The Definition of Compliance and the Determination of Similarity in the Context of the SONGS Mitigation Projects

Prepared for the staff of the California Coastal Commission by:

Richard Ambrose

Mark Page

Peter Raimondi

Daniel Reed

Russell Schmitt

Stephen Schroeter

March 2007

EXECUTIVE SUMMARY

The California Coastal Commission (CCC) has required Southern California Edison (SCE) and its partners to construct mitigation projects that provide adequate compensation for the loss of marine resources resulting from the operation of SONGS Units 2 and 3. The CCC is responsible for determining whether these projects are successful. Two related issues that reside at the core of this determination are: (1) the level and duration of performance by the mitigation projects that is needed to achieve compliance with specific conditions of the SONGS coastal development permit and (2) a methodological approach to determining whether the mitigation projects are performing similarly to naturally undisturbed reference sites. We address these two issues in the following sections of this document.

I. ASSESSING COMPLIANCE FOR THE SONGS MITIGATION PROJECTS

The conditions of the SONGS coastal development permit (6-81-330-A) were amended in 1991 to mitigate the adverse impacts of the operation of SONGS Units 2 and 3 on the marine environment. The conditions that were amended to the permit require SCE and its partners to (1) create or substantially restore a minimum of 150 acres of southern California wetlands (Condition A), (2) install fish barrier devices at the power plant (Condition B), and (3) construct an artificial reef large enough to sustain 150 acres of medium to high density kelp bed community (Condition C). A fourth condition (Condition D) requires SCE to fund the Commission's oversight of the mitigation and independent monitoring functions identified in and required by Conditions A, B, and C. Physical and biological standards are identified in conditions A and C that specify how the wetland and reef mitigation projects should perform and the timing and level of monitoring that is needed to evaluate their performance. The specific requirements for attaining compliance of these conditions are discussed in various sections throughout the permit. The purpose of this document is to provide SCE with clear and consistent interpretations of key terms in the SONGS coastal development permit, which provide the basis for assessing compliance of SONGS wetland and reef mitigation projects. We identify the specific sections in the permit that provide support for our interpretations, and provide schedules for the different levels of monitoring that are required to determine whether the wetland and reef mitigation projects are in compliance with Conditions A and C.

II. METHODS FOR DETERMINING SIMILARITY BETWEEN THE MITIGATION AND REFERENCE SITES FOR THE PURPOSE OF EVALUATING COMPLIANCE OF THE SONGS' MITIGATION PROJECTS

A requirement of the SONGS permit is that certain biological attributes (performance variables) of the mitigation sites (i.e., the restored wetland or

marine artificial reef) be "similar" to those at nearby natural reference sites. Evaluating compliance with these relative performance standards[1] requires the use of an objective approach to determine whether the values of the performance variables are similar between the reference and mitigation sites. Contract scientists working for the California Coastal Commission were charged with developing that method (Permit No. 6-81-330-A, Condition A.3.4). This document describes the *Hybrid Approach* the contract scientists will use to evaluate similarity between the mitigation and reference sites.

The Hybrid Approach draws on the strengths of two existing techniques but minimizes their weaknesses. The two techniques are called the "Range" test and the "Separate Confidence Interval" (SCI) test. For the range test, a particular performance variable would be considered similar if the average value at the mitigation site falls within the range of average values at the selected reference sites. While this is one of the simplest and least ambiguous ways of determining similarity, it has the major limitation of not considering measurement error for sampling at reference sites, and is thus likely to underestimate the true range of reference site values for a given performance variable. The SCI test does not have this limitation. It assesses similarity by determining whether the average value for a selected variable at the mitigation site falls within the range set by the upper confidence interval of the reference site with the highest average and the lower confidence interval of the reference site with the lowest average. In essence this is a more sophisticated range test that takes into account inherent natural variability within reference sites. Limitations of the SCI test arise when it is used by itself to evaluate similarity because it would allow the mitigation site to have a lower mean value than all the reference sites for *every* performance variable and still be in compliance. This is not consistent with the goals of the SONGS permit. For this reason, a Hybrid Approach will be adopted for evaluating similarity that includes two criteria that must be met to conclude that the mitigation and reference sites are similar: (1) the values of the performance variables at the mitigation site must be within the range of the SCI of the reference site for all performance variables, ***and*** (2) the mitigation site must not have the lowest value more often than expected by chance alone. Criterion 1 ensures that the values of each performance variable at the mitigation site will be greater than that of the lower confidence limit of the reference site with the lowest value. Criterion 2 ensures that the mitigation site behaves like the reference sites with respect to all the performance variables, eliminating the possibility of concluding that the mitigation site is in compliance when it has the lowest value for a disproportionately large number of performance variables.

When applied together, the two criteria of the Hybrid Approach ensures that the assessment of similarity is consistent with the SONGS permit requirement that all performance standards must be met without the unreasonable requirement that the mitigation project outperform all reference

[1] A glossary of terms is provided on page 21.

sites for every performance standard. Thus, the Hybrid Approach deals realistically with the inherent variability of nature in a manner that best serves the interests of the public and Southern California Edison.

I. ASSESSING COMPLIANCE FOR THE SONGS MITIGATION PROJECTS

The SONGS coastal development permit (6-81-330-A) requires SCE to create or substantially restore a minimum of 150 acres of southern California wetlands (Condition A), and to construct an artificial reef large enough to sustain 150 acres of medium to high density kelp bed community (Condition C). Physical and biological standards are identified in these conditions that specify how the wetland and reef mitigation projects should perform and the timing and level of monitoring that is needed to evaluate their performance is discussed. The purpose of this document is to provide consistent interpretations of key terms in the SONGS coastal development permit (6-81-330-A), which provide the basis for assessing compliance of SONGS wetland and reef mitigation projects. The specific sections in the SONGS permit that provide support for our interpretations are indicated by numerical superscripts in the text and are referenced below (see p. 6, **Permit language supporting CCC staff's interpretations on SONGS project compliance**).

