BASMAA Regional Monitoring Coalition

Regional Creek Status Monitoring

Draft Design and Sample Frame

Prepared for:

The Bay Area Stormwater Management Agencies Association

Prepared by:

EOA, Inc.
1410 Jackson St.
Oakland, CA 94612

March 28, 2011

TABLE OF CONTENTS

LIST OF TABLES & FIGURES

1. INTRODUCTION

The San Francisco Bay Regional Water Quality Control Board (Water Board) issued the Bay Area Municipal Regional Stormwater NPDES Permit (MRP) to 76 cities, counties and flood control districts (i.e., Permittees) on October 14, 2009. MRP provision C.8 requires Permittees to conduct water quality monitoring activities, including Creek Status Monitoring (C.8.c) and Pollutants of Concern and Long Term Trends Monitoring (C.8.e). The Bay Area Stormwater Management Agencies Association (BASMAA) has formed a Regional Monitoring Coalition (RMC) to coordinate and oversee regional monitoring efforts associated with the MRP. Representatives to the RMC have developed a Multi-Year Work Plan that identifies projects to assist Permittees in complying with MRP monitoring requirements.

The RMC has agreed to develop a probabilistic ambient survey design for regional creek status monitoring required under MRP Provision C.8.c. The ambient survey design includes the following tasks:

1. Confirm Management Questions and Scope for Regional Creek Monitoring
2. Establish Creek Status and Long-term Trends Sampling Frame Parameters
3. Acquire Applicable Data and Information and Develop GIS-based Sample Frame
4. Develop Master Sample
5. Develop Creek Status Sample Draw and Long-Term Monitoring Sites
6. Site Reconnaissance Training and Evaluation
7. Creek Status and Long-Term Trends Monitoring Design Report
8. Project Management

This memorandum describes the evaluation results of data layers and design factors that were considered during the selection of the sampling frame for the probabilistic ambient survey. Additionally, the memorandum documents the construct of the selected sample frame and serves as the deliverable for Tasks 2 and 3.

1.1. Applicability to MRP and Coordination with SWAMP

MRP Provision C.8.c requires Permittees to conduct creek status monitoring activities annually at a minimum number of sites (see Table 8.1 of the MRP). Of these activities, the following parameters will be included in the probabilistic design:

- Benthic Macroinvertebrate Bioassessment;
- Algae Bioassessment;
- Quantitative Physical Habitat Assessment;
- Nutrients/Cations/Anions/SSC; and,
- Chlorine.

On a county-by-county basis, Permittees are required to collectively monitor a minimum number of sites for these parameters to assess creek status (Table 1). The term creek status as

associated with these parameters is assumed to be synonymous with the condition of aquatic life.

In addition to sites associated with MRP requirements, the Water Board has also agreed to contribute sites to the probabilistic design via their regional Surface Water Ambient Monitoring Program (SWAMP).

Table 1. Annual creek status monitoring sites that will be monitored by RMC participants or the San Francisco Bay Regional Water Quality Control Board.

Permittees/Program	Annual # of Sites Committed to Probabilistic Design
Alameda Countywide	20
Santa Clara Countywide	20
San Mateo Countywide	10
Contra Costa Countywide	10
Cities of Fairfield -Suisun	4
City of Vallejo	4
R2 Water Board SWAMP	8
Napa Countywide	NA
Marin Countywide	NA
Sonoma Countywide	NA
San Francisco Countywide	NA
Total	**68**

1.2. Creek Monitoring Design Types and Terms

Site selection for creek monitoring can be "targeted" (i.e., where sites are selected to address specific study questions or focus on localized potential or existing problems) or "probabilistic", which provides information on the overall status or condition of a watershed, basin, or region. To meaningfully evaluate biological condition throughout a region with a targeted design, sampling locations must be similar enough to have similar biological expectations, which, in turn, provide a basis for comparison to benchmarks (e.g., Indices of Biological Integrity). Because targeted designs typically do not fully represent the range of conditions encountered in a region, as a group the sites typically present a somewhat skewed picture of regional conditions. In a probabilistic or random sampling regime, creek characteristics may be highly dissimilar among the sites, but the sites together provide a more accurate assessment of biological condition throughout the area than a targeted design. Selecting sites randomly provides for an unbiased assessment of the condition of the many waterbodies at a scale above the individual site or creek. Thus, a Permittee or group of Permittees can address questions at multiple scales.

