SWAMP
Surface Water Ambient Monitoring Program

CALIFORNIA NPS PROGRAM

Technical Report	2005

Probabilistic Assessment of the Biotic Condition of Perennial Streams and Rivers in California

August 2005

California Water Boards

www.waterboards.ca.gov/swamp

Probabilistic assessment of the biotic condition of perennial streams and rivers in California

Report to the State Water Resources Control Board

Aquatic Bioassessment Laboratory
Water Pollution Control Laboratory
California Department of Fish and Game
Rancho Cordova, California

August 2005

SUGGESTED CITATION

Ode, P.R. and A.C. Rehn. 2005. Probabilistic assessment of the biotic condition of perennial streams and rivers in California. Report to the State Water Resources Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory, Rancho Cordova, California.

ACKNOWLEDGEMENTS

The assessments presented in this report are entirely the products of data collected under the EPA's Western Pilot EMAP program (WEMAP) with financial and technical support provided to the California Department of Fish and Game under a Cooperative Agreement with the EPA's Office of Research and Development. This report was developed under Assistance Agreement No. CR82823801 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of the Aquatic Bioassessment Laboratory and EPA does not endorse any products or commercial services mentioned in this publication.

This report is part of an ongoing project (CMAP) developed in collaboration with the State Water Resources Control Board's Non-point Source Pollution Control Program (NPS) and Surface Water Ambient Monitoring Program (SWAMP) with funding from the US EPA's Region IX.

These analyses were based on methodologies developed by the US EPA's Office of Research and Development (ORD) for its EMAP program. We are especially indebted to Tony Olsen, who developed the probabilistic sampling designs used for this assessment, wrote the scripts for combining the four study designs and helped with interpretation of the output. We also thank Chuck Hawkins, who developed the new RIVPACS models for his assistance with calculating RIVPACS scores and interpreting model output, and to John Olson and Ryan Hill, graduate students at USU who provided assistance with the calculation of predictors used in the RIVPACS models. Development of the California RIVPACS predictive model by Dr. Chuck Hawkins of Utah State University was funded in part by the Pacific Southwest Region of the USDA Forest Service (Cooperative Agreement #PSW-88-0011CA, Cost Reimbursable Agreement # 03-CR-11052007-100, and Contract # 53-9A63-00-1T52). Will Patterson (DFG-ITB) provided helpful assistance with GIS grids.

The project owes much of its success to the efforts of ABL staff. James Harrington managed the ABL's collaboration with the EMAP program. ABL field staff, led by Shawn McBride with assistance from Michael Dawson and Jennifer York, were responsible for the extensive field reconnaissance and data collection efforts, ABL taxonomic staff (Dan Pickard, Doug Post, Brady Richards and Joe Slusark) were responsible for all taxonomic data, and Glenn Sibbald contributed watershed delineations for the RIVPACS and GIS analyses.

INTRODUCTION

States are required under the Clean Water Act §305(b) to report annually to Congress on the condition of their waterbodies, but satisfaction of this requirement has been confounded by the lack of resources (both technical and financial) to adequately assess these waterbodies. Partly because of this limitation, most regions of the nation have been unable to answer the most basic questions about their water quality, even 35 years after the passage of the act.

To address these constraints, the U.S. EPA established its Environmental Monitoring and Assessment Program (EMAP), a long term research program designed to develop the tools and techniques needed for cost-effectively answering the fundamental status and trends questions in the Clean Water Act. The EMAP studies are based on a probabilistic survey design in which each sampling location represents a known length of stream with known statistical confidence. This design permits the inference of stream conditions for large geographic regions with a relatively small investment in sampling. After completing an assessment of the condition of wadeable streams in the Middle-Atlantic states, the EPA initiated a similar assessment of streams in the western states (WEMAP), which included a high density of sites in California. Now that the first round of WEMAP studies has been completed, California has the ability to include the first statistically defensible condition assessments in its 305b reports.

Because they provide a direct measure of the biotic integrity, bioassessment data are a key component of EMAP water quality monitoring. Bioassessment, the science of describing the ecological condition of waterbodies from the assemblages of organisms they contain, is well established as a valuable tool for water resource management (Karr 1981, Yoder and Rankin 1995, Barbour *et al.* 1996, Wright *et al.* 2000, Bailey *et al.* 2004). Because assemblages of aquatic organisms (e.g., fish, benthic macroinvertebrates (BMIs) and algae) are comprised of taxa that are differentially responsive to different environmental stressors, bioassessments provide a direct means of measuring compliance with the goal of biotic integrity stipulated under the Clean Water Act. Although comprehensive condition assessments will ultimately include assessments of the physical and chemical conditions of waterbodies, condition assessments based on bioassessment data can stand alone as effective measures of the ecological condition of the state's waterbodies.

There are many different approaches to translating a list of organisms present at a site into an assessment of its ecological condition (Wright *et al.* 1984, Kerans and Karr 1994, Hawkins *et al.* 2000, Van Sickle *et al.* 2005, Ode *et al.* 2005). We have demonstrated elsewhere (Rehn and Ode 2004, Rehn and Ode 2005) the use of BMI data to produce regional 305(b)-type stream condition assessments using multimetric techniques to calculate site condition scores. These regional assessments were possible because we had previously developed benthic indices of biotic integrity (B-IBIs) for northern and southern coastal California; statewide assessments require a tool for scoring sites that can be applied statewide.

We present here a statewide condition assessment using predictive models based on the River Invertebrate Prediction and Classification System (RIVPACS, Wright 1984). Like multimetric approaches (Kerans and Karr 1994, Ode *et al.* 2005, Rehn and Ode 2005), predictive modeling techniques establish thresholds of ecological impairment based on a characterization of the biotic assemblages expected to occur under minimal human disturbance (Wright *et al.* 1984, 1989, 2000). However, predictive models compare assemblages at test sites to an expected taxonomic composition rather than expected metric values. Taxon-based models have seen widespread use since the first BMI models were created in Great Britain in the late 1970s (Norris and Georges 1993, Hawkins *et al.* 2000, Van Sickle *et al.* 2005) and have been promoted in the US (Hawkins *et al.* 2000, Hawkins and Carlisle 2001) as an alternative to the multimetric approach initially endorsed by the EPA (Barbour *et al.* 1999). For this analysis, we employed newly developed California RIVPACS models (C. Hawkins unpublished) that can be used to score sites throughout the state.

METHODS

Study Design/ Site Selection

The study designs used for this condition assessment were very similar to those used for our southern coastal California condition assessment (Rehn and Ode 2004). The probabilistic survey was a generalized random tessellation stratified (GRTS) design with reverse hierarchical ordering. There was no stratification in the design, but site selection weights were adjusted so that Strahler stream order categories (1^{st}, 2^{nd}, 3^{rd}, and 4^{th}+) were sampled in approximately equal proportions. We combined 4 separate survey designs for this analysis (Figure 1). Three of these were modifications of the main WEMAP sample frame: 1) the California statewide sites that were part of the larger WEMAP design, 2) the southern coastal California special interest sites, and 3) the northern coastal California special interest sites. A separate GRTS survey was created in 2003 to increase the representation of sites in the central coast region. In each of the designs, a list of potential sampling locations was generated randomly from the EPA's 1:100,000 RF3 hydrology layer. For analyses, each potential sampling site was assigned a weighting factor proportional to the number of stream kilometers it represented.

Site evaluation

Once the list of potential sampling coordinates was generated for each region, we conducted site reconnaissance to identify sites that were part of the population of streams of interest (natural channels with perennial flow). There are many reasons why potential sites were rejected during the reconnaissance phase. In the arid southwest, many streams that appear as perennial on USGS quadrant maps (and the 1:100,000 RF3 stream layer digitized from them) are, in fact, not perennial. Earlier analyses indicated that approximately 65% of stream length indicated as perennial in the southern coastal region was actually non-perennial (Rehn and Ode 2004). Underground pipelines, canals and aqueducts frequently cannot be distinguished from streams on the RF3 stream layer, and these also were rejected as non-target during reconnaissance. Also, some perennial sites were inaccessible due to physical barriers (e.g., access was too dangerous or required excessive backpacking). Private ownership further confounded site selection. When landowners denied access to a site, it was impossible to determine its target status, and it was categorized as "status unknown".

