

5.0 MEASURING CHANGE FROM REFERENCE CONDITION

5.1 Purpose

A primary goal of the MAP assessment is to measure the magnitude and direction of change of ecosystem responses to CERP implementation. In essence, this assessment is a measure of the recovery process. It is therefore appropriate that this section of the technical assessment guidance process address both the underlying concepts and assumptions of ecosystem recovery as they relate to measuring change from baseline/reference conditions and provide an overview of design and analysis strategies that have been used to address this topic. In addition, to assure these analyses are conducted in a systematic manner throughout the MAP, the following guidance is proposed for Principal Investigators:

- 1) Describe the methods used to estimate the direction and magnitude of change in performance measures from the reference state both annually and cumulatively over multiple years
- 2) Compare current status of the PMs with its desired trend or target
- 3) Evaluate consistency of monitoring results with map hypotheses
- 4) Determine if there are indications of unanticipated events
- 5) Describe how these events are affecting the desired outcome, the goal for the restoration.

5.1.1 Guidance for PIs and Module Groups

The focus of this section is on approaches to measure change from reference condition. The following outline the specific topics that PIs need to discuss as they conduct their assessments:

- 1) A description of the methods used to estimate the direction and magnitude of change in PMs from the reference state both annually and cumulatively over multiple years. Section 5 provides narrative descriptions of design and statistical approaches to examine the direction and magnitude of change, including the use of BACI (Section 5.3.4), univariate (Section 5.4.1), multivariate, weight of evidence approaches (Section 5.4.3), and environmental gradient analyses (Section 5.4.4). Section 5.4.6 provides information on where to learn more about time-series analyses, the use of a reference site as a covariate, intervention analysis, application of confidence intervals, decision theory, Bayesian methods, and complicated, non-parametric simulation-based tests of a hypothesis.
- 2) A comparison of the current status of the PMs and its desired trend or target. Section 5.2.1 describes the importance of understanding the temporal and spatial assumptions used in assessing ecosystem response. The characterization of the status of a PM needs to be made relative to the currently understood desired trend or target (if it exists) as described in the CERP System-wide PM Report

(www.evergladesplan.org/pm/recover/eval_team_perf_measures.cfm) and in the context of Section 5.2.1.

- 3) An evaluation of consistency of monitoring results with MAP hypotheses.
- 4) The degree and magnitude of consistency of monitoring results with MAP hypotheses are specific to the ecosystem and hypothesis being studied. Monitoring results at the PI level need to be characterized in the context of the MAP hypothesis as described in Section 9.
- 5) A determination of any indications of unanticipated events (See Section 10).
- 6) A description of how any unanticipated events are affecting the monitoring results, and thus the status of the respective MAP hypothesis of interest.

In characterizing the influence of an unanticipated event on monitoring results, it is important to understand whether, and to what extent, the unanticipated event had an ecologically meaningful influence on the PM, and subsequently, the MAP hypothesis. As part of this, it is important to be able to characterize the PM response rate to an unanticipated event such as an environmental disturbance (Fig. 5-1).

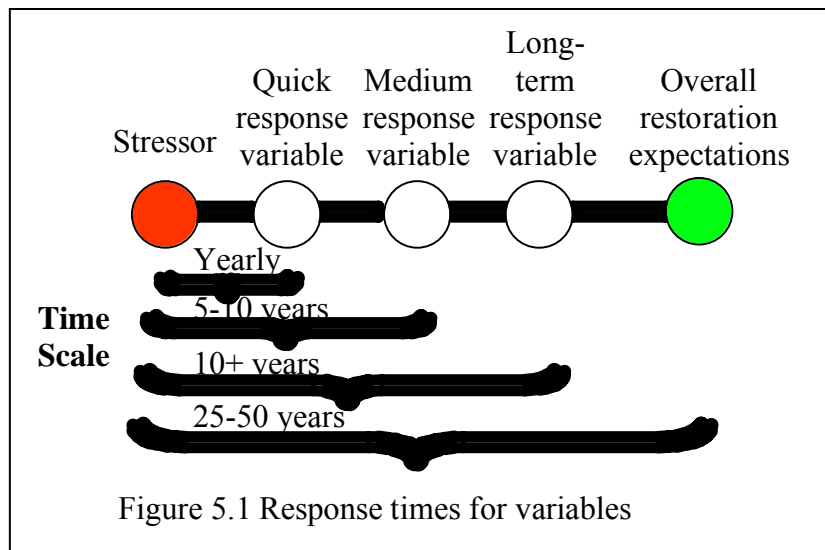


Figure 5-1: Response times for variables

While the intent of these five points is to provide guidance to the PIs, the burden remains on the investigator to use sound judgment (and document the steps in that judgment) in assessing the state of recovery of a biological or environmental resource. Underlying assumptions about equilibrium must be explicitly acknowledged and specified.

