



Independent Scientific Review Panel

for the Northwest Power & Conservation Council
851 SW 6th Avenue, Suite 1100
Portland, Oregon 97204
www.nwcouncil.org/fw/isrp

Memorandum (ISRP 2012-17)

November 20, 2012

To: Rhonda Whiting, Chair, Northwest Power and Conservation Council

From: Rich Alldredge, ISRP Chair

Subject: Follow-up review for the Confederated Tribes of the Umatilla Indian Reservation's *Biomonitoring of Fish Enhancement* project (#2009-014-00)

Background

In response to the Northwest Power and Conservation Council's October 15, 2012 request, the ISRP reviewed the document *Biological Effectiveness Monitoring and Evaluation Plan for Fisheries Habitat Enhancement in CTUIR Subbasins* (draft report, August 2012), which was produced for the Confederated Tribes of the Umatilla Indian Reservation's (CTUIR) *Biomonitoring of Fish Enhancement* project (#2009-014-00). As described in the CTUIR cover letter, this plan was developed to evaluate CTUIR fish habitat restoration projects throughout five subbasins: the Grande Ronde, John Day, Tucannon, Umatilla, and Walla Walla Rivers. Depending on the site, the monitoring plans include before-after-control-impact (BACI) or before-after (BA) analysis using a treatment and control. Field surveys will be an expansion of existing juvenile and adult sampling for spring Chinook, and summer steelhead.

This is a follow-up to the ISRP's review of the project in the Research, Monitoring and Evaluation and Artificial Production Categorical Review in 2010 ([ISRP 2010-44A](#)). The ISRP found that the project met scientific review criteria (qualified). The qualification was that the ISRP review the project's experimental design once it was completed. The Council concurred with the ISRP's recommendation and recommended that an "*implementation recommendation beyond FY 2012 depends on ISRP review of study design*" ([Final Council Decision, June 2011](#)).

The ISRP review of the project's experimental design, as described in the draft August 2012 report, follows below.

ISRP Recommendation

Meets Scientific Review Criteria (Qualified) with the qualifications identified below:

1. The proponents should prepare and submit an interim report to the ISRP for review within three years. The report should contain a summary of data collected, results, and program modifications.
2. The proponents should make a presentation to the ISRP during the upcoming Geographic Review in which information on the following items is provided and time is allowed for a clarifying discussion on issues raised in this review.
 - a. Documentation of the general condition and trend of the catchment above each treatment area or, if such data do not exist, a plan for implementing such assessment efforts. Knowing and integrating upslope conditions into the overall project is essential for project success.
 - b. Details on the extent and intensity of the individual restoration actions at each site and whether they were implemented and functioning as planned for each existing restoration site. For biomonitoring sites, where restoration actions have not been completed, details on planned restoration actions and plans for implementation monitoring.
 - c. Explanation of how the number of years of pre- and post-treatment monitoring at each site was determined; the number of years of monitoring before treatments are imposed is especially important in the BA analyses proposed.
 - d. Consideration of increasing the number of control sites based on the ISRP comments about design assumptions.
 - e. Connections between the CTUIR biomonitoring program, CHaMP, ISEMP, and the large scale effectiveness monitoring program identified in the proposal.
 - f. Links between specific limiting factors for focal species and restoration actions in a watershed context.
 - g. Details on how spatially explicit life-cycle modeling will be implemented.

- h. Description of the measurement of response variables in Table 14 to demonstrate the effectiveness of the biomonitoring program.
- i. Consideration of the advantages and disadvantages of possible alternatives to BACI, and especially the BA design, identified in the ISRP review.
- j. Details on data management strategy and coordination.

