Bay Area Municipal Stormwater Collaborative (BAMSC) Regional Bioassessment Report

Water Years 2012 - 2022

Submitted in Compliance with
NPDES Permit No. CAS612008 (Order No. R2-2022-0018)
Provision C.8.h.vi

Prepared for:

BAY AREA MUNICIPAL STORMWATER COLLABORATIVE

Prepared by:

EOA inc.

APPLIED marine SCIENCES

FINAL March 31, 2024

Table of Contents

List of Tables

List of Figures

Executive Summary

Biological assessment (bioassessment) is an evaluation of the biological, physical habitat and water quality conditions of a water body. In 2009, the Bay Area Municipal Stormwater Collaborative (BAMSC) Regional Monitoring Coalition (RMC) developed a bioassessment monitoring program to answer management questions identified in the Municipal Regional Stormwater National Pollutant Discharge Elimination System (NPDES) Permit (referred to as the Municipal Regional Permit or MRP):

- *Are water quality objectives, both numeric and narrative, being met in local receiving waters, including creeks, rivers and tributaries?*
- *Are conditions in local receiving waters supportive or likely to be supportive of beneficial uses?*

Bioassessment data and synoptic physical habitat and water quality data collected over the eleven years of RMC monitoring (2012-2022) are included in this report. The RMC's monitoring design addresses these management questions on a regional (San Francisco Bay Area) scale and across the five participating SF Bay Area counties (Alameda, Contra Costa, San Mateo, Santa Clara and Solano). Two study questions were developed to assist with addressing the management questions described above:

1) What are the biological conditions of urban streams in the region?
2) What stressors are associated with poor biological conditions?

The findings of this study are intended to help stormwater programs better understand the current condition of these water bodies and identify stressors that are likely to pose the greatest risk to the health of streams in the Bay Area.

KEY FINDINGS

- **Most streams in the five-county region are in poor biological condition.** The biological conditions of streams in the RMC area were assessed using two ecological indicators: benthic macroinvertebrates (BMIs) and algae. Results from the 2012 through 2022 study period indicate that urban streams in the RMC area are generally in poor biological condition. Based on BMIs, 89% of the urban stream-length was below the 10^{th} percentile of reference conditions for the California Stream Condition Index (CSCI). For algae indices (D ASCI and H ASCI), stream conditions were very similar to BMIs, with 92% and 90% of the streams ranking below the 10^{th} percentile of reference conditions. These findings should be interpreted with the understanding that the survey focused on urban stream conditions in the five-county region, and that these data represent current (baseline) conditions within these streams.
- **Poor biological conditions are strongly associated with physical habitat, water quality, and land use stressors.** The associations between biological indicators (CSCI and D ASCI) and stressor data were evaluated using random forest models. The study results showed that different biological indicators responded to different types of stressors. CSCI scores were strongly influenced by physical habitat variables (e.g., % fast water in the reach, % particles smaller than sand (<2mm), % coarse gravel and % fines) and land use factors (e.g., % impervious area within 5km upstream of the site), while D ASCI scores were moderately influenced by land use factors (e.g., impervious area and road crossings within 5km upstream of the site) and water quality variables (e.g.,

temperature and conductivity). Together, BMI and algae indices can be used to assess the overall biological conditions of water bodies and potentially identify the causes of poor (or good) conditions. In general, CSCI scores at urban sites were consistently low, indicating that degraded physical habitat conditions common in urban settings are impacting biological conditions in streams. D ASCI scores at urban sites were also relatively low, indicating that urban land use and degraded water quality may be impacting diatom conditions in urban streams.

- **The RMC monitoring design provides estimates for overall stream conditions in RMC area and urban stream conditions for each of the five counties.** Participating municipalities are primarily concerned with stormwater runoff impacts from urban areas, therefore the RMC focused sampling efforts on urban sites (approximately 85%) over non-urban sites (approximately 15%). As a result, non-urban sites are under-represented in the RMC dataset, resulting in lower overall biological condition scores in the region than would be expected for a spatially balanced (fully randomized) dataset.
- **Biological conditions are associated with channel type and to a lesser degree, flow status.** Urban sites with hardened channel beds had consistently low biological conditions. Bioassessment sites with soft channel bed, but containing some degree of modified/armored banks, generally had low biological conditions as well. Urban sites that were characterized as "natural channel" had a wider range of biological conditions, with nearly 25% of sites scoring above the 10^{th} percentile of reference conditions for CSCI. In general, biological index scores at perennial urban sites were lower than non-perennial (e.g., intermittent or ephemeral) urban sites. However, the strength of this relationship was highly variable across counties.

1 INTRODUCTION

1.1 BACKGROUND

This Bay Area Municipal Stormwater Collaborative (BAMSC) Regional Bioassessment Report was prepared by members of the BAMSC Regional Monitoring Coalition (RMC) in compliance with Provision C.8.h.vi of the Municipal Regional Stormwater National Pollutant Discharge Elimination System (NPDES) Permit (referred to as the Municipal Regional Permit or MRP; Order R2-2022-0018). The BAMSC[1] RMC is a consortium of San Francisco Bay Area municipal stormwater programs that joined together in 2010 to coordinate and oversee water quality monitoring required by the MRP. Members of the group include:

- Alameda Countywide Clean Water Program (ACCWP),
- Contra Costa Clean Water Program (CCCWP),
- San Mateo Countywide Water Pollution Prevention Program (SMCWPPP),
- Santa Clara Valley Urban Runoff Pollution Prevention Program (SCVURPPP), and
- Solano Stormwater Alliance (SSA).

The MRP was first adopted by the San Francisco Bay Regional Water Quality Control Board (SFBRWQCB or Regional Water Board) on October 14, 2009 as Order R2-2009-0074 (SFRWQCB 2009; referred to as MRP 1.0). On November 19, 2015, the Regional Water Board updated and reissued the MRP as Order R2-2015-0049 (SFBRWQCB 2015; referred to as MRP 2.0). The Regional Water Board subsequently updated and revised the MRP as Order R2-2022-0018 (SFBRWQCB 2022; referred to as MRP 3.0), which took effect on July 1, 2022.