5.2 Defining Recovery

Determining the extent to which recovery has occurred requires that the target condition, or the state of the fully “recovered” system, be defined. In an ecological sense, restoration can be defined as “the return of an ecosystem to a close approximation of its condition prior to disturbance” (National Research Council 1992). The goal is to “emulate a natural, functioning, self-regulating system that is integrated with the ecological landscape in which it occurs” (National Research Council 1992) or to reestablish “a critical range of variability in biodiversity, ecological processes and

structures, regional and historical context, and sustainable cultural practices” (Society for Ecological Restoration 1996).

Similarly, CERCLA (U.S. Code of Federal Regulations, Revised 2001, Volume 43, Sections 11:14 and 11:72) defines restoration as occurring when the disturbed resource reaches the level it would have been, had the system not been disturbed in the first place and the influence of disturbance-related factors is diminished to the point where the resource varies temporally and spatially in a natural way. Importantly, all these definitions recognize that natural variation and anthropogenic factors are also at play during the period recovery; the recovered state is not necessarily a previous condition but an expected condition based on the recovery process, natural variation, and anthropogenic influences. The inclusion of natural variability and uncertainty in the definition of recovery/restoration has important implications for the conduct of assessments.

5.2.1 Ecological Assumptions

Recovery assessments require assumptions about the temporal and spatial variability of the ecological parameter or process being studied. It is important that these assumptions are acknowledged or tested explicitly since the assumptions one makes about ecological variability determine the feasibility of assessing recovery (Green 1979, Wiens and Parker 1995). To illustrate this point, steady-state spatial equilibrium and dynamic equilibrium and their ecological assumptions will be described within the context of assessment procedures (Parker and Wiens 2005; Fig 5-2A).

In the strict sense, steady-state equilibrium implies that system features being considered do not vary in time. In practice, this assumption is usually applied to a single location (e.g., the area impacted by an oil spill or fire) or species. Levels of the resource, and the natural factors controlling them, have a constant mean value over time (Fig. 5-2A). The resource at a given location has a single long-term equilibrium to which it will return if perturbed. The vertical arrow indicates the accident, and the dotted line represents the condition of the affected system; the solid curve indicates the dynamics of the system in the absence of a perturbation, and the solid horizontal line represents the steady-state resource level. Recovery occurs when

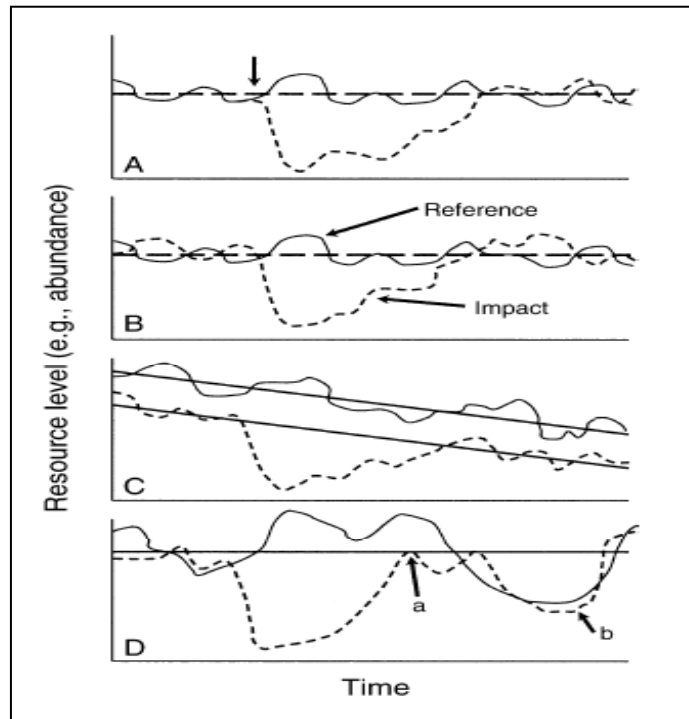


Figure 5-2 A-D: Ecological assumptions affecting recovery.

the impacted system returns to the steady-state equilibrium.

