SWAMP
Surface Water Ambient Monitoring Program

Final Technical Report | 2009

Recommendations for the Development and Maintenance of a Reference Condition Management Program (RCMP) to Support Biological Assessment of California's Wadeable Streams

March 2009

California Water Boards

www.waterboards.ca.gov/swamp

Recommendations for the development and maintenance of a reference condition management program (RCMP) to support biological assessment of California's wadeable streams

Report to the State Water Resources Control Board's
Surface Water Ambient Monitoring Program (SWAMP)

Peter Ode, SWAMP Bioassessment Coordinator
Aquatic Bioassessment Laboratory/
Water Pollution Control Laboratory
California Department of Fish and Game
2005 Nimbus Road
Rancho Cordova, CA 95670

Ken Schiff, Deputy Director
Southern California Coastal Water Research Project
3535 Harbor Blvd., Suite 110
Costa Mesa, CA 92626

March 2009

Technical Report 581

Table of Contents

EXECUTIVE SUMMARY

Direct measures of the ecological condition of waterbodies have received a recent surge in interest within California's water quality management and regulatory programs because biology-based assessments have several advantages over chemistry- or toxicity-based assessments. Biological assessments are more closely linked to the beneficial uses to be protected and chemistry- or toxicity-based criteria usually lack the predictive ability to infer biological condition. Ultimately, California needs to develop biology-based standards, or biocriteria, as a regulatory tool for monitoring and protecting aquatic life use.

Biological assessment tools, including biocriteria, attempt to objectively "score" the biological integrity at a given site. A crucial component to the development of assessment tools is understanding biological expectations at reference sites that consist of natural, undisturbed systems. These reference systems set the biological condition benchmarks for comparisons to the site(s) being evaluated. Two recent external reviews of the State Water Board's Surface Water Ambient Monitoring Program (SWAMP) affirmed the importance of a sound statewide reference condition program (i.e., TetraTech 2002, SPARC 2006).

In October 2007, the SWAMP bioassessment committee assembled a technical panel of statewide and national experts in bioassessment. The panel met for three days to develop a set of recommendations that the SWAMP program could use to establish and maintain a comprehensive reference condition management program (RCMP). The program accounts for biological variation caused by natural environmental gradients and balances statewide consistency with the flexibility needed to adapt to California's diverse regional settings. Furthermore, the plan allows for adaptive refinement over time.

The panel defined a general strategy for establishing the RCMP that has four components:

1. California will be divided into different geographic regions based on coarse biogeographic similarities in order to partition some of the natural variability among regions (these boundaries should be consistent with those used for the SWAMP Perennial Streams Assessment)
2. A pool of reference sites will be assembled within each region through a sequential process of identification and screening of candidate sites
3. The sites within each reference pool will be managed through iterative review of data to refine regional boundaries, ensure continued suitability of sites and ensure adequate representation of natural gradients
4. A monitoring design will be created for sampling this pool of reference sites to document the range of biological and physical condition at reference sites, and monitor for changes to this condition over time

The panel recommended identifying and screening candidate locations to create a pool of verified reference sites using either a "standard model" or an "alternative model". The standard model will cover the vast majority of the state where high quality sites are available. The alternate model will apply in those regions where an insufficient quantity of high quality sites exist and another strategy is required for selecting candidates for the reference pool. This may include regions such as the agriculturally dominated Central Valley or the intensely urbanized southern California coastal plain.

The standard model is a synthesis of widely used techniques for selection and screening candidate sites using a toolbox consisting of existing site data, GIS techniques, expert knowledge and site visits. The alternative approach consists of two general strategies: 1) modification of standard tools (e.g., lowering the GIS screening thresholds, collecting more intensive site data) and 2) use of non-standard approaches. The non-standard approaches include:

- Select best sites using existing biological indices
- Species pool approach
- Factor-ceiling approach
- Model taxon preferences for limiting environmental gradients

These different approaches are not mutually exclusive and several panel members recommended they be used in combination to provide weight-of-evidence that candidate sites are acceptable for the reference pool in these difficult locations.

The panel outlined a monitoring strategy for the RCMP, which included recommendations for sampling methods, sampling density and frequency, and the set of biological, chemical and physical attributes that should be collected at each reference site. The panel strongly recommended that the RCMP should be compatible with ongoing statewide monitoring programs such as the newly developed SWAMP Perennial Streams Assessment. For the monitoring design, the panel recommended both random and targeted sites. A probabilistic rotating panel was suggested for the random design because it provides an unbiased method for defining natural variability while still optimizing large-scale trend detection. Targeted repeated sampling designs are useful for detecting trends at specific locations; some of these sites have been sampled for years and provide a rich history that should not be lost.

To guide the SWAMP program as it implements the RCMP, the panel made a series of recommendations for prioritizing the elements of the plan. The panel recommended that the implementation begin by screening existing datasets for reference sites, followed by a combination of GIS screens and site visits to fill in gaps in regions with few reference sites.

FOREWORD

The recommendations in this document were developed by a technical panel composed of experts in bioassessment. The panel reflected a broad range of local, statewide, and national experiences with freshwater bioassessment, specifically with defining reference conditions for bioassessment and biocriteria. The panel met for three days on October 17-19, 2007 to outline the content of this document. The meeting followed a four-step process:

1) Defining the background of the problem
2) Establishing a set of guiding philosophies for the development of a reference site management plan
3) Providing general guidance by outlining an overall approach
4) Providing detailed guidance for specific technical issues

This document follows a similar format. This document captures all of the items agreed to by consensus of the group and attempts to point out diverging opinions or unresolved issues. On occasion, we expand on key concepts that were implicit to our discussions, but may not have been discussed directly. Where appropriate, we use sidebars, tables, and figures to illustrate key concepts or provide additional information. Thank you to Dr. Robert Hughes (Oregon State University) for additional document review.

Panel Members (from left to right): David Herbst (University of California at Santa Barbara, Sierra Nevada Aquatic Research Laboratory), Peter Ode (California Department of Fish and Game, Aquatic Bioassessment Laboratory), Raphael Mazor (Southern California Coastal Water Research Project), D. Phil Larsen (US EPA retired, Western Ecology Division), Andrew Rehn (California Department of Fish and Game, Aquatic Bioassessment Laboratory), Lenwood Hall (University of Maryland, Wye Research and Education Center), Terrence Fleming (US EPA Region IX, Office of Water), Charles Hawkins (Utah State University, Western Center for Monitoring and Assessment of Freshwater Ecosystems), Alan Herlihy (Oregon State University, Department of Fisheries and Wildlife), Kenneth Schiff (facilitator, Southern Coastal California Water Research Project).

CONTEXT: LINKING BIOASSESSMENT TO BIOCRITERIA[1]

Aquatic bioassessment is the applied science of interpreting the ecological condition of waterbodies directly from the organisms that inhabit them. Biocriteria are narrative or numeric standards that define whether the integrity of biological communities is impaired at a specific site. Water quality regulatory programs can receive many benefits from adopting biology-based standards as targets of their policies and management actions. The key to using biology-based methods effectively is the establishment of benchmarks that objectively define the biological expectations (or potential) of a given site. Reference conditions provide these objective benchmarks.

Why bioassessment?

The Clean Water Act (Section 101a) requires states to "restore and maintain the chemical, physical and biological integrity" of their waterbodies. For decades, most state water quality monitoring programs have focused on the chemical integrity (and to a lesser extent physical integrity) of waterbodies largely because these parameters are relatively simple to sample, relatively straightforward to measure and evaluate, and methods for developing chemical criteria are relatively standardized. While chemical/ toxicological and physical condition monitoring may provide indirect measures of ecological condition, exclusive focus on these measures is inadequate for protection of aquatic life uses, one of the primary beneficial uses of concern in water quality management. Because many chemical/ physical water quality thresholds are based on toxicity to aquatic organisms (USEPA WQS handbook, 2nd Edition 1994), these indirect measures are often surrogates for the beneficial use that is the target of protection efforts. Furthermore, biological integrity is frequently impaired by factors other than chemical contamination (e.g., hydrologic alteration, instream and riparian habitat alteration). Ultimately, ecological condition assessments provide the most appropriate assessment endpoint for protecting beneficial uses associated with aquatic life.

Why biocriteria?

Adoption of biology-based regulatory standards has the potential to provide significant enhancements to the protection of water resource integrity because biocriteria provide a regulatory mechanism for applying bioassessment's benefits to numerous water resource objectives.

The State Water Resources Control Board's Surface Water Ambient Monitoring Program (SWAMP) is supporting the biocriteria goal by developing tools for using benthic macroinvertebrates as indicators of the health of aquatic life in perennial streams. SWAMP's objective is to develop the bioassessment infrastructure (i.e., standardized methods, analytical tools, objective reference conditions, interpretive framework) that will enable water quality programs to employ biocriteria in a variety of regulatory applications.

---

[1] Much of the information summarized in this section was synthesized from several key sources: Barbour *et al*. 1996a, Karr 1995, 1997, Stoddard *et al*. 2006.

Importance of reference conditions to bioassessment and biocriteria

The development of chemical criteria for aquatic life follows a relatively straightforward process in which numerical standards are based on results from lab-based toxicity testing. For most chemical contaminants, management objectives are focused on keeping concentrations below these toxicity-derived numerical thresholds. In contrast, biological objectives are based on maintaining the integrity of an assemblage (or multiple assemblages) of organisms. The challenge in developing biocriteria is translating what is currently a narrative standard into an ecologically relevant numerical standard. Development of biological criteria, however, is complicated by the fact that the composition of stream communities varies naturally even in the absence of anthropogenic stress. Thus, biocriteria will require a fundamentally different approach to establishing the expectations for unimpaired waterbodies.

Reference conditions (based on reference sites) provide a widely accepted mechanism for defining appropriate expectations and accounting for this natural variability (Hughes *et al*. 1986, Barbour *et al*. 1996, Karr and Chu 1999, Bailey *et al*. 2004). Reference sites are sections of streams that represent the desired state of stream condition (*sensu* Meyer 1997) for a region of interest. Once suitable reference reaches have been identified, these are used to characterize the range of biotic conditions expected for minimally disturbed sites. Deviation from this range is then used as evidence that test sites are impaired. [2]

> **A standardized lexicon of terms used to define biological expectations (adapted from Stoddard *et al.* 2006):**
>
> **Reference Condition (RC(BI))** ~ Because this term has been used for a wide range of meanings, Stoddard *et al.* (2006) argue that the term should be restricted to meaning "reference condition for biological integrity … in the absence of significant human disturbance or alteration"
>
> **Minimally Disturbed Condition (MDC)** ~ stream condition in the absence of "significant" human disturbance. Assumes all streams have some anthropogenic stresses, but in most cases will approach true RC(BI)
>
> **Historical Condition (HC)** ~ stream condition at a specific point in time (e.g., pre-Columbian, pre-industrial, pre-intensive agriculture, etc.)
>
> **Least Disturbed Condition (LDC)** ~ the best physical, chemical and biological conditions currently available ("the best of what's left"). This definition is sufficiently flexible to establish biological expectations even in highly altered systems
>
> **Best Attainable Condition (BAC)** ~ the expected ecological condition of least disturbed sites given use of best management practices for an extended period of time. This definition is helpful for communicating the potential for improving ecological condition above the currently best available conditions

Tiered aquatic life use (TALU) framework

The potential for biocriteria to improve aquatic life beneficial use protection can be greatly enhanced by a flexible framework for interpreting beneficial use attainment in a variety of settings. The current system of aquatic life use designations in California is outdated and does not adequately take advantage of advances in our ability to assess aquatic life use attainment. The USEPA and other states (notably, Maine and Ohio) have recognized this problem and have

---

[2] Approaches to the selection of reference sites have been discussed extensively (Hughes and Larsen 1988, Hughes 1995, Rosenberg *et al.* 1999, Stoddard *et al*. 2006). Although there has been much debate about terminology used to describe expected biological conditions, the concept is flexible and can be applied either very narrowly (e.g., the condition of waterbodies before European invasions) or more broadly (e.g., the "least disturbed" or "best available" conditions currently found in a region of interest). The strategy in this document follows terminology usage recommended by Stoddard *et al*. 2006 (see text box).

developed a “tiered” system of aquatic life use designations, which utilize the power of biological information to develop graduated levels of protection.

> **Reference conditions play two distinct roles in the TALU framework**
>
> The y-axis in the TALU framework (see Figure 1) is biological condition, a scale that measures ecological integrity of a site. The upper limit of the biological condition axis is anchored by an idealized target that represents the natural state of ecological conditions, or RC(BI) in the strict sense of Stoddard *et al.* (2006).
>
> In addition, within each tier, there is some best attainable condition (BAC, *sensu* Stoddard *et al.* 2006) for waterbody classes in these tiers.

