyses referred to on p. 28 and in Appendix A are particularly hard to decipher.

- Abundance and productivity criteria: The overall approach can be summarized as follows:
  a) A risk target (<5% in 100 years) is set arbitrarily by policy albeit from recommendations from the NRC (1995) review. Ultimately, new information or approaches may supersede this target in the foreseeable future.
  b) A quasi-extinction threshold (QET)(spawning abundance <50 for 4 consecutive years) is set based on guidelines that consider known risks (inbreeding depression and demographic stochasticity) which then do not have to be modeled explicitly.
  c) Quasi-extinction risk is computed for each combination of productivity and abundance based on empirical data about productivity and environmental variability, an assumed reproductive failure threshold, and by assuming a hockey-stick model of density dependence. Isopleths for 1, 5, and 25% quasi-extinction risk demarcate categories for population viability.
  d) An additional constraint (minimum spawning abundance >500) is superimposed to address longer term concerns about the loss of genetic diversity (which depends primarily on absolute population size).
  e) The minimum abundance constraint is modified to avoid risks associated with low spawning density; it is increased for populations with larger historical spatial extent.

f) Constraints (c – e) jointly constitute three viability curves against which empirical data on geometric mean abundance and geometric mean productivity from each population are tested to designate population viability risk as very low, low, moderate, or high.

Three conceptual issues warrant further consideration or clarification.

First, the quasi-extinction threshold used seems inconsistent with advice from the Recovery Science Review Panel, judging by text on page 27, 2nd paragraph. Why 50 instead of 100? A total escapement over four consecutive years as low as 50 spawners could fail to trigger the quasi-extinction threshold of 50 if the spawning occurred all in one year. The probabilities of quasi-extinction should not be considered equivalent to the probability of biological extinction. Rather, the former should be interpreted as the probability of entering a state where the risk of extinction cannot be modeled but is considered to be unacceptably high. The true probability of extinction could be bounded by probabilities derived using quasi-extinction thresholds of 1 and 100.

Second, step (e) seems ad hoc, and apparently lacks the conceptual rigor of previous steps. For example, why are these adjustments different for Chinook and steelhead? If the concerns about risk at low density arise from demographic and ecological risks that cannot be modeled adequately, then perhaps they should have been addressed in step (b) by adjusting the quasi-extinction threshold to be a function of habitat area and species.

Third, if empirical data are sufficient to compute mean productivity (ln r = ln(Returns/Spawner)), then it seems likely that they would also be sufficient to estimate intrinsic productivity (r_{max}), or perhaps better, productivity standardized at the quasi-extinction threshold. Either of these approaches would control for the effect of density dependence, which geometric mean productivity does not. In the absence of human threats, a population with high intrinsic productivity that is moderately abundant at natural equilibrium would have high viability, yet it would have a mean productivity of only 1, and presumably it would score poorly against the viability curve in the test described. Comparisons in this document depend greatly on two assumptions: (i) the hockey-stick model adequately describes the underlying stock-recruitment (S-R) relationship and (ii) spawning escapements over the last 20 years have been less than the spawner breakpoint (SB). Under these assumptions, observed geometric mean productivity is a reasonable estimate of the A parameter. Observed geometric mean productivity would underestimate intrinsic productivity, $\ln(r_{max})$, for a Ricker S-R curve, and even more so, for a Beverton-Holt S-R curve. In these respects, the analysis of empirical data might be considered precautionary.

- Table 6, page 44 - incorporating uncertainty: In the first row (Option A), transitions from very low, to low, and moderate risk appropriately correspond to the 1, 5, and 25% viability curves. Why is the tolerance for being above the relevant viability curve relaxed to <50% probability for the moderate risk category, when it had been defined as <85% for the other two categories?

- Page 58 discusses whether phenotypic change is deleterious versus adaptive. This discussion warrants special care as it is ripe for “misinterpretation” or downright hand-waving.