Spatial equilibrium occurs when two (or more) sampling areas (e.g. disturbed and reference) have equal influence of natural factors on them and, consequently, similar levels of a resource (Fig. 5-2B). Spatial equilibrium implies that mean levels of a resource are equal among areas; in the absence of a disturbance event, differences in means are due to sampling error and stochastic variations. The unaffected system serves as a reference for the disturbed system, and recovery occurs when the disturbed system returns to the point where its dynamics are again similar to those of the reference site

The assumption of dynamic equilibrium recognizes variation in both space and time. Natural factors and levels of resources normally differ between two or more areas being compared, but the differences between mean levels of the resource remain constant over time (Fig. 5-2C). Mean levels can change over time for the areas, but such changes are similar among the areas. In the absence of a disturbance, the systems in different areas will track one another over time. Recovery occurs when the dynamics of the impacted system once again parallel those of the reference system, even though levels of the resource differ between the areas.

In this example (Fig. 5-2D), there is considerable natural variation relative to the long-term steady-state mean. If this long-term variation is not considered or is unknown, the impacted system may erroneously be deemed recovered when it is not (point a) or may be considered still to be disturbed when its dynamics in fact match those of an un-disturbed system (point b).

5.3 Measuring Change

For the purposes of this technical assessment guidance, analysis design and analytical strategies can be grouped into three broad categories: baseline, single year, and multiple year analyses (Parker and Wiens 2005). It is important to understand how each of these analyses addresses the issues of temporal and spatial variation which often confound the assessment process (See Table 5-1, after Parker and Wiens 2005, Appendix E).

5.3.1 Baseline Data Assessments

Baseline data assessment designs compare pre- and post-impact conditions with data from the impacted site only. These designs, which often rely solely on an assumption of steady state equilibrium, are of limited value for assessing recovery of biological resources in temporally varying environments. In these cases, recovery may be confused with temporal variations, leading to false conclusions. Used in conjunction with other single- and multi-year designs, however, baseline data may provide valuable insight into recovery processes. Murphy *et al.* (1997) used a baseline design for assessing recovery of seabirds, but only in concert with BACI, a multiyear design that relies on the more realistic assumption of dynamic equilibrium. In this case, the baseline studies were used more to gain insight into long-term trends in populations than to determine whether or not recovery occurred.

5.3.2 Single-Year Studies

Single year studies typically compare perturbed/impacted and reference areas where the investigator employs design and analytical strategies to control for spatial variation among the areas being compared (i.e., meet the assumption of spatial equilibrium). A disturbance/reference design is valid when sampling efforts are spread over a broad geographic area and conducting multi-year sampling is not feasible. Recovery is identified by the absence of significant differences between disturbed and reference areas; if sampling sites are randomly selected, results of disturbed/reference studies can be extrapolated to the total disturbed area. Single-year studies require acceptance of the assumption of spatial equilibrium, either through covariance analysis of concomitant variables (i.e., physical or chemical variables; Gilfillan *et al.* 1995) or by pairing sites with shared environmental factors (McDonald *et al.* 1995). Regressing measures of biota on a continuous measure of disturbance (e.g., gradient analysis; Wiens and Parker 1995) can also be used. Associated environmental variables should be sampled concurrently with the biological resources of interest; however, some of the ecologically important variables may be unknown and/or un-measurable, and different variables may be needed when multiple species are assessed.

When designing single-year studies, it is also prudent to consider the effects of spatial scale. Stressors are not distributed randomly across habitats, and reference samples are often drawn from a larger area than are the disturbed samples. Differing spatial scales can result in differing scales of environmental variation, further confounding tests of equality in resource levels between the disturbed and reference areas. Ideally, this problem can be addressed by:

- 1) Randomly sampling similar habitats in both the disturbed and reference areas
- 2) Using an equal number of samples at environmentally similar sites which will reduce the potentially confounding influences of differing scales
- 3) Using covariates to help reduce potential differences in variation between the two areas.

In general, single-year studies of this nature are not recommended for assessing components of MAP hypotheses unless these approaches are the only mechanism to provide insight into a particular performance measure.

5.3.3 Multiyear studies

Multi-year studies are necessary for taxa (or environmental factors such as soil phosphorus) that exhibit large temporal variations or have long recovery periods or for multiple taxa with different temporal dynamics. These studies are superior to single-year and baseline studies because they can reduce the confounding effects of natural variation and therefore provide an opportunity to observe the recovery process and evaluate the assumption of dynamic equilibrium.