DEFINITIONS

Monitoring Period: Post-construction monitoring will ensue upon completion of the reef construction and wetland restoration[1, 2]. The duration of such monitoring will last for a period not less than the full operating life of SONGS (defined below) plus years monitored without the project attaining compliance with permit standards[2, 3].

Compliance: The condition in which all performance standards are met.

*Compliance Perio*d: The number of years that a mitigation project is in compliance. The mitigation requirements will be fulfilled when the compliance period equals the total years of operation of SONGS Units 2 & 3, including decommissioning period to the extent that there is continuing entrainment or impingement or discharge of cooling water[3,4].

MONITORING EFFORT

Mitigation Reef (see Figure 1)

1) *Phase 1: Fully implemented monitoring*: Independent monitoring designed and conducted by CCC staff scientists will be done to evaluate the performance of the mitigation reef[5]. The sampling methodology, analytical techniques, and methods for measuring performance of the mitigation reef relative to the performance standards shall be described in the monitoring plan prepared for the mitigation reef[6]. Monitoring will ensue upon completion of the reef construction[2]. All performance standards must be met within 10 years[7,8]. The project will be considered successful when all the performance standards have been met each year for three consecutive years[9]. Hence, fully implemented monitoring will last a minimum of 10

years. All years that the project is in compliance will count towards the compliance period. The level of sampling effort may be reduced during this phase of monitoring if analyses of the data indicate that compliance of the performance standards can be adequately assessed using less sampling effort. Remediation may be required if the performance standards are not met within ten years and if three consecutive years of compliance has not occurred within 12 years [10, 11]. Note that the Executive Director could prolong this phase of monitoring or reinstate it if necessary following degradation of the artificial reef (resulting in a period of non-compliance) or remediation[12].

2) *Phase 2: Annual site inspections:* Monitoring can be reduced to annual site inspections[13,14], which will serve to identify noncompliance with the performance standards, when:
   a. The project has been in compliance with permit standards for at least three consecutive years, and
   b. The project has been evaluated for at least ten years post-construction.

The schedule for monitoring the mitigation reef project is shown in Figure 1.

Restored Wetland (see Figure 2)

1) *Phase 1: Fully implemented monitoring:* Independent monitoring designed and conducted by CCC staff scientists will be done to evaluate the performance of the wetland restoration project[5]. A description of the monitoring can be found in the wetland monitoring plan and details of the monitoring effort will be set forth in a work plan[15]. Monitoring will ensue upon completion of wetland construction[16]. Within 4 years of construction, the total densities and number of species of fish, macro-invertebrates and birds shall be similar to the densities and number of species in similar habitats in the reference wetlands[17]. All performance standards must be met within 10 years, which is the same amount of time required for the mitigation reef to meet all of the performance standards[7,8]. The wetland restoration project will be considered successful when all the performance standards have been met for each of three consecutive years[9]. All years that the project is in compliance will count towards the compliance period. Remediation may be required if all the performance standards are not met within ten years and if three successive years of compliance has not occurred within 12 years[18]. Note that the Executive Director could prolong this phase of monitoring or reinstate it if necessary following remediation or degradation of the wetland (resulting in a period of non-compliance)[12].

2) *Phase 2: Scaled back monitoring:* Upon determination that the project has been in compliance for three consecutive years, a scaled back phase of monitoring will ensue[14]. The scaled back monitoring program will be

designed and implemented by CCC staff scientists[5]. Reduction in effort will be based on analyses of data collected during the period in which the project was in compliance. Staff scientists will examine these data to determine the minimum effort that would have been necessary to assess compliance during the period**.** All monitoring, whether it is fully implemented or scaled back, must be sufficient for assessing compliance of the performance standards*.*

The schedule for monitoring the wetland restoration project is shown in Figure 2.

REMEDIATION

If the mitigation reef or restored wetland is not considered successful within 12 years post-construction or if the restored wetland has not met the biological community standard by year 4, then (at the discretion of the Executive Director):

1) The permittee shall fund an independent study to collect information needed to determine what remediation is required[19].
2) The permittee shall be required to implement any remedial measures determined necessary by the Executive Director in consultation with state and federal resource agencies and will provide funds for independent monitoring that evaluates the success of the required remediation[10,11,19]. Remediation monitoring may be different from the compliance monitoring required by the permit.

If the mitigation reef or restored wetland is in a period of reduced monitoring and if it falls out of compliance for a period of two consecutive years, then to determine if non-compliance is an artifact resulting from a reduction in monitoring effort, full monitoring (Phase1) may be re-established for those standards that are out of compliance. If resumption of full monitoring leads to the conclusion that the reduction in monitoring was responsible for non-compliance, then monitoring will remain at the full levels for the duration of the study or until the Executive Director concludes that reduced monitoring could be reinstituted[12]. CCC staff scientists will be responsible for designing and implementing the reduced monitoring program[5].

If resumption of full monitoring leads to the conclusion that non-compliance is due to poor performance of the mitigation project then:

1) The permittee shall be required to fund an independent study to collect the information necessary to determine what remediation is needed [19]
2) The permittee shall be required to implement any remedial measures determined necessary by the Executive Director in consultation with state and federal resource agencies and will provide funds for independent monitoring that evaluates the success of the required remediation[10,11,19]. Remediation monitoring may be different from the compliance monitoring required by the permit.

