I. Topic: Regional Strategy for Identifying Reference Conditions in Bay Area Streams

II. Background Information on Topic Area

This BAMBI Issue Paper addresses the topic of establishing reference conditions for biological assessment in streams of the San Francisco Bay Region. Reference conditions are the biological conditions expected to be found in streams exhibiting minimal human disturbance in a given region or stream class. Generally, minimal human disturbance is taken to equate to conditions prior to substantial European development (i.e. in California, conditions approximately 250 years ago). Although this standard is problematic because it disregards the fact that humans have been interacting with landscapes in the San Francisco Bay Area for thousands of years, the "human disturbance" criteria provides the best way to understand how humans have altered biological systems. Profound decreases in the biological integrity of water bodies have occurred in the past 100 years, especially as a result of urbanization and industrialization. Only by comparing biological communities from a gradient of altered habitats to those in minimally disturbed streams can we gain some appreciation for how past and present human practices have resulted in changes to biological integrity.

Why is it important?
Protecting and restoring biological integrity is a core value of the Clean Water Act. In order to monitor and improve biological water quality, standards and criteria need to be developed. Establishing reference conditions is an essential step towards developing biological water quality criteria (biocriteria), as reference conditions serve as the standard for detecting changes in biological integrity and beneficial use impairment.

What is the current status of the topic in the Bay Area?
No reference conditions have been developed for benthic macroinvertebrate communities in streams of the San Francisco Bay Region. Some potential reference sites have been identified, however. In the Regional Water Quality Control Board's (RWQCB) Surface Water Ambient Monitoring Program (SWAMP), potential reference sites were identified, when present, in each pilot watershed prior to biological sampling. Potential reference sites were evaluated based upon criteria including (1) upstream land use, (2) riparian and channel alteration, and (3) access. The applicability of these potential reference sites is further discussed in Section V. Other agencies conducting biomonitoring have compared biological assemblages at disturbed sites to conditions at relatively undisturbed sites. While reference sites were not necessarily designated, sites spanning a range of human disturbance were purposefully sampled to examine a range of biological condition. For example, in the San Mateo Countywide Stormwater Pollution Prevention Program's bioassessment study in the San Mateo Creek watershed (BioAssessment Services, 2002), biological assemblages at several less impacted sites had higher richness and diversity metrics and much lower tolerance than conditions at sites in more urbanized areas. The Contra Costa Clean Water Program selected sites from distinct land use classes. One upper watershed site (FC-2), representing open space land use, exhibited higher metric values than all others sites. In both cases, sites with little or no upstream human land use exhibited the highest biological metrics. These sites were taken to be examples of relatively undisturbed conditions against which impacted sites could be compared.

Establishing reference conditions

Four approaches are commonly used to establish reference conditions in freshwater streams (Gibson et al): (1) historical data, (2) regional reference sites, (3) simulation models, and (4) best professional judgment. Although historical records of biological condition are ideal for tracking changes, early biological surveys used different sampling methods than those used today. Historical data on fish abundances have been used to develop reference conditions in the eastern and mid-western United States, but very little if any historical benthic macroinvertebrate data exists for the San Francisco Bay Region. Regional reference sites (Hughes et al, 1986) are minimally impaired sites that collectively characterize a region's expected biological condition. Biological assemblages in test sites are then compared to measured data from representative reference sites. When historic or existing reference biological assemblages cannot be determined, simulation models or best professional judgment is used to establish reference conditions from first principles or extrapolations from other regions. Although any or all of these four methods can be used, the preferred approach is to develop regional reference sites, which represent achievable goals and can be regularly monitored to develop biocriteria (Gibson et al).

III. Objectives of Paper:

The goal of this paper is to provide context for and stimulate discussion of issues relating to the establishment of reference conditions in the Bay Area. I have identified three important issues that are vital to the development of reference conditions using regional reference sites. Although I originally was just going to discuss issues 1 and 2, issue 3 (classification) proved to be vitally important when considering the reference condition question. Based upon background research and analysis of bioassessment data in the region, I also propose solutions to these issues, outlined below and discussed in depth in this paper.