“Tiered aquatic life uses” (TALU), supported by numeric biocriteria, can be thought of as defining different management levels for biological condition across a quality continuum that ranges between “natural” conditions to complete loss of the natural biological community (Figure 1). In the TALU system, “tiers” represent classes of waterbodies that are grouped based on similarities in anthropogenic disturbance levels, resulting biological condition, and recovery potential (USEPA 2005). Under this flexible system, designated uses to support aquatic life can cover a broad continuum of biological conditions, with some waters being closer to the ideal of “natural” or “minimal human impact” than others. Biocriteria applied in a framework of TALU designations can help shift the regulatory focus from performance-based standards (e.g., limiting the number of chemical criteria exceedences) to impact-based standards (e.g., attainment of ecological condition targets).

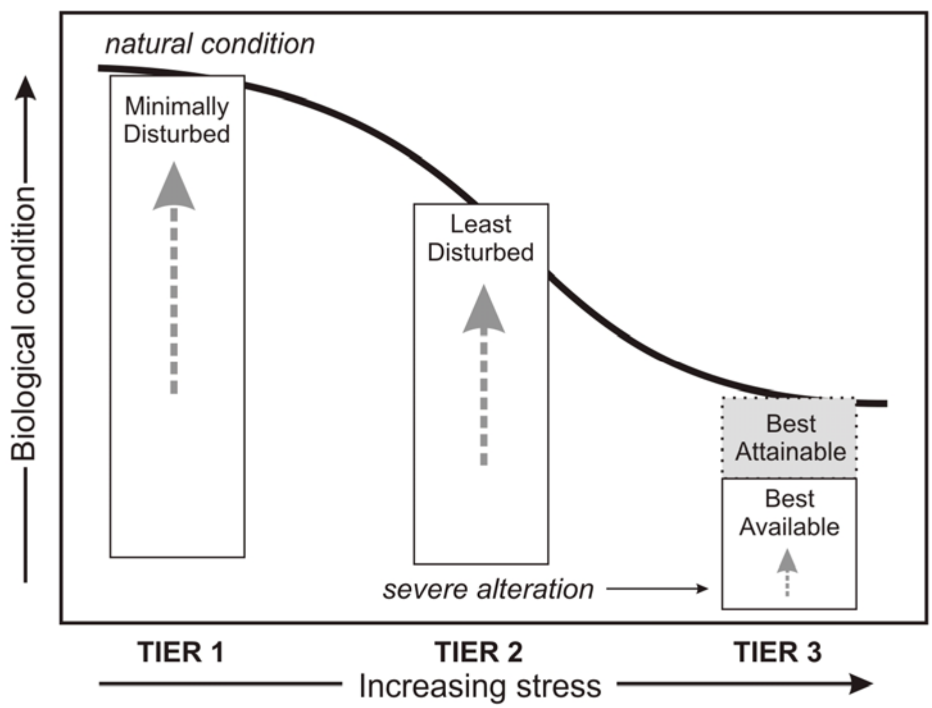

Figure 1. The biological condition gradient (BCG) used to define stream condition tiers in the TALU framework. Boxes indicate the expected range of biological condition scores at sites within each tier. Figure modified from Stoddard *et al*. 2006.

INTRODUCTION

General background

As the use of biological information in states' water quality regulatory programs has expanded across the US, these programs have followed a typical progression in which biosurveys (collection of biological samples, often as supplements to existing chemical monitoring) are followed by bioassessments (assessing ecological condition from biological data), finally progressing to full biocriteria (use of biological data to make regulatory decisions about aquatic life use condition).

As other programs proceeded along the path toward standardized interpretation of bioassessment data, they all recognized the need for grounding their programs with explicitly defined expectations for biological condition. Although criteria and procedures used to identify reference sites vary from program to program, the basic approaches used by most programs are quite similar. A partial review of water quality assessment programs in the North America (both state and federal programs), European Union (Water Framework Directive) and Australia (Water Reform Framework) revealed that many programs employed a similar GIS-based landscape-scale analysis to identify candidate watersheds, followed by site reconnaissance to evaluate reach-scale impacts (Barbour *et al*. 1996a, Whittier *et al*. 1987, Rosenberg *et al*. 1999, ANZECC and ARMCANZ 2000, Drake 2003, REFCOND 2003, Grafe 2004).

Reference sites manage natural variation

The composition of organisms at a site is a function of both natural and anthropogenic factors. These factors can be viewed as a series of "filters" that determine which taxa occur at a site (Poff and Ward 1990, Poff 1997, Statzner *et al*. 2001). For example, the pool of benthic macroinvertebrate taxa occurring within a large region like California's Sierra Nevada is a function of large scale processes (e.g., parent geology, climate and evolutionary history); the subset of taxa that occur at a given site at a given point in time is determined by a series of biotic and abiotic filters (e.g., life history traits, competition and predation, substrate composition, pH, thermal and hydrologic regimes, pollution tolerance) that further limit the occurrence of each taxon. The central challenge in bioassessment is to develop techniques that maximize the detection of signals of anthropogenic stress filters while minimizing the noise from natural filters. The identification of reference sites (that captures sources of natural variation) is a key component of most strategies for meeting this challenge (Hughes 1995, Wright and Li 2002, Bailey *et al*. 2004).

California's progress toward biocriteria implementation has followed a similar path. Since the early 1990s, bioassessment samples have been collected from more than 4000 sites by state and federal agencies alone (Figure 2). Some of these programs have been spatially extensive probability assessments of environmental condition such as the US EPA's Environmental Monitoring and Assessment Program (EMAP) and the California's Monitoring and Assessment Program (CMAP). Others are more directed studies to assess watershed-specific conditions or trends at locations of interest such as regional SWAMP monitoring, US Forest Service monitoring, and the US Geological Survey's National Water Quality Assessment Program (NAWQA). In addition, an abundance of additional sites have been sampled for National Pollutant Discharge Elimination System (NPDES) permit monitoring, and by citizen monitoring groups.

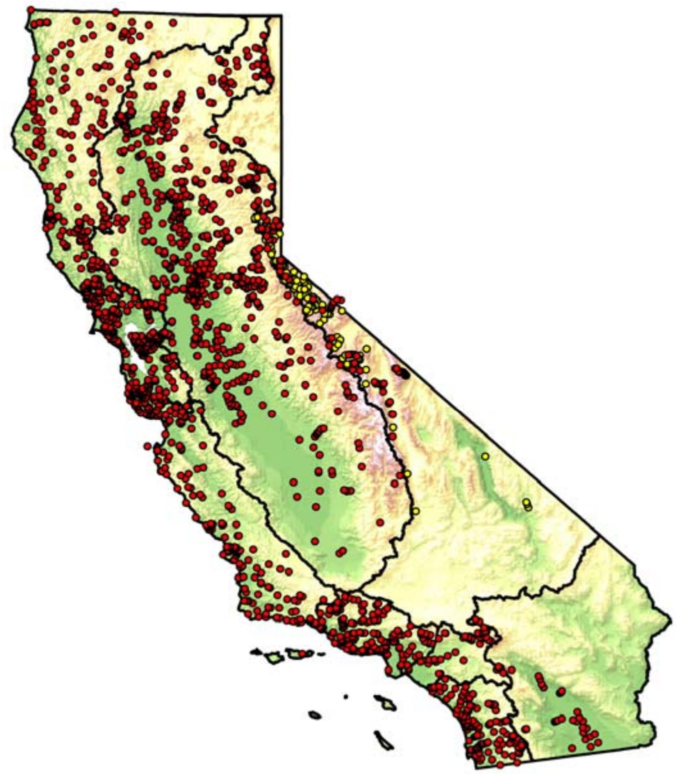

Figure 2. Approximately 3000 bioassessment sampling locations in California sampled between 1994 and 2007. Red circles represent sites processed by Aquatic Bioassessment Laboratory, yellow circles represent those processed by Sierra Nevada Aquatic Research Laboratory. More than 1000 other sites have been sampled by other state and federal agencies, permitted dischargers and citizen monitoring groups.

Because the early applications of bioassessment techniques in California were fragmented, the procedures for defining reference condition were largely *ad hoc* or project specific, with little or no attempt to apply consistent methods from project to project. Most of the reference or "control" sites used in early California bioassessment studies (e.g., point source enforcement cases, watershed specific bioassessments) were selected to define local expectations and were not selected using common criteria that would enable comparisons among projects.

Several large scale efforts to screen reference sites were undertaken in the early 2000's to support biological index development or as part of large state probability surveys: Western EMAP (2000-2003) and CMAP (2004-2007). In a concurrent effort, the USFS collaborated with scientists at Utah State University to identify over 200 reference sites on forest service lands in California between 1998 and 2000. Sites from these sampling programs were combined with other regional datasets to produce several of the main biotic indices used in California (statewide O/E models, North Coast IBI, South Coast IBI). Separate reference sites were used to develop the Eastern Sierra IBI (Herbst and Silldorf 2006).

In all of the large-scale studies between 1998 and 2007, both landscape scale and local scale factors were used for screening reference sites. Although common approaches were used to screen sites for most of these projects, little or no attempt was made to ensure consistency in screening among projects. This limits the utility of existing reference sites for statewide applications for several reasons. First, each project may use very different factors for selecting reference sites (e.g., one program may rely more on landscape scale factors while another may rely more on local scale factors). Second, some projects may use similar factors to select reference sites, but use different thresholds to screen sites (e.g., road density cutoffs or % upstream development cutoffs). Third, even when similar screening criteria are used for the same landscape or local scale factors, temporal variation in the reference site data has rarely been accounted for.

Why SWAMP needs an RCMP

The recent commitment by the SWAMP program to develop bioassessment/ biocriteria infrastructure provides us with an opportunity and impetus to standardize the reference site selection process statewide. The SWAMP program has long recognized this need, recently devoting a significant portion of its funding to developing reference condition datasets. Three recent peer reviews of SWAMP affirmed the importance of this effort:

1. In 2002, the SWAMP program funded an external review of bioassessment programs throughout California. That review was conducted by the lead author of the USEPA's bioassessment guidance document for streams and rivers.[3]

2. In 2005-06, the entire SWAMP program was peer-reviewed by an external "Scientific Planning and Review Committee" (SPARC), comprised of water quality experts from around the country.[4] The SPARC strongly recommended that SWAMP continue to develop its bioassessment program as a very high priority, specifically commenting that: a) the state board should consider revamping its entire standards program to make better use of biological endpoints (i.e., bioassessments) and b) the bioassessment program should focus particular attention on fostering consistency in its scoring indices.

3. In 2008, the USEPA (2009) conducted a Critical Elements Review of SWAMP's progress toward developing the technical elements to support biocriteria. The review stressed the fundamental importance of defining reference conditions and supported CA's reference condition strategy.

Establishing consistency in SWAMP's reference site selection process is clearly a key to effective implementation of biocriteria. However, identifying reference sites for California's perennial streams is complicated by its size (i.e., there are more than 300,000

---

[3] The external review, conducted by Dr. Michael T. Barbour and Colin Hill of Tetra Tech, Inc., produced a final report in January 2003 titled *The Status and Future of Biological Assessment for California Streams*, which may be viewed on the Internet at http://www.swrcb.ca.gov/swamp/reports.html

[4] The SPARC's final report is posted at: http://www.waterboards.ca.gov/swamp/docs/reports/sparc486_swampreview.pdf

stream kilometers), diverse ecological settings (12 Level III Omernik ecoregions are present in California, Figure 3), and anthropogenic settings (vast regions of the state are entirely converted to either agricultural or urban land uses). There are many natural gradients within each ecoregion. For example, the elevation in the Southern California Coastal Ecoregion extends from sea level to 8,000 feet encompassing cold water, high gradient mountain streams, but also includes warm water, low gradient streams in the flood plain. To complicate matters further, there are extreme natural temporal cycles of dry and wet years, which may not occur in all regions of the state during the same year. This is compounded by the episodic natural disturbance of flooding and fires. Finally, human-dominated landscapes can be so pervasive in locations such as urban southern California and the agriculturally dominated Central Valley that no undisturbed reference sites may currently exist in these regions. A statewide framework for consistent selection of reference sites must account for this complexity.