Site reconnaissance continued until a pool of approximately 60 target sites each was identified and sampled from the northern coast, the southern coast and statewide and 30 sites were sampled from the central coast region. During the reconnaissance process, we evaluated 1140 sites, keeping careful records of each site's target status, and if applicable, reasons why sites were eliminated from the target pool for use in later analyses. We sampled over 200 study reaches throughout California between April and September of 2000 through 2003, sampling southern sites at the beginning of the sampling season and progressing north later in the year.

Field Methods

Once target sites were identified, we sampled sites according to standard WEMAP field methods (Peck *et al.* 2004). A sampling reach was defined as 40 times the average stream width at the center of the reach, with a minimum reach length of 150m and maximum length of 500m. We collected two BMI samples from each reach: 1) a reachwide composite sample (multiple habitat) consisting of 11 one ft^2 samples taken from equally spaced locations throughout the reach and 2) a targeted riffle sample consisting of 8 one ft^2 samples taken from fast water habitat units within the reach (Hawkins *et al.* 2001). Only the targeted riffle sample was used in these analyses. Fish and algae samples were collected according to Peck and others (2004). Water chemistry samples were collected from the mid-point of each reach and analyzed using WEMAP protocols (Klemm and Lazorchak 1994). Field crews recorded physical habitat data using EPA qualitative methods (Barbour *et al*. 1999) and quantitative methods (Kaufmann *et al*. 1999).

Lab Methods

All BMI samples were processed at DFG's Aquatic Bioassessment Laboratory in Chico, CA. A 500 organism subsample of each BMI sample was processed and identified according to WEMAP standard taxonomic effort levels (CSBP II, www.dfg.ca.gov/cabw/camlnetste.pdf). All taxonomic data were entered into an MS Access database (CalEDAS) that allowed us to produce standardized taxa lists at different standard effort levels. Five percent of taxa were re-identified for quality assurance and archived vials of all samples are housed at the Chico facility.

Calculation of RIVPACS scores

The goal of RIVPACS is to compare the list of taxa observed at a site (O) to the list of taxa predicted to occur at a given site in the absence of human disturbance (E). The approach has four components: 1) reference sites are classified according to degree of taxonomic similarity, 2) environmental variables associated with each class are identified, 3) discriminant functions analysis (DFA) is used to predict class membership of new test sites based on the values of their environmental predictor variables, 4) the observed list of taxa is compared to the expected list to calculate the O/E ratio.

The most recently derived RIVPACS models for California streams were completed in June 2005 (Hawkins unpublished presentation). Preliminary attempts to create one model for California resulted in relatively imprecise models, but an initial classification step using precipitation and temperature variables produced 3 separate sub-models with better performance.

To apply the new RIVPACS models to our WEMAP data, we prepared separate files of taxa and predictor variables for each of the 3 sub-models. Note that we did not include any targeted EMAP reference sites in the condition assessments. Taxonomic lists generated from CalEDAS were modified for compatibility with the formats used in the RIVPACS models by: 1) eliminating ambiguous taxa, 2) using a rarefaction subroutine to subsample 300 organism counts from the original 500 count samples, 3) converting the final taxonomic names to the operational taxonomic names (OTUs) used in the models (converting chironomid midges to subfamily), and 4) cross-tabulating the taxonomic list into a taxon by site matrix.

We determined the values of six map-based predictor variables for each site: 1 and 2) geographic coordinates (latitude and longitude) were obtained from the original study design file, 3) watershed area was calculated by delineating upstream watershed boundaries for each site in using a GIS, 4) log mean "normal" precipitation was estimated by overlaying sites on a GIS grid of mean monthly precipitation (1961-1990) obtained from the Oregon Climate Center (OCC, www.ocs.orst.edu/prism), 5) mean "normal" temperature was estimated from mean monthly temperature grids (1961-1990) also obtained from the OCC, 6) percent sedimentary geology was estimated from an unpublished GIS geology classification of the western United States derived by John Olson, (PhD student at Utah State University) from a generalized geologic map of the coterminous US (Reed and Bush, pubs.usgs.gov/atlas/geologic/).

Once predictor variables were determined for each site, we used precipitation and temperature data to assign each site to one of the three classes based on the following criteria. Sites with mean monthly temperatures (Tmean) less than 9.9ºC were assigned to Class 3, sites with temperatures greater than 9.9ºC were assigned to Class 2 if they had log mean monthly precipitation values (logPPT) less than 2.952, and to Class 1 if logPPT was greater than 2.952. The three sub-models required different sets of predictor variables: Class 1 used latitude, log watershed area, and mean temperature; Class 2 used longitude, percent sedimentary geology and mean precipitation; Class 3 used log watershed area and mean temperature.

The three sets of site files were uploaded to the web interface containing the California models at the Western Center for Monitoring and Assessment of Freshwater Ecosystems (http://129.123.10.240/WMCPortal/DesktopDefault.aspx?tabindex=2&tabid=27). The model output included the probability matrix, O/E scores, and taxon sensitivity scores.

Condition Assessments

Estimation of Stream Condition

The statistical program "R" (Version 2.1.1; www.r-project.org), was used to combine the four design models and adjust site weights to reflect their true percent contribution to the target population (see Rehn and Ode 2004 for more detailed discussion). Adjusted weights were used in conjunction with RIVPACS scores calculated for each site to estimate the percentage of stream miles in "Non-Impaired", "Impaired" and "Very Impaired" ecological condition. We used thresholds of 1.5 and 3 standard deviations below an O/E score of 1.0 (the score expected under no impairment) to set the boundaries between Non-Impaired and Impaired (O/E <0.77), and Impaired and Very Impaired (O/E <0.55), respectively. Although we could have used separate thresholds for each of the three models based on their respective standard deviations, we used the average standard deviation for the three sub-models (0.15) because they were nearly identical.

Associate NPS stressors with Biotic Condition
Because each site represents a known number of stream kilometers, we can evaluate the association between biotic impairment and various watershed stressors by estimating the percentage of stream length associated with these stressors. We used box plots to compare the extent (or concentration) of 18 anthropogenic stressors among study reaches in each of the 3 biological condition classes (non-impaired, impaired, very impaired) to evaluate these associations. Most of the 18 attributes can be directly or indirectly altered as a result of human activity and have been known to have harmful effects on stream biota (EPA 2000). Relative bed stability is a measure of whether a stream has too much or too little sediment (Kaufmann *et. al.* 1999); increasingly negative numbers on a logarithmic scale indicate "fining" of the sediment, (i.e., the median particle size is much smaller than the stream can transport at bankfull flow). Increasingly positive numbers indicate "armoring" of the substrate, which is solidification of the channel bottom when the stream is sediment starved. W1_HALL is an index created by Phil Kaufman (EPA-ORD) of all human related activities noted in a sampling reach weighted by their proximity to the stream channel (EPA 2000).

A stressor was considered to have a moderate association with biological condition (O/E score) if the means and quartiles of the 'very impaired' and 'unimpaired' classes did not overlap, or a strong association with biological condition if the quartiles did not overlap. When the association was moderate, we used the mean of the 'very impaired' distribution to define a stressor threshold; when the association was strong, we used the point of separation between the quartiles of the 'very impaired' and 'unimpaired' classes to define a stressor threshold. We then calculated the percent of stream miles classified as either 'very impaired' or 'impaired' that also had stressor intensities/concentrations that exceeded the thresholds defined above.

RESULTS

<u>RIVPACS results</u>
Model Outliers
Only 12 sites out of the 195 had environmental predictor values that were outside the experience of the models. With the exception of two sites in the Imperial Valley and one site in the Central Valley, all of these sites were within geographic areas that were well represented in the model.

Model Output
The web interface for running the California RIVPACS models permits the calculation of O/E scores either with credit given for all taxa expected to occur with any probability (including very rare taxa) or with taxa expected to occur with > 50% probability (observed taxa credit given only for common expected taxa). All condition assessments were based on the p>0.5 output because model output with rare taxa excluded tends to be more repeatable (Hawkins personal communication, Rehn and Ode in prep)

Condition Assessments
Site Evaluations
Approximately 37 percent of total stream length (~35,000 km) in the California portion of the RF3 was determined to be not part of the target population of this study (**NT**, Figure 3). Most of this stream length was rejected from the target population either because sites were non-perennial or because the stream channels were diverted to aqueducts or underground pipelines. This proportion was highest in the southern and central coast sites and lowest in the northern coast sites.