In a probabilistic design, there are several terms that are used to describe specific elements of the design, including sample universe, sample frame and sample draw. Before sample sites can be selected from the creek network layer, all the possible sample units must be identified. This population of sample units (i.e., assessment areas) is referred to as a **sample universe**. The Geographical Information System (GIS) data layer that is required to identify the sample universe is called the **sample frame**. The set of sampling sites that are selected for a particular survey is known as the **sample draw**. These terms will be used throughout the remainder of this memorandum.

1.3. Existing Statewide and Regional Probability Surveys

1.3.1. SWAMP Perennial Streams Assessment (PSA) Program

Since 2000, California has conducted three successive statewide probability surveys of its perennial streams and rivers, each with a primary focus on biological endpoints. The EPA's Environmental Monitoring and Assessment Program (EMAP) provided California with a foundation of four years (2000-2003, approximately 200 sites) of probability-based stream condition data with its Western Pilot Monitoring Program (EMAP-West). While EMAP-West was nearing completion, two State Water Board monitoring units, the Surface Water Ambient Monitoring Program (SWAMP) and the Non-Point Source Program (NPS) collaborated with their counterparts in the EPA Region IX to develop the California Monitoring and Assessment Program (CMAP), which produced an four additional years (2004-2007, approximately 200 sites) of data. In 2008, the SWAMP program built upon these two previous surveys (EMAP and CMAP) to develop its ongoing Perennial Streams Assessment Program (PSA). These surveys are now combined and managed collectively by the Surface Water Ambient Monitoring Program's (SWAMP) under its Perennial Streams Assessment (PSA) program. In 2010, SWAMP's Perennial Streams Assessment (PSA) conducted the State Water Board's 11th continuous year of probability monitoring. To date, the program has collected biological data (invertebrates, algae) and associated chemical and habitat data from approximately 850 sites statewide. A number of these sites are located in the San Francisco Bay region.

1.3.2. Southern California Watershed Monitoring Program

The Stormwater Monitoring Coalition (SMC) is a coalition of stormwater management agencies and Regional Water Quality Control Boards (RWQCBs) from Ventura to San Diego. Since 2008, the SMC has implemented a regional monitoring program using a probabilistic design for southern California's coastal streams and rivers. The design includes stratification of 15 different management units, defined as watershed areas, and three different land uses (i.e., urban, agricultural, and open). At each sample location, multiple indicators are used to assess the ecological health of the stream, including water chemistry, aquatic toxicity, benthic macroinvertebrate community structure, periphyton community structure and biomass, and physical and riparian habitat.

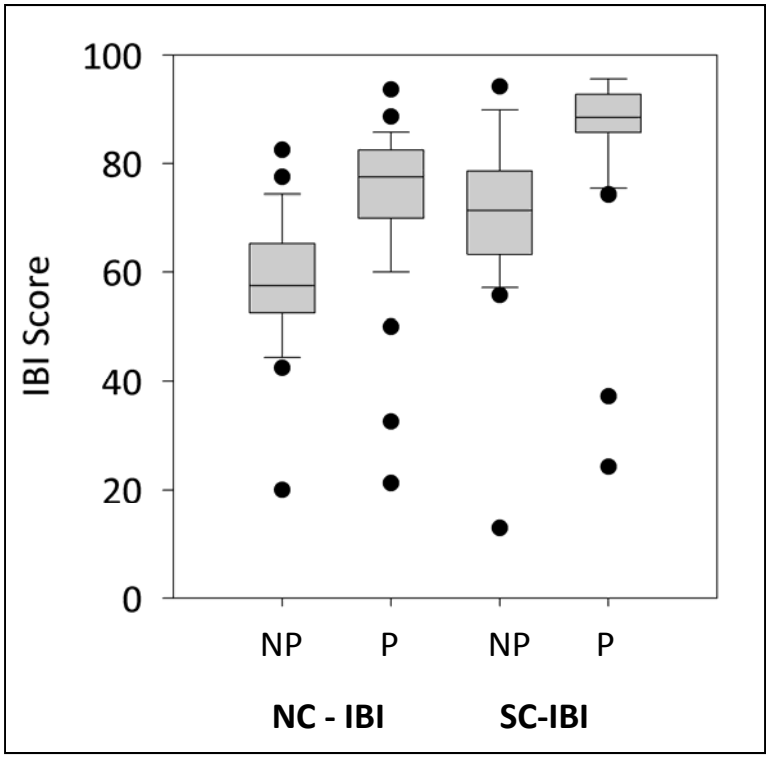
2. REGIONAL CREEK STATUS MONITORING DESIGN

The sample universe and the sample frame that were selected by the RMC for regional creek status monitoring and evaluation process that was conducted are described in this section. The sample draw will be discussed in the deliverable to Task 7 of this project (i.e., *Creek Status and Long-Term Trends Monitoring Design Report*).