MRP 1.0 and 2.0 required bioassessment monitoring in accordance with Standard Operating Procedures (SOPs) established by the California Surface Water Ambient Monitoring Program (SWAMP), including sampling of benthic macroinvertebrates (BMIs), benthic algae (i.e., diatoms and soft algae), and water chemistry, and the characterization of physical habitat. Bioassessment monitoring was intended to address two broad management questions:

- *Are water quality objectives, both numeric and narrative, being met in local receiving waters, including creeks, rivers and tributaries?*
- *Are conditions in local receiving waters supportive or likely to be supportive of beneficial uses?*

Consistent with the requirements of MRP 1.0 and 2.0, the RMC developed a probabilistic monitoring design to address the management questions on a regional scale and compare monitoring results across counties. Under MRP 1.0 and 2.0, municipal stormwater programs were required to assess a minimum number of stream/channel sites based on their relative population. As a result, Santa Clara and Alameda counties were required to each sample 20 sites and San Mateo and Contra Costa counties were required to each sample 10 sites, on an annual basis for both permit terms. Fairfield-Suisun and Vallejo (now

---

[1] The BAMSC was organized by the Bay Area Stormwater Management Agencies Association (BASMAA) Board of Directors to continue the information sharing and permittee advocacy functions of BASMAA in an informal manner after BASMAA's dissolution.

joined under the SSA) were required to sample 8 and 4 sites, respectively, during each of the two permit terms. In addition, the Regional Water Board collaborated with the RMC by monitoring additional sites in non-urban areas in each of the counties between 2012 and 2015.

The RMC previously conducted a regional analysis of bioassessment data collected at sites within the five participating San Francisco Bay Area (Bay Area) counties (Alameda, Contra Costa, San Mateo, Santa Clara and Solano) over the first five years of monitoring (2012-2016). The study assessed the biological condition of streams in the region and identified stressors that are likely to pose the greatest risk to the health of streams in the Bay Area. A final report was submitted to the Regional Water Board in March 2019 (BASMAA 2019) and a fact sheet was developed in August 2019.

MRP 3.0 Provision C.8.h.vi requires municipal stormwater programs to collectively submit to the Regional Water Board by March 31, 2024 a comprehensive analysis of bioassessment data collected by the RMC during MRP 1.0 and 2.0 (2012-2021). The RMC continued bioassessments into WY 2022, which was the last year of the MRP 2.0; therefore, this report includes an analysis of bioassessment data collected over an eleven-year period (2012 – 2022).

1.2 Previous Studies

A comprehensive analysis of bioassessment data collected in San Francisco Bay streams was recently completed (Brown and Mazor 2023). The study was conducted by the Southern California Water Research Project (SCCWRP) and funded by the Regional Water Board. The report includes an analysis of bioassessment data collected over the past 20 years at more than 1,500 sites in Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano, and Sonoma Counties. Findings from the report indicate that most streams in the San Francisco Bay area are in poor biological condition based on biological indices that evaluate BMI and benthic algae (i.e., diatoms and soft algae) communities sampled at each site. The indices include the California Stream Condition Index (CSCI) for BMIs and Algal Stream Condition Indices for diatoms (D ASCI) and hybrid of diatoms and soft-bodied algae (H ASCI).

The study also evaluated the influence of stream flow duration (perennial vs non-perennial) and channel type (e.g., modified vs. natural) on all three biological indices. The study concluded that flow status has an important influence on CSCI scores. Specifically, the CSCI scores at reference, non-perennial sites were consistently lower than expected. As a result, CSCI scores at natural, intermittent sites may incorrectly indicate degraded conditions. For these sites, the authors propose that alternative thresholds should be used to distinguish between reference and non-reference conditions. However, existing thresholds should still apply to sites with intermittent flow conditions caused by anthropogenic activities (e.g., flow regulation below diversions and dams). Stream flow duration did not appear to influence the benthic algae index scores (D ASCI, H ASCI).

Biological index scores were generally low at sites with modified channels (both hard and soft bottom). The findings suggest additional thresholds at these sites might be useful for setting appropriate management goals. Potential thresholds based on the best observed score for each channel type were identified in the report.

The study concludes that the association between flow status and channel modification with biological index scores may be an important factor to consider when evaluating sites with degraded conditions and determining where and what type of management actions may be most appropriate.

1.3 Project Goal

The primary goal of this project was to compile and evaluate bioassessment data collected over the two cycles of monitoring conducted by the RMC (2012 – 2022). The evaluation was designed to address two main questions, consistent with the overarching questions in the MRP:

1) What are the biological conditions of urban streams in the region?
2) What stressors are associated with poor biological conditions?

The entire 11-year dataset includes bioassessment data collected at both probabilistic and non-probabilistic (i.e., targeted) sites. The probabilistic data was used to create cumulative distribution function (CDF) curves showing biological conditions for urban streams within the RMC study area, as well as urban streams within each county. Data collected at targeted sites were incorporated with data collected at probabilistic sites to evaluate stressor association with biological conditions. In addition, supplemental data characterizing stream flow duration (perennial vs non-perennial) and channel type (e.g., modified vs. natural) was developed and evaluated for potential association with biological condition.

The findings of this report are intended to help municipal stormwater programs better understand the current condition of these water bodies, prioritize stream reaches in need of protection or restoration, and identify stressors that are likely to pose the greatest risk to the health of streams in the Bay Area.

This project was implemented by a Project Team comprised of EOA, Inc. and Applied Marine Sciences, Inc. (AMS). A Project Management Team (PMT) consisting of representatives from BAMSC municipal stormwater programs and associated municipalities provided oversight and guidance to the Project Team.