Permit (No. 6-81-330-A) language supporting CCC staff's interpretations on SONGS project compliance

1. (III.A.3.4). Upon completion of construction of the wetland, monitoring shall be conducted to measure the success of the wetland in achieving stated restoration goals (as specified in restoration plan) and in achieving performance standards, specified below.

2. (III.B. 2.4). Following completion of construction the mitigation reef shall be monitored for a period equivalent to the operating life of SONGS.

3. (III.A.3.0). Monitoring, management (including maintenance), and remediation shall be conducted over the "full operating life" of SONGS Units 2 and 3**.** Full operating life" as defined in this permit includes past and future years of operation of SONGS units 2 and 3 including the decommissioning period to the extent there are continuing discharges. The number of past operating years at the time the wetland is ultimately constructed, shall be added to the number of future operating years and decommission period, to determine the length of the monitoring, management and remediation requirement.

4. (III.B 2.4). The permittee shall insure that the performance standards and goals set forth in this condition will be met for at least the length of time equivalent to the full operating life of SONGS Units 2 and 3…."Full operating life" as defined in this permit includes past and future years of operation of SONGS Units 2 and 3, including the decommissioning period to the extent there are continuing discharges.

5. (III.C.1.0). Personnel with appropriate scientific or technical training and skills will, under the direction of the Executive Director, oversee the mitigation and monitoring functions identified and required by conditions II-A through C. The Executive Director will retain approximately two scientists and one administrative support staff to perform this function.

This technical staff will oversee the preconstruction and post-construction site assessments, mitigation project design and implementation (conducted by permittee), and monitoring activities (including plan preparation); the field work will be done by contractors under the Executive Director's direction. The contractors will be responsible for collecting the data, analyzing and interpreting it, and reporting to the Executive Director.

6. (III.B.2.4. A monitoring plan for the mitigation reef shall be developed by the Commission staff scientists pursuant to Condition D. The monitoring plan shall be completed within six months of approval of a coastal development permit for the mitigation reef proposed in a final plan developed pursuant to this condition.

The monitoring plan shall provide an overall framework to guide the monitoring work. The monitoring plan shall describe the sampling methodology, analytical techniques, and methods for measuring performance of the mitigation reef relative to the performance standards identified below.

7. (III.B.2.4). The independent monitoring program for the mitigation reef shall be designed to assess whether the performance standards have been met. If these standards are met after ten years following the completion of construction, then monitoring can be reduced to annual site inspections.

8. (III.B.2.4). If the standards listed above are not met within ten years after reef construction, then the permittee shall undertake those remedial actions the Executive Director deems appropriate and feasible.

9. (III.C.3.0). The mitigation projects will be successful when all performance standards have been met each year for a three-year period. The Executive Director shall report to the Commission upon determining that all of the performance standards have been met for three years and that the project is deemed successful.

10. (III.B.2.4). The permittee shall undertake necessary remedial actions based on the monitoring results and annual site inspections for the full operating life of the SONGS Units 2 and 3.

11. (III.B.2.4). If the standards listed above are not met within ten years after reef construction, then the permittee shall undertake those remedial actions the Executive Director deems appropriate and feasible.

12. (III.C.3.0). If subsequent monitoring shows that a standard is no longer being met, monitoring may be increased to previous levels, as determined necessary by the Executive Director.

13. (III.B.2.4). The independent monitoring program for the mitigation reef shall be designed to assess whether the performance standards have been met. If these standards are met after ten years following the completion of construction, then monitoring can be reduced to annual site inspections.

14. (III.C.3.0). If the Commission determines that the performance standards have been met and the project is successful, the monitoring program will be scaled down, as recommended by the Executive Director and approved by the Commission. A public review shall thereafter occur every five years, or sooner if called for by the Executive Director.

15. (III.A.3.1). A monitoring and management plan will be developed in consultation with the permittee and appropriate wildlife agencies, concurrently with the preparation of the restoration plan, to provide an overall framework to guide the monitoring work. It will include an overall description of the studies to be conducted over the course of the monitoring program and a description of management tasks that are anticipated, such as trash removal. Details of the monitoring studies and management tasks will be set forth in a work program.

16. (III.A.3.4). Upon completion of construction of the wetland, monitoring shall be conducted to measure the success of the wetland in achieving stated restoration goals (as specified in restoration plan) and in achieving performance standards.

17. (III.A.3.4.b.1). *Biological Communities.* Within 4 years of construction, the total densities and number of species of fish, macroinvertebrates and birds shall be similar to the densities and number of species in similar habitats in the reference wetlands.

18. (III.A.3.4). The permittee shall be fully responsible for any failure to meet these goals and standards during the full operational years of SONGS Units 2 and 3. Upon determining that the goals or standards are not achieved, the Executive Director shall prescribe remedial measures, after consultation with the permittee, which shall be immediately implemented by the permittee with Commission staff direction. If the permittee does not agree that remediation is necessary, the matter may be set for hearing and disposition by the Commission.

19. (III.B.2.4). Executive Director may also use any other information available to determine whether the performance standards are being met. If information from the annual site inspections or other sources suggests the performance standards are not being met, then the permittee shall be required to fund an independent study to collect the information necessary to determine what remediation is needed. The Executive Director shall determine the required remedial actions based on information from the independent study. The permittee shall be required to implement any remedial measures determined necessary by the Executive Director in consultation with state and federal resource agencies, as well as provide funds for independent monitoring that evaluates the success of the required remediation. As described under the funding option (Condition D) of this permit, the cost of remediation shall not be limited if the permittee elects to implement the mitigation reef.

Figure 1. Idealized monitoring schedule for the mitigation reef showing the minimum time periods for the two phases of monitoring: (1) Fully implemented monitoring and (2) annual site inspection. The actual time periods for each phase may be longer, depending on the performance of the project.