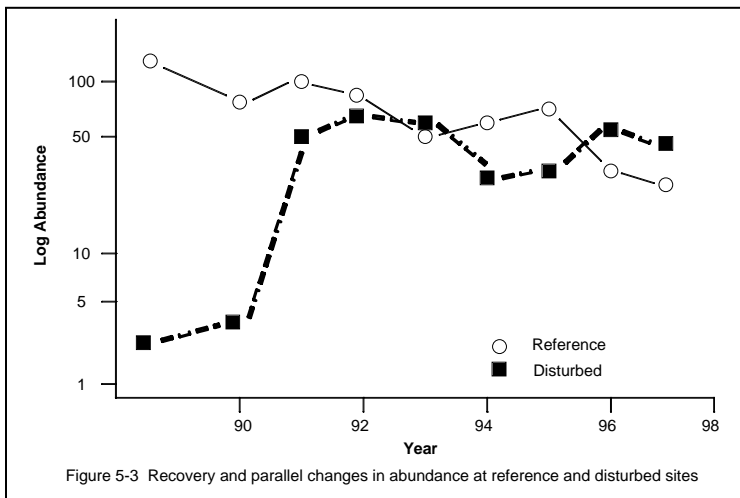


Figure 5-3: Mean abundance for disturbed and reference sites (Skalski *et al.* 2001)

If environmental factors change similarly over time at both disturbed and reference areas then the assumption of dynamic equilibrium holds. Assessing parallel trends in means has the advantage of requiring sampling data for only the ecological resource/attribute being assessed and no measures of concomitant environmental variables are needed.

The following example of a multi-year recovery study illustrates parallelism in temporal trends for two populations of intertidal snails (*Littorina sitkana*) (Skalski *et al.* 2001). In this study, the authors assessed recovery by testing for equal trends in the means of population abundance between a disturbed and reference sites. Figure 5-3 illustrates the mean abundances of the snails for both disturbed and reference sites over a period of seven years (1989-1996). The mean population abundances of the reference site show a general decline over the course of the observation period. Between 1989-1992 the snails repopulated the disturbed area and after 1991 the abundances showed similar trends at the two areas supporting the assumption of dynamic equilibrium. Skalski *et al.* (2001) assessed parallelism between reference and disturbed areas with an ANOVA test of interaction and trends in the ratio of reference means to disturbed means over time. Based on these analyses, Skalski *et al.* concluded that recovery occurred within two to five years after the spill, that is, within the 1992-1994 timeframe.

The BACI-type design (see examples below) is a good example of a multi-year analysis that relies on the assumption of dynamic equilibrium (Stewart-Oaten *et al.* 1986, Green 1979).

5.3.4 Before-After -Control-Impact (BACI) Studies

As discussed above, one of the most important uses of monitoring programs is to evaluate the changes or impacts to ecological systems from a particular event. In fact, the goal of this assessment guidance is to develop a valid measure of the magnitude and direction of changes in performance measures, both temporally and spatially, that result from CERP projects.

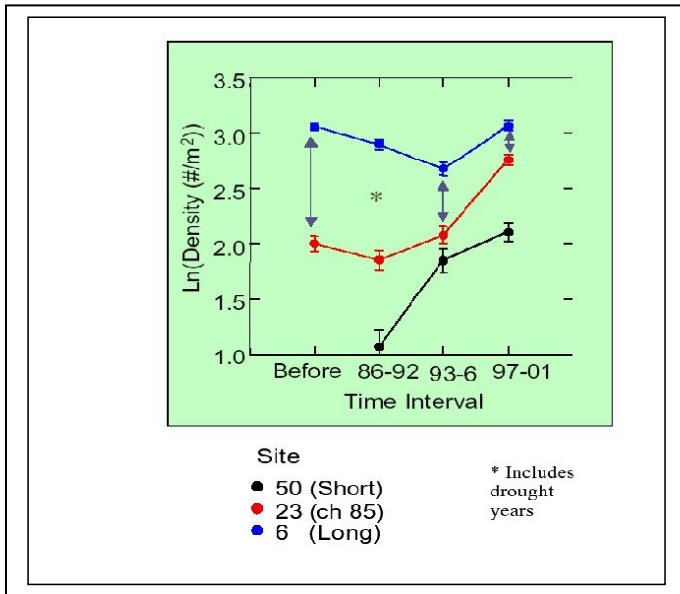


Figure 5-4: Shark River Slough fish data (Trexler *et al.* 2003)

Stewart-Oaten *et al.* (1986) coined the phrase BACI to describe a basic approach which has undergone modification by Underwood (1994) and Downes (1992). While there are limitations that need to be addressed when using this, or any other, design (Hulbert 1984, Smith *et al.* 1993, Underwood 1997), it remains one of the best methods for impact assessment (Smith *et al.* 1993). Smith *et al.* (1993) points out these designs can show that observed differences in ecological variables between the control and impact sites are neither artifacts of sampling nor due to temporal trends unrelated to the impact. Most importantly, the strength of the inference is directly related to the design issues (e.g., frequency of sampling, number of control sites, etc.) that are directly under the control of the managers and researchers.