Sections of this report are organized according to the following topics:

- **Section 1.0** – Introduction, including project goals;
- **Section 2.0** – Methods including monitoring survey design, site evaluation procedures, field sampling and data analysis;
- **Section 3.0** – Results summarizing biological conditions and stressor association with conditions;
- **Section 4.0** – Discussion organized by the management questions and goals; and
- **Section 5.0** – References cited in the Report.

2 METHODS

2.1 STUDY AREA

The study area for RMC creek status monitoring consists of the perennial and non-perennial streams, channels and rivers within the portions of the five participating counties (San Mateo, Santa Clara, Alameda, Contra Costa, Solano) that overlap with the San Francisco Bay Regional Water Quality Control Board (Region 2) boundary, and the eastern portion of Contra Costa County that drains to the Central Valley region (Region 5). The RMC bioassessment sample frame consists of the urban and non-urban portions of the stream network flowing through the RMC area. The source dataset used to create the sample frame was the 1:100,000 National Hydrography Dataset (NHD).

2.2 SURVEY DESIGN AND PROBABILISTIC SAMPLING SITES

Bioassessment monitoring sites were selected based on a probabilistic survey design consisting of a master draw of 5,740 sites (approximately one site for every stream kilometer in the sample frame). The selection procedure employed the U.S. EPA's Generalized Random Tessellation Stratified (GRTS) survey design methodology (Stevens and Olson, 2004). The GRTS approach generated a spatially-balanced distribution of sites covering the majority of the San Francisco Bay Area. It should be noted that the sample draw of 5,740 sites did not account for land use designations or other emphases (i.e., County) and therefore, the master draw of sample sites was weighted towards commonly occurring conditions (i.e., non-urban sites), with less common conditions (i.e., reference and urban sites) being less represented due to their lower relative abundance in the sample frame (i.e., there are fewer urban creek miles in the study area than there are non-urban creek miles).

The RMC sampling design focused on the population of accessible streams with flow conditions suitable for sampling (i.e., adequate flow during spring index period). A random set of potential monitoring sites (i.e., the master draw) was established, with each site having an equal, non-zero weight, proportional to the inverse of its selection probability. Thus, all sites were assumed to have an equal probability of selection throughout the sample frame. The weights represent the amount of stream length encompassed by each site in the overall population of all possible sites.

Once the master draw was established, the list of monitoring sites was separated into 19 categories to facilitate site evaluations and implement bioassessment monitoring (Table 1). The following attributes were used to generate the categories:

- County (n=5): San Mateo, Santa Clara, Alameda, Contra Costa, Solano (source: California Department of Forestry and Fire, 2009);
- Water Quality Control Board Region (n=2): Region 2, Region 5 (source: San Francisco Regional Water Quality Control Board, undated);
- Land use Category (n = 4): Urban or nonurban in all counties, except Solano ('urban_V' and 'urban_FS' in Solano County). Urban land use was defined as a combination of US Census (2000) areas classified as urban, and areas within Census City boundaries. This definition of urban land

use results in some relatively undeveloped areas and parks along the fringes of cities to be classified as urban. Urban sites therefore represent a broad range of developed (i.e., impervious surface) conditions. Non-urban area was defined as all remaining area in the RMC boundary not classified as urban.

Table 1. Number of sites and stream length from the master draw in each post-stratification category.

<table>
<tr><th rowspan="2">County</th><th colspan="2">Urban</th><th colspan="2">Non-Urban</th><th colspan="2">Total</th></tr>
<tr><th>Sites</th><th>Stream Length (km)</th><th>Sites</th><th>Stream Length (km)</th><th>Sites</th><th>Stream Length (km)</th></tr>
<tr><td>San Mateo</td><td>222</td><td>234</td><td>528</td><td>556</td><td>750</td><td>790</td></tr>
<tr><td>Santa Clara</td><td>542</td><td>571</td><td>1,376</td><td>1,449</td><td>1,918</td><td>2,020</td></tr>
<tr><td>Alameda</td><td>454</td><td>478</td><td>842</td><td>887</td><td>1,296</td><td>1,365</td></tr>
<tr><td>Contra Costa (Region 2)</td><td rowspan="2">587</td><td rowspan="2">618</td><td>363</td><td>382</td><td>845</td><td>890</td></tr>
<tr><td>Contra Costa (Region 5)</td><td>349</td><td>368</td><td>454</td><td>478</td></tr>
<tr><td>Solano (Vallejo)</td><td>12</td><td>13</td><td rowspan="2">386</td><td rowspan="2">407</td><td rowspan="2">477</td><td rowspan="2">502</td></tr>
<tr><td>Solano (Fairfield-Suisun)</td><td>79</td><td>83</td></tr>
<tr><td colspan="5">Overall Total</td><td>5,740</td><td>6,045</td></tr>
</table>

To maintain a spatially-balanced pool of monitoring sites, sites were evaluated in the order that they appeared in the master draw list (with a few exceptions). Sites were evaluated for sampling using both desktop and field reconnaissance. Field crews attempted to locate a reach suitable for sampling within 300 m of the target coordinates. Sites without a suitable reach were rejected for sampling. Reasons for rejection included physical barriers, lack of flowing water, refusal or lack of response from landowners, unwadeable (i.e., >1 m deep for at least 50% of the reach) and inappropriate waterbody types (e.g., tidally influenced). Sites with temporary inaccessibility, unsafe/hazardous or permission issues (e.g., construction, lack of response from landowners) were re-evaluated for sampling in subsequent years. All program participants were instructed to use a standard set of codes to identify the reason behind exclusion of sites.

Additionally, at the outset, each countywide Program agreed they would attempt to assess up to 20% of their required sites in non-urban areas.

2.3 Targeted Sampling Sites

During MRP 1.0, the RMC Monitoring Program focused biological assessments at probabilistic sites determined by the RMC sample frame. During MRP 2.0, the pool of remaining urban probabilistic sites for some of the counties was fully depleted (i.e., all remaining sites were evaluated for sampling). As a result, these municipal stormwater programs began to include targeted (i.e., non-probabilistic) sites to meet the required annual minimum number of sites and answer specific questions associated with these sites (Table 2).