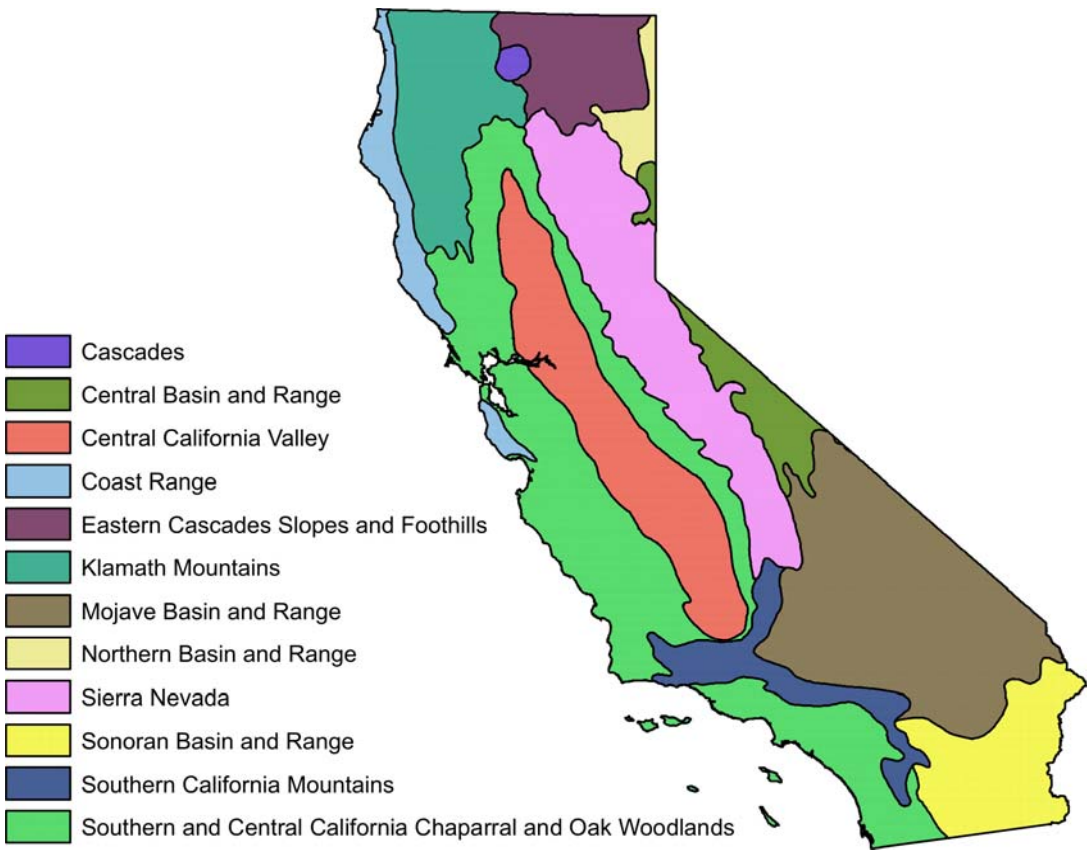

Figure 3. Boundaries of 12 Level III Omernik ecoregions present in California.

GOALS AND OBJECTIVES

This document summarizes recommendations to SWAMP for the development and maintenance of a Reference Condition Management Program (RCMP) that will support its regulatory biological assessment programs. The goal of the SWAMP RCMP is to provide an objective system for defining the expected biological and physical condition for wadeable streams and rivers in California. This system will identify pools (populations) of verified reference sites and outline procedures for sampling them to determine the range of biological expectations in these pools.

The monitoring objective

Data collected from reference sites will be used to answer a primary question: "what is the expected natural composition of lotic freshwater organisms in each of the major biogeographical regions of California"? The answer needs to be determined with sufficient rigor to serve as the basis for setting defensible numeric biocriteria. Our primary focus is on establishing expectations for benthic macroinvertebrate assemblages in perennial wadeable streams, but we expect that the approach will allow similar assessments of algal and fish assemblages as well as instream habitat condition and riparian condition.

Accounting for natural variability

An extension of the central monitoring question is: "what is the range of biotic measures (e.g., taxonomic composition, individual metrics and biological indices) in high quality sites and which natural environmental gradients (both spatial and temporal) are most strongly related to this variation." Ultimately, the goal is to identify the major sources of natural variability for all biological response measures (Figure 4). To account for these gradients, reference sites should be distributed to represent the full gradient range.

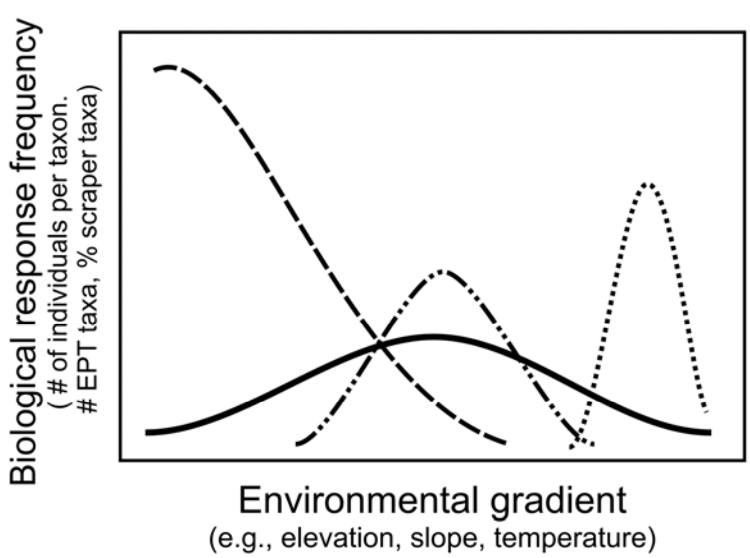

Figure 4. Hypothetical frequency distribution relationships between biological responses and environmental gradients.

GUIDING PHILOSOPHIES

In order to guide the development of the RCMP, the panel agreed upon a set of basic philosophies. These philosophical principals were used to guide their decision-making:

- **Use natural condition as the desired state whenever possible -** The panel's goal was to identify sites in natural or near-natural conditions whenever possible. However, the panel recognized that there are regions in the state where an insufficient number of sites in near-natural condition were likely to be found. The panel agreed that setting biological expectations were no less important in these regions. Therefore, the panel endeavored to identify the best attainable condition in these suboptimal regions of the state.

- **Balancing statewide consistency with regional flexibility** - The panel agreed that the reference strategy should balance a set of desirable, but sometimes naturally conflicting, traits: objectivity, consistency and flexibility. For example, a reference program that works for all of California can't be both perfectly consistent and perfectly flexible. This strategy aims to balance the competing demands of statewide consistency with the flexibility needed to adapt to unique regional conditions.

- **Reference site management is an iterative process -** The management of a reference site network is an ongoing and iterative process. The monitoring program should be responsive to new information and perspectives gained from selecting and monitoring reference sites. The general strategy should build in analysis of data to optimize selection strategies (process of selecting sites) and management design (e.g., how many sites, regional boundaries, which natural gradients to account for).

- **The RCMP should be transparent -** The technical process of determining reference conditions should be transparent to external review. As the state moves toward implementation of biocriteria, transparency and comprehension of the RCMP process will improve stakeholder confidence and provide structure for discussions about setting objective standards.

- **These recommendations are a starting point -** The panel understood that their recommendations provide a starting point for evaluating reference condition rather than an exhaustive set of operating procedures for selecting reference sites. This document is written assuming that SWAMP will develop a technical workplan that details a more refined program as the RCMP is implemented.

GENERAL GUIDANCE

The general approach for establishing the SWAMP reference site network has four components:

1. California should be divided into different geographic regions based on coarse biogeographic similarities in order to partition some of the natural variability among regions
2. A pool of reference sites should be assembled within each region through a sequential process of identification and screening of candidate sites
3. The reference pools should be managed through iterative review of data to refine regional boundaries, ensure continued suitability of sites and ensure adequate representation of natural gradients
4. A monitoring design should be created for sampling this pool of reference sites to document the range of biological and physical condition at reference sites, and monitor for changes to the condition of reference sites over time

All but the second component, site selection, apply equally to all regions of the state. The site selection process has two versions depending on the availability of high quality reference sites. We refer to the two versions in this document as: 1) the "standard model", which applies to regions with a sufficient number of reaches with relatively low levels of anthropogenic stress; and 2) the "alternate model", which applies to regions that do not have a sufficient number of high quality reaches. The vast majority of California should be able to apply the standard model.

Component I: Partitioning CA into biogeographic regions

Two general schemes are available for delineating California's ecoregions (Omernik 1995 and Bailey *et al.* 1994). We follow Omernik's divisions here because the boundary delineation decisions were generally based on a broader range of geology, climate and zoogeography than Bailey's. Omernik Level III ecoregions have been delineated for all of North America (Omernik 1995), with 12 Level III ecoregions falling in California (Figure 3).

Partitioning the state into different regions based on habitat similarities has some precedence in California bioassessment. The SWAMP Perennial Streams Assessment (PSA) has relied on a combination of Omernik ecoregions and regional board boundaries to partition the state for assessment purposes (Figure 5). Because these definitions include significant ecological gradients that contribute to natural variability in biological assemblages, and because they comprised existing assessment units, the panel agreed that these delineations were appropriate to use as initial boundaries for the reference network. However, the panel also stressed that ecoregions do not always adequately capture natural gradients that are key drivers of aquatic assemblages (insert references here, Hawkins and Norris 2000). Thus, data analyses must address the suitability of these boundaries as the program collects more data.

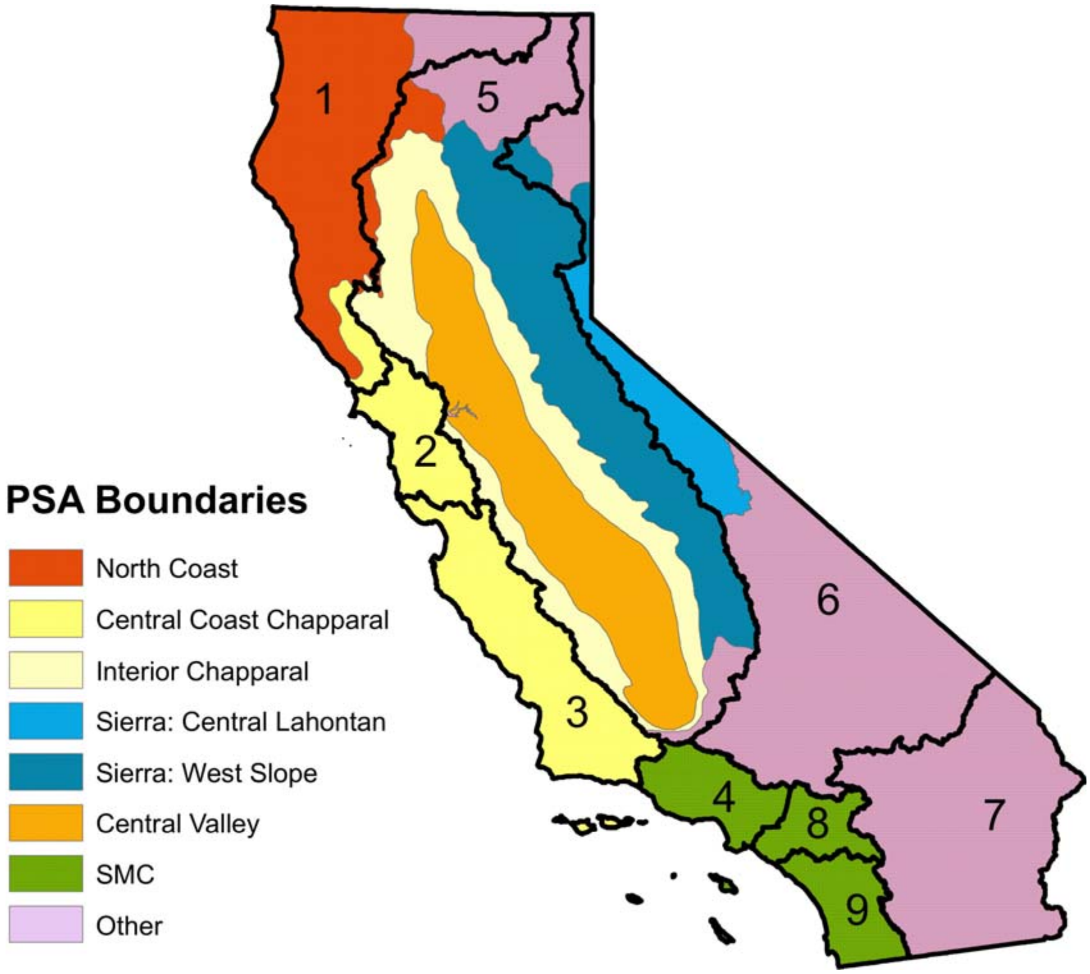

Figure 5. Boundaries used for defining the regional subunits of the SWAMP Perennial Stream Assessment (PSA) survey. SWAMP regional board boundaries one through nine are indicated by thick lines. SMC=Southern Coastal California Stormwater Monitoring Council.

Component II (a): Selecting sites: the "standard model"

The second step in the general approach is the most resource intensive and technically challenging: to develop a large pool of reference sites within each ecoregion. The ability to precisely establish biological expectations within each region is a function of the number of sites that are sampled and natural variability within each region. Therefore, the pool of reference sites should be large enough to provide a robust characterization of natural variability. Furthermore, reliance on a small number of reference sites is risky because it increases the consequences of catastrophic failure of individual sites. The size of the site pool in each region will depend on the number of major environmental gradients in each region (e.g., elevation, temperature, etc.) and the strength of influence of these gradients on biotic assemblages.

The panel recommended a sequential approach for assembling a set of candidate reference sites and screening suitable sites for the final reference pools within each region. The process includes: 1) screening data from previous site visits to identify candidate sites, 2) application of remote sensing and point-source GIS data screens of all potential stream reaches (combining landscape and local scale) to identify candidate sites, 3) use of best professional judgment/ local knowledge to add sites to the candidate pool.