1. Do sufficient minimally disturbed sites exist in the region in order to use regional reference sites to develop biocriteria?

Conclusion: Many minimally disturbed sites exist, but the vast majority of these sites come from low-to-mid order mountainous, forested streams. These sites are not representative of all streams in the region, however, and separate biocriteria are needed for other stream types. To establish interim expectations of biological integrity, the upper tail tendency of the best available sites from other stream types will have to be used. Future monitoring efforts must identify and sample best available sites in order to further the development of realistic biocriteria for areas where minimally disturbed sites are not available.

2. What methods best identify regional reference sites and rank human disturbance?

Conclusion: Candidate reference sites can be identified using GIS techniques and/or other land use data. Although water quality data and detailed habitat assessments are useful for predicting biological water quality impairment, the time and resource requirements of these methods are enormous. Rapid field reconnaissance of candidate sites should be performed to assess channel and riparian disturbance, local land use, and historical

disturbance. These factors, along with upstream land use, should be included in a final ranking of human disturbance to identify minimally disturbed and best available sites.

3. How should streams be classified into different types, each with separate reference conditions and biocriteria?
Conclusion: Classification of stream types by climate, vegetation community, channel gradient, or confinement may be necessary to distinguish between expected biological condition. Biological communities in potential reference sites suggest that major differences in metric scores, especially taxonomic diversity, exist between mountainous forested watersheds and Mediterranean grassland/chaparral watersheds. Less variation is observed between sites with different drainage area or gradient.

IV. Current Issues

Preliminary resource assessment to determine feasibility of using regional reference sites
Do minimally disturbed sites exist in the region? Reference sites must be minimally disturbed by past and present human action. If sufficient minimally disturbed sites are identified and monitored, the observed range of biological metrics in reference sites can be directly used as reference conditions for the stream type of interest. If sites are more than minimally disturbed, the best available sites can be used to estimate the upper tail tendency of the biological metric values. The upper tail tendency of the data is then used as an interim expectation of biological integrity in biocriteria. The availability of minimally disturbed sites depends in part on the classification system used. In the San Francisco Bay Area, human land use is concentrated along the bay margin and flatlands, where practically no minimally disturbed streams exist. If low gradient, bayside streams are classified as a distinct group of sites, moderately disturbed sites must be used to develop interim best available conditions for biocriteria.

Methods for the selection of regional reference sites
Regional reference sites are "minimally disturbed sites" that are representative of the biological communities expected to be found at other similar sites if no impairment exists. To identify minimally disturbed sites, methods should be chosen that best rank levels of human disturbance. In addition to minimally disturbed sites, "best available" sites need to be identified and monitored. Therefore, a quantitative ranking system needs to be used to judge human disturbance. The essential qualities of an effective ranking system include (1) an efficient estimation of human disturbance as it relates to biological integrity, (2) little or no additional data collection requirements, and (3) applicability towards sites that have already been monitored in addition to potential future sites. In this section I review potential methods for ranking human land use disturbance.

1. Best Professional Judgment
Traditionally, reference sites were chosen by knowledgeable field biologists familiar with the watersheds and streams of a region. Reference sites were selected based upon previous experience and expectations for how undisturbed streams should look. Recent efforts have focused on selecting reference sites in a more objective, quantified manner.

2. GIS Land Use Analysis

David Herbst (UCSB/SNARL) and Peter Ode (CDFG-ABL) developed a quantitative method to select regional reference sites that is being used in the Sierra foothills and northern California Coast (Ode and Herbst, 2001). This process consists of (1) defining the region of interest and a set of stream types to be evaluated (classification), (2) identifying major land use disturbances using GIS, (3) ranking and identifying least-disturbed candidate areas or sites, and (4) ground-truthing candidate sites to define refined list of reference sites. Once candidate reference sites are identified, sites are sampled and biological metrics are chosen to use in an IBI for each stream type. The GIS Analysis was performed with ArcView using Analytical Tools Interface for Landscape Assessments (ATtILA). The data layers required include elevation, land use, road networks, stream networks, and population, all of which are available for the San Francisco Bay Region. I performed a similar land use analysis on bioassessment sites from SWAMP and the Alameda and Contra Costa Stormwater agencies with the help of Jeff Kapellas, GIS Specialist at the RWQCB. Because of time constraints, however, I only analyzed land use data such as urban area and agricultural area. Other variables, especially road density, may ultimately prove to be more effective in predicting human disturbance. The percentage of upstream urban area was a decent predictor of biological condition (inferred from a relative ranking of 10 biological metrics) in urbanized watersheds of the San Francisco Bay Area (Figure 1). Sites with greater than 40% urban land use upstream all had very low biological metric scores. Several sites with very little upstream urban development also scored low on the biological index, however, indicating that habitat condition or localized disturbances are also important indicators of biological condition. Moreover, a few sites with significant urban development (10-30%) had moderately high biological scores.