Table 2. Number of probabilistic and targeted (i.e., non-probabilistic) sites sampled in each County between 2012 and 2022.

County	Bioassessment Site		
	Probabilistic	Targeted	Total
Alameda	196	33	229
Contra Costa	115	0	115
San Mateo	98	22	120
Santa Clara	178	54	232
Solano	30	2	32
TOTAL	**617**	**111**	**728**

Targeted bioassessment sites were selected for several reasons:

- Re-sample previously sampled probabilistic sites to evaluate variability in biological conditions over time.
- Conduct biological assessments to evaluate changes in biological integrity following the implementation of stream habitat enhancement projects.
- Provide additional information for Stressor Source Identification Projects.
- Increase spatial density of sites within a watershed of interest.
- Evaluate relationships between biological metrics calculated for bioassessment data (i.e., CSCI and ASCI) with fish survey data collected by Valley Water.

For this report, only data collected at probabilistic sites were used to evaluate overall biological conditions (i.e., as represented by CDF curves) for Bay Area streams within each of the five participating counties. The data collected at targeted sites were incorporated into the analyses investigating association between biological conditions and stressor data. For sites that were sampled more than once, only the first sample event was used in the analysis since the first sample event is also used in the biological condition analyses using data collected at the probabilistic sites.

2.4 Sampling Protocols/Data Collection

Biological sample collection and processing was consistent with the BASMAA RMC Quality Assurance Project Plan (QAPP)[2] (BASMAA 2020) and Standard Operating Protocols (SOPs) (BASMAA 2016) which were developed to be consistent with the current SWAMP Quality Assurance Program Plan (QAPrP) and SOPs. Bioassessments were conducted during the spring index period (approximately April 15 – June 30) with the goal to sample a minimum of 30 days after any significant storm (defined as at least 0.5-inch of

[2] The RMC QAPP and SOP documents were initially developed in 2012 (Version 1.0), revised in 2013 (Version 2.0) and 2016 (Version 3.0).

rainfall within a 24-hour period). The 30-day period allows diatom and soft algae communities to recover from peak flows that may scour benthic algae from the bottom of the stream channel.

2.4.1 Biological Indicators

Each monitoring site consisted of an approximately 150-meter stream reach divided into 11 evenly spaced transects placed perpendicular to the direction of flow. Benthic macroinvertebrate (BMI) and algae (i.e., diatom and soft algae) samples were collected at each transect using the Reach-wide Benthos (RWB) method described in Ode et al. (2016). The algae composite sample was also used to collect chlorophyll a and ash free dry mass (AFDM) samples following methods described in Ode et al. (2016).

Biological samples were sent to laboratories for analysis. The laboratory analytical methods used for BMIs followed Woodward et al. (2012), using the Southwest Association of Freshwater Invertebrate Taxonomists (SAFIT) Level 1a Standard Taxonomic Level of Effort, with the additional effort of identifying chironomids (midges) to subfamily/tribe instead of family (Chironomidae). Soft algae and diatom samples were analyzed following SWAMP protocols (Stancheva et al. 2015). The taxonomic resolution for all data was standardized to the SWAMP master taxonomic list.

2.4.2 Physical Habitat

Both quantitative and qualitative measurements of physical habitat structure were taken at each of the 11 transects and 10 inter-transects at each monitoring site. At the outset of the monitoring program in 2012, physical habitat measurements followed procedures defined in the “BASIC” level of effort (Ode 2007), with the following exceptions as defined in the “FULL” level of effort: stream depth and pebble count + coarse particulate organic matter (CPOM), cobble embeddedness, and discharge measurements. In 2016, the entire “FULL” level of effort for the characterization of physical habitat described in Ode et al. (2016) was adopted, consistent with the reissued MRP 2.0 (SFBRWQCB 2015). Physical habitat measurements include channel morphology (e.g., channel width and depth), habitat features (e.g., substrate size, algal cover, flow types, and in-stream habitat diversity) and human disturbance in the riparian zone (e.g., presence of buildings, roads, vegetation management). In addition, a qualitative Physical Habitat Assessment (PHAB) score was assessed for the entire bioassessment reach. The PHAB score is composed of three characteristics for the reach, including channel alteration, epifaunal substrate, and sediment deposition. Each characteristic is individually scored on a scale of 0 to 20, with a score of 20 representing good condition.

2.4.3 Water Quality

Immediately prior to biological and physical habitat data collection, general water quality parameters (dissolved oxygen (DO), pH, specific conductance (SpCond) and temperature) were measured at or near the centroid of the stream flow using pre-calibrated multi-parameter probes. In addition, water samples were collected for nutrients and conventional analytes analysis using the Standard Grab Sample Collection Method as described in SOP FS-2 (BASMAA 2016).

2.4.4 Stressor Variables

Physical habitat, land-use, and water quality data were compiled and evaluated as potential stressor variables for biological condition. Land-use variables were calculated in GIS by overlaying the drainage area for sample locations with land use and road data. The variables included percent urbanization, percent impervious, total number of road crossings and road density at three different spatial scales (1 km, 5 km, and entire watershed).

Physical habitat metrics were calculated using the SWAMP Bioassessment Reporting Module (SWAMP RM). The SWAMP RM output includes calculations based on parameters that are measured using EPA's Environmental Monitoring and Assessment Program (EMAP) for freshwater wadeable streams (Kaufmann et al. 1999), as well as parameters collected under the SWAMP protocol (Marco Sigala, personal communication, 2017). The SWAMP RM produces a total of 176 different metrics based on data collected using the SWAMP "FULL" habitat protocol. Ten metrics that responded well across a stressor gradient (Andy Rehn, CDFW, personal communication) were selected to analyze the physical habitat data.

General water quality (e.g., DO, SpCond) and chemistry (e.g., nitrate and phosphorus) data collected at the bioassessment sites were also included. Some of the water chemistry variables were calculated from the analytes that were measured. These include Total Nitrogen (sum of Nitrate, Nitrite and Total Kjeldahl Nitrogen) and Unionized Ammonia (calculated using pH and temperature).