Landowner denial **(LD)** was a major factor in site elimination, leaving about 15% of stream length unassessed (Figure 3). Because we can't determine in most cases whether the sites on these private lands were part of the target pool, they were not included in either the Target or Non-Target category and were not included in these estimations. A similarly large proportion of sites (representing ~12% of total stream length) was not sampled because of physical access constraints (**PB**, e.g., sites that were at bottom of inaccessible canyons, sites that required impractical hiking times). Since these had previously been determined through field reconnaissance to meet the target criteria, these sites were included in the target population. In addition, some sites known to be part of the target population through reconnaissance were not used because the required number of sites was sampled without using them (**TNS,** ~8% of total stream length).

Contribution of the Four Surveys to Statewide Estimates
Because the four separate survey designs used in this assessment represent different geographical regions of the state, their contribution to the overall assessment varied with the relative proportion of stream length in each region (Figure 4). For example, although nearly as many sites were assessed in the central coast region (26) as the entire statewide set (35), the central coast sites contributed a very small proportion of the overall assessment of stream condition.

Statewide Condition Assessments
Although condition assessments from probabilistic surveys are commonly expressed as pie charts showing the percentages of the target population in various condition classes (e.g., good, fair, poor), cumulative distribution functions (CDFs) provide a more comprehensive way of presenting the relationship between stream length and biotic condition. The results of the condition assessments are presented in tabular (Table 1), CDF (Figure 5), and pie chart (Figure 6) form. Based on the two preliminary cutoffs used here (1.5 and 3.0 sd below the expected RIVPACS score of 1.0), 67% (±6%) of stream kilometers in perennial streams had no impairment of the BMI assemblages, 23% (±5%) had impaired BMI

assemblages and 10% (±2.5%) had very impaired BMI assemblages statewide (Figure 6). Estimates for the separate surveys are based on smaller sample sizes than the combined statewide estimate (and are therefore much less precise), but are included for comparison purposes. The upper and lower 95% confidence limits on all estimates is based on two standard deviations above and below the estimates and is provided for both percentages of stream length and absolute stream length (Table 1).

Effect of Changing Assumptions on Stream Length Estimates

The total stream length estimated under these condition assessments is a function of assumptions about sites that were part of the target population but not sampled. Figure 7 illustrates the effect of 3 alternate ways that total estimates could be stated. In the least conservative example (Figure 7a), if we assume that all target sites (sampled or not) and all landowner denial sites have the same proportion of impaired and non-impaired sites as the sampled sites, then the condition assessments in this report can be extrapolated to 59,807 km (~2/3 of all streams in California). In the most conservative example (Figure 7c), if we make no assumptions about landowner denial sites or other target sites that were not sampled, 26,415 km can be extrapolated with the remainder left as unknown. Most of the analyses presented in this demonstration are based on the intermediate level of assumptions (i.e., that TS, PB, and TNS sites had similar proportions of impairment), which allowed us to extrapolate to 45,496 km. Note that the different assumptions only affect the stream length that can be estimated; the ratio of impaired to unimpaired streams is constant because it is always based on only the sampled sites.

<u>Stressor Associations</u>

Seven of the 18 stressors that we included in our preliminary analyses showed good discrimination between sites that were in 'very impaired' and 'unimpaired' biological condition (Figure 8). Two of these, mean embeddedness and percent fines and sand, were highly correlated with log relative bed stability (LRBS) and were omitted from further analyses. Total riparian disturbance had the greatest association with poor biological condition: 47% of stream miles that were biologically impaired (i.e., that were either 'impaired' or 'very impaired') also had riparian disturbance scores that exceeded the thresholds defined above (Figure 9). Only 15% of 'unimpaired' stream miles had riparian disturbance scores that exceeded the threshold. Chloride concentration, percent urban at both watershed and local scales, and excess sediment (LRBS) also had strong associations with biological condition.

DISCUSSION

The assessment of perennial streams presented here is the first attempt to make broad statements about the biological condition of an entire class of waterbodies in California. Because of the nature of the probability design used here, we were able to extrapolate the condition of nearly 75 percent of all perennial streams in California. While the accuracy of these extrapolations is ultimately limited by a number of assumptions, it represents the first statistically defensible attempt to address the broad question of the state of California's waterbodies.

Although the fundamental methodologies used here are well established (http://www.epa.gov/nheerl/arm/designpages/monitdesign/survey_overview.htm), the actual estimates of impaired stream length reported here should be interpreted as an example of the estimates that can be made from these results. Ultimately, their value depends on the degree to which water quality regulators are involved in their interpretation. For example, we used impairment thresholds based on the statistical properties of the RIVPACS models, but these thresholds are not static and should be defined by water quality regulators on the basis of how they will be used in a regulatory context. The CDFs shown here provide a means of visualizing how changes in thresholds affect the condition assessments.

One of the most significant limitations to our assessments of is the extent to which we were limited by landowner denial of site access. Improvement of access to private lands is one of the biggest keys to improving the accuracy of these assessments. Even a moderately conservative approach to extrapolation (e.g., Figure 7b) results in a reduction in the total stream length that can be assessed by about 30%. To the extent that landowner denial sites are in better or worse shape than the sampled sites, the overall impairment estimates will be over or under estimates. Physical barriers may also contribute to biased estimates if difficulty getting to a site is correlated with higher site quality (it's unlikely that the relationship is reversed). Both of these factors clearly are good candidates for increased resources, but the strength of these potential biases could be tested directly by sampling some of these sites to see if biases exist. Whatever the result of these tests, this issue illustrates one of the most important features of the statistical approach: the effects of key assumptions can be made transparent to the end user.

Comparisons with previous condition assessments

Although the condition assessments for the four separate sub-regions are based on relatively few sites, they allow us to compare these results with other reported condition assessments based on B-IBIs. The condition assessment for southern and central coastal California (Rehn and Ode 2004) based on our B-IBI (Ode *et al.* 2005) had a nearly identical assessment of stream condition to these results based on RIVPACS scores. The condition assessment for northern coastal California based on RIVPACS scores suggests greater biotic impairment than did the assessment based on our northern coastal B-IBI, using the same BMI data (40% vs 6%, Rehn and Ode 2005). The impairment threshold in the NorCal IBI was set at 2 standard deviations below the mean score of the reference population. When an impairment threshold of 1.5 standard deviations below the mean of the reference pool (equivalent to that used for the RIVPACS results) was applied to the NorCal IBI, 15% of sites were classified as impaired. The BMI assessments were very similar to those of our assessments based on fish assemblages (Rehn and Ode 2005). Two potential explanations for the apparent discrepancy are that 1) this could be evidence that a multi-metric tool (B-IBI) is less sensitive to impairment than a taxon based tool like RIVPACS (perhaps due to the effect of species replacements), or 2) the discrepancy could be related to the way we defined reference thresholds. In a similar assessment, Herlihy and others (2005) reported impairment percentages for B-IBI scores in Oregon headwater streams that were similar to those reported here for RIVPACS scores (31% slight impairment, 6% severe impairment).

Next Steps:

CMAP: Improving Associations with NPS classes

The California Environmental Monitoring and Assessment Program (CMAP) is a collaboration between the State Water Board's SWAMP program and its NPS Monitoring program. The program is now collecting the second year of data of a five year continuation of the WEMAP effort. The sampling design is similar to that described here, except that it is stratified to include approximately equal representation of sites in landcover classes of Agriculture, Urban and Forest. Although the program will provide the state with the continued ability to produce yearly condition assessments (based on five year rolling averages), it will also allow the water board to make statements about how the condition of perennial streams relates to various landuse practices. The next steps in advancing this work involve improving the tools for associating landuse practices with biotic condition. Future efforts will be focused on associating timber harvest activities and agricultural practices with physical and biological condition of streams.

RIVPACS models

The RIVPACS models used here represent the state of the art for predictive modeling of BMI assemblages in California and they compare well with other regional RIVPACS models (Hawkins personal communication). However, there is room for potential improvement in the models with the addition of reference sites from underrepresented regions of California. Construction of new models with an expanded set of reference sites is currently in the planning stages.