The following example (Fig. 5-4) illustrates the application of the BACI approach to fish data collected in northeast Shark River Slough (Trexler *et al.* 2001). A monitoring initiative using throw traps to sample fish and macro-invertebrates was begun in 1977. This study ultimately included two reference sites; a long-hydroperiod site (#6) and a short-hydroperiod site (#50) and an impact site (#23). Beginning in 1985, water management practices for Northeast Shark River Slough were altered and the minimum annual water depth increased from a depth similar to site 50 to a depth more similar to site 6.

The density of all fish collected by throw trap is one possible performance measure to assess this change. In the accompanying figure (Figure 5-4), site 6 (blue) is the long hydroperiod reference site, site 50 (black) is the short-hydroperiod reference site, and site 23 (red) is the impact site where hydrological management was changed. The length of the vertical double-headed arrow on the left side of the figure represents the "before" magnitude of difference in fish density between the reference and impact sites for BACI. Note that there is a definite change (decrease) in the magnitude of the arrow as well as direction of the trajectory after the management change has taken place (double-headed arrows on the right side of the figure). This indicates that the management change to lengthen the hydroperiod at site #23 had the desired result, that is, to increase the fish densities to those of the long hydroperiod reference site (#6). As Trexler notes, the short-hydroperiod reference site also converged on the long-hydroperiod reference in the 1993-1996 interval. Thus, only the change in 1997–2001 provides evidence of successful management.

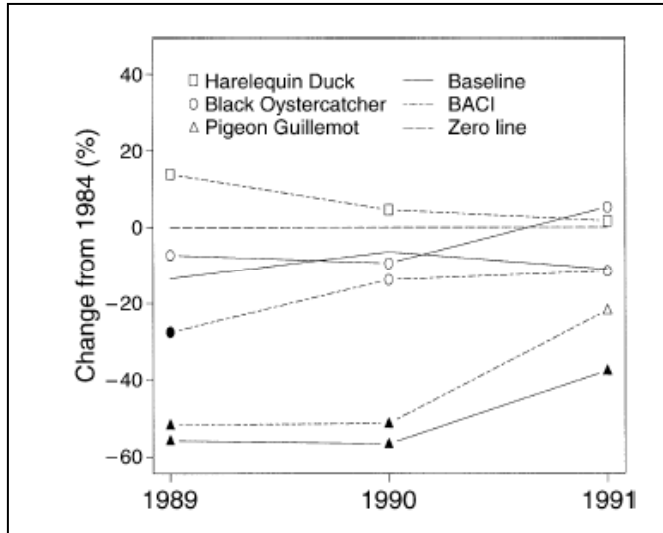


Figure 5-5: Comparison between baseline and BACI analysis.

An analysis of variance can be used to statistically validate the change by testing for a significant interaction between the time (before vs. after impact) and site (reference sites vs. impact site).

A second example of a multi-year BACI analysis is that of Murphy *et al.* (1997) as discussed in Parker and Wiens (2005). Murphy *et al.* (1997) used a baseline study for assessing the recovery of 12 species of seabirds from the *Exxon Valdez* oil spill. For impacted sites, densities for post-spill years (1989 1990 1991) were each compared to pre-spill densities in 1984. To illustrate the difficulties with using baseline studies to assess recovery, they compared the baseline results with those of a BACI analysis (Murphy *et al.* 1997; Tables 1 and 2). Figure 5-5 shows a comparison of percentage changes in density from 1984 under steady-state (baseline study) and dynamic (BACI) equilibrium assumptions. Significant effects are noted by solid symbols.

For Harlequin Ducks (*Histrionicus histrionicus*), BACI analysis showed nominal, non-significant decrease throughout the pre- and post-disturbance period while the baseline remained unchanged. For Black Oystercatchers (*Haematopus bachmani*), the BACI analysis showed a significant decrease in 1989 and subsequent recovery while the baseline analysis showed no statistically significant changes throughout the pre- and post disturbance periods. Baseline and BACI results for Pigeon Guillemots were similar,

showing significant negative impacts through 1990. The absence of significant differences was shown from a BACI analysis in 1991, suggests recovery, even though baseline showed continuing negative effects. When these results were examined with the context of a longer time frame (1984-2001) they showed that the apparent temporal invariance in abundances of Harlequin Ducks is short-lived and therefore misleading in assessing recovery (Wiens *et al.* 2004).