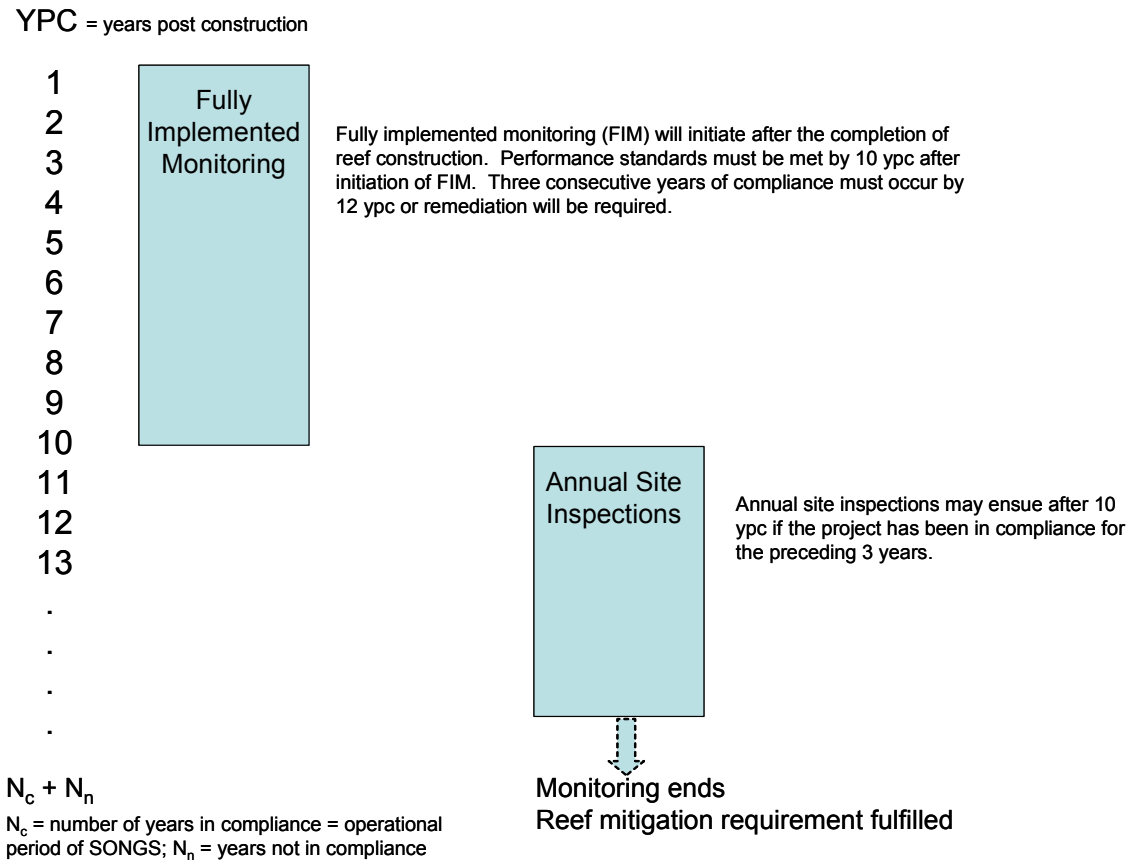
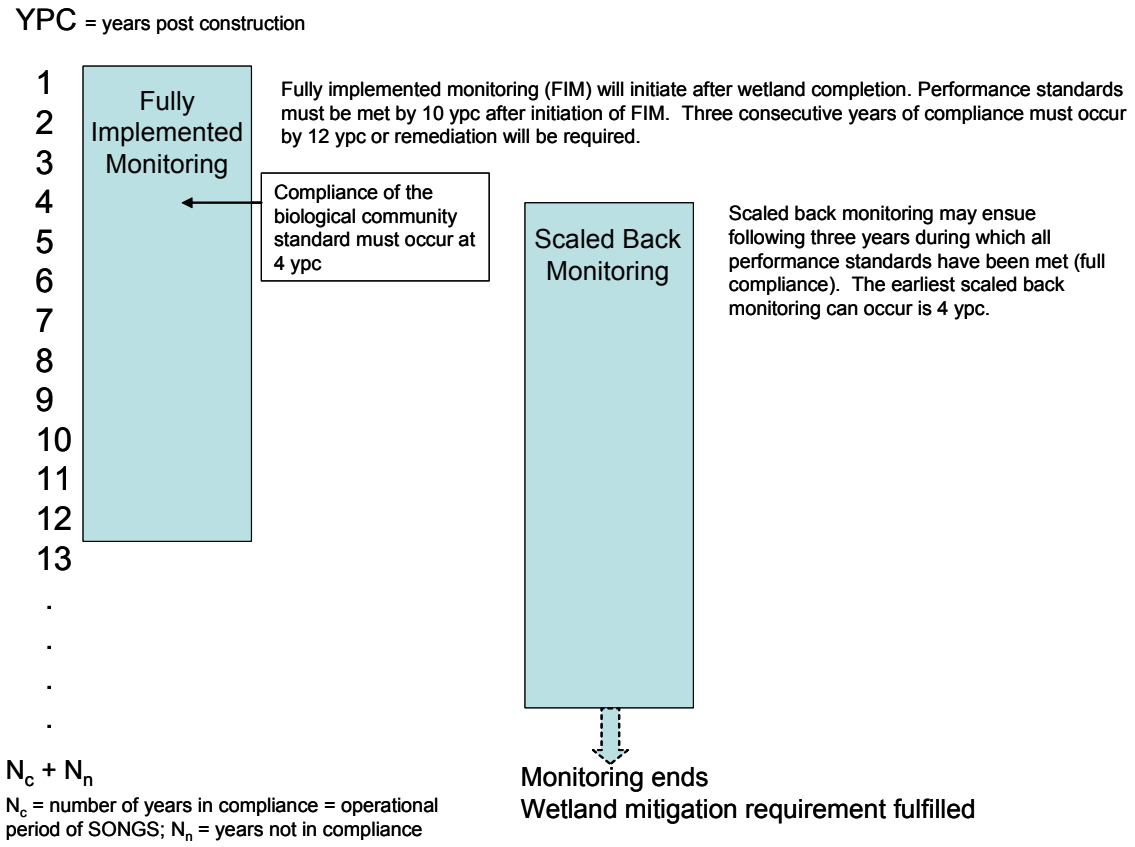