ISRP Comments

The ISRP supports and encourages biomonitoring efforts in the region. The proposal under consideration describes an ambitious undertaking, and it is apparent that a good deal of thought and organization has gone into it. The proposal clearly identifies the species that would be used as indicators of ecosystem function in the monitoring effort. The proposal clearly states the objectives of the study and identifies the biomonitoring sites, selection process, and the experimental designs to be used. The great challenge associated with the extensive spatial area, varying conditions, multiple species, differing restoration actions, confounding biological influences, out of basin effects, fish use at multiple scales, and bias in site selection, due to constraints such as landowner cooperation and site accessibility, is daunting.

The intent to create a monitoring program that is complementary to other monitoring efforts in the Columbia River Basin is clearly articulated. The focus of the program is to monitor the effectiveness of CTUIR restoration actions in terms of improvement in fish numbers to guide future restoration decisions. NOAA's Viable Salmonid Population (VSP) parameters were considered in selecting project objectives. This biomonitoring program is primarily designed to determine the biological response to floodplain/riparian area and in channel restoration actions. Limitations in the scope of the plan mean that direct measurement of changes in habitat, in the target areas, their relation to individual or combinations of restoration treatments and their persistence will depend on other monitoring efforts. Integration of this biomonitoring program with other monitoring studies is essential for increasing knowledge of restoration effects most efficiently.

The two primary biomonitoring questions (page 9) are:

1. What are the effects of the habitat improvement/restoration actions on fish abundance and distribution at multiple scales?
2. What particular habitat restoration action(s) have had a positive effect on species of concern?

A number of factors, not addressed in the proposal, make unambiguous answers to the questions unlikely. If these factors are not considered, especially when using “before/after” (BA) monitoring sites, incomplete or inaccurate conclusions on the effects of habitat improvement actions in the treatment area may result. The proponents have shown much care in choosing monitoring sites in an attempt to control for a number of confounding effects that have been anticipated. Nevertheless there are a number of areas in the proposal that, without additional work and revision, are likely to preclude fully meeting project objectives. First, there has been very little work done to address the condition and trend of the watersheds upstream of the treatment areas. Given that increased sediment and stream temperature are listed as limiting factors for nearly all project areas, the condition and trend of the catchment, above the treatment area will likely play a dominant role influencing changes in habitat condition. This situation may confound interpretation in a BA design and in a BACI design in which control sites are located on different streams than the treatment area. Consideration of any upstream and/or upslope treatments or change in management to address either of these limiting factors is needed. Additionally, monitoring of sediment transport/deposition or water temperatures to complement the biomonitors would be extremely useful.

Most treatment areas have, or, will be treated with a variety of restoration actions. Details are needed on the extent and intensity of each of the treatments at each site and whether the restoration actions were implemented and functioning as planned. Without this information it is unlikely that effects of treatments can be assigned to an individual or combination of restoration treatments, and whether the combinations of the same treatments at different sites are comparable.

Information regarding any major changes in land and resource management practices that may have accompanied restoration in the treatment area, or in the watershed upstream of the treatment area, should be documented to better understand the effectiveness of restoration efforts.

Section 3

It is stated (Section 3, page 10) “When the collective findings have been evaluated in the context of the River Vision, they should provide resource managers across the basin with a basis for involving not only agencies but also tribal members in management dialogues to help improve the ecological functionality of the CTUIR subbasins.” Since all of the restoration work to be monitored will occur in valley bottom and stream channel locations, and given that there is no discussion of upslope processes and conditions, meaningful discussion about ecological functionality at either watershed or subbasin scales seems improbable.

The focus of the project on monitoring CTUIR restoration projects that are not being monitored by other programs is clear, but additional discussion is needed concerning how the CTUIR biomonitoring results will be combined with Intensively Monitored Watershed (IMW) results and input from the Integrated Status and Effectiveness Monitoring Program. The hope that collective findings among monitoring efforts will result in management dialogues should be actively facilitated through planned, structured initiatives such as regularly scheduled CTUIR sponsored workshops.

The connection between the proposed CTUIR biomonitoring program, the Columbia Habitat Monitoring Program (CHaMP), and the need identified in the proposal for another large scale habitat effectiveness monitoring program to supplement CHaMP should be provided.