3. Biological data

Biological assemblages and metrics serve as the basis for judging the health and biological integrity of streams. However, the purpose of the regional reference site process is to compare conditions in minimally disturbed (control) sites and disturbed (test) sites. When sites are chosen on the basis of biological scores, no underlying relationship between human disturbance and biological integrity is developed. In addition, biological attributes cannot be used to identify sites prior to sampling. Given these shortcomings, biological data is extremely useful for confirming minimally disturbed, high quality biological assemblages.

4. Chemical water quality

Water quality data such as toxicity, nutrients, metals, organics, and basic parameters are essential indicators of the health of water bodies and aquatic communities. Along with habitat quality, water quality is one of the most important factors affecting biological integrity. When chemical water quality data is available for streams or sites of interest it can be very useful for judging impairment. This type of data is not available for the majority of bioassessment sites, however. One of the benefits of aquatic bioassessment is to provide an quick, integrated assessment of water quality, so that numerous expensive chemical water quality tests do not need to be completed. Requiring water quality data to assess biological data subverts the convenience of aquatic bioassessment.

5. Integrated Index

Don Huggins and Michael Johnson (UCD) developed a watershed index that includes both process (causal) and outcome (bio-metrics) factors to determine reference sites (Huggins and Johnson, 2002). Eleven measured 'core factors', including physical habitat, land use, land cover, water chemistry, pollution sources, altered hydrologic regime, biological metrics, and biological assemblages, were used to create a watershed index ranking of human disturbance and habitat quality. Integrated indices such as this one provide a comprehensive rating of ecological health, but require large amounts of data that is generally not available in most watersheds.

6. Urban Streams Habitat Index

Jim Carter and Steve Fend (USGS) developed a habitat index for urban streams designed to quickly categorize habitat condition in a standardized manner. It consists of 8 measured or visually estimated scores, including bank and channel modification, riparian vegetation composition, riparian corridor width, canopy cover, embeddedness, sediment siltation, turbidity, and land use, which are given scores between 1 and 4. Although not specifically designed to assess human disturbance or identify reference conditions, many of the measurements are useful indicators of channel and habitat disturbance.

7. Rapid Human Disturbance Index

I developed the Rapid Human Disturbance Index (RHDI) as a quick and simple index that integrates the major human disturbance factors in the San Francisco Bay Region affecting stream benthic macroinvertebrate community integrity. It was developed to meet a perceived need for a method that requires little or no additional data collection, but instead relies upon information that can be obtained from maps and rapid field reconnaissance. It is loosely based upon the Carter/Fend Urban Streams Index and the Huggins/Johnson Watershed Index discussed above. RHDI is a semi-quantitative ranking of watershed and local land use, channel disturbance, riparian disturbance, and water quality. These five metrics are ranked on a 1-8 scale, and points are subtracted from the total for point sources of pollution and hydrologic alterations. The scoring system is shown in Figure 2. As an initial test of the index, I scored sites from the Stevens Creek and Permanente Creek watersheds in the Santa Clara Basin, and compared the scores to biological metric relative ranking scores[1] (Figure 3). RHDI effectively incorporates several key types of human disturbance in the form of one simple to use index. The scoring system may prove too subjective for the tastes of some users, however. In the future I would like to test and incorporate a road density metric to provide an additional indicator of human disturbance.