Figure 2. Idealized monitoring schedule for the wetland restoration project showing the minimum time periods for the two phases of monitoring: (1) Fully implemented monitoring and (2) scaled back monitoring. The actual time periods for each phase may be longer, depending on the performance of the project.



II. METHODS FOR DETERMINING SIMILARITY BETWEEN THE MITIGATION AND REFERENCE SITES FOR THE PUROPOSE OF EVALUATING COMPLIANCE OF THE SONGS' MITIGATION PROJECTS

INTRODUCTION

Many of the performance standards used to assess compliance of the SONGS mitigation projects require various attributes of the restored wetland and artificial reef mitigation sites to be similar to those of nearby relatively undisturbed, natural reference sites. The method used for determining similarity between the mitigation and reference sites for these "relative performance standards" is to be specified in a work program developed by contract scientists working for the California Coastal Commission (Permit No. 6-81-330-A, Condition A.3.4). The SONGS permit provides two examples of statistical methods that might be used to determine similarity: (1) the range test (within the range of the means of the reference sites), and (2) the 95% confidence interval (CI) of the mean of the reference sites (Box 1)

Contract scientists working for the Commission used the range test and a confidence interval test (i.e. the 95% CI centered on the mean of the reference sites) to evaluate the performance of the different artificial reef designs tested during the experimental phase of the reef mitigation project (referred to as the Universe Approach and Sample Approach, respectively; Reed et al. 2005). However, use of such range and confidence interval tests for evaluating similarity in the context of SONGS permit compliance may be problematic due to some inherent limitations of these methods, as described below. In the following sections, we describe the conceptual basis of the range and confidence interval tests and discuss their pros and cons for determining similarity in the context of SONGS permit compliance. We conclude by introducing a new approach to determining similarity (the Hybrid Approach) that draws on the strengths of these two methods and minimizes their weaknesses.

Note that the permit requires the mitigation site to be "similar" to reference sites but does not address the issue of the performance variable values being above versus below the values at the reference sites. In the examples given in this document, we focus on mitigation values below the reference values for simplicity, but it is possible that a mitigation site would be judged dissimilar because the value of some attribute was higher than that at the reference sites.

METHODS FOR DETERMINING SIMILARITY

Although the permit mentions two possible approaches that could be used to assess similarity, there are other possibilities. There is a well developed literature concerning methods for assessing similarity of ecological communities. For example, similarity could be assessed using multivariate statistical techniques (such as cluster analysis or non-metric multi-dimensional scaling; see Edwards and Proffitt 2003) or established indices such as the Jaccard Index of Similarity (Jaccard 1901). Most of these approaches focus on assessing the similarity of communities; they evaluate how similar the species identities and abundances are in two or more communities. Because the SONGS performance standards do not require similar communities *per se*, these types of methods are not appropriate for evaluating permit compliance of the SONGS mitigation projects. The most appropriate approaches for SONGS mitigation are based on the range of reference site values or some type of confidence interval around the mean of a reference value (Box 1), and these are discussed below.

Box 1. Confidence Intervals. For practical reasons populations are usually sampled rather than censused to obtain a measure of central tendency of the population such as the mean or median. Sampling error is typically associated with such sample statistics and the use of confidence intervals is a common way to account for such error. A confidence interval is a range of values that has a defined probability of including the sample statistic of interest (e.g. mean). For example, if the 95% confidence interval of the population mean is the interval between 25 and 75, then there is a 95% percent probability that the true population mean (not the sample mean) lies between 25 and 75. Another way to use confidence intervals is to define a range of acceptable values. This is the model used to assess compliance with mitigation standards that are derived from values from reference sites.

Although relatively few published studies address how mitigation sites should be judged relative to reference sites, all of the approaches discussed here have been used previously. For example, Craft et al. (2003) used the range of means of natural wetlands to judge the development of restored wetlands of different ages. Kentula et al. (1993) were early proponents of evaluating restoration success by using reference sites and assessed similarity to reference sites by comparing the mean at a mitigation site to confidence intervals calculated from data collected at *multiple* reference sites (an approach we call the Composite Confidence Interval [CCI]). In this case the restoration was judged to be successful if the mean value of the mitigation site was within the confidence interval of the reference sites. The basis for using such an approach is: (1) the reference sites are considered to be independent replicates of the natural condition, (2) measurements are made with error and, (3) use of the confidence interval based on reference sites as replicates ensures a realistic target for mitigation that accounted for measurement error. Evans and Short (2005) used a version of the CCI, although they evaluated similarity using standard deviations rather than a confidence interval. Proper use of the CCI requires that the confidence interval be calculated around the mean of the reference sites, with replication based on the

number of reference sites. When reference sites are few and variable the use of the CCI for assessing similarity is ineffective because it produces undesirably large confidence intervals with the lower limit often close to or including zero. Because of this trait, the unit of replication sometimes has been based on sub samples within sites (e.g. transects) in cases where there have been few reference sites. Such improper use of the CCI makes the invalid assumption that the sub samples are independent of site[2]. The proper use of the CCI for evaluating similarity for the SONGS mitigation projects would result in undesirably large confidence intervals with little power to discriminate because only two or three reference sites are envisioned. For this reason the CCI is not considered a useful approach for determining similarity for the SONGS mitigation projects.