Multiple Indicators/ Multiple Waterbodies

Statewide estimates of stream conditions (such as those made in support of annual 305(b) reports) should ultimately be based on a fully-integrated ecological assessment of multiple communities (BMIs, fish and periphyton) and assessments of physical and chemical conditions; the analyses presented here are one component of EMAP's survey of multiple biological indicators.

Techniques similar to those used here could be adapted for answering similar condition questions in other waterbodies in the state. These probabilistic designs have been successfully used on a wide variety of waterbodies (e.g., SCCWRP 1998, papers in Wright *et al.* 2000, Didonato *et al.* 2003, Hill *et al.* 2004, Richards *et al.* 2004) and predictive models have successfully been used as indicators for wetlands, lakes and estuarine habitats (Hawkins and Carlisle 2001). There is also a pressing need for similar work on non-perennial streams, which make up nearly 40 percent of the stream network in California.

Why probability surveys?

This report is intended to provide a demonstration of the use of statistical surveys to assess the ecological condition of streams and to associate stressors with ecological condition. The statistically defensible probabilistic survey approach provides an objective view of stream quality throughout California that can not be obtained by non-statistical methods. As such, it provides a critical tool for viewing the overall state of waterbodies. This perspective provides a logical foundation of water quality monitoring because it establishes the full range of biotic condition (i.e., how good is good and how bad is bad). Thus, it gives water quality managers a mechanism for determining where targeted sites fit on this scale and to answer questions related to the extent of impairments (e.g., "is a given site better or worse than most sites in my region", "are most of the biologically impaired sites in my region associated with physical or chemical impairment"). This also serves as a means of monitoring response to management (e.g., "am I putting money where the biggest problems are?", "am I protecting my best sites?"). These are questions that can not be answered without a statistical sampling design. The approach demonstrated here could easily be extended to smaller regions than were used here (e.g., regional water boards) and these designs could be nested into the ongoing statewide condition monitoring (CMAP) to provide a valuable framework for interpreting ongoing targeted monitoring data.

LITERATURE CITED

Bailey, C.R., R.H. Norris and T.B. Reynoldson. 2004. *Bioassessment of Freshwater Ecosystems: Using the Reference Condition Approach*. Kluwer, The Netherlands.

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

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Revision to rapid bioassessment protocols for use in stream and rivers: periphyton, BMIs and fish. EPA 841-D-97-002. U.S. Environmental Protection Agency. Washington DC.

Didonato, G.T, J.K. Summers, T.H. Roush. 2003. Assessing the ecological condition of a coastal plain watershed using a probabilistic survey design. Environmental Monitoring and Assessment 85(1):1-21.

Hawkins, C.P. and D. M. Carlisle. 2001. Use of predictive models for assessing the biological integrity of wetlands and other aquatic habitats. pp. 59-84 in: R. B. Rader, D.P. Batzer and S.A. Wissinger, eds. Bioassessment and Management of North American Freshwater Wetlands, John Wiley and Sons, New York.

Herlihy, A.T., W.J. Gerth, J.Li, J.L.Banks. 2005. Macroinvertebrate community response to natural and forest harvest gradients in western Oregon headwater streams. Freshwater Biology 50: 905-919.

Hill, B.H., D.W. Bolgrien, T.R. Angradi, and E.W. Schweiger. 2004. Design and indicator considerations for a probabilistic survey of U.S.A. great rivers: Missouri, Mississippi, Ohio Presented at the NABS Annual meeting, Vancouver, British Columbia.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6: 21- 27.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger and D.V. waters: quantifying physical habitat in wadeable streams. Research and Development. EPA/620/R-99/003.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-the Tennessee Valley. Ecological Applications 4: 768-785.

Norris, R. H. & A. Georges, 1993. Analysis and interpretation of benthic macroinvertebrate surveys. In Rosenberg, D. M. & V. H. Resh (eds), Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, New York: 234–286.

Ode, P.R., A.C. Rehn and J.T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. Environmental Management 35(4): 493-504.

Peck, D.V., J.M. Lazorchak, and D.J. Klemm (editors). 2004. Environmental Monitoring and Assessment Program -Surface Waters: Western Pilot Study Field Operations Manual for Wadeable Streams. EPA/XXX/X-XX/XXXX. U.S. Environmental Protection Agency, Washington, D.C.

Rehn, A.C. and P.R. Ode. 2004. Condition assessment of coastal streams in southern and central California. Report to the State Water Resources Control Board. California Department of Fish and Game Aquatic Bioassessment Laboratory, Rancho Cordova, California.

Rehn, A.C. and P.R. Ode. 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, California.

Rehn, A.C. and P.R. Ode. Comparison of reachwide and targeted riffle methods for bethic macroinvertebrates. *In prep.*

Richards, C., N. Danz, L.B. Johnson, T.P. Hollenhorst, V.J. Brady, D.H. Brenneman, J.J.H. Ciborowski, and T.R. Hrabik. 2004. Developing indicators for use in probabilistic survey designs in Great Lakes coastal ecosystems. Presented at the NABS Annual meeting, Vancouver, British Columbia, 2004.

Southern California Coastal Watershed Research Project. 1998. Southern Califonira bight 1994 pilot project. www.sccwrp.org

Tang, Z., B.A. Engel, B.C. Pijanowski and K.L. Kim. 2005. Forecasting land use change and its environmental impact at a watershed scale. Journal of Environmental Management 76: 35-45.

Van Sickle, J., C.P. Hawkins, D.P. Larsen, and A.T. Herlihy. 2005.A null model for the expected macroinvertebrate assemblage in streams. Journal of the North American Benthological Society: Vol. 24, No. 1, pp. 178–191.

Wright, J.F., D. Moss, P.D. Armitage and M.T. Furse. 1984 A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology, 14: 221-256.

Wright, J.F., P.D. Armitage, M.T. Furse and D. Moss. 1989. Prediction of invertebrate communities using stream measurements. Regulated Rivers: Research and Management 4: 147-155.

Wright, J.F., D.W. Sutcliffe and M.T. Furse. 2000 Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, June 2000.

Yoder, C.O. and E.T. Rankin. 1995. Biological criteria program development and implementation in Ohio. In: W.S. Davis and T.P. Simon (Eds.). Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, FL. 109-144.