Once a set of candidate sites is assembled, each candidate site should receive an on-site visit to evaluate its suitability. The exact type of data collected for evaluations during this stage will vary by region, but at a minimum should include: observations of local landuse activities, instream and channel habitat condition, riparian condition, evidence of recent natural disturbance. Some regions may require additional chemical data (water column or sediment) or toxicological data to confirm site suitability.

Sites that pass both the remote sensing and field reconnaissance screens become part of the reference pool for that region.

Component II (b): Selecting sites: the "alternate model"

The panel recognized that the standard model is not likely to work in all regions of California. The conversion of natural landscapes to agricultural and urban land uses is so extensive in some parts of the state that the entire region is devoid of waterbodies that could be used to define reference condition. Most regions of California should be able to use the standard model; the alternate model should only be used when the standard model is not feasible.

The panel defined the following criteria as triggers for acceptable use of alternate site selection strategies (both criteria must apply):

1) Insufficient high quality sites are available within one of the main regions (or a large section of one of the main regions) to adequately characterize ecological potential. Suitable stream reaches are unavailable for one or more of the following reasons:
   a) Anthropogenic landuse is a dominant factor in all watersheds within the region (or subregion)
   b) Normal flow is modified (e.g., flow diversions, dams, withdrawal or augmentation)
   c) Natural channels are altered (e.g., all or most channels converted to conveyances, irrigation supply/drains)
   d) Riparian corridors are impacted throughout the region (e.g., concretized riparian or surrounding landscape modified)

2) No comparable region exists from which to draw inference about biological expectations. That is, the areas are unique in their biological expectation so regions with few reference sites are not able to incorporate sites from another region.

This situation is not unique to California streams and many large programs have recognized the need to deal with regions with insufficient reference sites (REFCOND 2003, Stoddard *et al*. 2005, Paulsen *et al*. 2006). National guidance for developing state biocriteria programs highlighted the need for special treatment of these conditions (Barbour *et al*. 1996a,b). While the unique needs of these regions are widely recognized, the approaches for establishing ecological potential for reference-poor regions are far from standardized.

The RCMP panel outlined a general strategy for approaches to explore in reference-poor regions. The RCMP panel did not take any strong position on the relative strengths of these alternatives nor how different approaches should be combined to define expected conditions in reference-poor regions. Some of the alternative strategies included:

1. Use a modified version of the standard approach (e.g., use lower thresholds, emphasize local condition measures)
2. Alternate approaches
    a. Use existing tools to screen sites
    b. Species pool approach
    c. Factor-ceiling approach
    d. Model taxon preferences for key environmental gradients

These alternative strategies are not mutually exclusive and, when appropriate, should be used as multiple lines of evidence to reinforce an objective definition of biological expectation in regions without reference sites. In the "specific guidance" sections of this document (see Alternate Strategies for Selecting Sites) we describe these approaches and discuss strategies for applying them to California's challenging landscapes.

Component III: Managing the regional site pools

After the site pools have been assembled for each region, the RCMP requires an ongoing evaluation of data from these sites to address several key management questions. There are two major components to managing the reference pools: 1) evaluation of the regional representation of natural gradients and 2) periodic review of sites to evaluate changes to their suitability.

The ability to effectively understand natural sources of biological variation is fundamental to establishing sound biocriteria[5]. Therefore, the RCMP must directly assess the reference pools to ensure representation of regionally important natural gradients. This review should include a periodic review of the suitability of the initial regional boundaries proposed here.

The second aspect to site management is periodic review of sites in the reference pools to assess their continued suitability as reference sites. Conditions within stream reaches and in their upstream drainages can change over time (e.g., timber harvest, conversion of natural landscapes to agricultural or urban/suburban/exurban uses). Furthermore, we may discover sources of stress that were unknown when sites were initially added to the reference pools (e.g., discovery of nonpoint source discharges, mines, flow withdrawals/diversions, small-scale placer mining, etc.). Sites that fall into this category may be monitored to measure the impacts of these stressors, but they should be removed

[5] See discussion on p. 5.

from the reference site pools. In contrast, natural disturbances (e.g., forest fires, catastrophic flooding or landslides) can also alter the biological condition at sites and they should be excluded for sampling temporarily, but should remain in the reference site pool[6].

Component IV: The monitoring strategy

The panel recommended an integrated probabilistic and targeted sampling design for the RCMP. The probabilistic approach will sample a rotating subset of randomly-selected (rotating panel design) sites from within the reference pool each year to estimate average biological condition. A subset of the randomly-selected sites should be sampled annually to measure year-to-year variability at sites and improve SWAMP's ability to detect drift in reference condition within each region over time. This design provides an unbiased assessment of natural variability with enhanced trend detection.

Targeted sampling is comprised of fixed sites near locations of special interest, but this should be supplemental to the probabilistic sampling effort. Fixed sites provide additional power to detect trends, but suffer from its inability to extrapolate to other locations. However, many agencies already monitor reference sites and, provided they meet the RCMP selection criteria, these sites have the added benefit of years of historical data. As SWAMP extends its reference monitoring program through collaboration with other state and federal programs, it should retain the ability to incorporate these sites.

The panel emphasized sampling more probabilistically selected sites over targeted sites, but did not make any recommendations about relative proportion of each type. This decision should reflect the relative importance to the SWAMP program of estimating current biological expectation versus detecting changes in the reference state. Changes in the reference state may become increasingly important due to factors such as climate change.

---

[6] A special study of natural disturbance recovery could be especially enlightening with regard to understanding natural variation.

SPECIFIC GUIDANCE

1.0 Site Selection: Assembling the reference candidate pool

The panel recommended a sequential approach for assembling a pool of potential reference sites using a series of tools to identify candidate sites (Figure 6). The toolbox components included: 1) use of existing data from previous site visits, 2) GIS data screens of all potential stream reaches using databases of stressor data (combining landscape and local scale), 3) expert selection of site locations based on regional experience.

1.1 Use of existing sites

Previously sampled sites are an excellent source of candidate reference sites and where available in sufficient numbers, can constitute a ready-made pool of reference sites. However, previously sampled sites vary widely in the amount of information associated with them, and they fall into two categories: 1) sites with a large amount of associated environmental data that is sufficient to evaluate without additional data collection, 2) sites that require additional data collection to produce adequate evaluations. Several programs in the state have collected sufficient data to meet the first condition (e.g., EMAP, Central Valley WEMAP, CMAP, SNARL, some regional board programs), but most sampled sites fall into the second class.

The current distribution of existing candidate sites in California is illustrated in Figure 7. Sites were pre-screened from ABL and SNARL databases and sorted into one of three tiers based on the availability of different types of screening data. Under the RCMP, Tier 1 sites would pass to the pool of verified reference sites if they passed a BPJ screen (see following section), sites in other tiers would be placed in the candidate pool and be subjected to the full site screening process (Figure 6).

1.2. GIS data screens of all potential stream reaches using databases of stressor data [7]

If regions do not have sufficient existing sites to fill the final pool of fully screened reference sites (steps 1 - 3 of the general guidance), then new candidate sites should be identified through use of geographic information systems (GIS) techniques for screening remote sensing data and GIS databases of point source stressors. GIS-based searches for candidate reaches are expected to contribute the majority of sites in many regions.

Ode (2002) described a GIS based method for identifying candidate stream reaches using a series of remote sensing data filters. Under this approach, candidate watersheds are identified for a region with GIS techniques and then stream reaches within these watersheds are targeted for reconnaissance to verify reference quality characteristics. The RCMP generally follows this approach, which consists of the following steps:

---

[7] GIS techniques are used at two different stages of the RCMP process: 1) searching for potential new reference streams (described in this section) and 2) quantifying impacts to existing sites (described in the following section). The techniques are very similar, but differ somewhat in their application. The search phase is a relatively coarse screen of candidate watersheds while the verification phase is site specific and allows for multiple spatial scales of GIS analysis (see Figure 8).

1.2.1 Assemble GIS layers of important landuse disturbances

The list of potential impacts to stream condition is very long and includes multiple point and non-point sources of disturbance. Quantitative measures of many human or human-influenced activities are available in digital spatial (GIS) formats from various state and federal agencies (see Tables 1 and 2), but there is a very large amount of variation in the degree to which datasets are accurate, current, and consistent across wide geographical ranges.

1.2.2 Determine appropriate reporting units (areas of analysis) and create necessary GIS layers~ Current GIS applications for locating least disturbed waterbodies in a region (see ATtILA text box) calculate summary stressor metrics (e.g., % urban landuse, road density) for each reporting unit (typically watersheds) in the region of interest. Candidate stream sites are then selected from within these watershed areas. It is recommended that the RCMP use a modified version of watershed polygons developed by the national NHD+ program.[8]

1.2.3 Use ATtILA extension to calculate stressor metrics using remote sensing and point source datasets (see ATtILA text box)~ ATtILA produces summary output in a spreadsheet containing multiple stressor metrics for each candidate watershed (i.e., % agricultural landuse, % impervious surface, # of mines, # road crossings/stream km).

1.2.4 Analyze distribution of stressor metrics and select appropriate thresholds

Screening thresholds for GIS stressor metrics can be set using a variety of approaches: 1) visual inspection of frequency histograms for natural breaks in distributions, 2) statistical criteria[9] (e.g., eliminate watersheds with road densities greater than 1.5 standard deviations above the mean for all watersheds in the region, or eliminate all but the lowest 25th percentile of all road densities), 3) established (i.e., literature based) impact thresholds. At this stage in the screening process, the RCMP panel recommended the use of fairly liberal screening thresholds since GIS data are often inexact and impacted sites can be screened during later stages of the site verification process.

1.2.5 Eliminate watersheds that fail GIS screens

Because of the large number of stressor variables that are quantified in this step, there will be a large number of metrics to evaluate. The panel discussed two options for how to combine the information from these different screens:

---

[8] With funding from the SWAMP program, CSU Chico's Geographic Information Center (GIC) has developed a method for creating nested watersheds from the native polygons available from the NHD+ program. The NHD+ polygons are limited in their utility as reporting units because they are non-overlapping. Thus, 2nd order watershed boundaries in NHD+ do not include their tributary 1st order basins. The GIC's modification creates new watershed polygons that are aggregates of all upstream polygons (e.g., 4th order watersheds contain all upstream 3rd, 2nd and 1st order polygons).

[9] Effective of statistical properties of distributions to define thresholds depends on a normal distribution of scores. Some distributions (e.g. highly skewed or bimodal) may be better interpreted by looking for natural breaks or using literature based criteria.

> **ATtILA extension for GIS Landscape Analysis**
> http://www.epa.gov/esd/land-sci/attila/intro.htm
>
> To quantify landuse activities occurring upstream of sites, the Ebert and Wade (2004) developed a user friendly interface that accepts a range of GIS data layers and produces summary statistics for areas defined by the user. The extension, Analytical Tools Interface for Landscape Analysis (ATtILA), is a plugin to ESRI's ArcView® (version 3.x) GIS software (ESRI Products) and takes advantage of ESRI's Spatial Analyst extension to run the spatial calculations.
>
> - The ATtILA extension calculates the percentages of various landuse activities occurring in specified areas (urban; forested; agricultural-row crops; agricultural- orchards/vineyards; agricultural-total), other correlated measures of human activity (population density; road length; road density; road crossings/stream mile; percent impervious surface), and estimated nitrogen and phosphorus loadings.
> - ATtILA can use polygons of any spatial extent as reporting units (e.g., entire upstream basin, local buffers)
> - In 2007, the SWAMP program provided funds for a project to adapt the ATtILA extension to meet the GIS needs the RCMP process. Specific enhancements being developed include the ability to add custom stressor coverages, summarize point source data, and facilitate rapid adjustment of stressor thresholds for screening candidate sites. The project will be coordinated with the implementation of the RCMP
> - It is expected that the capabilities of the modified ATtILA extension will expand as the RCMP process develops over time.

a) Screens could be applied as a series of filters, with failure in any metric resulting in elimination of the watershed from the candidate pool.

b) Alternately, a multi-metric index of stressors could be used to create a composite score for each candidate site and low scoring watersheds would be removed from the candidate pool.

The panel recommended the use of a hybrid approach, in which the multi-metric scoring would be used to screen watersheds, but "kill-switches" would be employed to eliminate watersheds that exceeded high impact thresholds for particular stressors (e.g., eliminate watersheds with > 10% urban landuse).

As an additional consideration, the panel recommended that the RCMP explore quantitative methods for deciding which impacts to use for selection. For example, some stressors may have a greater effect than others and, thus, should be weighted more heavily than relatively benign influences. A corollary would apply to data sets with different levels of confidence. For example, information about mine locations may be available, but not about which are actively contributing contaminants to streams.

1.2.6 Identify candidate stream reaches within candidate watersheds[10]

After eliminating watersheds using GIS screens, the remaining watersheds represent potential candidates for the reference pool. These areas may be able to be further refined to further isolate candidate stream reaches (see Figure 8).