Section 4

Selection criteria for identifying treatment and control sites for monitoring are clearly presented in the proposal. Considerable effort went into selection criteria for appropriate monitoring sites. Additional factors, not included in the selection criteria, could play an important role in understanding the influences of restoration on fish populations and should be considered. First is the general condition and trend of watershed condition upstream of the treatment and control sites. This could be addressed by looking at vegetative cover types and a general suite of land use/management categories and perhaps road densities. Second, some assessment of whether there have been any major disturbances in the near past and their likely influence on stream processes and function identified. Not addressing these factors will likely limit the ability to draw conclusions regarding the actual success of individual or groups of restoration treatments between otherwise similar watersheds.

It is possible that the criteria for selecting “treatment sites” from the list of restoration sites are so restrictive (“eliminated numerous restoration efforts we had hoped to study”) that the selected sites are no longer representative of the majority of restoration sites. Excessive restriction could limit the power of inference from the monitoring sites to the full suite of activities. Criteria involving feasibility, access, or the availability of data do not affect the expectation of success, but two of the criteria might. For example, criterion 2 (one or more primary species are *expected to benefit* from the restoration action) and criterion 8 (low incidence of hatchery fish releases and introgression) might bias the outcomes towards detecting benefits that might not occur elsewhere.

Section 5

A total of eight sites that differ in management actions will be monitored (Table 5). One of the objectives is to try and tease out which management actions are most beneficial, but the overlap in action items among the sites will make this difficult. For example, action A3 occurs in

all sites but one; but action D1 occurs in only 1 site. To help visualize this overlap, a matrix of sites x treatments would be useful to see which actions nearly always occur together or only appear in one site. Additionally, a description of the extent and intensity of the treatments at each site is needed to allow better comparisons between sites.

Experimental designs are available for investigating the effects of many factors (e.g. Plackett–Burman designs and similar designs). Some investigation should be made to see if these designs provide a better distribution of management actions to ensure that effects of the individual actions can be separated from each other.

Finally, identification of limiting factors is very general and lacks the specificity to link them with the design of restoration treatments. Statements like “Summer steelhead: spawning and juvenile rearing (temperature and habitat) and Spring Chinook salmon: holding and juvenile rearing (temperature and habitat)” are common. These descriptions provide a general context to guide restoration actions but are not sufficient to prioritize and develop individual or combinations of restoration actions. Improved identification of these factors and specific links to existing or proposed restoration actions, or groups of actions, would be helpful as a foundation for monitoring and for better understanding their effectiveness.

Section 6

Hypotheses 1 to 4 predict responses in juvenile behavior, growth and survival that would likely occur quickly after the restoration activity is completed, and thus would likely lead to an abrupt “step” in the time series of monitored values, consistent with assumptions in the BA and BACI models described. In contrast, hypothesis 5 involving a change in fish distribution (presumably due to changes in the abundance of adults returning, and of the number of progeny subsequently recruited to the treatment sites) predicts a more gradual change over generations, leading to a ramped response in the variables that is not consistent with the proposed BA and BACI analyses. Values for adult returns and subsequent fry recruitment are also less likely to be independent among sites due to straying and more likely correlated due to the shared influence of out-of-basin factors.

In Section 6 some justification for selecting 30% increase in juvenile density and survival as objectives should be provided. The last three objectives do not have a similar specific size of effect determined. In particular, it is unclear how differences in spatial structure (the last objective in Section 6) will actually be measured.