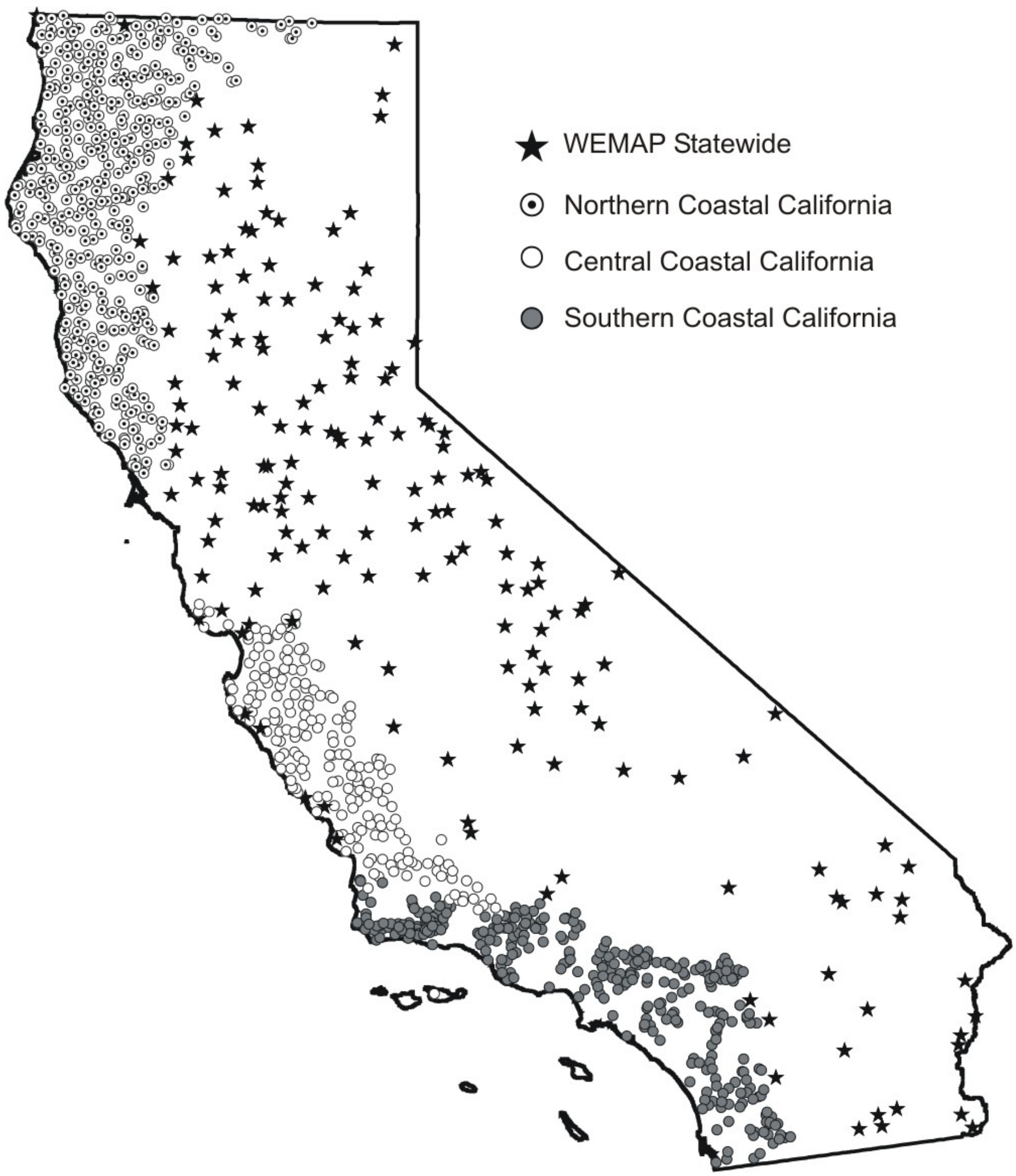

Figure 1. Distribution of all potential sampling locations under the four survey designs combined for this analysis.

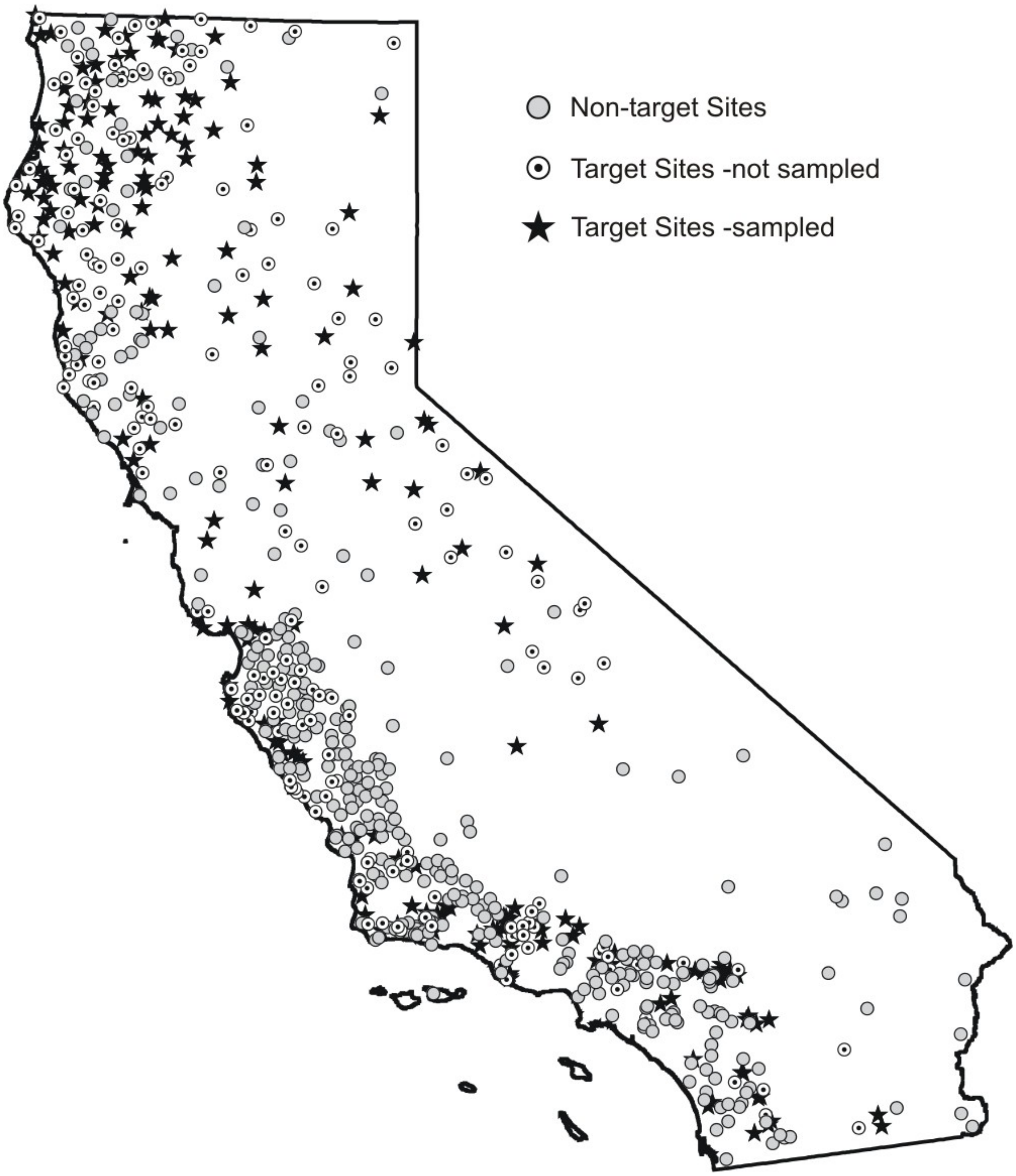

Figure 2. Distribution of sampling locations for 728 sites evaluated for potential inclusion in the project.