Stream classification for the development of biocriteria

Reference sites must be representative of the streams they are being compared to. Although no two streams have exactly the same characteristics, the biological condition of streams in watersheds with comparable climate, topography, and terrestrial biology tend to be more similar than streams draining watersheds from different regions (Hughes et al, 1986). A successful classification system acknowledges regional variability, but uses the fewest number of classes to effectively predict conditions. Numerous distinct classes, each with its own IBI and biocriteria, require monitoring an immense number of reference sites and test sites. In the San Francisco

[1] Relative ranking score is the site metric score minus the mean metric score divided by the standard error of the mean for five biological metrics.

Bay region, the question boils down to 'Should one set of reference conditions and biocriteria apply to the entire region, or are subregions diverse enough to warrant separate standards?'

Classification can either be done prior to or subsequent to the selection of candidate reference sites. In *a priori* classification, sites are grouped by channel and landscape characteristics that are believed to significantly affect biological communities in natural systems. Variables such as stream order, gradient, ecoregion, and climate are commonly used to classify sites. *A posteriori* classification attempts to identify variables that result in variation in biological assemblages within a defined population of reference sites. Until minimally disturbed sites are available representing all stream types in the region, it is difficult to perform complete *a posteriori* classification using biological data, because it is impossible to distinguish between natural and human-caused variation in biological signals. Given a large dataset of reference sites, however, some *a posteriori* classification should be possible. Both *a priori* and *a posteriori* classification can be used in order to optimize the effectiveness of the classification scheme.

Analytical approaches

Reference sites are used to establish biocriteria using either multimetric or multivariate methods. Multimetric approaches utilize biological metrics to compare reference sites and test sites, while multivariate approaches compare taxonomic data in reference and test sites based upon probabilities of predicted taxa occurrences (Reynoldson et al, 1997). Although not discussed in this paper, an important consideration for future monitoring is the analytical approach that will be used. In general, multivariate approaches require environmental measurements (such as channel dimensions, substrate, and water quality data) that are not available for the majority of sites already sampled in the San Francisco Bay Region. In this paper, reference conditions are discussed primarily within the context of biological metrics.

V. Options for improvement

Preliminary resource assessment to determine feasibility of using regional reference sites

Do minimally disturbed sites exist in the region? Based upon several years of bioassessment data collected by SWAMP, there are many minimally disturbed sites that exhibit high biological integrity. Prior to sampling, RWQCB staff identified potential reference sites. These sites, from undisturbed and minimally disturbed areas, exhibit high metric values compared to non-reference sites (Figure 4). The clustering of the data suggests that these sites are representative of minimally disturbed reference conditions. Nearly all of the identified reference sites, however, are located in mountainous, redwood and hardwood forested areas. Until minimally disturbed sites from other sub-regions are identified and monitored, best available conditions must be used as an interim expectation for these stream classes.

Methods for the selection of regional reference sites

A unified method for selecting reference sites and ranking best available conditions is needed to develop reference conditions in the Bay Area. In accordance with the Ode/Herbst method discussed above, it is recommended that the method utilize both (1) GIS/topo maps to rank potential reference sites according to local and upstream land use, dams or diversions, and (2)

site visits to assess physical habitat and local conditions. Based upon the analysis of methods discussed above, I reached the following conclusions:

1. A completely integrative method that incorporates water quality, physical habitat, and land use is ideal. To quantitatively measure and rank all of these factors is difficult, time-consuming, and costly, however.
2. Land use analysis can distinguish between levels of human disturbance.
3. Biological values are the best indicators of biological integrity, but cannot be obtained until after sampling, and provide no link to human disturbance.
4. Simple ranking methods based upon both BPJ and quantitative data can be extremely effective at predicting human disturbance, and can identify disturbance that other methods may miss. For example, the "Gibson Score", essentially a BPJ-based ranking of human disturbance, was effectively used in the development of an IBI for San Diego. Also, the Ode/Herbst method, although utilizing GIS land use data to determine candidate areas, relies upon BPJ to refine the list of potential sites.