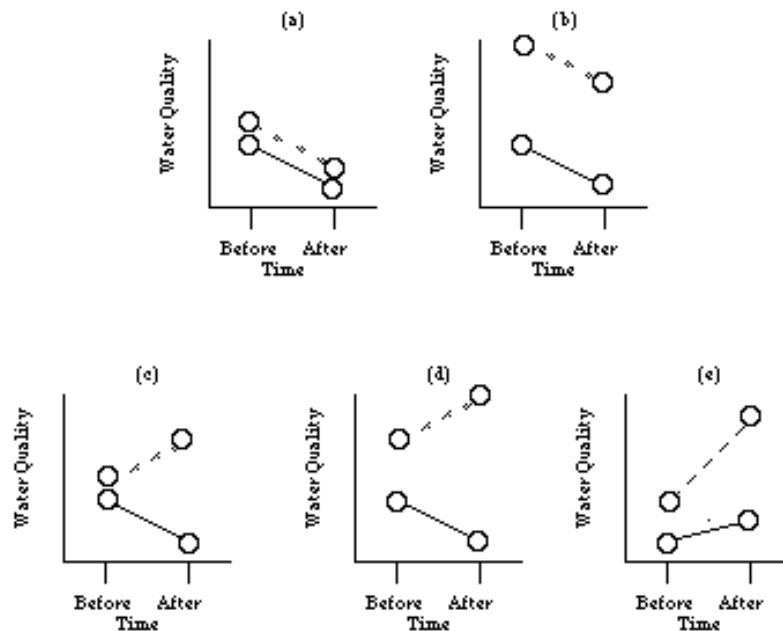
Section 7

The proponents emphasize here, and in section 11, that life-cycle modeling is required to assess how habitat restoration actions affect the target populations at larger scales. The proposal should include a more complete description of how this modeling would be accomplished.

With only one treatment and one control site, the conclusions will be limited to those particular sites chosen. For example, suppose the control site just happens to respond differently than other possible control sites due to site-specific factors. BACI designs often try to replicate at least control sites (e.g. multiple controls for each treatment site) to expand the inference past those the chosen sites.

It is not necessary for the treatment and control sites to have a common mean before treatment occurs – indeed BACI designs work quite well even if the initial means are different, as long as the systems are in steady state over time, that is, the mean prior to treatment is equal over time, followed by a potential shift in the mean after treatment begins. Note that because the population means can differ among sites before treatment, the criteria for choosing treatment sites could be relaxed and more control sites could be added. Alternatively, relaxing this assumption permits comparison of treatment sites in different tributaries and subbasins, so that opportunities for a joint, hierarchical analysis should be explored. In any case, the proposal should discuss explicitly how results from different tributaries and subbasins would be considered jointly.

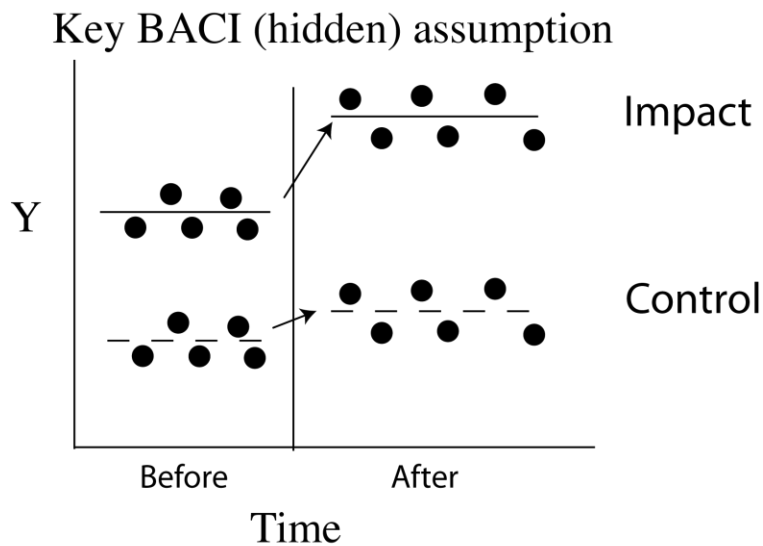
The following conceptual diagrams are offered as an aid in clarifying ISRP review comments.



In both cases in the first row no treatment effect is declared even though the treatment and control sites may differ. All the cases in the bottom row are indicative of a treatment effect because of the non-parallel response.