Similarity has also been assessed by comparing the mitigation site mean to the confidence interval calculated from data collected at a *single* reference site (Zedler and Callaway 1999, Burdick et al. 1997). This Separate Confidence Interval (SCI) test was also used by Ambrose et al. (2006) to evaluate similarity at five different mitigation sites, using two reference sites per mitigation site.

The Range test

Use of the Range test to determine similarity relies on comparing the mean values of a performance variable at the mitigation and reference sites. With the Range test, the mitigation site would be considered similar to the reference site for a given performance variable if the mean of the performance variable at the mitigation site were to fall between the means at the reference sites. For example, in Figure 3a the mean of the mitigation site is between the means of the reference sites in scenarios A through E, but not in scenarios F through H. In these hypothetical examples, the mitigation site would be judged to be similar to the reference sites in scenarios A through E and dissimilar in scenarios F through H.

The underlying assumption of the range test is that the selected reference sites represent the full range of suitable sites for evaluating the mitigation project; in other words, the selected reference sites represent the “universe” of possible reference sites. The Range test is one of the simplest and least ambiguous ways of determining similarity between the mitigation and reference sites.

A major limitation of the Range test is that it does not consider sampling and measurement errors within reference sites and thus could underestimate the true range of reference site means for a given variable.

Separate Confidence Interval (SCI) test

[2] Transects were used as the unit of replication when applying the CCI to evaluate the suitability of different reef designs during the experimental phase of the SONGS reef mitigation program (Reed et al. 2005). This improper usage, however, had no consequences to the recommendations made regarding the suitability of designs for the mitigation reef because conclusions based on the range test and the CCI did not differ.

Use of the SCI test to determine similarity between the mitigation and reference sites involves calculating separate confidence intervals for each reference site using within-site sub samples as replicates. The range of values used to determine similarity for a given performance variable is set by the upper confidence limit of the reference site with the highest mean value for the performance variable and the lower confidence limit of the reference site with the lowest mean value for the performance variable. In essence it requires the mitigation site to perform at least as well (in a statistical sense) as the lowest performing reference site. For example, in Figure 3b (which uses the same set of scenarios as Figure 3a) the mitigation site would be judged similar to the reference sites for all scenarios but G. This result differs from the range test (Figure 3a) in that scenarios F and H would be judged similar by the SCI, but dissimilar by the Range test. Unlike the CCI, the SCI does not assume that all reference sites can be characterized by a single mean value for a given performance variable, nor does it allow the mean value of any reference site to fall outside the range of values used to assess similarity.

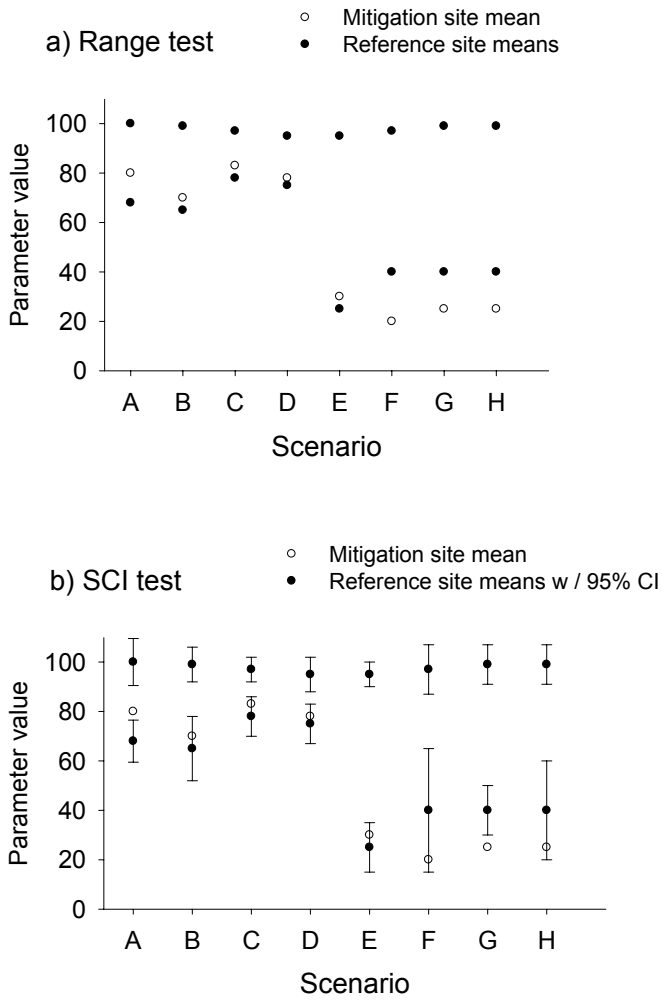


Figure 3. Hypothetical scenarios for determining similarity involving one mitigation and two reference sites using (a) the Range test and (b) the Separate Confidence Interval (SCI) test. The parameter values in (a) and (b) are the same for each scenario.

The issue of statistical power and within-site variability can be critical for decisions regarding similarity when the mean of the mitigation site falls outside the range of the means of the reference sites. For example, in scenarios G and H, the mean values for the reference and mitigation sites are identical, but the variability around the reference site with the lowest mean value differs (variability is greater in scenario H). Consequently the value of the performance variable for the mitigation reef would be judged to be similar to the reference sites for scenario G, but not for scenario H. Factors that increase the SCI of the

reference sites, whether due to natural variability or sampling design, will make it easier for a mitigation site whose mean is outside the range of the reference site means to nonetheless be judged similar to the reference sites.