[10] An alternative strategy is to select candidate stream segments directly using analytical tools designed to work with the NHD+ datasets. Under this approach, confluence points would be the the reporting unit and NHD+ tools would summarize all upstream landuses. Errors in the current version of NHD+ (primarily problems with flowline connectivity) currently limit the effectiveness of this approach, but it may become more useful as NHD+ improves. The RCMP should remain open to both approaches and revisit this issue as new versions of NHD are released.

1.3 Use of local knowledge to add sites to the candidate pool

Although existing data and GIS searches will contribute the majority of sites to the candidate pool, a few sites may be added to the candidate pool on the basis of local knowledge. Local knowledge can sometimes help in identifying candidate sites because GIS datasets are imperfect and GIS screens may pass over good sites because of inaccurate or outdated disturbance information. These sites, however, should be critically evaluated because subpar sites based on local knowledge will dilute the quality of the reference pool. More rigorous evaluation of these sites should include examination of existing data.

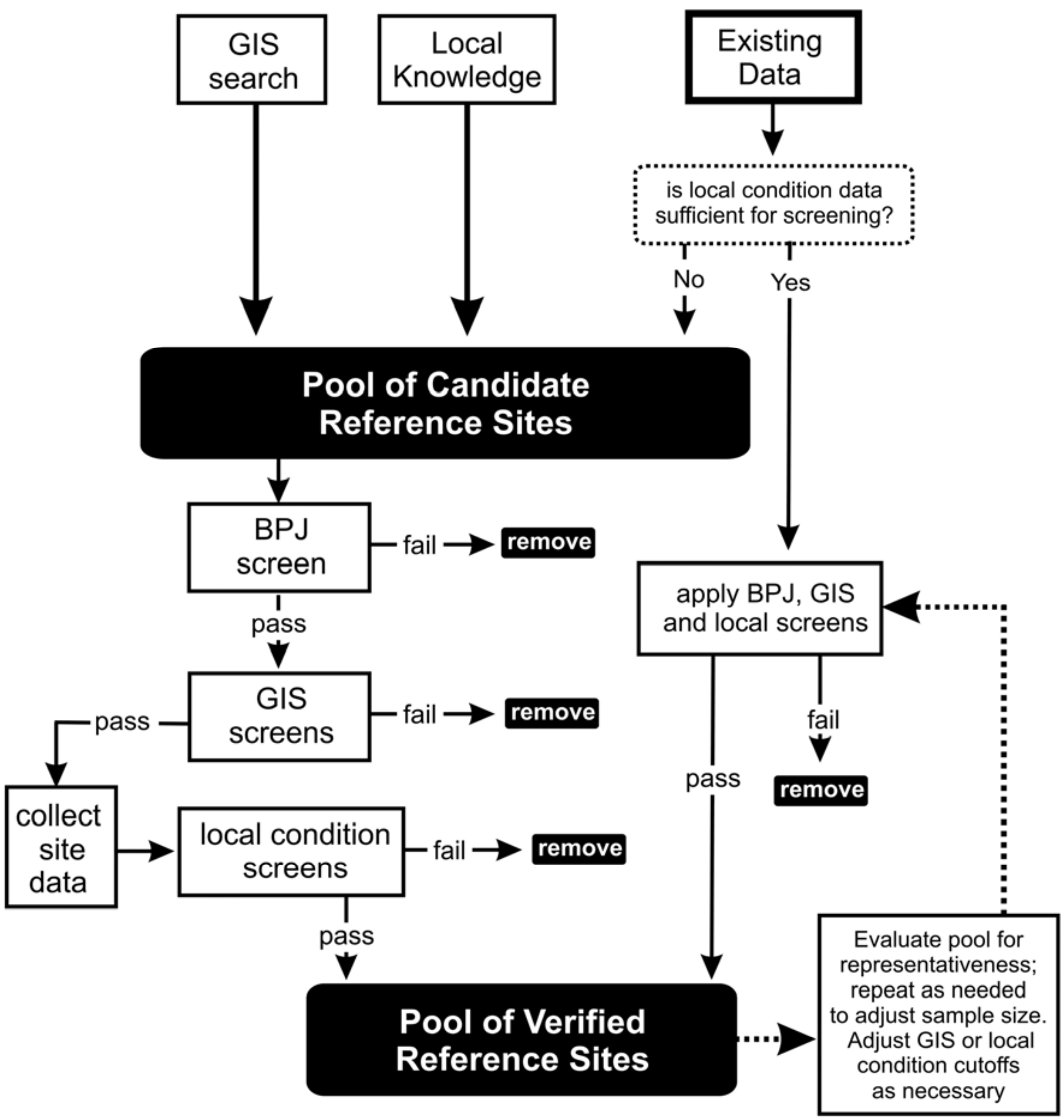

Figure 6. Schematic of the standard reference site selection and verification process.

Table 1. Potential GIS data coverages for nonpoint sources.

NON POINT-SOURCE COVERAGES			
Information Type	**Data Source(s)**	**Notes**	**Coverage**
Landuse/Landcover	National Landcover Dataset (NLCD), MRLC	1992, 2001 satellite imagery, allows for 9-yr landcover change assessments	Statewide
Impervious Surface	NLCD, Others	Quality varies regionally	NLCD statewide, others patchy
Road Density	USFS, TIGER		Statewide, but patchy
Timber Harvest	CDF, THPs		
Vegetative Change/ Vegetative Change Cause (LCMMP)	USFS/CDF		Not Statewide
Population Density	Census Blocks, CDF	Produced in conjunction with decadal population censuses; censuses can be combined to estimate population change	Statewide
Grazing	Cattlemen's Association		Not Statewide
Fire History	CDF, USFS		Best for FS lands

Table 2. Potential GIS data coverages for point sources.

POINT-SOURCE COVERAGES			
Information Type	**Data Source(s)**	**Notes**	**Coverage**
Mining	USGS	Possibly outdated	Statewide
NPDES	EPA	Prone to inaccuracies	Statewide
303(d) listed streams	SWRCB	Every three years	Statewide
Water Diversions/ Extractions	USGS, NHD+	Possibly outdated	Statewide
Dams	CalWater	Doesn't include overflow info	Statewide
Stormwater Inputs	NHD+, Counties	Uneven coverages	Patchy
POTW	EPA	Prone to inaccuracies	Statewide
Landslide Datasets	CalTrans		Statewide

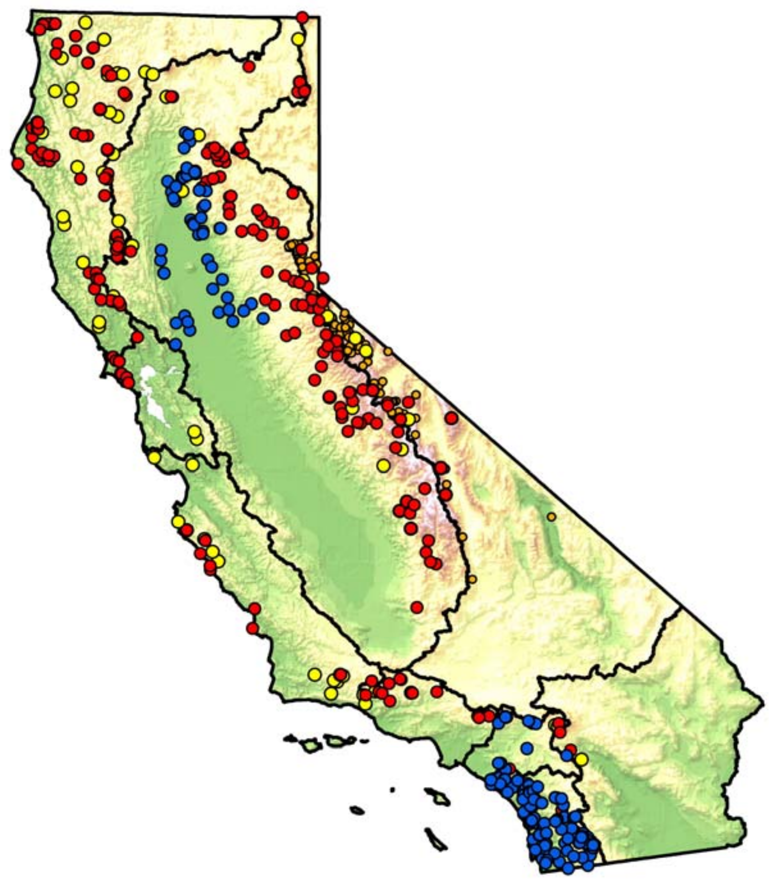

Figure 7. Partial set of bioassessment sites available for initial screens assigned to one of three tiers. Tier 1 sites (yellow circles) are EMAP and CMAP sites that passed a full suite of screens based on the most complete data for evaluation. Chemical and habitat thresholds were based on Stoddard *et al*. (2005) and landuse thresholds were based on Ode *et al*. (2005) and Rehn *et al*. (2005). Tier 2 sites (red circles) are USFS and Regional Water Board sites that have passed a less stringent screening process, but might very well be reference and need additional data before they either passed into Tier 1 or eliminated from the candidate pool. Tier 2 sites were screened based on land use, less extensive physical habitat data and limited or no chemical data. Tier 3 sites (blue circles) are cases in the Sacramento Valley, Sierra Nevada foothills and southern coastal California that probably need an alternative reference screening process (e.g., the factor ceiling approach). SNARL sites (orange circles in Eastern Sierra Nevada) used different screening thresholds, but are likely equivalent to Tier 1 sites.

2.0 Site Selection: Screening the candidate pool

Once a large set of sites is selected for the candidate pool, sites in the pool undergo a series of screening steps to either validate sites as appropriate reference sites or eliminate them from the pool. The major screening tools are: 1) expert opinion (BPJ), 2) landscape screens (GIS), and 3) local condition screens.

2.1 BPJ screens

While BPJ can play a role in identification supplementing the pool of candidate sites, it plays a bigger role in eliminating candidate sites. Sites should be eliminated on the basis of BPJ knowledge that there are known problems that aren't accounted for in GIS datasets. For example, GIS datasets may miss recent development, known pollutant spills, or nonpoint sources. This step should include coordination with local watershed groups, landowner groups and other stakeholders to eliminate inappropriate sites. The rationale for rejection should be documented.

2.2 Landscape scale screens (GIS)

Just as GIS techniques are essential for adding sites to the candidate pool (Figure 6), they also play a crucial role in reference site verification. The datasets and techniques used in this step are essentially the same as those used in searching for candidate watersheds/ stream segments, but the application of the tools differs somewhat. Whereas the GIS analyses were applied at a fairly coarse spatial scale in Section 1.2, GIS tools can be applied at multiple spatial scales during the screening stage.

The first step in the second GIS stage is to convert candidate watershed areas into specific sampling sites by selecting a common point on the stream segments in each watershed (e.g., the downstream confluence point), making them equivalent to other sites in the candidate pool (as in Figure 8a).

The chief benefit to the two-stage application of GIS techniques is that it gives us the opportunity to identify multiple sampling locations within reference watersheds. While sites would normally be screened using stream confluence points as the candidate site locations, site locations could be moved to other points in the watersheds to identify additional reference sites within good watersheds or to avoid portions of the watershed with undesirable sources of human disturbance (Figure 8b).[11]

Using watershed delineation tools and local site buffering tools currently available for use with GIS software, polygons should be created to represent different spatial scales upstream of each site (e.g., the entire watershed draining to the site, the upstream area within a 5 km radius of the site, the area within a 200m buffer on either side of the stream within 1km upstream). Once created, these areas can be used as reporting units for

[11] Although the two stage application of GIS techniques gives us greater flexibility to identify multiple candidate stream reaches within each candidate watershed, an alternative strategy would be to eliminate the coarse search for watershed described in Section 1.2 and go straight to the more refined screening analysis indicated in Figure 8a.

ATtILA analyses. Metrics calculated for the different spatial scales can be screened as in Section 1.2.5.

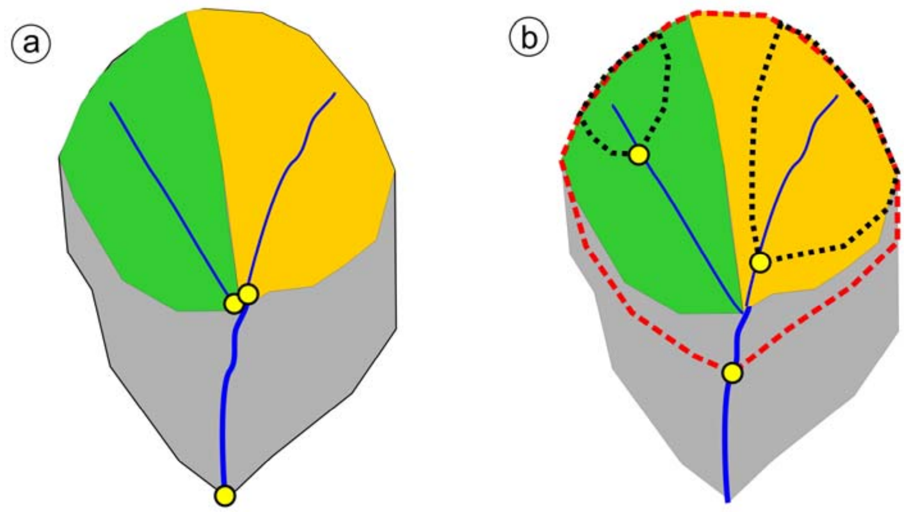

Figure 8. Illustration of alternative applications of the second stages of GIS analysis in the RCMP using a hypothetical second order watershed containing two first order watersheds: a) normal site locations represented by yellow circles, b) alternate site locations and their watershed boundaries (represented by dotted lines).