I recommend a five step process for identifying reference sites and ranking human disturbance, modified after the Ode/Herbst method discussed above:

1. Sites are nominated by workers in the region
2. Access and permission is confirmed
3. Human disturbance ranking is obtained from both (1) land use data and (2) site conditions observed during field recon
4. If the site scores above a threshold for its stream class, it is sampled by CSBP protocol.
5. If biological data confirms that the site is minimally disturbed or best available conditions, as compared to existing data, the site is officially designated "reference site" or "best available conditions."

<u>Stream classification for the development of biocriteria</u>

Attempts at *a posteriori* classification with existing data suggest that there is little variation in expected condition between headwater (1st order) and mid-gradient (4th-5th order) streams in the coastal mountains of Marin and San Mateo Counties. There appear to be significant differences in biological metrics between minimally disturbed forested sites from coastal (Marin and San Mateo) and inland (East Bay Hills) mountainous sites, however, suggesting that climatic differences are important. Also, minimally disturbed grassland sites exhibited lower metric values than sites in nearby forested areas. Additional *a posteriori* classification is limited by the lack of minimally disturbed sites elsewhere and the strong correlation between urban development and hillslope gradient in the Bay Area.

The most common *a priori* approach to stream classification for the development of biocriteria is the ecoregion approach (Hughes et al, 1986). I propose an initial ecoregion classification of the region based largely upon vegetation community and climate. Significant differences in stream morphology, hydrology, and climate exist between mixed conifer/hardwood forests, typical of the northern California Coast, and Mediterranean oak grassland/chaparral areas typical of much of central and southern California. In the San Francisco Bay Area, both of these vegetation communities exist and intermingle. Coastal mixed conifer/hardwood forests are common in the Coast Ranges of Marin County, the Santa Cruz Mountains of San Mateo County, and in the

Oakland-Berkeley Hills. Elsewhere, vegetation is dominated by oak grasslands and chaparral. In addition to vegetation community, regional climatic differences play an important role in affecting biological condition. Both east-west and north-south gradients in precipitation and temperature result in vastly different environments around the region. I have identified three potential stream types in the region:

1. Coastal forested
2. Grassland/chaparral
3. Inland forested

Existing data for minimally disturbed sites from the coastal forested class is shown in Figure 5. Because numerous preliminary reference sites have been sampled from the coastal mountain stream type, future efforts should focus on identifying best available conditions from the rest of the region, representing Mediterranean conditions. Spring 2003 SWAMP sampling in the Mt. Diablo Creek and upper San Mateo Creek watersheds should provide further insight into conditions in drier non-forested areas.

VI. Next Steps

Although the ultimate goal for many of us is to develop biocriteria for the San Francisco Bay Region in the most scientifically rigorous and defensible way possible, the limited financial and human resources available to complete this task must be acknowledged. I believe it is possible to develop reference conditions, and eventually, biocriteria, in a timely, effective, and rigorous manner. However, we must pool our limited resources in an effective manner, focusing on (1) correctly defining reference conditions and (2) sharing ideas and knowledge. Several key issues may facilitate the process of biocriteria development.

1. Adopt a method for identifying reference sites and ranking human disturbance.
2. Adopt a preliminary classification system for Bay Area streams.
3. Establish a committee for nominating reference sites to be sampled. Develop a priority list for monitoring of reference sites to fill data gaps. Pool resources for BAMBI funded monitoring.
4. Obtain funding for BAMBI projects, including a Bay Area bioassessment coordinator and a central location for a bioassessment database.
5. Obtain data from organizations outside the Bay Area to help with establishing reference conditions.
    a. North Coast RWQCB
    b. Central Coast RWQCB