MULTIPLE PERFORMANCE STANDARDS

The examples described above have focused on assessing similarity for a single performance standard. The SONGS permit contains multiple performance standards for both the reef and the wetland mitigation projects, all of which must be met in order for the projects to be in compliance with the SONGS permit (see "I. Assessing Compliance of the SONGS Mitigation Projects" above).

In evaluating multiple performance standards, consideration must be given to how the means of the mitigation site and reference sites vary due to chance alone. For example, imagine there are three reference sites, all of which are from the same population of sites. If we measure a number of different variables at these three sites, then we would expect that, on average, the mean of any given reference site would have the lowest value for one-third of the variables. We would expect a mitigation site that truly is similar to natural reference sites to have similar performance. It follows that if there were two reference sites and one mitigation site, then one would expect the mitigation site to have the lowest value for one third of the variables if it was similar to the reference sites[3]. Requiring that the value at the mitigation site be higher than any reference site for 100% of the variables would constitute a different and more severe threshold for the mitigation site than for the reference sites. Conversely, allowing a mitigation site to have the lowest value for more than one-third of the variables would likely not result in full compensation for the resources lost due to SONGS operations. Thus, relying solely on the SCI approach to evaluate similarity would be problematic because it would allow the mitigation site to have a lower mean value than all the reference sites for *every* performance variable and still be in compliance. This would not be consistent with the goals of the SONGS permit.

THE HYBRID APPROACH

Because neither the Range nor the SCI tests are satisfactory by themselves for determining similarity for the SONGS mitigation projects, we propose a hybrid approach that avoids the limitations of these tests. The Hybrid Approach requires that the value of the mitigation site be within the range of the SCI of the reference sites for all performance variables ***and*** that it not be the lowest more often than expected by chance alone.

To illustrate this approach, imagine a hypothetical scenario in which the success of a mitigation site is based on nine performance variables and two reference sites. Compliance using the Hybrid Approach would require: (1) the

[3] Generally, the probability of the value for a single performance variable being lower at the mitigation site than at any of the reference sites by chance alone will be 1/(x+1), where x is equal to the number of reference sites.

value of the means for all nine variables at the mitigation site be within the SCI of the two reference sites, and (2) that the mitigation site not have the lowest value for more than 1/3 of the variables (in this case, no more than 3 variables). Criterion 1 ensures that the values of all the performance variables at the mitigation site will be greater than that of the lower confidence limit of the mean of the reference site with the lowest value. Criterion 2 is based on the null hypothesis that the resource value of the mitigation site represents a sample from the same population as the reference sites and that the mitigation site behaves like the reference sites with respect to all the performance variables. If this is true, then in the example above it follows that each of the three sites (i.e., the mitigation site and the two reference sites) has an equal one-third chance of having the lowest value for any performance variable. This criterion eliminates the possibility of concluding that the mitigation site is in compliance when it has the lowest value for a large number of the performance variables. When applied together, the two criteria of the Hybrid Approach guarantee that the assessment of similarity is consistent with the SONGS permit requirement that **all** performance standards be met for the mitigation project to be in compliance without requiring the mitigation site to outperform all the reference sites for every performance variable.

THE IMPACT OF THE SCI ON SAMPLING DESIGN

The use of the SCI as the first criterion in the Hybrid Approach has dramatic and insightful implications for the sampling designs used to assess compliance which stem from the interplay among (1) the size of the confidence interval based on a defined significance level (Type I error = α), (2) the desired statistical power (1-Type II error = 1-β), (3) the percentage difference between the value of a performance variable at the mitigation site and the lowest performing reference site (effect size), and (4) sample size. Generally speaking, for a given sampling effort and target effect size, minimizing the Type I error (i.e., minimizing the chance of incorrectly concluding that the mitigation site is not similar to the reference site) tends to increase the Type II error (i.e., incorrectly concluding that the mitigation and reference sites are similar) and vice versa. Committing a Type I error when assessing similarity would be unfair to SCE because they would not be given credit for a successful mitigation project. On the other hand committing a Type II error would be a disservice to the public because lost resources would not be adequately compensated for. Increasing sample size (**n**) will reduce both types of errors, but comes with added monetary costs. Thus, in the interest of fairness to the public and the utility company it is desirable to strike a reasonable balance in setting α, β, and effect size when implementing the SCI to evaluate the success of the SONGS mitigation projects.

The conventional value used for α in the ecological literature is typically 0.05. However, given the variability inherent in natural wetland and reef ecosystems, it will likely be difficult to achieve an α = 0.05 without excessive sampling effort and/or reduced statistical power. A reasonable approach to balancing the risks of Type I and Type II errors in evaluating the SONGS

performance standards is to set $\alpha = \beta$ (as recommended by Peterman 1990). Allowing $\beta > 0.20$ (i.e. power < 0.80) is unconventional (Cohen 1988), and would result in an unacceptably high probability of failing to detect a difference between the mitigation and reference site. Hence, we recommend setting both α and β equal to 0.20.

Although the SONGS permit does not specify an effect size for the purpose of determining similarity, its intent clearly is to fully replace coastal resources lost due to the operation of the power plant. Ideally, the mitigation project should have the same value as the reference site. In practice, there will be uncertainty about the true value at the mitigation and reference sites due to sampling error, and requiring the mitigation site to be identical to the value at the reference site would be unreasonable. In designing the monitoring program, we propose using the "effect size" to accommodate this uncertainty. Given this uncertainty, it seems reasonable to assume that a SONGS mitigation project estimated to yield less than 80% of the resource value of natural reference sites will not result in full compensation. Hence monitoring programs used to evaluate the performance standards should be designed to have the capability to detect a difference between the mitigation and reference sites of at least 20%. As per the recommendation of Mapstone (1995), effect size will be used as a criterion for developing the sampling designs used to evaluate the permit standards, not as an explicit criterion for determining permit compliance.