2.1 Sample Universe

The sample universe (i.e., project area) for RMC regional ambient creek survey will include creeks and rivers within the boundary of the San Francisco Bay Water Board and the eastern portion of Contra Costa County that drains to the Central Valley region. These creeks include those that flow year around (i.e., perennial) or only part of the year (i.e., non-perennial). Perennial and non-perennial creeks serve as valuable natural resources and aquatic life within these systems is inherently different under least disturbed conditions (Figure 1). For these reasons, the RMC has agreed to include both types of creeks in the sample universe.

Figure 1. Northern NC) and Southern (SC) Coastal California Index of Biotic Integrity (IBI) scores for minimally disturbed perennial (P) and non-perennial (NP) reference sites in the San Francisco Bay Area. Box plots represent the 25th to 75th percentiles and whiskers represent 10th and 90th percentiles, with outliers noted.



The MRP only requires monitoring to be conducted by RMC participants within Alameda, Contra Costa, Santa Clara, San Mateo and Solano counties. That said, the RMC agreed to expand the sample universe to include the remaining land area within the boundaries of the Water Board. This expansion includes Marin, Napa, San Francisco and Sonoma counties.

Lastly, the RMC is also interested in evaluating the condition of aquatic life in creeks draining urban or non-urban areas in the sample universe. The portions of creeks running through or adjacent to urbanized areas are required to be assessed under the MRP. The RMC agreed that non-urban areas should be assessed to provide context for understanding the condition of urban sites.

2.2 Sample Frame

Several existing GIS data layers were evaluated for use in the development of the sample frame, including hydrography, land use and management units (i.e., counties). These data layers are discussed in more detail below.

2.2.1 Hydrography Data

Several existing hydrologic data sources are available for use in developing creek sample frames for the Bay Area. Many of these, however, are incomplete or do not cover the entire project area. Two datasets that provide adequate coverage of the project area were evaluated:

- National Hydrography Dataset (NHD) – This dataset is managed by the USGS and is the surface water component of The National Map. The NHD is a digital vector dataset used in GIS and contains features such as lakes, ponds, streams, rivers, canals, dams and streamgages. The NHD can be viewed at the 1:24,000 or 1:100,000 scale. The 1:100,000 scale NHD dataset was used by SWAMP in the PSA sample frame development.
- Bay Area Aquatic Resource Inventory (San Francisco Estuary Institute) - BAARI is a highly detailed base map of the Bay Area's aquatic features that includes all wetlands, open water, streams, ditches, tidal marshes and flats, and riparian areas. BARRI is currently under development and has a high resolution (1:5,000).

The advantages and disadvantages of using each data layer in the development of the RMC creek status sample frame were developed through the evaluation of each dataset. The results are summarized in Table 2.

Table 2. Comparison of two hydrography datasets considered for use in developing the RMC sample frame.

National Hydrography Dataset (NHD)		Bay Area Aquatic Resources Inventory (BAARI)	
Advantage	**Disadvantage**	**Advantage**	**Disadvantage**
• Complete Regional Coverage • Consistent with SWAMP and SMC • Attributes in "plus" version include flow regime, stream order	• Moderate Resolution (1:100,000)	• High Resolution (1:5,000)	• Gaps in Regional Coverage • Currently not attributed

In summary, the NHD dataset is currently the only existing creek data layer that covers the entire geographical area addressed by the RMC. Additionally, it was used by both SWAMP and the SMC to develop their sampling frames. That said, the moderate resolution of the NHD was a noted disadvantage. Specifically, if NHD was selected as the hydrography data layer for the RMC sample frame, higher order creeks would likely not be within NHD and therefore would not be part of the frame. To further assess this potential issue, the resolution of the NHD was evaluated by plotting 276 existing targeted bioassessment sampling sites located within four major counties currently applicable to the RMC (Alameda, Contra Costa, Santa Clara and San Mateo) onto NHD. A total 31 of the 276 sites (11%) of the sampling sites occurred on creeks that were not represented in the NHD. Based on this evaluation, the RMC agreed that the advantages of using the NHD outweighed the disadvantages, and the NHD was selected as the hydrography data layer that will be used to develop the RMC sample frame.

The proportions of channel types in the NHD data layer for the RMC geographical area are shown in Table 3. Intermittent and perennial creeks account for 50% and 26%, respectively, of channel types for the entire region. The remaining 24% of channel types were artificially constructed water conveyance channels.

Table 3. Total miles and percentage of water conveyance types as attributed in NHD within the five counties applicable to current RMC participants.