VII. Figures

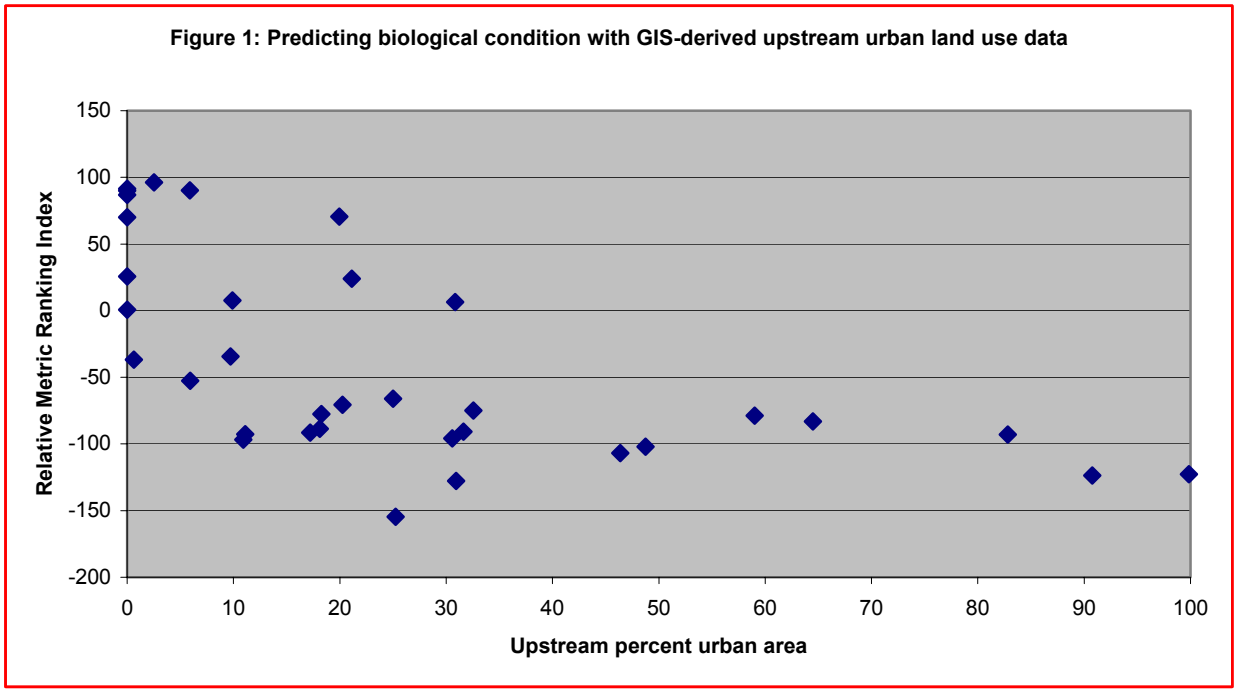
Figure 1: Predicting biological condition with GIS-derived upstream urban land use data
Relative Metric Ranking Index
150
100
50
0
-50
-100
-150
-200
0
10
20
30
40
50
60
70
80
90
100
Upstream percent urban area

Figure 2: Rapid Human Disturbance Index

	Local land use	Upstream land use	Channel	Riparian	Visual Water Quality
1	Urban industrial, or open mines or quarries	>80% human land use (urban + agriculture)	Concrete trapezoidal or "v" shaped channel	None: concrete, gravel, barren or tilled earth	Extremely toxic (no life observed)
2	Urban commercial, or golf course, or row crops	>50% human land use	Constructed ditch, dirt channel	Lawn, or heavily grazed with little remaining vegetation	Toxic, collect sample with ten-foot pole
3	Urban or suburban residential, or rural w/ confined animals	>25% human land use	Concrete banks, natural bed; or bed and banks heavily trampled by cows	Native vegetation removed: sparse planted trees and shrubs	Toxic, wear gloves to collect samples
4	Light residential or rural residential, or urban park	>10% human land use, or heavy grazing, or heavy logging	Rip-rap banks, natural bed	Landscaped residential garden	Wash hands after water contact
5	Rural, or heavy grazing, or heavy logging	>3% human land use, or light grazing or logging	One bank reinforced, otherwise natural; or natural with culverts at road crossings	Mostly undisturbed, but non-natives present (ivy, etc.)	Edible fish or shellfish (given presence)
6	Rural, light grazing	any human land use, or significant historical land use	Mostly natural, but bridge abutments or other structures	Native, <10 feet wide, or compromised by dirt or paved roads	Swimable or wadeable, given deep enough water
7	Protected parkland with minor historical land use	Protected parkland with minor historical land use	Natural, historical geomorphic disturbance	Native, <25 feet wide, or altered by trails	Drinkable with filter or tablets
8	Protected, undisturbed parkland	Protected, undisturbed parkland	Undisturbed	Native, undisturbed	Drinkable

Subtract one point for each (1) golfcourse, (2) open mine, (3) dam, (4) diversion, (5) wastewater treatment discharge, (6) pollution point source and (7) quarry. If within one stream mile of site, subtract two points.