2.3 Local Condition Screens

Sites that have passed BPJ and GIS screens are then subjected to an evaluation of site scale stressors. Some of the local scale information can be obtained from aerial photography of sites, but the majority of this information will come from site visits and in some cases collection of water quality data.

2.3.1 Site scale data: Aerial photography

Aerial photography provides a unique view of potential site scale stressors. Digital orthophoto quadrangles (DOQs) are available for the entire state of California (DFG). Google Earth is another source of digital satellite imagery. DOQs and other sources of aerial photographic images can provide excellent information about local stressors not available through other sources, but are subject to the same timeframe limitations as other digital sources.

2.3.2 Site scale data: Site visits

The panel strongly recommended site visits as a crucial component of reference site verification. Once candidate list have been narrowed down to sites that meet BPJ, GIS and DOQ screens, land ownership should be determined for each site and owners contacted to obtain access permission and or sampling permits as needed. Site owners can also be contacted at this point to determine if there are any reasons for rejecting sites.

Field visits should be used to collect both qualitative (e.g., presence of obvious disturbances) and quantitative data (e.g., % intact riparian zone). Quantitative measures should focus on data that can be collected and analyzed cost-effectively.

2.3.3 Qualitative data

Visual assessments of site suitability should include a minimum set of observations:

- Upstream impoundments, or evidence of water withdrawal or diversion
- Evidence that the site is non-perennial
- Evidence of recent fire, flooding or landslides
- Local grazing impacts
- Presence of significant anthropogenic use (e.g., campgrounds, etc.)

2.3.4 Quantitative data

At a minimum, site visits should include characterization of physical habitat using the SWAMP Physical Habitat Procedures (Ode 2007) and conventional water chemistry. Physical habitat characteristics should include measures of both instream and riparian condition. SWAMP habitat procedures may be supplemented with riparian condition measures collected with the California Rapid Assessment Method (CRAM) for riverine wetlands. Water chemistry analyses should include the following analytes: chloride, turbidity, pH, total nitrogen, total phosphorus, conductivity, and alkalinity. Some chemical analytes may not be needed in all regions. For example, sulfate (a good indicator of mining activity) is not likely to be informative in xeric regions. One recommendation was to create a checklist of activities by region. Another option is to supplement with sediment and/or water column toxicity. While these tests may be expensive, they are less expensive than a screen for a long list of toxic constituents.

2.3.5 Combining site data for screening decisions

As with GIS screens (Section 1.2.5), there are many ways to combine site data to make determinations. The panel again recommended use of a hybrid approach in which site scale data is combined to calculate a multi-metric site condition score. The use of kill switches was also recommended for excessively high or low scores for individual habitat or chemistry.

3.0 Alternate strategies for selecting reference sites

While most regions of California can follow the standard approach for selecting reference sites, there are at least two large regions in California that lack sufficient high quality sites. The first is the Central Valley where natural landscapes have been almost entirely converted to agricultural and urban land uses. Most natural stream reaches in this region have been channelized or otherwise modified to support irrigation and flood control. The second is in coastal southern California (elevations below 1200 ft – upper elevations can follow standard model) where conversion to urban and suburban land uses has led to the channelization of most stream reaches. Recent studies in these regions demonstrate that at least some waterbodies in highly modified regions can support fairly rich BMI assemblages, even under considerable alteration and agricultural development (Griffith *et* *al*. 2003, deVlaming *et al*. 2004, deVlaming *et al*. 2005, Ode *et al*. 2005). Thus, there is enough range in biotic condition to differentiate degrees of impairment in these regions.

The panel recognized the unique limitations of these regions and recommended that a separate set of approaches be developed for them. Despite the differences in methodology, the goal of the alternate strategy is the same as the standard approach: to characterize the best attainable biological condition in these regions. This section outlines a set of approaches that the RCMP could follow. These fall into two general categories:

- Use a modified version of the standard approach
- Explore non-standard approaches

3.1 Modified use of standard approach

The first option is to use the set of techniques described for the standard approach, but to modify the way the techniques are applied. Modifications fall into two general types: 1) much greater emphasis on reach scale screening data, 2) use of less stringent criteria for rejecting sites.

One of the panel's philosophies is that potential reference sites in highly modified regions need a much larger amount of supporting data to verify their status than in less modified regions. In both the Central Valley and southern coastal California lowlands, streams exist in a landscape matrix with a universally high level of unnatural land uses. Furthermore, many streams have extensive flow manipulation, including water diversion, re-introduction, and inter-basin transfers that render watershed based tools irrelevant. For both these reasons, watershed based stressor analyses are less informative screening tools. Accordingly, much greater reliance should be placed on data collected from direct site visits than on remote sensing data. The panel recommended increased emphasis on riparian condition, instream habitat condition, and water column chemistry. In some cases, additional data (e.g., sediment and or water column toxicity) will be necessary to verify sites.

Selective relaxation of screening thresholds may also be an effective means of identifying the best available sites in a region. For example, acceptable road densities are likely to be much higher in southern coastal California than in other regions of the state. Likewise, acceptable local agricultural landuse percentages and acceptable levels of fine sediments are likely to be higher in the Central Valley than in less modified regions. While less stringent thresholds may help identify some of the best sites in highly modified regions, the use of kill switches is an essential safeguard against accepting unacceptably low thresholds. Specific cutoffs such as >10% local impervious surface, or toxin concentrations greater than the standards set by the California Toxic Rule may be more appropriate in these heavily modified landscapes.

A version of this modified standard approach was applied to search for reference sites in the Central Valley (Ode *et al*. 2005). Remote sensing data (e.g., landuse percentages) and other GIS datasets (e.g., pesticide application rates) was used as a coarse screening tool, but this data was de-emphasized in favor of riparian condition and instream habitat scores. This study identified approximately 20 potential reference creeks in the Sacramento Valley (see Figure 7), but these still need to be screened for water chemistry and toxicity before they are acceptable.

3.2 Non-standard approaches

Although modified use of the standard techniques can go a long way toward providing the data needed to adequately characterize biological expectations in these areas, it is unlikely to resolve the entire problem of identifying a sufficient number of candidate reference sites. The panel recommended the exploration of several different alternative, non-standard techniques:

- Select best sites using existing biological indices
- Species pool approach
- Factor-ceiling approach
- Model taxon preferences for limiting environmental gradients

All of the non-standard strategies suffer to a greater or lesser degree from circularity since the establishment of a biological reference site is being established with biological data. However, the extreme lack of reference sites in these regions requires us to consider accepting some circularity while adding additional steps to guard against the risks of circularity. The best way to guard against these risks is to use independent datasets to select the biotic response metrics.[12]

3.2.1 Use of existing indices to select sites with high quality biology

A straightforward alternate approach is to use existing biological assessment tools from the same region to identify sites that could be used to establish biological expectation in problem regions.[13] High scoring sites would be assumed to represent the "least disturbed" sites in the region. The method assumes that BMI assemblages in the target region have similar responses to anthropogenic stress as the region(s) for which the indices were created. Issues with circularity are mitigated by the fact that the scoring tools were derived objectively using independent datasets.

A variation on this approach is possible in regions where only a few reference sites can be identified (either using the standard methods or the modified standard described above). Under this variation, a model (either MMI or O/E) would be created using a small number of reference sites. Then new sites with similar BMI assemblages would be added to the reference pool and the model recalculated. This recursive approach results in more explanatory power because it is based on a larger number of reference sites, but it is inherently circular because the new sites are not chosen based on independent information.

---

[12] Note also that some have argued that the circularity concern is less of a problem in highly modified systems than more pristine systems because relationships between metrics and stressors are simpler (Karr and Chu 1999).

[13] Examples of existing biological assessment tools include the Southern California IBI (Ode *et al*. 2005), northern California IBI (Rehn *et al*. 2005) and the California RIVPACS models (Hawkins unpublished).

3.2.2 Species pool

Another option is the species pool approach, which uses the total faunal diversity of a region (i.e., central valley or southern California coastal urban lowlands) to establish a biological condition axis. The process involves assembling a pool of all BMI taxa ever collected from the region, then using taxonomic richness as the measure of biological integrity at test sites. The inventory could be compiled from existing data sets, historical records (i.e., museums or other voucher collections), or directed field surveys. This technique assumes that richness is a good measure of condition, that there hasn't been extensive extinction of native fauna and that the constituent species in the pool are all potential colonists of any test stream.

The utility of this approach could be enhanced in at least two ways. The number of richness metrics could be increased by breaking richness out by taxonomic groups (midges, worms, mayfies, etc.), isolating the different information content in these groups. Further, the species pool could be modeled to associate expected taxa with key environmental gradients (i.e., substrate composition, elevation, etc.) and the proportion of taxa present at reference sites could be a potential target for attainment of reference state. If this approach were taken, then the species pool concept should be tested first in a region where identifying reference sites are not problematic as proof of concept.

3.2.3 Factor-ceiling approach

Carter and Fend (2005) developed a technique for defining a range of biotic expectation that takes into account the decrease in biotic condition caused by physical modification along an axis of increasing urbanization. In their example, a simple statistical technique (partitioned least squares regression, OLS) was used to identify the highest biotic scores along an urbanization gradient. Upper values define the range of expected biotic conditions for the region. Since a full urbanization gradient was used to take into account decreasing biotic potential with increasing urbanization, the resulting range of expected conditions is a conservative estimate of biotic potential for the region. While this approach could be used in both the Central Valley and southern coastal California lowlands, the method would work especially well in the Central Valley because the agricultural impact gradient is not as strongly confounded by elevation or other longitudinal gradients as the urban ones studied by Carter and Fend (2005).

The first step is to identify key measures of physical modification (hydrologic modification, channel modification, streambed modification) and to combine these into a multifactor axis of agricultural modification (i.e., the primary axis in a PCA of these stressors). The second step would be to identify appropriate metrics for detecting biotic impairment in valley streams.

3.2.4 Modeling taxon preferences for limiting environmental gradients

The final alternate strategy involves modeling taxon preferences for key environmental gradients, or limiting environmental differences (LED) and then using these relationships to select the most appropriate sites for setting biological benchmarks. Different habitat features (e.g., climate, channel morphology, water chemistry, substrate characteristics) can be thought of as acting as "filters" that select for particular species traits (Poff 1997).

This conceptual framework provides a way of accounting for the influence of both natural and anthropogenic factors on species distributions. Chessman and others (Chessman 1995, 2006, Chessman and Royal 2004, Chessman *et al*. 2008) recently developed a technique for using the tolerance or preference of individual taxa for key environmental filters (e.g., water temperature range, substrate composition, flow regime) to predict the assemblage of taxa that could be expected to occur at any test site under minimal human stress. Deviation from that expectation is used to infer degradation just as it is in predictive models (e.g., RIVPACS).

This is a promising approach; even the primitive assignment of taxa to simple preference classes used by Chessman and Royal (2004) resulted in stronger associations between their water quality assessments and independent measures of human disturbance than did the Australian predictive models developed from reference sites. They achieved similar results when applying the technique to fish assemblages (Chessman *et al*. 2007).

To adapt this to California's heavily modified regions, there is a need to develop models of the environmental affinities of Central Valley and southern coastal California lowland BMI taxa. It will likely take several years to collect enough samples to characterize individual BMI responses across key environmental gradients, but some of this data has already been collected and could be worked with now.

3.3 Combining approaches

The alternatives described in this section are not mutually exclusive; the RCMP could use more than one in each region. It is possible that not all approaches will work equally well in all regions and, as a result, different alternatives might be used in different regions. The panel was silent on which approaches, or which combinations of approaches should be prioritized.

The panel cautioned that using these non-standard approaches would require significant effort. Since these non-standard approaches have been used sparingly elsewhere, and essentially not at all in California, pilot studies looking into their applicability was recommended. The first step in the panel's recommendation was to evaluate existing datasets to determine if historical data exists for implementing any of these approaches. As mentioned in section 3.2.2, these approaches should be tested in a location where reference sites exist. Developing any non-standard approach needs to be ground-truthed before widespread use of the tool should be applied. Once this proof-of-concept occurs, then targeted data collection in one of the reference-poor regions can be initiated.