5.4 Statistical Approaches

As noted above, there are a variety of statistical methods that have been used to assess ecosystem response. While no specific method is recommended for the PI, the investigator must have in-depth knowledge of the selected statistical method to ensure that critical assumptions are met, and that the method has been documented in the peer-reviewed literature as appropriate for the purpose for which it is used. For purposes of illustration we provide a brief discussion of both univariate and multivariate statistical techniques below within the context of studies of the recovery (Gilfillan *et al.* 1995). Here, we also discuss the use of a weight of evidence approach used by Wiens *et al.* (2004) for examining the recovery of shorebirds. Finally, Appendix A provides an example of the statistical approaches used to analyze a 30-year record of pelagic Total Phosphorus.

5.4.1 Univariate Analysis

One application of univariate analysis is to test the null hypothesis that there is “no disturbance effect” by using a model that adjusts the means for important environmental parameters (i.e. concomitant variables). In this case, the disturbance effects on any biological variable of interest (e.g., species abundance, diversity index, etc.) can be estimated on the basis of differences among the model means that have been adjusted among the concomitant variables for the disturbance levels of interest. If the null hypothesis is rejected (i.e., there is a difference), the direction (positive or negative) of the difference from baseline/reference is noted.

The goal of univariate analysis is to use a model that satisfies the appropriate statistical assumptions and has high power to detect differences. For example, a “site mode” that uses among-site variability as an error term can be run to test for disturbance effects. The results from this model can be compared to within-site (transect) variability. If the among-site variability is not significantly greater than the within-site variability, a “transect model” can be run in which the within and among site variability can be pooled into a single error term. Disturbance effects can then be tested against this pooled error term. When this modeling approach was applied to the recovery of shoreline species impacted by the *Exxon Valdez* oil spill, Gilfillan *et al.* (1995) found that the transect model always had greater power to detect differences and was found to be appropriate for about 78 percent of the target species of concern.

It is important to note that testing disturbance effects for individual species (abundance, biomass, etc.) was first done with normal-theory methods on log-transformed data using

one-way analysis of covariance (ANCOVA) data (Gilfillan *et al.* 1995). If the residuals were not normally distributed then non-parametric methods were employed. Non-normal methods can be conducted with General Linear Interactive Models (GLM). This method is analogous to ANCOVA except that the differences in abundances are tested using Poisson, binomial, or negative binomial distributions.

5.4.2 Multivariate Analysis

As an alternative to univariate modeling, multivariate techniques can be used to simultaneously model multiple dependent variables such as individual species abundances. These techniques are useful when individual species abundances have inconsistent or contradictory trends with respect to the disturbance of interest.

One multivariate approach often used to assess disturbance effects on community structure is Detrended partial Canonical Correspondence Analysis (DCCA). This is an ordination technique that arranges samples along axes defined in terms of weighting factors for each parameter in the dataset. In DCCA, the ordination is performed after removing the effects of other physical or environmental variables. In contrast with techniques such as Principle Components Analysis (PCA), the detrending removes the effects of artifacts that tend to distort the scores along the second ordination axis. Another important feature of DCCA is that it does not require sample data to be normally distributed. Multivariate analyses are not limited to biological data. Ordination techniques, for example, are often used for examining suites of water quality data.

5.4.3 Weight of Evidence

Equilibrium assumptions for multi-year data may be uncertain at times. Erratic non-parallel trends in impact and reference values would trigger concerns of uncertainty. Dramatic regional declines in species abundance, shifts in geographic centers of abundance, or zero mean abundances for some years would also call into question equilibrium assumptions. Wiens *et al.* (2004) encountered all these elements of uncertainty and used a weight-of-evidence approach to assess recovery of 25 seabird species over a 12 year period from 1989 to 2001.

Wiens *et al.* synthesized results from four separate analyses, taking into account effects of potentially confounding habitat variables. The four analyses treated ecological assumptions differently, as described here. (1) For each year, densities were regressed on an index of shoreline oiling for 10 bays. It was assumed that the 10 bays were in spatial equilibrium for each year. (2) Covariates of important habitat variables were used in the same regression analyses used in (1). Covariates reduced the potentially confounding effects of spatial non-equilibrium due to environmental differences among the bays by subtracting out covariate effects before testing for continuing effects of oiling. (3) Densities from 1984 were used in a repeated-measures design, which assumed dynamic equilibrium in natural factors among categories of moderately-to-heavily and none-to-lightly oiled bays. (4) Plots of trends in density by oiling category were visually examined. In interpreting plots, either steady-state or dynamic equilibrium was assumed,

depending on whether one considered oiling categories separately over time or relative to each other. Using a weight-of-evidence approach based on all of these analyses strengthened the possibility of finding a continuing effect of oiling under different assumptions.