Figure 3: Rapid human disturbance index (RHDI) and biological metric relative ranking

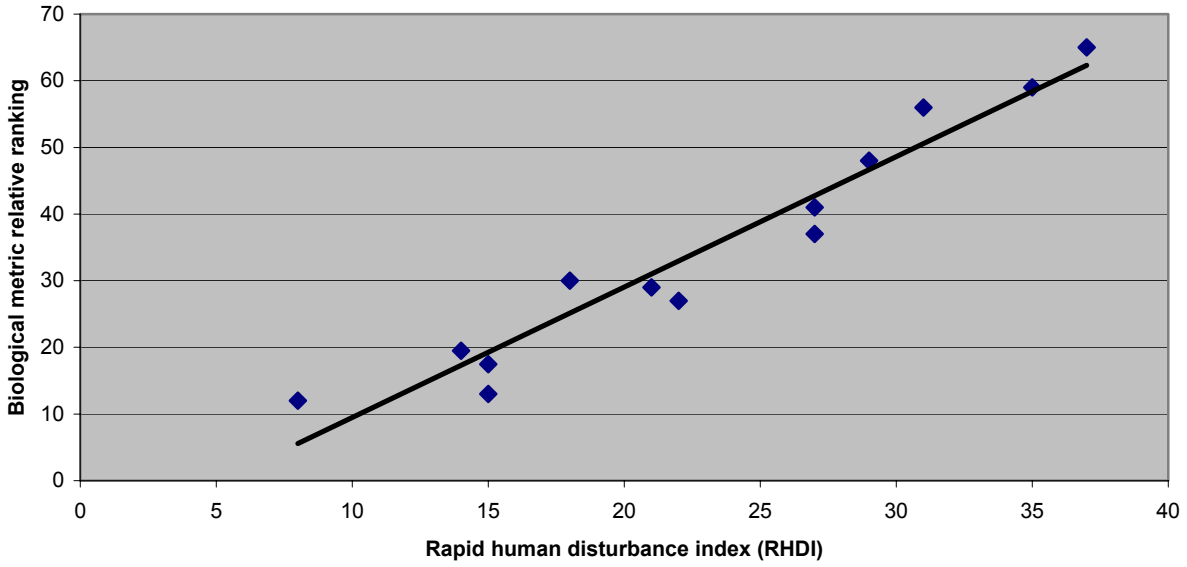

Figure 4: Biological metrics at sites ranked by human disturbance (BPJ Method)