REFERENCES

Ambrose, R. F., R. R. Vance, and N. Diaz. 2006. Wetland Restoration Monitoring Report for Navy Base Ventura County, Mugu Lagoon: July 2001 to September 2005. Report to the U.S. Navy, Naval Base Ventura County. December 4, 2006.

Burdick, D.M, M. Dionne, R.M. Boumans and F.T. Short. 1997. Ecological responses to tidal restorations of two northern New England salt marshes. Wetlands Ecology and Management 4: 120 144.

Cohen, J. 1988. Statistical power analysis for the behavioural sciences. 2nd edition. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Craft C., Megonigal P., Broome S., Stevenson J., Freese R., Cornell J., Zheng L. and Sacco J. 2003. The pace of ecosystem development of constructed *Spartina alterniflora* marshes. Ecological Applications 13: 1417-1432.

Edwards, K.R. and C.E. Proffitt. 2003. Comparison of wetland structural characteristics between created and natural salt marshes in southwest Louisiana, USA. Wetlands 23: 344–356.

Evans, N.T. and F.T. Short. 2005. Functional Trajectory Models for Assessment of Transplanted Eelgrass, *Zostera marina* L., in the Great Bay Estuary, New Hampshire. Estuaries 28: 936–947.

Jaccard, P. 1901. Bulletin del la Société Vaudoise des Sciences Naturelles 37: 241-272.

Kentula, M.E., R.P. Brooks, S.E. Gwin, C.C. Holland, A.D. Sherman and J.C. Sifneos. 1993. An approach to improving decision making in wetland restoration and creation. C.K. Smoley/CRC Press, Boca Raton, Fl.

Mapstone, B. D. 1995. Scalable decision rules for environmental impact studies: Effect size, Type I, and Type II Errors. Ecological Applications 5: 401-410.

Pacific Estuarine Research Laboratory (PERL). 2002. Tijuana river national estuarine research reserve: Annual report on ecosystem monitoring. Annual Report for Monitoring Conducted July 1, 2001 - June 30, 2002. NOAA Technical Memorandum NOS SRD.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47, 2-15.

Reed, D.C, S.C. Schroeter, and D. Huang. 2005. Final report on the findings and recommendations of the experimental phase of the SONGS Artificial Reef Mitigation Project. Prepared for the California Coastal Commission August 1, 2005. Available at http://www.coastal.ca.gov/

Whitney, K. 2005. Appendix 6. Estimating bird species richness and density in coastal wetlands of southern California: application to monitoring birds in the San Dieguito Lagoon Restoration project. In: Monitoring Plan: The S.O.N.G.S.

Wetland Mitigation Program. California Coastal Commission, 45 Fremont, Suite 2000, San Francisco, California 94105.

Zedler, J.B. and J.C. Callaway. 1999. Tracking wetland restoration: Do mitigation sites follow desired trajectories? Restoration Ecology 7: 69-73.

Glossary

Confidence interval: The interval between the upper and lower confidence limits of the mean and specified by a percentage ranging from 0 to 100, which indicates the probability that the true mean lies within the confidence interval. A 95% confidence interval is often used.

Composite Confidence Interval (CCI): In the context of monitoring the performance of a mitigation project, the confidence interval centered on the mean calculated from data averaged across all reference sites.

Effect size: The difference between the values of two means. In the context of monitoring the performance of a mitigation site, the difference between the mean values of a performance variable at a reference and a mitigation site, often expressed as a percentage.

Hybrid Approach: A method for assessing similarity between the mitigation and reference sites that requires the mitigation site be within the range of the Separate Confidence Interval of the reference sites for all performance variables ***and*** that it not have the lowest value for the performance variables more often than expected by chance alone .

Performance Standards: Specific requirements in a permit that specify how a mitigation site should perform. Two different types of performance standards are included in the SONGS permit. *Fixed performance standards* require that performance be judged against a predetermined fixed value (e.g., "The total aerial extent of the mitigation reef (including the experimental reef and all larger artificial reefs) shall be no less than 150 acres."). *Relative performance standards* require that performance be judged relative to reference sites (e.g., "The resident fish assemblage shall have a total density and number of species similar to natural reefs within the region"). Compliance with relative performance standards is based on similarity to the reference sites.

Performance variable: The variable of interest used to assess a performance standard (e.g., the density of resident fish).

Power (1-β): The probability of correctly rejecting the alternative hypothesis. In the context of monitoring the performance of a mitigation project, the probability of correctly concluding that the value of a performance variable at the mitigation site differs from (i.e., is less than) the value at the reference site with the lowest value.

Separate Confidence Interval (SCI): In the context of monitoring the performance of a mitigation project, the confidence interval set by the upper confidence limit of the reference site with the highest mean value for a given performance variable and the lower confidence limit of the reference site with the lowest mean value for that performance variable

Type I error (α): Probability of incorrectly rejecting the null hypothesis. In the context of monitoring the performance of a mitigation project, the probability of incorrectly concluding that the value of a performance variable at the mitigation site is less than that of the reference site with the lowest value.

Type II error (β): Probability of incorrectly rejecting the alternative hypothesis. In the context of monitoring the performance of a mitigation project, the probability of incorrectly concluding that the value of a performance variable at the mitigation site does not differ from (i.e., is not less than) that of the reference site with the lowest value.