2.4.5 Flow Data

A recent study of biological conditions in San Francisco Bay streams indicates that flow status may be an important factor that influences CSCI scores (Brown and Mazor 2023). Specifically, the CSCI scores at natural, intermittent sites may incorrectly indicate degraded conditions. For this reason, an evaluation of flow status and biological condition was conducted for the 11 years of bioassessment data collected under the MRP.

Flow status (perennial versus non-perennial) was identified by assessing flow conditions during the dry season. Between 2012 and 2019, flow observations were made during field visits conducted at each bioassessment location during late summer/early fall months following the spring season sampling event. Sites were assessed using following categories:

- Perennial flow: Included three subcategories- wet flowing, majority flowing (>25% channel wet), and wet trickle.
- Non-perennial: Included two subcategories - no water and minority flowing (<25% channel wet).

In 2019, field crews stopped determining flow status following guidance from California Department Fish and Wildlife staff indicating that there were no significant differences in biological conditions for perennial versus non-perennial sites. As a result, all sites that could be sampled during the spring index period were treated the same.

For the purposes of this analysis, flow status for all sites sampled after 2019 was estimated using best professional judgement.

It should be noted however, that defining flow status, whether through direct field observations or professional judgement, comes with uncertainty. Many sites were observed to be wet one year and dry the next due to inter-annual variability in precipitation. There are likely many factors related to flow influencing biological conditions, including the timing, duration and extent of dry channel conditions. Thus, two classifications for flow status are likely an oversimplification.

2.4.6 Channel Classification

A recent study of biological conditions in San Francisco Bay streams indicates that biological index scores were generally low at sites with modified channels (both hard and soft bottom) (Brown and Mazor 2023). For this reason, an evaluation of channel type and biological condition was conducted for the 11 years of bioassessment data collected under the MRP.

Channels were classified into 6 categories, Natural, Constructed (not historically occurring), Hard bottom, Soft bottom with 1 hardened bank, Soft bottom with 2 hardened banks, and Soft bottom with no hardened banks (Brown and Mazor 2023). These classes were then grouped into three main categories: Natural, Hard (includes artificial and hard bottom); and Soft Beds (including soft bottom with 0, 1, or 2 hardened banks). Channel classification was assigned using existing data sources. Data sources included reach photos taken during biological sampling; qualitative rank scores assigned during bioassessments that assessed the level of channelization in the sampled reach; visual observations of the channel using aerial photos on Google Earth; and prior knowledge of the sampled reach.

2.4.7 Rainfall Data

For evaluation of climate, a representative rainfall dataset was collated for the region. The total accumulated rainfall in each water year[3] during the period of 2002-2022 was calculated. Rainfall measured at the San Francisco Airport was used to represent rainfall in the region.

2.5 Data Analyses

All statistical, tabular, and graphical analyses were conducted in R Studio, running R version 4.3.1 (R Core Team 2023). For analyses involving water quality data, censored results (i.e., below the method detection limit) were substituted with 50% of the method detection limit (MDL). Generally, analytical sensitivity was good, with the only two variables with > 20% non-detects: nitrite (34%) and ammonia (25%).

2.5.1 Biological Condition Indices

The California Stream Condition Index (CSCI) is an assessment tool developed by the State Water Board to support the development of California's statewide Biological Integrity Plan. The CSCI translates BMI data into an overall measure of stream health. The CSCI was developed using a large reference data set that represents the full range of natural conditions in California and site-specific models for predicting biological communities. The CSCI combines two types of indices: 1) taxonomic completeness, as measured by the ratio of observed-to-expected taxa (O/E); and 2) ecological structure and function, measured as a predictive multimetric index (pMMI) that is based on reference conditions. The CSCI score is computed as the average of the sum of the O/E and pMMI.

CSCI scores for each station are calculated using a combination of biological and environmental data following methods described in Rehn et al. (2015). Biological data consist of the BMI data collected and

[3] Most hydrologic monitoring occurs for a period defined as a Water Year, which begins on October 1 and ends on September 30 of the named year. For example, Water Year 2023 (WY 2023) began on October 1, 2022 and concluded on September 30, 2023.

analyzed using the protocols described in the previous section. Environmental predictor data are generated in GIS using drainage areas upstream of each BMI sampling location.

The State Water Board and SCCWRP developed the Algae Stream Condition Index (ASCI) which uses benthic algae data as a measure of biological condition for streams in California (Theroux et al. 2020*)*. The ASCI uses pMMIs to evaluate ecological conditions. There are three versions of the ASCI pMMI: an index for diatoms (D ASCI), one for soft-bodied algae (S ASCI) and a hybrid index using both assemblages (H ASCI). Using a statewide data set, all three indices were evaluated by Theroux et al. (2020) for precision, accuracy, responsiveness, and regional bias. The diatom and hybrid indices were found to be the most sensitive to anthropogenic stressor gradients.

2.5.2 Biological Indicator Thresholds

Existing thresholds for CSCI scores (Mazor 2015) and hybrid (H) and diatom (D) ASCI scores (Theroux et al. 2020) were used to evaluate the BMI and algae data analyzed in this report (Table 3). The thresholds for both indices were based on the distribution of scores for data collected at reference calibration sites located throughout California. Four condition categories are defined by these thresholds: “likely intact” (greater than 30th percentile of reference site scores); “possibly altered” (between the 10th and the 30th percentiles); “likely altered” (between the 1st and 10th percentiles); and “very likely altered” (less than the 1st percentile) (Figure1).

A CSCI score below 0.795 is referenced in the MRP as a threshold indicating a potentially degraded biological community, and thus should be considered for a Stressor Source Identification Project. The MRP threshold is the division between the “possibly altered” and “likely altered” condition categories (below the 10th percentile of reference site scores) described in Mazor (2015). Further investigation is needed to evaluate the applicability of this threshold to sites in highly urban watersheds and/or modified channels that are common throughout the Bay Area.