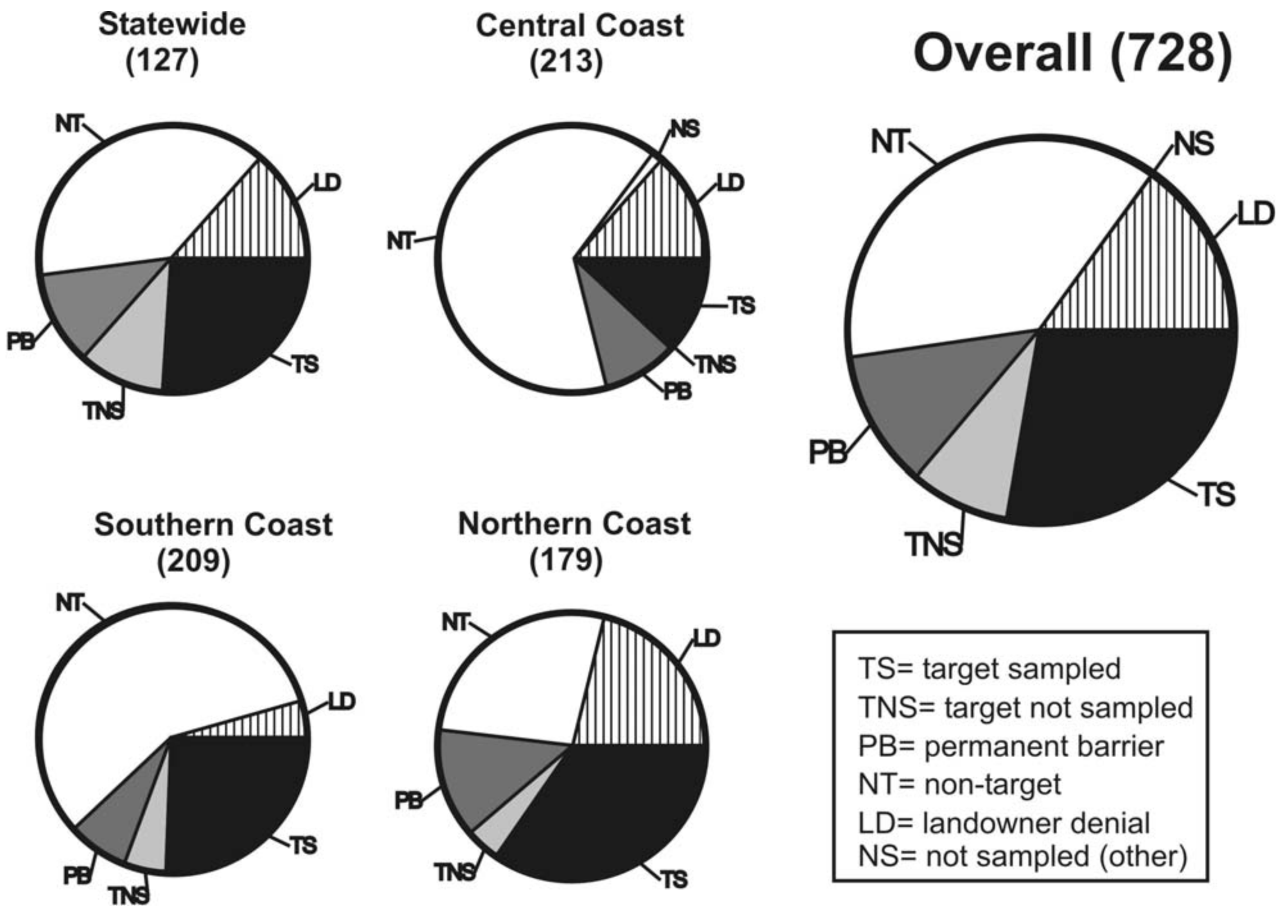

Figure 3. Survey fates of the 728 sites evaluated for the overall assessment with the site status percentages broken out for the four individual surveys.

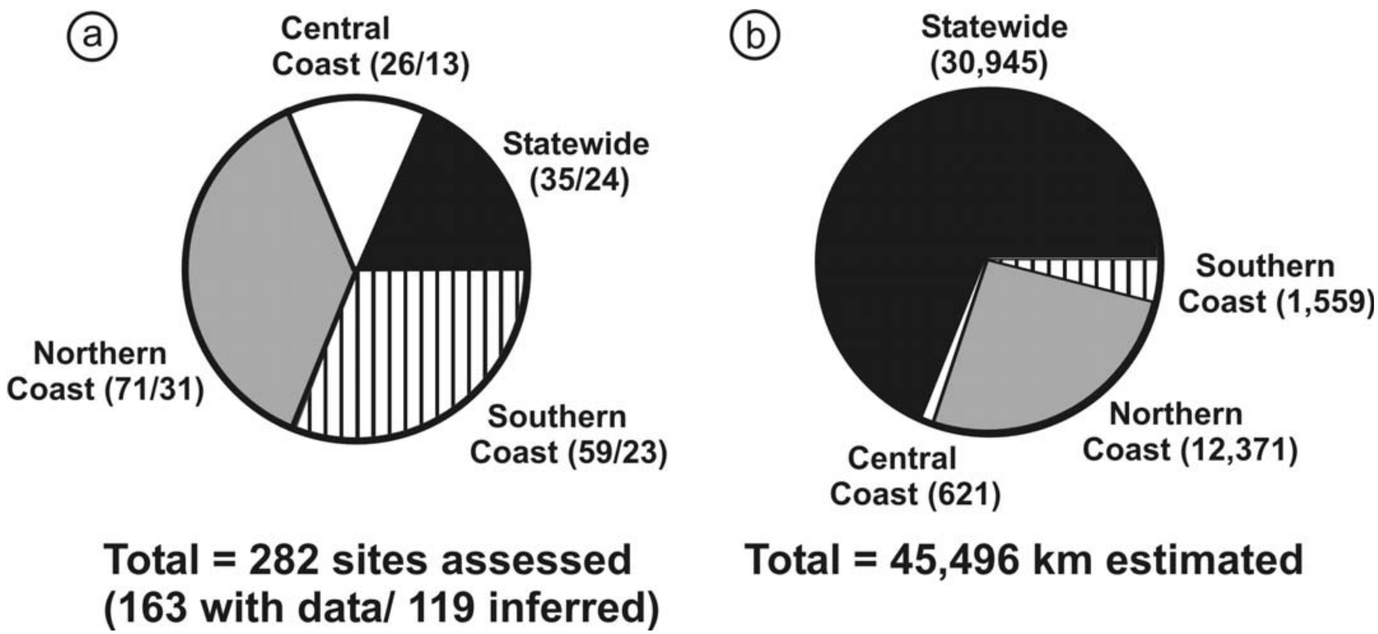

Figure 4. Contribution of the four component surveys to the statewide condition survey: a) number of sites used to make the assessment (numbers in parentheses indicated number of sites with RIVPACS data), b) stream length estimated

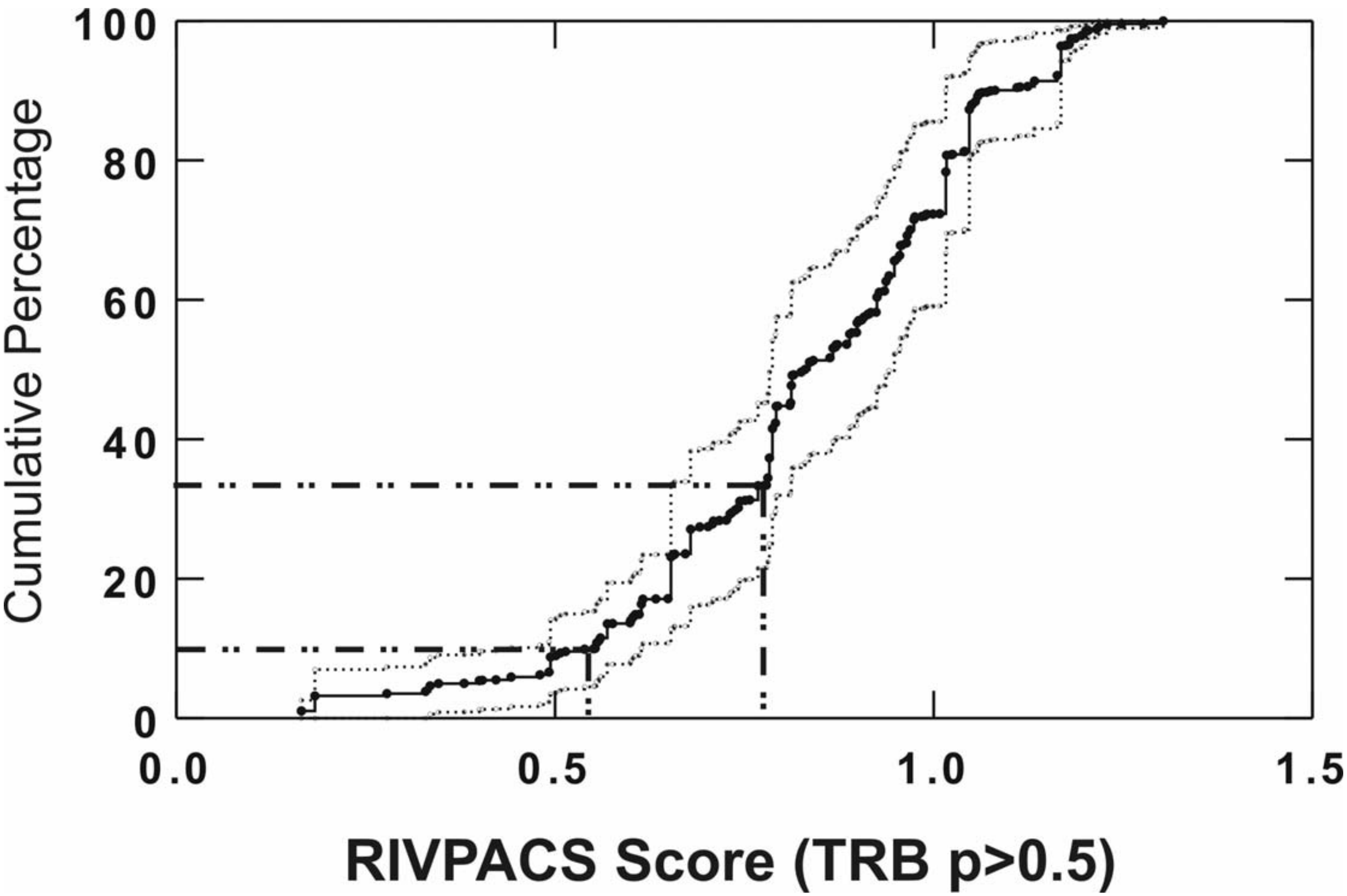

Figure 5. Cumulative distribution function of RIVPACS scores calculated from targeted riffle samples. Dotted lines represent upper and lower 95% confidence limits and bold dashed lines represent an impairment threshold based on 1.5 and 3.0 sd below the expected RIVPACS score of 1.0.