- Table 13, page 62 – risk associated with change in genetic variation: The logic for including this factor is reasonable, but in the end, is anything accomplished by adding a metric for which absolute thresholds cannot be assigned? If the distinction between no, low, or moderate change in genetic variation can only be judged uniquely for each case, then why not skip the metric and judge risk level directly based on the underlying concerns?

2. Implementation Issues

Can the criteria be assessed using the types of data that are, or could reasonably be, collected for the Interior Columbia River Salmonid ESUs?

The types of data required to complete this exercise certainly exist for at least some of the populations; however, whether the data available for all populations are extensive enough for valid application needs to be addressed.

The criteria will be challenging to implement, and knowledge gaps will likely preclude full quantitative application. Even so, provided the conceptual framework is sound, the criteria will likely have considerable value as a checklist to ensure:

- broad consideration of potential risk factors
- consistency across populations and ESUs
- transparency
- documentation of professional judgements
- a basis for recommending further work to address knowledge gaps

3. Clarity of Writing Issues

Is the report clearly written?

The greatest problem with the present document is a lack of clarity in some sections. Many sections are very well written, but some sections are confusing and difficult to read because of awkward writing style and extensive use of undefined jargon and undefined acronyms. This is particularly true in the early sections of the document where terminology and abbreviations are used before they have been explained. This practice makes reading and comprehension unnecessarily difficult. There are also a number of typos, several missing references, and at least one incorrectly numbered figure. A list of some of the editorial deficiencies we encountered is provided below.

> The flow charts in Figures 9, 10, 14 are very effective for illustrating the use of metrics to evaluate risk under each criterion; the document might be improved by using flow charts rather than tables wherever possible for other criteria.

Are the methods described in sufficient detail for a reader to understand and replicate what was done?

> For the most part the approach is described sufficiently well for readers to be able to understand and successfully carry out the process. There does need to be some clarification of the method used for adapting the viability curves with respect to spawning and rearing habitat.

Are the assumptions and uncertainties about the analyses clearly described?

> The authors seem to have worked quite hard to address the issues of transparency, explicitly laying out assumptions, and providing for uncertainties both in science and data.

Overall, from a scientific point of view, this report is in good order. It does need a substantial amount of editing, but there seem to be no major issues beyond improving the clarity.

References

ISAB. 2003. ISAB Comments on Draft NOAA Technical Recovery Team Documents Identifying Independent Salmonid Populations Within Evolutionarily Significant Units.

ISAB. 2007a. Climate Change Impacts on Columbia River Basin Fish and Wildlife.

ISAB. 2007b. Human Population Impacts on Columbia River Basin Fish and Wildlife.

NRC. 1995. Science and the Endangered Species Act. Committee on scientific issues in the endangered species act. Commission on Life Sciences. National Research Council. National Academy Press. 271p.

Specific Comments on Report Presentation

- **Key Uncertainties** identified in report (could these be summarized in the M&E section along with those listed) and highlight immediate, ongoing, or future research needs:
    - Page 9: “We do not have sufficient information on movement or exchange rates among ICB populations to directly model MPGs or ESUs as metapopulations.”

  - Page17: "We hypothesize that the increased protection against catastrophic loss provided by ….independent populations within an ESU."
  - Page 22: "More detailed information on the spatial structure of the Stanley Basin…efforts progress."
  - Page 27: "The impact of repeated parent spawning years …is a major uncertainty."
  - Page 32: "We encourage the development of metrics at other life stages, including juvenile productivity."
  - Page 48: "There is a good deal of uncertainty…population and metapopulation viability."
  - Page 51: "Future monitoring should be structured to assess occupancy more rigorously."

- *Page 2, 1st paragraph*: MPG not defined (until page 4).

- *Page 9, viability criteria for MPGs*: clarity could be improved by numbering the paragraph headings to link text description to bullets.

- *Page 9, last paragraph*: Ruckelshaus et al. (2003, 2004) not in references.

- *Page 10, 2nd paragraph*: Maintained population not defined (until page 11).

- *Page 10, 3rd paragraph*: add "required" to read "achieving viability goals for the *required* minimum number of populations…"

- *Page 14, 1st sentence of text*: add "Salmon" to read "*Salmon* ESUs that contain only one MPG…" The statement is not generally true (across species) but is likely true for salmon.