Common Mergansers (*Mergus merganser*) provide an example of how weight of evidence was used to detect initial spill-caused impacts, followed by recovery. Regressions indicated negative relationships with oiling level that remained after habitat factors were included as covariates in the analysis (Table 1 in Wiens *et al.* 2004). BACI comparisons with the 1984 baseline data yielded weak, nonsignificant indications of reduced abundance in oiled bays in 1989 and 1991. The oiled/unooled abundance plot (Fig. 5) showed that numbers of mergansers were low in oiled bays in 1984–1991 and that many more mergansers were seen in previously oiled bays in 1996–2001. Wiens *et al.* concluded that merganser habitat use was negatively affected by the spill but had recovered by 1996. This example illustrates that a weight-of-evidence approach is useful for assessing multi-year data for biological resources that do not meet equilibrium assumptions.

Weight-of-evidence approaches are valuable for environmental data as well. For example, a weight-of-evidence-approach was used to sift through potential alternative hypotheses explaining high phosphorus values in the A.R.M. Loxahatchee National Wildlife Refuge in 2005 (http://www.sfwmd.gov/org/ema/toc/archives/2005_10_18/may-june_workgroup_evidence_handout.pdf). The approach assisted in determining whether unanticipated events had a meaningful influence on the environmental parameter of interest.

5.4.4 Sampling Environmental Gradients

Another approach for analyzing trends in recovering ecological systems is to study the changes in the response of ecological attributes along a known stressor gradient (e.g., hydrologic, water quality, etc.). This approach may be particularly suited for cases where the temporal extent of a database (i.e., baseline/reference) is not sufficient to detect statistically significant trends and changes beyond the limits of background variability. This methodology, which would incorporate sampling along environmental stressor gradients, can supplement other approaches to evaluate CERP induced changes.

In this approach, stations should be arrayed in a manner that enables detection of change in an ecological attribute's response along the stressor gradient of interest. Researchers should recognize that the domain necessary to detect change is scale-dependent, with specific spatial and temporal limits set by the particular stressor. The interpretation of data from gradient designs often requires a reference site that has the same physical, chemical, and ecological attributes as the gradient sites but is sufficiently remote so as not to be affected by the stressor of interest. To satisfy this requirement, a sufficiently broad spatial domain has to be incorporated into the sampling design. Alternatively, if a reference site cannot be identified, the interpretation of change will be limited to the sampling stations and be constrained by the threshold for the response of the ecological

attribute. Changes in the response along the gradient can be detected by analyzing the slope or other characteristics of the gradient response curve. The analysis of environmental gradients is not meant to supplant the approaches for determining trends of pre and post-CERP sampling protocols established in this guidance. However, it is offered as an alternative when data are insufficient to conduct BACI, statistical, or other types of analysis.

5.4.5 Other Designs

The examples described above provide sampling and analytical strategies that have demonstrated their usefulness for assessing recovery from environmental disturbance in temporally and spatially varying environments. Which design to apply in order to reduce the confounding effects of varying natural factors depends on several issues, including the availability of data for pre-disturbed conditions, and the financial and personnel resources needed.

Other sampling designs, such as time-series analysis, reference as covariate, and intervention analysis may prove equally useful or better for measuring changes due to CERP implementation, depending on the nature of the disturbance and the reasonableness of the ecological assumptions one makes (see Stewart- Oaten and Bence 2001). In addition to standard tests of hypothesis, assessments of recovery can be performed with at least three other parametric statistical methods, including confidence intervals, decision theory, and Bayesian methods. Confidence intervals provide useful information on how confident one can be of the true state of nature being close to the hypothesized null state.

Decision theory (DeGroot 2004) allows the investigator to adjust type I and II error based on penalties for wrong decisions. Fisheries and wildlife managers use decision theory, in which associated economic and social penalties (e.g., suboptimal fishery yields) are quantifiable. Decision theory would be practical for assessing recovery given penalties for making the wrong decision were quantifiable. In the absence of known or estimable economic, social, and/or ecological value, it is difficult to objectively quantify penalties. Bayesian statistics (Box and Tiao 1973) rely on prior probabilities (for examining the truth of the null hypothesis). Each dependent variable would likely need its own prior probability, making it difficult to use Bayesian methods for multiple species, e.g., 25 species of seabirds. Such prior probabilities and their influence on results would be difficult to justify, especially for such high-profile assessments as the Everglades restoration.