FCODE	Description	Total Length (miles)	Percent of total
46003	Creek/River - Intermittent	3,241	50
46006	Creek/River - Perennial	1,712	26
33600	Canal/ditch	1,138	18
55800	Artificial Path – no attribute	210	3
42803	Pipeline	171	3
Total		**6,472**	**100**

2.2.2 Land Use

Two land use datasets were evaluated for use in the sample frame – the National Land Cover Dataset (NLCD) and the Association of Bay Area Governments (ABAG) land use dataset. The NLCD is a nationwide dataset that was developed by the Multi-Resolution Land Characteristics Consortium of government agencies. The NLCD data were created using digital imagery and distinguish 16 different classes of land use. The ABAG (2005) dataset provides land use information for the nine counties in the San Francisco Bay Area. The ABAG dataset is a hierarchical classification system containing four levels of subdivision used to express various levels of detail of map data. The ABAG land use data were developed using a variety of data sources, including the NLCD, county assessor data, aerial photography and local knowledge.

The NLCD (2001) was used by both SWAMP and the SMC to develop their sample frames; however, the ABAG dataset has a higher resolution. SWAMP assigned land use information to the sample frame by overlaying the NLCD land use grid over the NHD data set. A lower resolution (300 meter) NLCD grid was used to subdivide the NHD data layer into smaller segments (or sites) and assign a landcover class to each site. Sixteen different landcover codes in the NLCD data were reclassified into four classes: urban, forest, agriculture and other. Water cover was not included in the reclassification.

To evaluate the consistency in the two datasets with regard to creek status monitoring, land use information from the SWAMP statewide sample frame in the four main RMC counties was compared to the land use data provided in ABAG. For the purposes of this analysis, the ABAG land uses were reclassified as urban and non-urban using best professional judgment. Figure 2 illustrates the comparison of the two land use data sets.

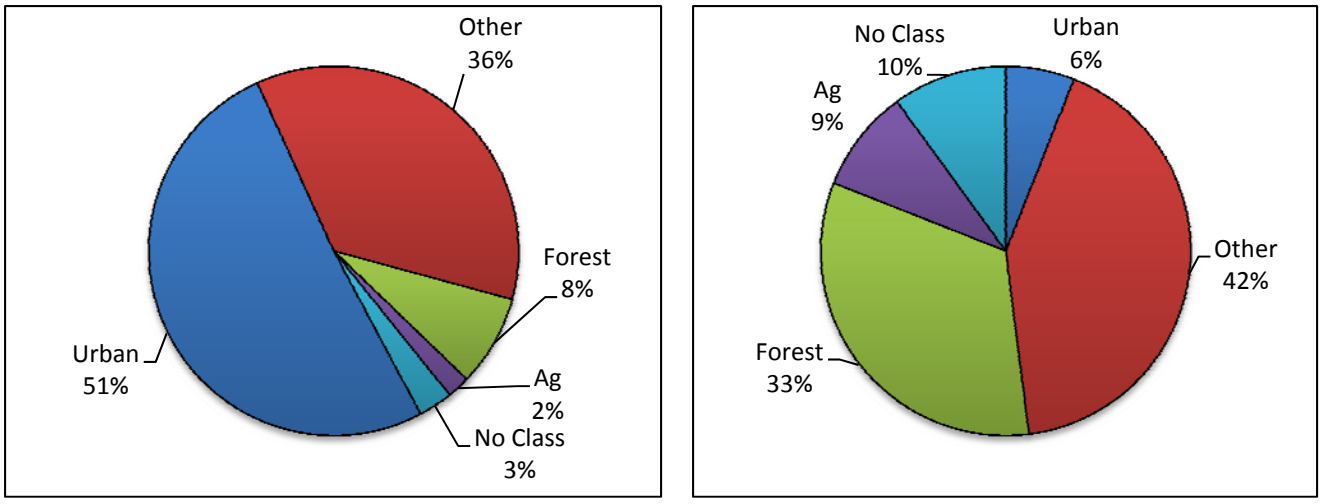


Figure 2. Percentages of creek miles in the SWAMP sample frame designated as urban, other, agriculture, forest, or no class with the urban (left) and non-urban (right) creek miles defined using ABAG land use dataset.

Approximately thirty percent of the stream miles (1308 miles) in the SWAMP sample frame occurred in urban land uses as defined by ABAG. About 51% of these miles (661 miles) were also classified as urban in the SWAMP sample frame. The remaining stream miles were defined as other (36%), forest (8%), agriculture (2%) and no class (3%).