Table 3. Biological condition indices, categories and thresholds.

Index	Likely Intact	Possibly Altered	Likely Altered	Very Likely Altered
Benthic Macroinvertebrates (BMI)				
CSCI Score	≥ 0.92	≥ 0.79 to < 0.92	≥ 0.63 to < 0.79	< 0.63
Benthic Algae				
D ASCI Score	≥ 0.94	≥ 0.86 to < 0.94	≥ 0.75 to < 0.86	< 0.75
H ASCI Score	≥ 0.94	≥ 0.86 to < 0.94	≥ 0.75 to < 0.86	< 0.75

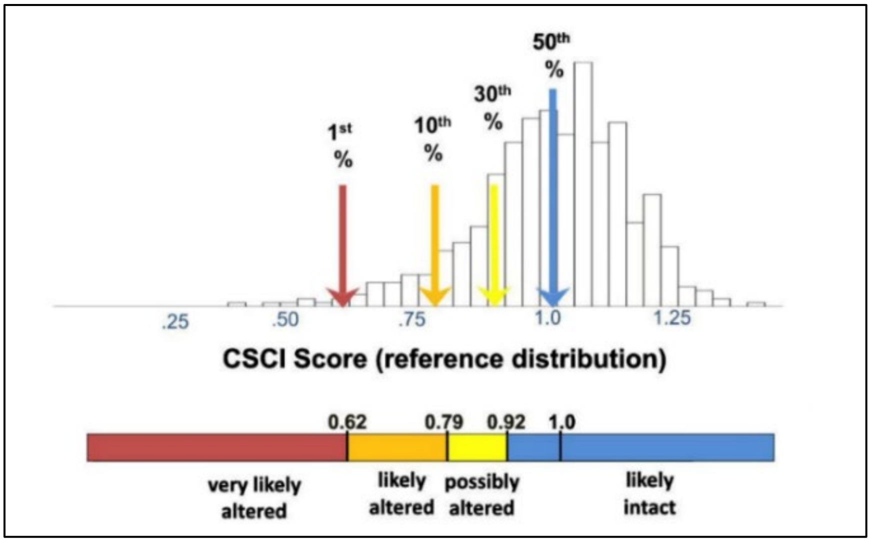

Figure 1. Distribution of CSCI scores at reference sites with thresholds and condition categories used to evaluate CSCI scores in California streams (from Rehn et al. 2015). Note: colors in this figure differ from other figures in this report.

2.5.3 Estimating Extent of Healthy Streams in SF Bay Area

To estimate overall extent of biological conditions in streams within the five-county RMC area, cumulative distribution functions (CDFs) of biological condition scores were generated. Because the survey focused significantly more effort in urban areas compared to non-urban areas, sample weights were re-calculated as the total stream length in the sample frame and divided by the stream length evaluated in each land use category. Therefore, sites contribute a proportional amount of stream length to the extent estimates, based on the number of sites assessed in each land use category. Sites without evaluations, primarily non-urban sites, were excluded from the analysis. The adjusted sample weights were used to estimate the proportion of stream length represented by CSCI, D ASCI, and H ASCI scores both regionwide and for urban sites only. Estimates for non-urban streams were not calculated separately due to the lower number of monitoring events at non-urban sites and differences in prioritization in sampling of urban sites among programs. Condition estimates and 95% confidence intervals were calculated for all urban sites sampled sites in the RMC sample frame. Post-stratification of the urban sites by County was also performed. All calculations were conducted using the R-package *spsurvey* (Kincaid and Olsen 2016). See Section 2.4 for further discussion of the RMC sample design.

2.5.4 Spatial Analyses and Mapping

Evaluation classes and biological conditions of streams (CSCI, D ASCI, and H ASCI) were plotted for Alameda, Contra Costa, San Mateo, Santa Clara, and Solano counties in ArcGIS Pro (3.1.3). Biological condition scores were grouped by color where green represents likely intact conditions, yellow represents possibly altered conditions, red represents likely altered conditions, and purple represents very likely altered conditions. Urbanization and hydrographic layers were added to illustrate the differences in the biological conditions between urban and non-urban streams. Mapping of condition scores were also illustrated using the channel classification categorization. CSCI, D ASCI, and H ASCI thresholds were plotted with the channel classifications to illustrate the biological conditions of the three major channel types.

2.5.5 Evaluating the Importance of Stressors

Stressor association with biological condition scores was evaluated using random forest statistical analyses. Random forest analysis is a non-parametric classification and regression tree (CART) method commonly applied to large datasets of multiple explanatory variables. Recent papers describe its use for stressor identification in stream bioassessment studies (e.g., Maloney et al. 2009, Waite et al. 2012, Mazor et al. 2016). Random forest models use bootstrap averaging to determine splits of numerous trees (Elith et al. 2008) for reducing error and optimizing model predictions. Model outputs provide an ordered list of importance of the explanatory variables that can be applied to a new or validation dataset for prediction.

Random forest models were developed using the R-package *randomForest* to determine a list of explanatory variables related to biological condition scores (CSCI or D ASCI score). Only urban sites were considered in the analysis. The stressor data consisted of 26 variables, related to (1) water quality; (2) habitat; and (3) land use factors that could potentially influence condition scores (Appendix 1, Table A). Subsequently, the data were partitioned into training (80%) and validation (20%) sets for model testing. A random selection of samples was generated by sub-sampling from within each RMC County to maintain a regional balance of samples within the partitioned datasets. The training dataset had 496 sites, while the validation data encompassed 128 sites across the five counties.

First, several iterations of the model procedure were performed with the training data set to optimize the random forests, including tuning the model to the maximum number of predictors per branch, the number of trees to build, and validation of the predictions. The final set of models evaluated a maximum of 6 predictor interactions, and 1000 trees. Two variable importance statistics were used to estimate the relative influence of predictor variables: (1) % Increase in Mean Square Error (MSE) = percent increase in mean-square-error of predictions as a result of variable values being permuted; (2) Increase in Node Purity = difference between the residual sum-of-squares before and after a split in the tree. More important variables achieve larger changes in MSE and node purity. K-fold cross validation of the selected models was performed to assess prediction error, by evaluating residual error and R-squared differences.