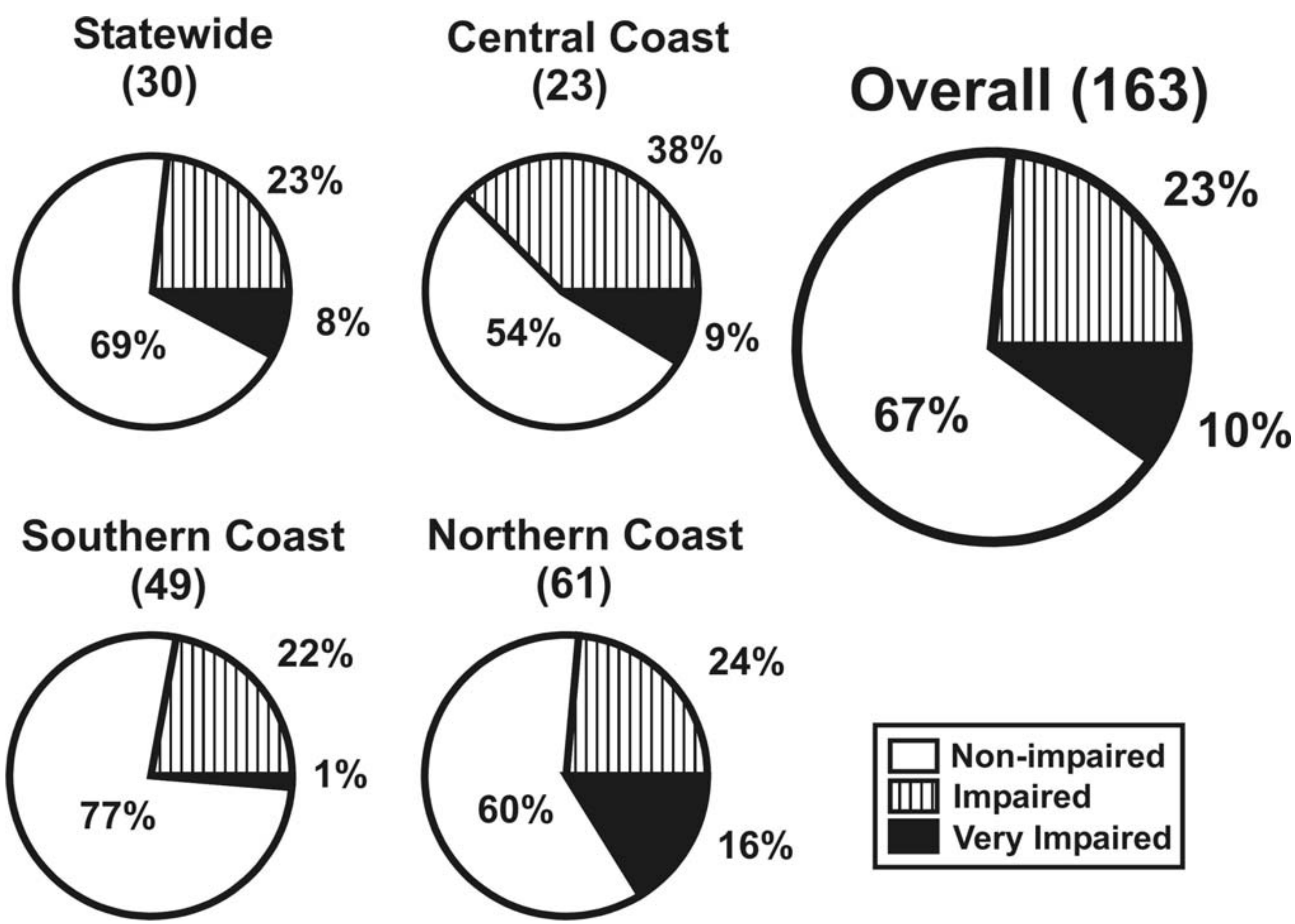

Figure 6. Proportion of stream length in non-impaired, impaired and very impaired biotic condition estimated from benthic macroinvertebrate RIVPACS scores. Condition categories were based on an impairment thresholds of 1.5 sd (Impaired <0.77) and 3.0 sd (Very Impaired <0.55) below a RIVPACS score of 1.0. Numbers in parentheses indicate number of sites used in estimates.

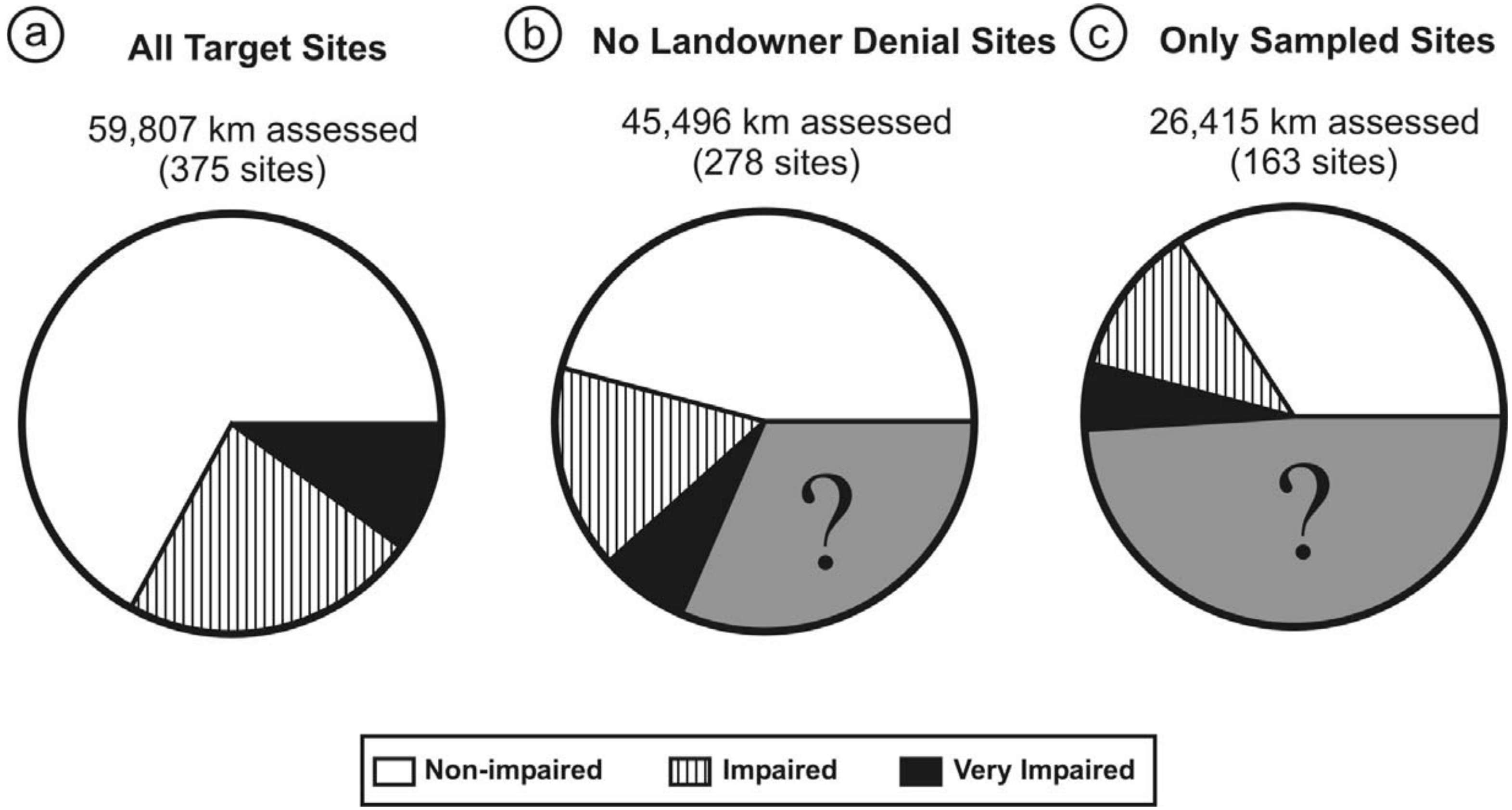

Figure 7. The effect of accepting different assumptions on the amount of stream length that can be extrapolated with this study.