The assumption of no serial correlation is too restrictive. Serial correlation often occurs because of external factors such as weather that operate on longer than a yearly basis, for example, an extended drought. In such situations the abundance could be depressed for several years in a row (serial correlation). However, by measuring both the treatment and control sites for several years in a row, the effect of these external forces will “cancel” and most of the serial correlation will be removed.

It is crucial however, that the system be in steady state prior and after response.



There may be variability around the respective means in each period, but the means must remain constant over time and the treatment effect must be a simple shift in the means so the non-parallel difference is detected. Material in Section 10, which indicates that treatment actions may take several years to have an impact, violates the key assumptions that the change must be immediate and remain in effect.

If the treatment effects are slow to take effect, other designs are available that incorporate gradual changes over time (Wiens and Parker 1995¹). Another alternative to the BACI model (equation 1, p. 61), involves taking the difference in the response between treatment and

¹ Wiens, J.A., Parker, K. R. (1995). Analyzing the Effects of Accidental Environmental Impacts: Approaches and Assumptions. *Ecological Applications* 5, 1069-1083. <http://dx.doi.org/10.2307/2269355>.

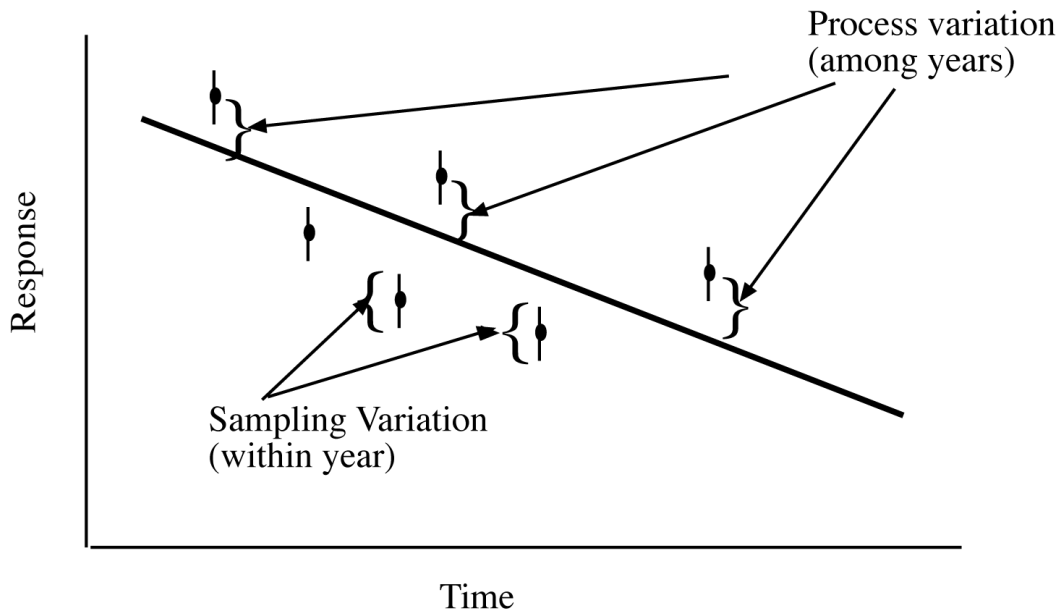
control sites. This reduces the data to one value per year. Then a two-sample t-test can compare if the mean difference before treatment begins is the same as the mean difference after treatment begins. Note it is not necessary to assume that the mean difference is zero before treatment begins – the treatment and control sites can have different means – but the difference must be consistent.

For the before-after design, a simple t-test could be used to see if the mean differs before and after treatment begins. In this case, serial correlation is a real concern because no control sites are available and it is not possible to “eliminate” any local year effects, especially those that span several years. This lack of controls can have extreme unforeseen consequences. For example, if all the pre-treatment years occur during a drought, while all of the post-treatment years occur during a wet-period, then the change in mean due to treatment is completely confounded with the changes in moisture regime. The ISRP urges the proponents to revisit the ability to measure control sites.