MANAGING THE REFERENCE POOLS

Accounting for natural variation

Classification of streams according to natural gradients can help partition natural sources of variation in biological assemblages and thereby improve our ability to detect deviation from reference condition (see Hughes 1995 for a review of the history of stream classifications). The RCMP needs to ensure that the regional reference site pools are representative of the most important regional gradients. The best way to test the representation of these gradients is through ordination of BMI datasets to determine which natural gradients explain most BMI variation in each region. Assessment of natural variation should include a periodic review of the suitability of the initial regional boundaries. The initial boundaries may either expand or contract and regions may need to be subdivided or merged as we gain more detailed information about the drivers of natural biological variation in each region.

However, since most regions do not have many reference sites to begin with, these analyses will have to take place iteratively as the program builds up a sufficient number of sites in each region. As initial guide, the panel recommended that the RCMP attempt to distribute sites to represent the following natural gradients:

- Stream size (stream order, discharge volume, etc.)
- Geology (with special attention to gradients in calcareous composition)
- Climate (temperature and precipitation)
- Elevation
- Reach slope (an important driver of stream morphology and substrate composition)
- Conductivity and natural nutrient gradients (associated with alkalinity)

The second component to site management is periodic review of sites in the reference pools to assess their continued suitability as reference sites. Conditions within stream reaches and in their upstream drainages can change over time (e.g., timber harvest, conversion of natural landscapes to agricultural or urban/suburban/exurban uses). Furthermore, we may discover sources of stress that were unknown when sites were initially added to the reference pools (e.g., discovery of point source discharges, mines, flow withdrawals/diversions, small-scale placer mining, etc.). Sites that fall into this category may be monitored to measure the impacts of these stressors, but they should be removed from the reference pools.

Dealing with natural disturbance

Natural disturbances such as forest fires, catastrophic floods and landslides can have a significant impact on biological assemblages and physical habitat conditions. As such, they can contribute considerable noise to reference distributions, thereby reducing the precision of biological assessment tools based on these distributions.

There are several competing philosophies for how to handle sites with recent natural disturbances. For example, Idaho's program flagged sites affected by natural disturbance to assess in parallel with other reference sites (Grafe 2004). In contrast, Oregon explicitly included these sites with other reference sites, as a means of incorporating natural disturbance as a component of natural variability (Drake 2003). The RCMP will keep these sites in the reference pools, but will not sample them after the disturbance. The appropriate time to avoid sampling disturbed reference sites is not currently known and should be the subject of targeted research or special study.[14]

[14] The San Diego Regional Water Quality Control Board has funded a multi-year project with the ABL to track biological assemblage recovery in reference and test sites following two large scale forest fires events in 2003 and 2007.

MONITORING STRATEGY

Monitoring Design

The primary question to be answered from the monitoring of the RCMP is "what is the expected natural composition of lotic freshwater organisms in each of the major biogeographical regions of California"? In order to answer this question, the panel agreed it is most important to gather information from a large number of sites in order to capture the full range of natural variability within a region. To collect this information in a spatially balanced and unbiased fashion, the panel advocated a probabilistic sampling design. Probabilistic designs were used in the REMAP, WEMAP, CMAP and PSA surveys in order to get unbiased estimates of stream condition and the approach for this design would be similar. In this case, the regional reference pool would represent the sample frame where sites would be selected at random for sampling. As in the PSA, these randomly drawn sites could be stratified to ensure the spatial distribution across natural gradients such as stream order, elevation, slope, geology, precipitation, or other factors.

An important secondary component to answering the monitoring question is to assess how the range of natural conditions changes over time. Certainly year-to-year variability can alter the distribution and abundance of organisms based on climatic conditions (i.e., wet vs. dry year, warm vs. cold year, etc.). Revisiting sites is the most powerful way to gather this type of temporal information. Two designs lend themselves to answering this question. The first would be to revisit a subset of the probabilistic sites. The panel favored this type of design, termed "rotating panel", because it provides both temporal and spatial variance terms. Urquhart and Kincaid (1999) and Larsen *et al*. (2004) describe the rotating panel strategy in more detail. However, a large number of potential reference sites are already being monitored on a regular basis. Provided these sites can pass the large- and local-scale screening criteria, the panel recommended sampling these sites as a cost-effective method to gain trends information at specific locations of interest. The main drawback to the targeted design, however, is the lack of ability to extrapolate to other reference locations.

Indicators and methods

Once the reference site pools are established, they can be sampled to meet the needs of a variety of programs. However, the panel agreed that a base program should monitor those indicators that are currently being used for SWAMP's statewide assessments (see PSA text box). These indicators include BMIs, physical habitat quality and basic water quality measurements. In some instances, enhancement of the indicators in certain regions or at certain sites may be needed to

Indicators sampled for the SWAMP Perennial Stream Assessment (PSA)

Biological
- BMIs
- Algae (diatoms, soft algae)
- CRAM riverine wetland methods

Physical Habitat
- SWAMP instream and riparian condition (derived from EMAP field protocols)

Chemical
- Nutrients (SRP, NO_2, NO_3,TP, TN, Si)
- Major ions (Cl^-,$S0_4$)
- SSC, turbidity
- pH
- Hardness, alkalinity, conductance

address local concerns. Region-specific enhancements were deemed acceptable as long as the base program is not handicapped to implement the enhancements. For example, additional biological indicators such as fish have been used by others (Hughes *et al*. 2005; Brown and Moyle 2005). Field and laboratory methods and quality assurance measures should also be consistent with SWAMP.

Number of reference sites

The appropriate number of sites to sample in each region depends on the extent of variation related to natural gradients, which is currently unknown for most regions. The panel therefore could not provide specific guidance on sample size. Instead the panel made two recommendations:

1. The RCMP should sample approximately 50 sites in each region to support assessments of natural variability. Intensification of sampling in initial years was recommended to establish the reference baseline, with potentially reduced intensity in later years.

2. The RCMP should conduct power analysis to determine the optimal sample size for assessing confidence in the statistical parameters of the distribution of biological metrics (Figure 9). For example, an assessment of variance at reference sites within a region can be calculated based on existing data (although not all regions have enough sites to support this at present). The inflection point of this power curve represents an efficient sample size where additional sites provide little improvement in confidence, yet fewer sites might dramatically broaden the confidence limits.

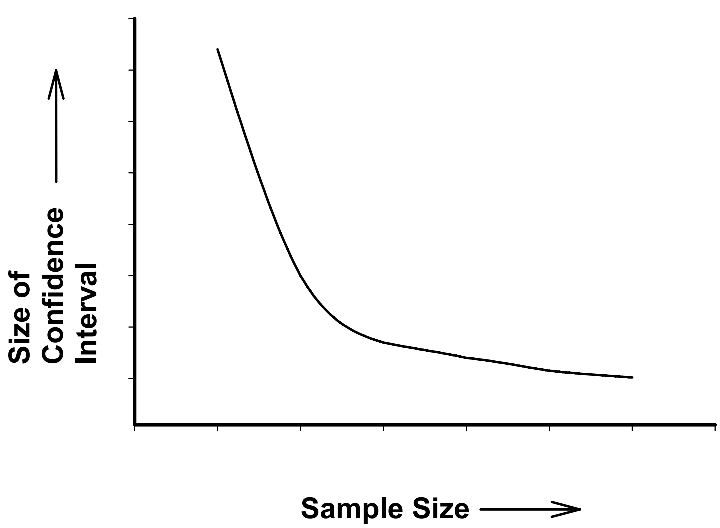

Figure 9. Example power curve defining sample sizes relative to site variability.

Sampling frequency
Sampling frequency affects trend detection. The optimal sampling frequency for trend detection is a function of sampling design. Trend detection as part of the probabilistic design is a function of number of sites (spatial variability), sampling frequency (temporal variability), amount of change to be detected, and other factors. The panel recommended a subset of probabilistic sites be sampled once within the appropriate index period for the region (should be consistent with the index period used for the SWAMP PSA). The recommended index period should capture a time frame where benthic macroinvertebrate communities are sufficiently stable to produce repeatable results, but prior to stress from late season flow reductions. Revisiting a subset of probabilistic sites each year will provide an estimate of interannual variability, thus improving large-scale trend detection. The proportion of revisited sites was not addressed specifically by the panel, but could be optimized using power analysis.

The panel agreed that targeted sites were an efficient way to assess long-term trend detection. Sampling frequency at targeted sites is a function of variability in the biological metrics, the amount of time required to detect a trend, and the amount of detectable change. The panel recommended that the RCMP use power analysis to establish the optimal sampling frequency (Figure 10). Once again, this could possibly be accomplished using data from existing sites that have been sampled for a number of years.

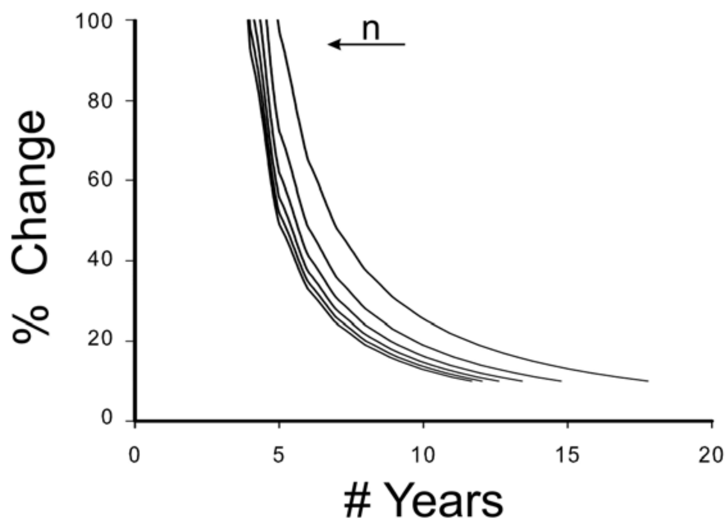

Figure 10. Theoretical power curves describing the relationship between the number of samples collected and the magnitude of detectable change at fixed sites. Individual curves represent different numbers of samples per year, with higher numbers toward the left of the figure.

ADDITIONAL RECOMMENDATIONS

Funding

Defensible bioassessment techniques and biocriteria require a reference condition program that can document both spatial and temporal variation. While the panel did not recommend a minimum level of funding, they advised that funding will need to be long term and stable. Several cost-effective strategies are available, but options discussed included trade-offs between probabilistic and targeted sites, optimizing sample size using power analysis (see previous section on sample size and frequency), and finding additional partners to help support the RCMP (see section below on collaboration). Regardless, SWAMP should prioritize some sampling effort every year to document annual variation in reference condition.

Inter-regional consistency

The RCMP should continue to focus on the issue of fostering consistency among the various regions of the state. Statewide assessments and comparisons among regions require a common currency for interpreting statewide assessments, and for inter-regional comparisons. However, this goal is complicated by the need for regional specific reference selection criteria. While the panel did not deliberate extensively on this topic, it recognized the importance of the issue and provided some initial guidance to help focus the thinking of the program. The main advice from the panel was that the objective of inter-regional consistency can probably not be resolved by the reference site selection process itself, but rather must be dealt with through data analysis and interpretation.

Development and application of assessment tools can be based on either regional reference pools or combined statewide reference pools. Regionalized assessment tools provide sensitivity to local environmental gradients and are more likely to pick up sites that deviate from the regional expectation. In contrast, statewide assessment tools would judge all of the state's sites on the same basis, but may reduce responsiveness to locally important gradients. Furthermore, we may have to accept that the performance of statewide analytical tools may vary regionally depending on the quality of the respective regional reference pools.

An example of an analytical solution is a hybrid approach in which both the regional and statewide indices are built and both tools are used to score test sites. Where both tools agree, there is relative certainty in the assessment of that site (i.e., both tools indicate reference-like or both indicate impacted). Where the tools disagree, a greater degree of relative uncertainty exists and additional information may be required to help interpret the status of that site (i.e., other indicators, additional sampling).

Collaborations/Coordination

Consistent with its policy to coordinate with other state and federal water quality monitoring efforts, SWAMP should seek opportunities to build partnerships with other state and federal agencies. Many of these entities have current reference programs (e.g., USFS, EPA, USGS), while others would benefit from joining an established reference program (e.g., Non-point Source Monitoring, State Parks, Irrigated Lands Program, Agricultural Coalitions, National Park Service, etc.). In addition SWAMP should explore ways to combine its bioassessment RCMP with other program components that would benefit from reference condition (e.g., CRAM, wetland monitoring, nutrient and sediment criteria monitoring).

The panel recommended exploration of ways to improve the types and quality of data used in GIS analyses. For example, the program could seek opportunities to coordinate with other state/federal/university efforts to enhance base layers like the NHD+ and stressor layers for quantification of grazing, timber harvest, pesticide application, etc. Further, the RCMP would should explore research efforts designed to improve prediction of specific stressor impacts and efforts to develop models that can be used to assess impact components that are not easily summarized by the ATtILA model. For example a model predicting sediment load (AnnAGNPS sedimentation model, USDA 2000) was applied by the University of Nevada, Reno. Other needs include estimating mining impacts, pesticide impacts and a means for summarizing the intensity of water manipulation within candidate areas.