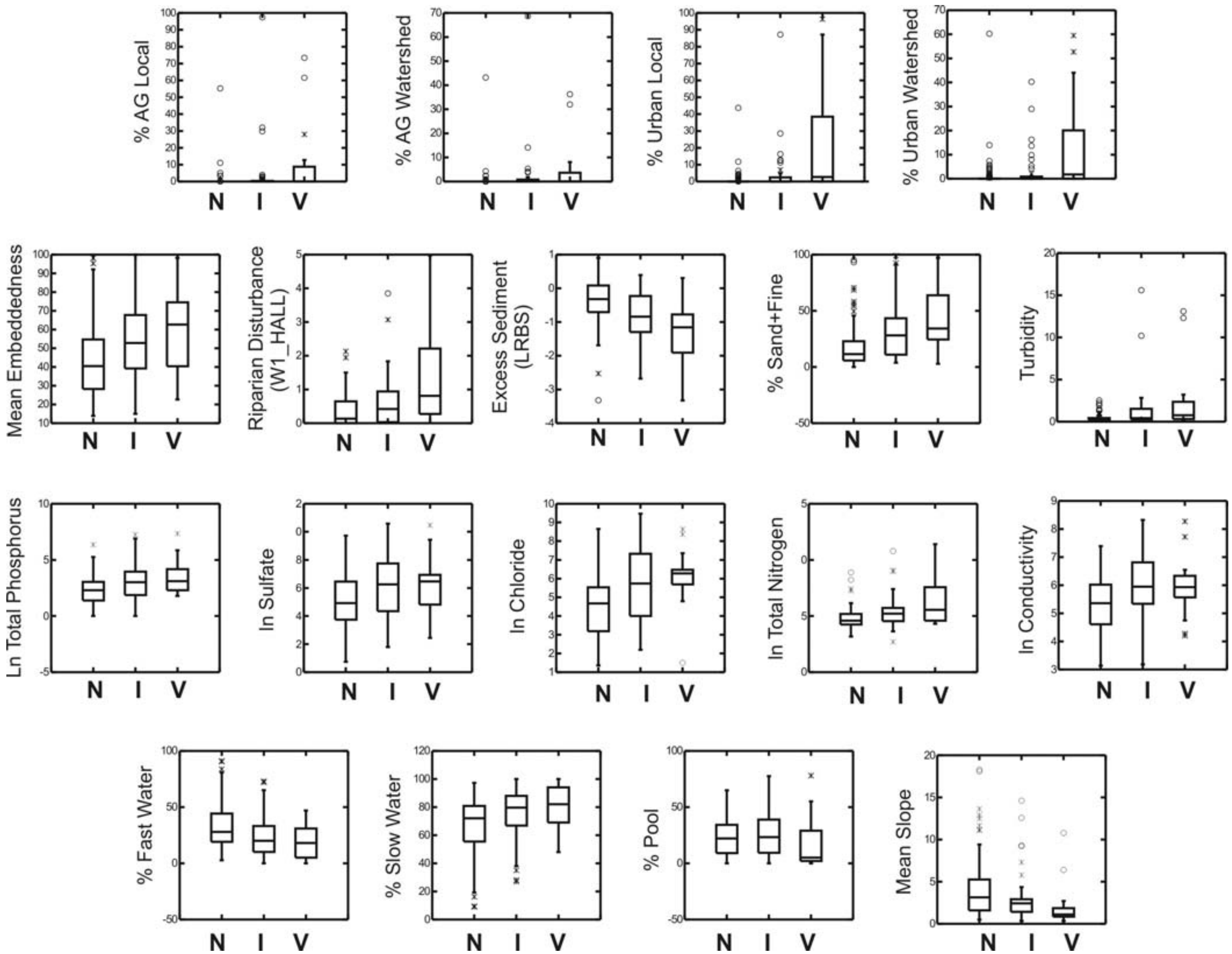

Figure 8. Boxplots of the distribution of site values (reach attributes or potential stressor values) for non-impaired (N), impaired (I) and very impaired (V) sites.

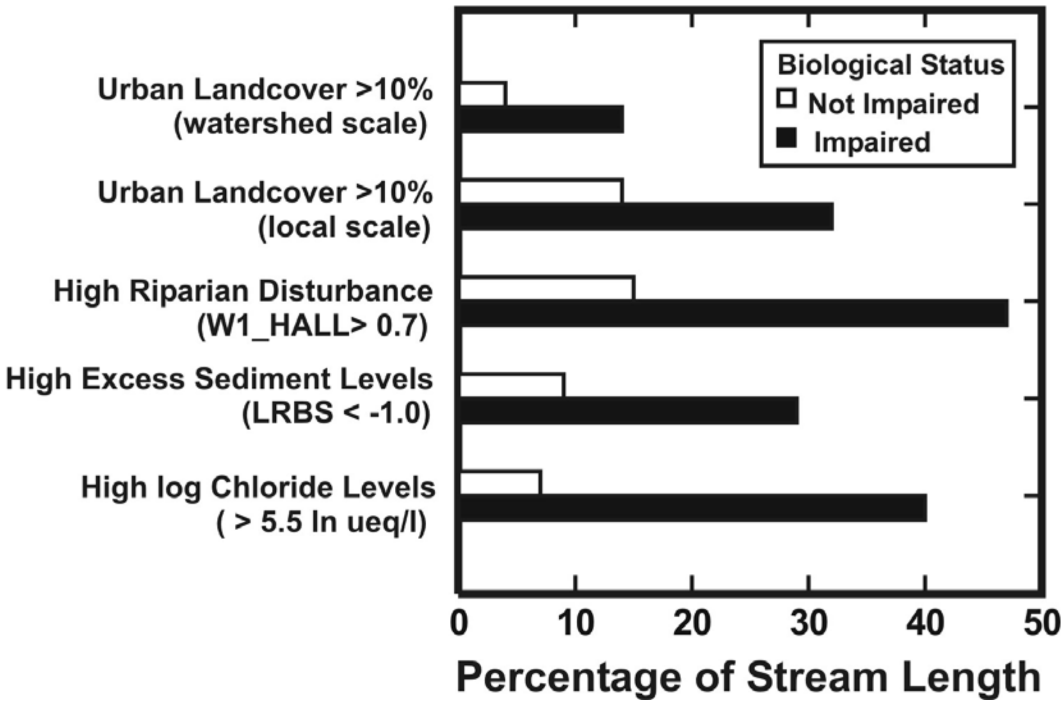

Figure 9. Stressor association histograms showing the percentage of impaired (includes very impaired) or non-impaired stream length associated with five stressors

Table 1. Stream condition estimates for the statewide assessment based on targeted riffle samples, with breakdown of condition by four sub-surveys.

Survey	Category	Sites	Percent	Lower 95% CL	Upper 95% CL	Stream Length (km)	Lower 95% CL	Upper 95% CL
Overall	Non_Impaired	102	67	55	78	30325	25179	35470
	Impaired	43	23	13	34	10666	6077	15255
	Very_Impaired	18	10	5	15	4500	2236	6764
	Total Scored	**163**	**100**			**45490**		
Statewide	Non_Impaired	20	69	52	86	21203	15952	26454
	Impaired	7	23	8	39	7158	2376	11940
	Very_Impaired	3	8	0	16	2421	41	4800
	Total Scored	**30**	**100**			**30782**		
North Coast	Non_Impaired	33	60	50	71	6886	5668	8103
	Impaired	16	24	16	31	2694	1849	3538
	Very_Impaired	12	16	9	23	1848	1086	2610
	Total Scored	**61**	**100**			**11427**		
Central Coast	Non_Impaired	13	54	35	73	715	460	969
	Impaired	8	38	19	56	499	249	750
	Very_Impaired	2	9	1	17	117	11	222
	Total Scored	**23**	**100**			**1331**		
South Coast	Non_Impaired	36	77	65	88	1494	1265	1723
	Impaired	12	22	14	31	432	266	597
	Very_Impaired	1	1	-1	3	26	-13	65
	Total Scored	**49**	**100**			**1951**		