Some justification should be provided for identifying the 10-year duration of the study. Also, it is not clear how many years of monitoring would occur at each site before restoration begins (especially given comments about contingency planning on page 70). A single pre-treatment year would be insufficient due to the confounding of measurement error and process error.

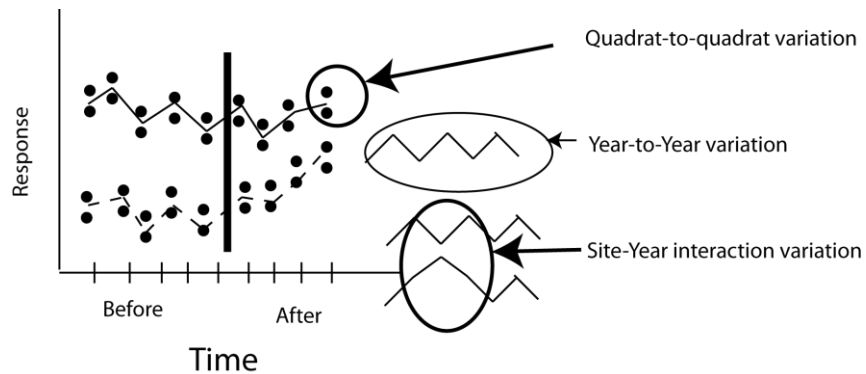
In section 7.1.1 the claim that power curves will be established based on the data from a single year seems incorrect. Insufficient information is available to determine the important variance components that are needed for a power analysis without at least two-years (and preferable more) of data. The key problem is that there are two sources of variability in the responses – sampling error (how precisely is the population parameter for this particular year known) and process error (how much does the yearly parameter vary from the long-term mean even if perfect information is available). With only a single year of monitoring, both sources of uncertainty are confounded and cannot be separated.

Consider the analogous problem of detecting a trend over time as illustrated in the following diagram.



In any one year, an estimate of abundance is obtained with a measure of precision (the dots and the bars on the graph). This is sampling error. However, even if perfect information were available, the abundances would still not lie on the straight line – this is process error. The process error is often the limiting variation that determines how much sampling effort (number of years) is needed to detect a trend. With only a single year of data there is no information on the size of the process error, and using the uncertainty of the single estimate in a power analysis would be incorrect.

With a BACI-P design, a similar breakdown of the sources of variation can be made:



The above diagram uses multiple measurements on the same site in a year, but the logic is the same. Information about the uncertainty of the measurement in one particular site-year combination (the se of the estimate) is available; the pairing ensures that the year effects will “cancel” in the analysis and so information on the year-to-year effects is not needed, but the process error is the site-year interaction – the fact that both sites do not respond in exactly the same way in each year. Once again, to estimate this variance component (which is most often

the limiting feature of the design) will require at least two (and preferably) more years of measurement before treatment occurs to be estimated. Unfortunately, there is very little control over this variance component – doing more sampling in a site-year reduces the uncertainty of each estimate but has no effect on the process error.

Section 8

The response variable “temporal density” for both juveniles and adults is inadequately explained. Is this variable intended to detect any changes in the timing of migration and spawning? Other response variables worth considering include age at outmigration and spawning, and species composition and diversity within the stream sites.

Section 8.2 has a summary (Table 14) of the response variables that can be connected to the objectives of Section 6. However, more details are needed describing how these response variables in Table 14 will be measured.