- *Page 16, 2nd paragraph*: viability curves not yet defined;
  - also not clear why spawners would be more directly linked to population genetic characteristics and demographics than other life history stages. Presumably the main reason for focusing on spawners is that populations are completely segregated at that time and *population identification is unambiguous.*

- *Page 16, last two paragraphs:* awkward and unclear.
  - What is "historical *intrinsic* potential"? Why not just historic potential? (needs definition at first use)
  - Description of weighting scheme for habitat is complicated and unclear. Isn't the metric simply the summed estimate of spawning area weighted by quality?
  - Last sentence, continuing on next page, seems incomprehensible.

- *Page 17, 2nd paragraph*: Not clear what the two methods are.
  - Use "spawning areas" instead of "production areas"

  - Add “no fewer than” to read “to support *no fewer than* 50 spawners…”
  - add a sentence to clarify that branches were further classified as major or minor.

- *Page 17, last paragraph*: could improve clarity of last sentence by rephrasing to read “...assigned to a size category based on the weighted sum of spawning habitat and given a complexity rating based on the historic spatial configuration of spawning habitat…”

- *Page 18, Table 2a*: does (ext) after Asotin = extirpated? Are ratings based on current status, some running average, or other?

- *Table 2*: Define MaSA and MiSA.

- *Page 22, 1st and 3rd paragraph*: awkward and confusing.

- *Page 22, last paragraph*: Why is there no mention of kokanee in Redfish Lake and their relationship to sockeye salmon?

- *Page 23, bottom*: The text here and in Appendix A is confusing because only the variable terms “productivity” and “abundance” are used without indicating how they relate to parameters of the underlying population dynamics model. It seems that the viability curve describes the combinations of INTRINSIC productivity (parameter A in the hockey stick model) and abundance AT CARRYING CAPACITY (parameter SB in the hockey stick model) that provide an adequate probability of persistence.

- *Page 27, last paragraph*: QET value is missing.

- *Page 42, first two paragraphs*:
  - Description is awkward and seems unnecessarily complicated. Perhaps it would be simpler to state that recent estimates of population parameters for the population in question must satisfy 3 conditions: (1) geometric mean abundance must exceed the corresponding minimum abundance threshold, (2) intrinsic productivity and carrying capacity must exceed thresholds defined by the corresponding viability curve. Then the discussion can focus on how the population parameters are estimated.
  - delete “intrinsic” in “However, as stocks approach rebuilding target levels, direct estimates of ~~*intrinsic*~~ productivity can be affected by carrying capacity.” Intrinsic productivity is not affected by carrying capacity.

- *Page 42, last paragraph*: Several sentences repeated verbatim from page 26.

- *Page 48. 1st full paragraph*: Last two sentences are awkward and vague; the meaning of “conditions expressed by natural-origin fish” is unclear.

- *Page 48, last paragraph*: Goal 1 is stated inconsistently with other parts of the document, e.g., "maintaining" on page 48, and "allowing" in Table 7 and page 53.

- *Page 51 (bottom), bullets on page 52, and first sentence following bullets*: Presumably the intention is to rank factors that determine the level of uncertainty, but this point is not made clearly, and the grammar is awkward.

- *Page 52, 2nd last and 3rd last paragraphs*: add "any" in two instances to read "…certainty for *any* one of the following characteristics:"

- *Page 57, 2nd paragraph*: first sentence is awkward and unclear.

- *Table 11*: the distinction between "non-negligible change in pattern of variation" versus "significant change" is ambiguous and not defined.

- *Page 68, last paragraph*: add "negative" to read "…will cause a *negative* exponential change in the value, ultimately reaching an asymptote..."

- *Page 76, 3rd paragraph*: what are "direct metrics"?

- *Page 78*: add subheading "Maintained Populations" before last paragraph.

- *Page 80, 1st paragraph*: spell out or define "M & E."

- *Page 80*: are knowledge gaps listed in order of priority? If so, this should be stated; if not, group by species.