The results show that a relatively large proportion of creeks (468 miles) classified as "other" in the SWAMP sample frame occurs in urbanized areas (as defined in the ABAG data). The "other" land use category also had the largest proportion of non-urban creek segments in the sample frame. Thus potential sample sites identified as other will need to be further evaluated after the sample draw to determine if the correct land use class is actually urban or non-urban.

2.2.3 Management Units

Management units in the RMC design are represented by a combination of county and Water Board boundaries. Specifically, the portions of the nine Bay area counties that fall within the Water Board boundaries form the management units of the RMC regional creek monitoring design, with the exception of Contra Costa County's management unit, which extends into the Central Valley Regional Water Quality Control Board's region. Although the RMC only includes the Solano County jurisdictional areas of the cities of Vallejo, Fairfield, and Suisun City, the entire portion of Solano County within the Water Board's boundary was used as the Solano County management unit for the sample frame. The reason for expanding the area beyond the city boundary is that the available stream miles in these three cities is too limiting to achieve the desired number of sampling sites for these programs. Management units for each RMC participant are listed in Table 4.

Table **4**. **M**anagement units used in **RMC** sample frame.

MRP Permittee Associated County	**Management Unit**
Alameda	Alameda County within the boundaries of Region 2
Contra Costa	Contra Costa County
San Mateo	San Mateo County
Santa Clara	Santa Clara County within the boundaries of Region 2
Solano	Solano County Santa Clara County within the boundaries of Region 2

2.3 Sample Frame Strata

When populations vary or direct comparisons between subpopulations (stratum) are anticipated, it is advantageous to sample each subpopulation independently. Stratification is the process of dividing members of the population into subgroups before sampling. With regard

to the RMC creek status ambient survey, participants agreed that at a minimum, comparisons between urban and nonurban land use sites are needed. Therefore, consistent with the SWAMP and SMC surveys, stratification of the RMC sample frame will occur based on land use. Additionally, because the MRP-required sample sites cannot be distributed between counties, stratification based on management unit (county) is also warranted. The RMC creek status ambient survey strata and associated GIS data layers are presented in Table 5.

Table 5. RMC creek status ambient survey strata.

Sample Frame Strata	GIS data layer
Land Use	SWAMP sample frame using NLCD landcover class attribute
Management Unit	County boundary included within SF Bay Regional Water Board boundary, except for Contra Costa (entire county)

3.0 SUMMARY AND NEXT STEPS

In summary, the SWAMP sample frame was determined to be suitable for use in developing the RMC regional creek status monitoring design for the following reasons:

- Complete range of sampling units (i.e., assessment areas) for the RMC geographical area
- Maintains consistency with the SWAMP's PSA Program; as a result the RMC may be able to leverage sites sampled in the San Francisco Bay Region by the SWAMP PSA Program.
- Provides land cover data that can be used to stratify the sample universe and meet MRP constraints related to minimum number of sites per management unit. (Please note that sampling locations assigned the "other" land cover category will need to be further evaluated to determine if sites represent urban or non-urban conditions).
- County boundaries within the Water Board area provide strata for the sampling frame that are necessary to achieve the number of sampling locations required in the MRP.

The following tasks summarize next steps to develop the sample draw:

- Agree on the temporal extent of sampling that will occur under the RMC creek status sampling design. A five-year horizon is recommended to increase the statistical confidence in the condition assessment that will occur as an outgrowth of the data collection.
- Determine best approach to weighting of sample frame strata to obtain desired sample draw. Highest weight will be assigned to sites classified as urban, moderate weight

given to sites classified as “other”, and lowest weight given to sites with forested and agriculture classes. RMC participants will need to consider how weighting affects the confidence in condition assessments.

- Develop the sample draw by implementing the GRTS design using the R statistical software system (using version 2.12 of the "psurvey.analysis" library package). The sample draw will target a minimum of 30 – 50 sites per land use stratum, to be tentatively covered over the five-year assessment period of RMC creek status monitoring under this design. These sample sizes will be targeted to maintain optimal confidence intervals[1] for monitoring data. A confidence interval conveys the certainty (margin of error) associated with data. A GRTS sample draw produces a set of points that are ordered so that the sequential use of the points will maintain spatial balance in the design. Target sample sites that cannot be assessed should be replaced by the first available oversample point. For this reason, it is important to keep a close record of the sites that are assessed and those that cannot be assessed, with detailed notes on the reasons why a site is being skipped.

[1] It is expressed as a percentage and represents how often the true percentage of the population represented by a sample who would have the value within the associated confidence interval. A 95% confidence interval means you can be 95% certain; the 99% confidence level means you can be 99% certain. Most researchers use the 95% confidence level.