Random forest models were developed with all variables included (N = 26), retaining the top 5 variables in the variable relative importance list ranked by % increase in MSE. The top-5 variable list was scrutinized by evaluating the corresponding variable importance scores, partial dependency plots, and the change in R^2 once the variable was excluded. No variable with less than 15% influence on CSCI or D ASCI predictions was retained in the final models.

2.5.6 Evaluating the Importance of Flow Regime and Channel Classification

Environmental factors such as flow status (perennial vs non-perennial streams) and channel type (hard, soft, natural) were examined as additional factors that may covary with biological condition. Although flow status appeared to have some correlation with biological index scores, the association was relatively weak compared to other variables. Condition categories for each of the biological indices were grouped by channel conditions for both the region and county spatial scales. Analysis of variance (ANOVA) tests were conducted to investigate the contribution of these variables and potential interactions. Significance of model parameters was assessed at $p < 0.05$.

3 Results

3.1 Probabilistic Site Evaluation Results

A total of 617 probabilistic sites were sampled in the RMC region between 2012 and 2022. These are identified as “Accepted” sites in Figure 2 and Table 4. Of the probabilistic sites, 527 were classified as urban (85%) and 92 as non-urban (15%) (Table 4). The greatest number of non-urban sampling locations were in Santa Clara (n=31) and San Mateo Counties (n=29). Samples were collected at 7 to 14 non-urban sites for each of the other counties. Site evaluation classes were designated as:

1. Accepted – A site that met the established water body type, flow, and access criteria and was assessed.
2. Rejected (Non-target) – A site that was rejected due to lack of flow, incorrect water body type (e.g., was not a stream), tidal-influence, or non-wadeable and was not assessed.
3. Rejected (Not Accessible) – A site that met the established criteria but was not assessed due to physical barriers to access or lack of permission.

The population of 617 monitored sites was obtained through the evaluation of 2,443 unique sites, which equates to a rejection rate of 75% for the entire RMC region over the 11-year period. Solano County had the highest rejection rate (87%) and San Mateo County had the lowest (67%). The most common reason for site rejection (75% of all evaluated sites) was that a site did not present the physical requirements to support monitoring within a 300-meter radius of target coordinates. These “rejected” sites were not assessed for several reasons, including lack of flowing water, site was not a stream (e.g., aqueduct or pipeline), tidally influenced, or non-wadeable. Lack of flow was the most common reason for rejection.

Another reason for site rejection was the inability to access the reach selected for sampling (e.g., physical access or obtain private land/permission). A total of 24% of the sites rejected were located on private land in non-urban areas where permissions were not granted and/or where steep, highly-vegetated conditions prevented access. Obtaining access to sites in urban areas was variable by county. For example, most of the streams in the urban area of San Mateo County are privately owned, while most of the urban sites in Santa Clara County are owned by municipal jurisdictions and water district agencies, making permissions more easily obtained.

Table 4. Number of probabilistic sites per county in each site evaluation class.

County	Accepted (Assessed)		Rejected (Non-target; Not Assessed)		Rejected (Not Accessible; Not Assessed)		Total by County
	Non-Urban	Urban	Non- Urban	Urban	Non- Urban	Urban	
Alameda	31	115	193	158	14	182	691
Contra Costa	35	80	49	212	7	108	493
San Mateo	49	75	10	71	29	69	303
Santa Clara	50	63	109	320	32	146	720
Solano	44	6	109	47	8	22	236
Total	**209**	**339**	**470**	**808**	**90**	**527**	**2,443**
% of Total	9%	15%	19%	33%	4%	22%	