In addition, nonparametric methods can also be used to test hypotheses, especially for biological data. Bootstrap, jackknife, permutation tests, and other simulation-based tests of hypothesis provide ways to avoid making assumptions about the distribution of residual error. In the examples cited previously, Gilfillan *et al.* (1995, shoreline ecology) and Wiens *et al.* (2004, seabirds) found generalized linear models (McCullagh and Nelder 1989) on members of the exponential family (e.g. negative binomial and Poisson) to be more realistic of natural conditions than computer-intensive nonparametric methods.

5.5 Summary

Recovery from an environmental disturbance involves re-establishment of the physical environment to a non-disturbed state and restoration of complex biological systems. The greater the disturbance and complexity of the system, the longer the anticipated recovery time. Natural environmental variation and the passage of time enlarge the probability field of what a non-disturbed resource would have been had the injury not occurred. Consequently, the longer the time to recovery, the less likely the recovered state will be what it was before the injury occurred.

Moreover, as Pimm (1991) has noted, population variability increases with the length of time considered, increasing the likelihood that one or another of the equilibrium assumptions will be violated. Defining the recovered state of a biological resource and inferring when a recovered state has been reached depends on what one can reasonably assume about natural variation of *both* populations and the environments they occupy.

The appropriateness or inappropriateness of the different assumptions about equilibrium is scale-dependent. This means that there is no single “best” study design or statistical analysis that will fit all situations. The risks of reaching conclusions based on incorrect assumptions about equilibrium are likely to be variable at different scales of investigation. The scales in turn depend on the species, communities, or environments being considered, the temporal and spatial scales of their natural dynamics, and the scale of the environmental disturbance itself. The scale of the impact is usually well defined, whereas scales of interdependencies among biological resources and their environments and of natural dynamics are generally unknown, at least with any degree of precision.

The burden is therefore on the investigator to use sound judgment (and documenting the steps in that judgment) in assessing the state of recovery of a biological or environmental resource. Underlying assumptions about equilibrium must be explicitly acknowledged and specified. There should be a clear *a priori* statement of what represents an acceptable level or end point for recovery. Given natural variation in the environment and the resources of interest, this means that one must consider the width of the envelope of natural variation, which in turn helps to define the appropriate “recovery zone” for the system.

5.6 Additional Applicable Resources

DeGroot, M. H. 2004. Optimal statistical decisions. John Wiley and Sons, New York, New York, USA.

Gilfillan, E. S., D. S. Page, E. J. Harner, and P. D. Boehm. 1995. Shoreline ecology program for the *Exxon Valdez* oil spill. Part 3—biology. Pages 398–443 in P. G. Wells, J. N. Butler, and J. S. Hughes, editors. *Exxon Valdez* oil spill:fate and effects in Alaskan waters. STP 1219. American Society for Testing and Materials, Philadelphia, Pennsylvania, USA.

Gilfillan, E. S., and K. R. Parker. 2003. Multivariate analysis of community structure over ten years following the *ExxonValdez* oil spill. Pages 559–567 in Proceedings of the 2003 Oil Spill Conference. American Petroleum Institute, Washington, D.C., USA.

Hoening, J. M., and D. M. Heisey. 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. *American Statistician* **55**:19–24.

McDonald, L. L., W. P. Erickson, and M. D. Strickland. 1995. Survey design, statistical analysis, and basis for inferences in coastal habitat injury assessment: *Exxon Valdez Oil Spill*. Pages 296–311 in P. G. Wells, J. N. Butler, and J. S. Hughes, editors. *Exxon Valdez oil spill: fate and effects in Alaskan waters*. STP 1219. American Society for Testing and Materials, Philadelphia, Pennsylvania, USA.

Murphy, S. M., R. H. Day, J. A. Wiens, and K. R. Parker. 1997. Effects of the *Exxon Valdez* oil spill on birds: comparisons of pre- and post-spill surveys in Prince William Sound, Alaska. *Condor* **99**:299–313.

Parker, K., and J.A. Wiens. 2005. Assessing recovery following environmental accidents: environmental variation, ecological assumptions, and strategies. *Ecol. Applications* **15**:2037-2051.

Stewart-Oaten, A., and J.R. Bence. 2001. Temporal and spatial variation in environmental impact assessment. *Ecological Monographs* **71**:305-339.

Wiens, J. A., R. H. Day, S. M. Murphy, and K. R. Parker. 2001. On drawing conclusions nine years after the *Exxon Valdez* oil spill. *Condor* **103**:886–892.

Wiens, J. A., R. H. Day, S. M. Murphy, and K. R. Parker. 2004. Changing habitat and habitat use by birds after the *Exxon Valdez* oil spill 1989–2001. *Ecological Applications* **14**:1806–1825.