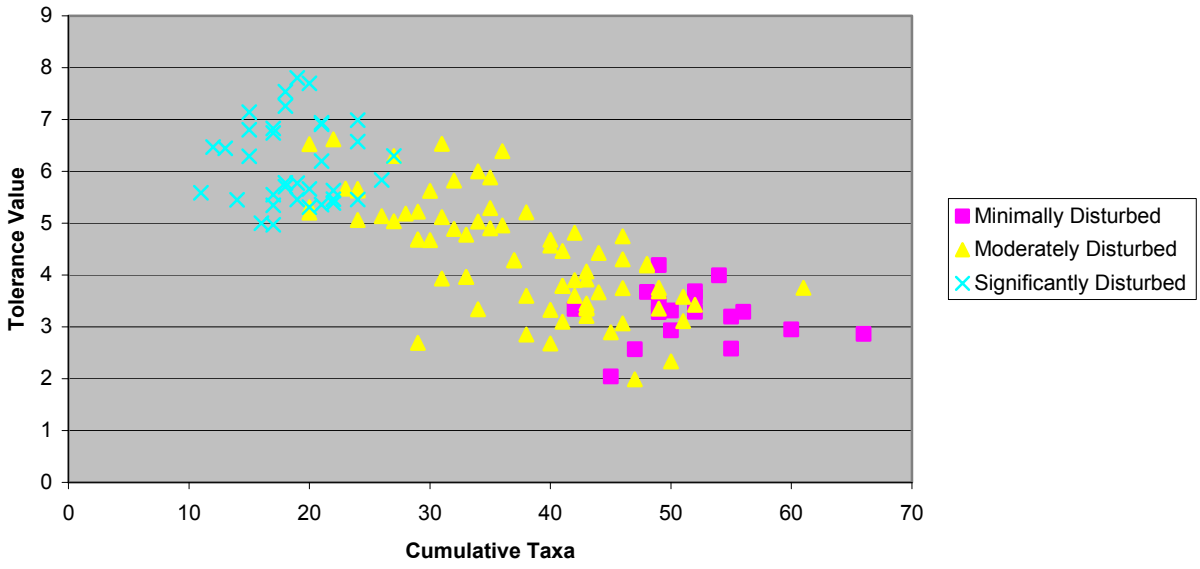

Figure 5: Biological metrics for preliminary reference sites for coastal forested streams

		Cumulative Taxa	Percent Dominant Taxa	Sensitive EPT Index	Cumulative EPT Taxa	Shannon Diversity	Tolerance Value	Percent Intolerant	Percent Tolerant
100%	Maximum	66.0	36.7	67.0	32.0	3.3	4.2	68.2	11.9
75%	Quartile	55.0	22.6	44.9	27.3	3.1	3.6	45.6	1.4
50%	**Median**	**51.0**	**18.4**	**33.9**	**25.0**	**2.9**	**3.3**	**35.7**	**1.1**
25%	Quartile	48.8	15.6	28.8	22.0	2.7	2.9	30.4	0.4
0%	Minimum	42.0	13.4	12.0	17.0	2.5	2.0	17.1	0.0
	Mean	51.7	20.3	36.1	24.7	2.9	3.2	37.9	1.5
	Std Dev	5.5	6.9	13.2	3.7	0.2	0.5	12.1	2.6
	Std Err Mean	1.3	1.6	3.1	0.9	0.1	0.1	2.8	0.6
	N	18	18	18	18	18	18	18	18

VIII. References

BioAssessment Services, 2002, Benthic Macroinvertebrate Assemblages of San Pedro Creek, San Mateo County, Folsom, CA.

Gibson, G.R., Barbour, M.T., Stribling, J.B., Gerrirsen, J., and Karr, J.R., Date Unknown, Biological Criteria: Technical guidance for streams and small rivers, revised edition,

Huggins, D., and Johnson, M., 2002, Reference Conditions and Metrics for Bioassessment, presentation given at the Bay Area Aquatic Bioassessment Information Network meeting, February 2002.

Hughes, R.M., Larsen, D.P, and Omernik, J.M, 1986, Regional Reference Sites: a Method for Assessing Stream Potentials, Environmental Management 10, 629-635.

Ode, P., and Herbst, D., 2001, A quantitative framework for reference sites selection, presentation given at the California Aquatic Bioassessment Workgroup meeting, November 2001, Sacramento, CA.

Reynoldson, T.B., Norris, R.H., Resh, V.H., Day, K.E., and Rosenberg, D.M., 1997, The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates, J. N. Am. Benthol. Soc. 16, 833-852.

The Bay Area Macroinvertebrate Bioassessment Information Network (BAMBI) includes scientists, watershed managers and regulators interested in local applications of *bioassessment* --the use of biological community data for assessing the condition of waterbodies and watersheds. BAMBI's focus is on *benthic macroinvertebrates* (bottom-dwelling animals without backbones, visible to the naked eye) which are present in most aquatic environments and are useful indicators of ecosystem function because their community composition responds to a wide range of ecosystem variables. BAMBI Issue Papers provide background and discussion of technical areas important to the development and improvement of bioassessment in the Bay Area. These are provisional workproducts for BAMBI discussion in January 2003, contributed by members of the Bay Area Stormwater Management Agencies Association (BASMAA), and the SFBay RWQCB.
Forward comments or questions to BAMBI c/o watersheds@acpwa.org