- How will fry/km² (Table 14) be estimated? Will this density be estimated using multi-pass electro fishing? How will the appropriate denominator be computed?
- How will peak-spawners be estimated – multiple redd visits and area-under-the curve? Suppose there are two years in which the same number of fish return, but in one year all the fish return within a 1-week period and in a second year, all fish return over a two-year period. The peak-spawner counts will be quite different? Is this what is wanted? Perhaps a better measure is total number of spawners.
- How will smolt out-migration be measured? Using RST? How will smolts from the treatment and control sites be distinguished from smolts outside of the sites, for example further upstream?
- How will seasonal growth rates be measured? To determine growth rates some fish need to be measured twice. Is this one reason for the PIT-tagging?
- How will size at outmigration be measured? This implies that these fish must be sampled and measured as they leave. Will fish be identified as belonging to a particular site through use of PIT tags?
- How will juvenile survival be measured? Presumably this comes from the PIT tagging of juveniles and detection as they cross the PIT arrays?
- How will fry-to-smolt survival be estimated? It will be necessary to either mark individual fry and recapture them as smolt, or get numbers of each for a common area. What is to prevent the fry in one location dying and being replaced by smolts from another location?
- How will smolts/redd be determined. Redd surveys will give the number of redds, but how will the number of smolts be determined?
- How will percent of habitat occupied be determined? What does “occupied” mean for a mobile species such a fish? How is the total habitat available determined?

A summary of how these response variables will be measured should be made available. A matrix showing what data comes from what part of the monitoring plan, and how it is used to

estimate the variables in Table 14 which in turn answer the objectives of Section 6, would be very helpful.

Similarly, on page 68, the proponents state that "... and spatial structure ... will be evaluated using ANOVA" which implies that a single summary value of spatial structure will be constructed, but it is not clear what this value is.

The proponents intend to use the Barker robust-design (page 67) to estimate survival. This design usually is intended for long-term capture-recapture designs with multiple years of data (population not closed between years) and with multiple samples within each year (population assumed closed within years) and augmented with additional recaptures outside of the mark-recapture study area. It is not clear at all how this will be used to estimate fry-smolt survival. How will fry be marked? What do the primary and secondary periods of the robust design correspond to? If elastomer marks are used, this will be batch marking, so the individual capture-history of a fry will be unknown, and so it seems the standard mark-recapture methods cannot be used.

Additionally, two major assumptions "Proposed restoration actions will be of sufficient size to produce a measurable biological response and that changes in metrics between the treatment and control sites result from restoration actions (causal relationship) found in Section 8, page 69, do not seem likely to be met.

Section 9

More details on data management are needed. How will data management relate to other data management activities in the region? Many different types of data will be collected, but little information is presented on how to index and archive these data so they can be accessed.

Section 10

The comment here about a gradual response is a concern if a BACI paradigm is being adopted (see above). The proponents may wish to consider the level-by-time or trend-by-time designs of Wiens and Parker (1995; see footnote 1, page 8 in this report for the citation).

Other comments:

Page 8 – Treatment types: There is no category for treatments that occur in the edge of the floodplain or on upslope areas. Actions such as road restoration (maintenance/repair, storm-proofing or obliteration), fuel treatment, erosion control, and/or silvicultural activities would seem to be an important part of restoring ecological processes affecting the general watershed condition and the health of aquatic habitat and should be identified.

Section 3, page 10 – in the listing of other monitoring efforts the long term PNAMP monitoring coordination project is not mentioned.

Section 5, pages 17-18 – Identification of limiting factors by species and life stage is discussed, and it is stated, “These are the critical conditions that, if changed by restoration actions, are most likely to significantly affect the population of interest. Sampling and analyses will focus on these site specific limiting life stages.” On many sites, limiting factors such as stream temperature and/or increased sediment transport/deposition are listed but restoration actions either do not address them or are unlikely to significantly affect them due to the limited nature and scale of restoration treatment. In the case of increased sediment, the lack of any upslope treatment of roads or other erosion prone areas are likely to preclude meaningful changes as a result of reach scale, instream/floodplain restoration.

The listed assumptions (p. 60) are not quite correct. The distribution of normality refers not to the actual measurements, for example log (abundance), but rather to the residuals after the model is fit.

Page 67, Section 6.1.1.2. – First bullet point. The power plays no part in an ANOVA.

It is good that absolute effects of biological interest (Section 6) have been developed rather than relying of rules of thumb of relative effect sizes (Section 7.1.1).