Involving stakeholders in the process

It is often desirable to select sampling locations that occur on publicly owned land or land with easy access. Since it is important to sample streams from a truly representative set of sites within an area, it is often necessary to sample from reaches running through privately owned land. Reasonable efforts should be taken to obtain permission from landowners before rejecting candidate sites. This stage is very important and the quality of the final data set (and the ability to make inferences about reference conditions in the region of interest) will depend on the ability to obtain a representative set. The degree to which this stage is important varies regionally since some areas have more private ownership than others (e.g., western Sierra Nevada has many more publicly-owned lands than the interior chaparral).

Building effective relationships with local stakeholders (regional boards, watershed groups, landowner group, tribal groups, etc.) is clearly a critical part of making this reference site selection methodology work, especially in regions with a large degree of private ownership. To this end, implementation of this RCMP should include efforts to promote transparency in methods, encourage feedback and participation and explore opportunities to improve access to important privately held reference sites.

CONSIDERATIONS FOR OTHER FLOWING WATERS

The following section is not intended to be an exhaustive review of issues for defining reference conditions for these waterbodies, but a summary of the panel's preliminary guidance regarding issues that are likely to be important in these systems.

Large Rivers/ Non-wadeable streams
Large rivers are likely to require non-standard approaches to defining biological expectations because there are relatively few non-wadeable streams/rivers in the state and most receive the cumulative impacts of all human activities in their watersheds. Furthermore, several panelists suggested that standard chemical and physical habitat screening was unlikely to work in these systems. Screening criteria should include quantification of hydromodification, distance downstream from dams or other stressors.

Several of the alternative strategies could apply to these systems. Another alternative would be to target sampling at points along river just before they experience significant increases in sources of anthropogenic stress (e.g., where rivers in the western Sierra Nevada descend into the Central Valley).

Non-perennial streams
Non-perennial streams tend to have more variable biological assemblages than perennial streams. The standard approach should work for most of these systems statewide, but special attention should be given to classification of non-perennial streams by their degree of "intermittent-ness" in both space and time. The panel suggested that the RCMP should take advantage of current statewide vegetative mapping efforts to explore the potential for classifying non-perennial streams.

LITERATURE CITED

ANZECC and ARMCANZ. 2000. Australian and New Zealand guidelines for fresh and marine water quality. Volume 2. Aquatic ecosystems- rationale and background information. Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Australian Government, Department of the Environment and Heritage, Canberra, Australia.

Arnold, C.L.J. and C.J. Gibbons. 1996. Impervious surface coverage: The emergence of a key environmental indicator. *Journal of the American Planning Association* 62:243–258.

Bailey, R.G., P.E. Avers, T. King, and W.H. McNab, (eds.). 1994. Ecoregions and subregions of the United States (1:7,500,000 scale map, with supplementary table of map unit descriptions, compiled and edited by W.H. McNab and R.G. Bailey). USDA Forest Service. Washington, DC.

Bailey, R.C., M.G. Kennedy, M.Z. Dervish, and R.M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. *Freshwater Biology* 39:765-774.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. Bioassessment of Freshwater Ecosystems: using the Reference Condition Approach. Kluwer Academic Publishers. Norwell, MA.

Barbour, M.T., J.B. Stribling, J. Gerritsen, and J.R. Karr. 1996a. Biological Criteria: Technical guidance for streams and small rivers. Revised Edition. EPA 822-B-96-001. USEPA Office of Water. Washington, DC.

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996b. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15:185-211.

Barbour, M.T. and C. Hill. 2003. The status and future of biological assessment for California streams. Report to the California State Water Resources Control Board. Tetra Tech, Inc. http://www.swrcb.ca.gov/swamp/reports.html

Brown, L.R. and P.B. Moyle. 2005. Native fishes of the Sacramento-San Joaquin Drainage, California: a history of decline. Pages 75-98 in: J.N. Rinne, R.M. Hughes, and B. Calamusso (eds.) Historical Changes in Large River Fish

Assemblages of the Americas. American Fisheries Society, Symposium 45. Bethesda, MD.

Carter, J.L and S.V. Fend. 2005. Setting limits: the development and use of factor-ceiling distributions for an urban assessment using benthic macroinvertebrates. *American Fisheries Society Symposium* 47:179-191.

Chessman, B.C. 1995. Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family-level identification, and a biotic index. *Australian Journal of Ecology* 20:122-129.

Chessman, B.C. 2006. Prediction of riverine fish assemblages through the concept of environmental filters. *Marine and Freshwater Research* 57:601-609.

Chessman, B.C. and M.J. Royal. 2004. Bioassessment without reference sites: use of environmental filters to predict natural assemblages of river macroinvertebrates. *Journal of the North American Benthological Society* 23:599-615.

Chessman, B.C., M. Muschal, and M.J. Royal. 2008. Comparing apples with apples: use of limiting environmental differences to match reference and stressor-exposure sites for bioassessment of streams. *River Research and Applications* 24:103-117.

deVlaming, V., D. Markiewicz, K.L. Goding, T. Kimball, and R. Holmes. 2004. Macroinvertebrate assemblages in agriculture- and effluent-dominated waterways of the lower Sacramento River watershed. Report to the Central Valley Water Quality Control Board. UC Davis Aquatic Toxicology Laboratory. Davis, CA.

deVlaming, V., D. Markiewicz, K.L. Goding, A. Morrill, and J. Rowan. 2005. Macroinvertebrate assemblages of the San Joaquin River watershed. Report to the Central Valley Water Quality Control Board. UC Davis Aquatic Toxicology Laboratory. Davis, CA.

Drake, D. 2003. Selecting Reference Condition Sites: An Approach for Biological Criteria and Watershed Management. Oregon Department of Environmental Quality, Watershed Assessment Division. Portland, OR.

Grafe, C.S. 2004. Selection of Reference Condition for Small Streams in Idaho: A Systematic Approach. Department of Environmental Quality, Surface Water Quality Program. Boise, ID.

Griffith, M.B., P. Husby, R.K. Hall, P.R. Kaufmann, and H.H. Brian. 2003. Analysis of macroinvertebrate assemblages in relation to environmental gradients among lotic habitats of California's Central Valley. *Envirnomental Monitoring and Assessment* 28:281-309.

Hall, L.W. and W.D.Killen. 2001. Characterization of benthic communities and physical habitat in an agricultural and urban stream in California's Central Valley. University of Maryland-Agricultural Experiment Station. Queenstown, MD.

Hawkins, C.P. and R.H. Norris (eds.). 2000. Landscape classifications: aquatic biota and bioassessments. Journal of the North American Benthological Society 19.

Herbst, D.B. and E.L. Silldorff. 2006. Development of an Index of Biological Integrity (IBI) for Stream Assessments in the Lahontan Region of California. DRAFT Report to the Lahontan Regional Water Quality Control Board.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 in: W.S. Davies and T.P. Simon (eds.), Biological assessment and criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers. Ann Arbor, MI.

Hughes, R.M. and D.P. Larsen. 1988. Ecoregions: an approach to surface water protection. *Journal of the Water Pollution Control Federation* 60:486-493.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: A method for assessing stream potentials. *Environmental Management* 10:629-635.

Hughes R.M., J.N. Rinne, and B.Calamusso. 2005. Historical changes in large river fish assemblages of the Americas: A synthesis. Pages 603-612 in: R.M. Hughes, J.N. Rinne, and B.Calamusso(eds.), Historical Changes in Large River Fish Assemblages of the Americas. American Fisheries Society, Symposium 45. Bethesda, MD.

Karr, JR. 1995. Protecting aquatic ecosystems: clean water is not enough. Chap. 2. Pages 7-13 in: W.S. Davis and T.P. Thomas (eds.), Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Inc., Boca Raton.

Karr, JR. 1997. Measuring biological integrity. Essay 14A. Pages 483-485 in: G.K. Meffe and G.R. Carroll (eds.), Principles of Conservation Biology, 2nd edition. Sinauer, Sunderland, Massachusetts.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press. Washington, DC.

Larsen, D.P., P.R. Kaufmann, T.M. Kincaid, and N.S. Urquhart. 2004. Detecting persistent change in the habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 61:283-291.

Meyer, J.L. 1997. Stream health: incorporating the human dimension to advance stream ecology. Journal of the North American Benthological Society 16:439-447.

Ode, P.R. 2002. A quantitative framework for reference site selection: case study from the Sierra Nevada Foothills ecoregion. Sacramento River Watershed Program. California Department of Fish and Game, Water Pollution Control Laboratory. Rancho Cordova, CA.

Ode, P.R. 2007. Standard operating procedures for collecting macroinvertebrate samples and associated physical and chemical data for ambient bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 001. Sacramento, CA.

Ode, P.R., D. P. Pickard, J. P. Slusark, and A.C. Rehn. 2005. Adaptation of a bioassessment reference site selection methodology to creeks and sloughs of California's Sacramento Valley and alternative strategies for applying bioassessment in the valley. Report to the Central Valley Regional Water Quality Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory, Rancho Cordova, CA.

Omernik, J.M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49-62 in: W.S. Davies and T.P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making. Lewis Publishers, Ann Arbor, Michigan.

Paulsen, S.G., J.L. Stoddard, S. Holdsworth, A. Mayio, and E. Tarquino. 2006. Wadeable streams assessment: A collaborative survey of the nation's streams. EPA 841-B-06-002. USEPA Office of Research and Development/Office of Water. Washington, DC.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16: 391-409.

Poff, N.L. and J.V. Ward. 1990. Physical habitat template of lotic systems: recovery in the context of historical pattern of spatiotemporal heterogeneity. Environmental Management 14: 629-645.

Reference Condition Working Group 2.3 (REFCOND). 2003. Rivers and Lakes – Typology, Reference Conditions and Classification Systems. Guidance Document No. 10. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Luxembourg.

Rehn, A.C., P.R. Ode, and J.T. May. 2005. Development of a benthic index of biotic integrity (B-IBI) for wadeable streams in northern coastal California and its application to regional 305(b) assessment. Report to the State Water Resources

Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory. Rancho Cordova, CA.

Rosenberg D.M., T.B. Reynoldson, and V.H. Resh. 1999. Fraser River Action Plan: Establishing Reference Conditions for Benthic Invertebrate Monitoring in the Fraser River Catchment. Environment Canada No. DOE-FRAP 1998-32. British Columbia, Canada.

Statzner, B., A.G. Hildrew, and V.H. Resh. 2001. Species traits and environmental constraints: entomological research and the history of ecological theory. *Annual Review of Entomology* 46: 291-316.

Stoddard, J.L., D.V. Peck, S.G. Paulsen, J. Van Sickle, C.P. Hawkins, A.T. Herlihy, R.M. Hughes, P.R. Kaufmann, D.P. Larsen, G. Lomniki, A.R. Olsen, S.A. Peterson, P.L. Ringold, and T.R. Whittier. 2005. An Ecological Assessment of Western Streams and Rivers. EPA 620/R-05/005. US Environmental Protection Agency, Washington, D.C.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R. Johnson, and R. Norris. 2006. Setting expectations for ecological condition of streams: the concept of reference conditions. *Ecological Applications* 16:1267-1276.

Urquhart, N.S. and T.M. Kincaid. 1999. Designs for detecting trend from repeated surveys of ecological resources. *Journal of Agricultural, Biological, and Environmental Statistics* 4:404-414.

US Environmental Protection Agency (USEPA). 2005. Use of Biological Information to Better Define Designated Aquatic Life Uses in State and Tribal Water Quality Standards: Tiered Aquatic Life Uses. EPA-822-R-05-001. USEPA Office of Water, Washington, DC.

USEPA. 2009. Evaluation of the California State Water Resource Control Board's Bioassessment Program. Prepared by C. Yoder and R. Plotnikoff, Center for Applied Bioassessment and Biocriteria, Midwest Biodiversity Institute, Columbus, OH, for USEPA Office of Science and Technology. Washington, DC.

Wright, K.K. and J.L. Li. 2002. From continua to patches: examining stream community structure over large environmental gradients. Canadian Journal of Fisheries and Aquatic Sciences 59: 1404-1417.

Waite, I.R., A.T. Herlihy, D.P. Larsen, N.S. Urquhart, and D.J. Klemm. 2004. The effect of macroinvertebrate taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands, U.S.A. Freshwater Biology 49: 474-489.

Whittier, T.R., D.P. Larsen, R.M. Hughes, C.M. Rohm, A.L. Gallant, and J.M. Omernik. 1987. The Ohio Stream Regionalization Project: A compendium of results. EPA/600/3-87-025. Corvallis, OR.

Yoder, C.O. and E.T. Rankin. 1995. Biological criteria program and implementation in Ohio. Pages 109-144 in: Davis W.S. and T.P. Simon (eds), Biological Assessment and Criteria: Tools for Water Resource Planning. CRC Press. Boca Raton, FL.