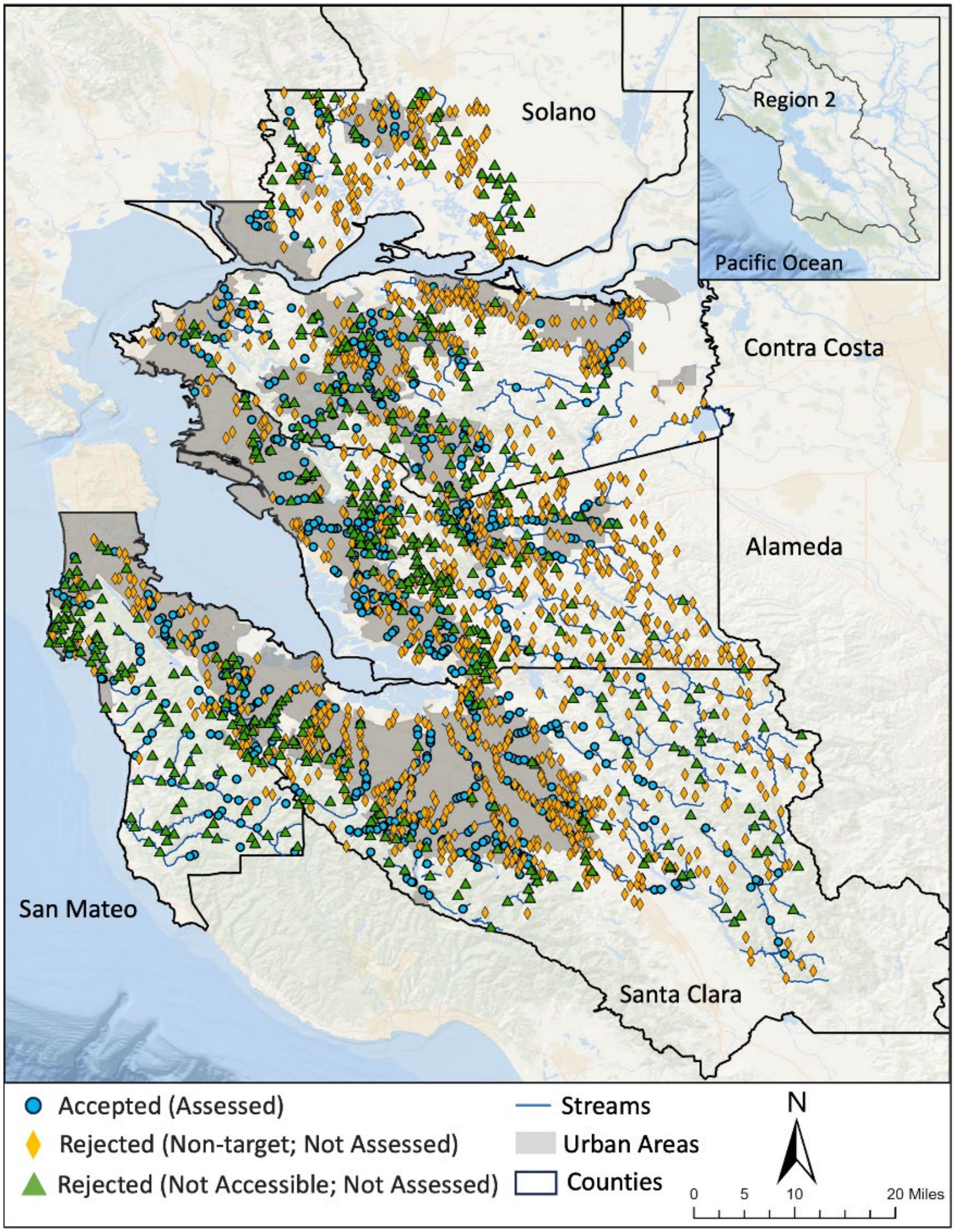

Figure 2. Regional bioassessment sites identified by evaluation class; accepted, rejected (non-target), or rejected (not accessible).

Figure 3 presents annual rainfall for the period of Water Year 1946 – 2022 at the San Francisco Airport. During the sampling period (2012 – 2022), rainfall was generally below the long-term average, with 9 of the 11 years drier-than-normal. The three driest years (all below 10 inches of precipitation) in the record occurred during WY 2014, WY 2020 and WY 2021, with WY 2021 one of the driest on record. Because biological condition index scores can vary naturally due to multi-year climatic patterns, it is important to note that the 11-year period of record may not be representative of the long-term biological condition.

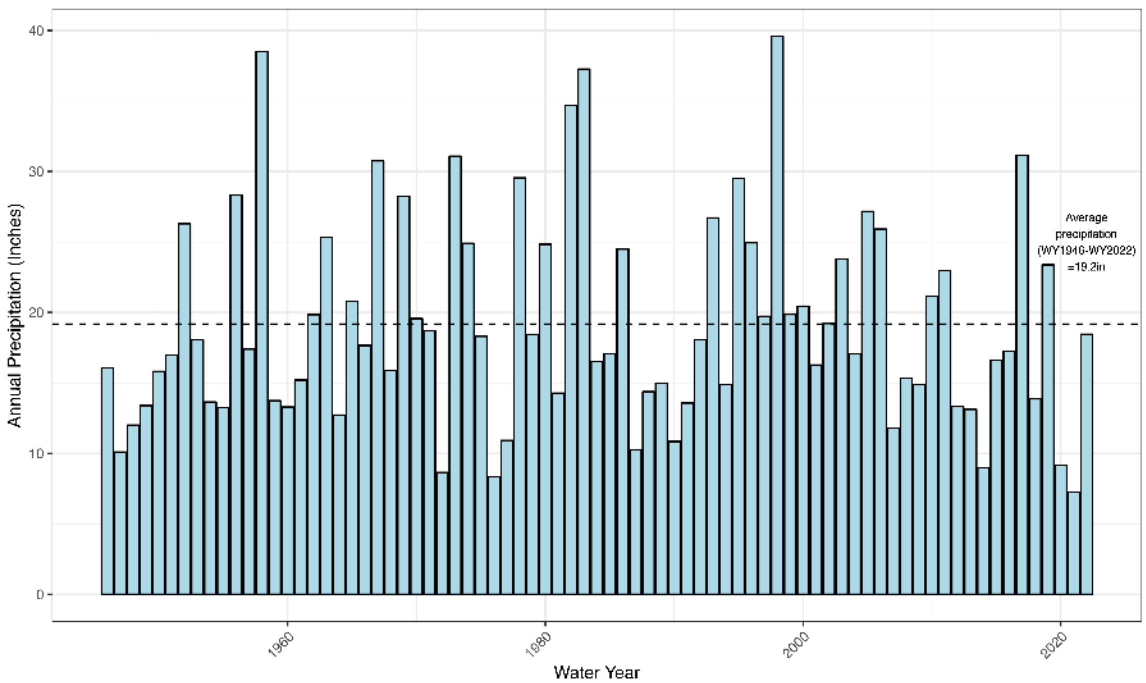

Figure 3. Annual precipitation at San Francisco Airport (1946-2022).

3.2 Biological Condition of Bay Area Streams

3.2.1 Condition Assessment for Urban Streams (Probabilistic sites only)

The distribution of biological index scores suggests that a majority of the urban streams in the five-county RMC area do not exhibit healthy biological conditions. Cumulative distribution functions of the three biological index scores (CSCI, D ASCI, H ASCI) for urban sites in the entire SF Bay region are presented in Figures 4-6. The two lowest biological condition classes (Very Likely Altered and Likely Altered) fall below the 10th percentile of reference conditions, which is shown as a vertical line in the graphs. For streams within the SF Bay Region, 89% of the stream-length was below the 10th percentile of reference conditions for CSCI (Figure 4). Both of the algae index scores (D ASCI and H ASCI) exhibited a very similar pattern, with 92% and 90%, respectively, of the stream-length falling below the 10th percentile of reference conditions (Figure 5 and 6).

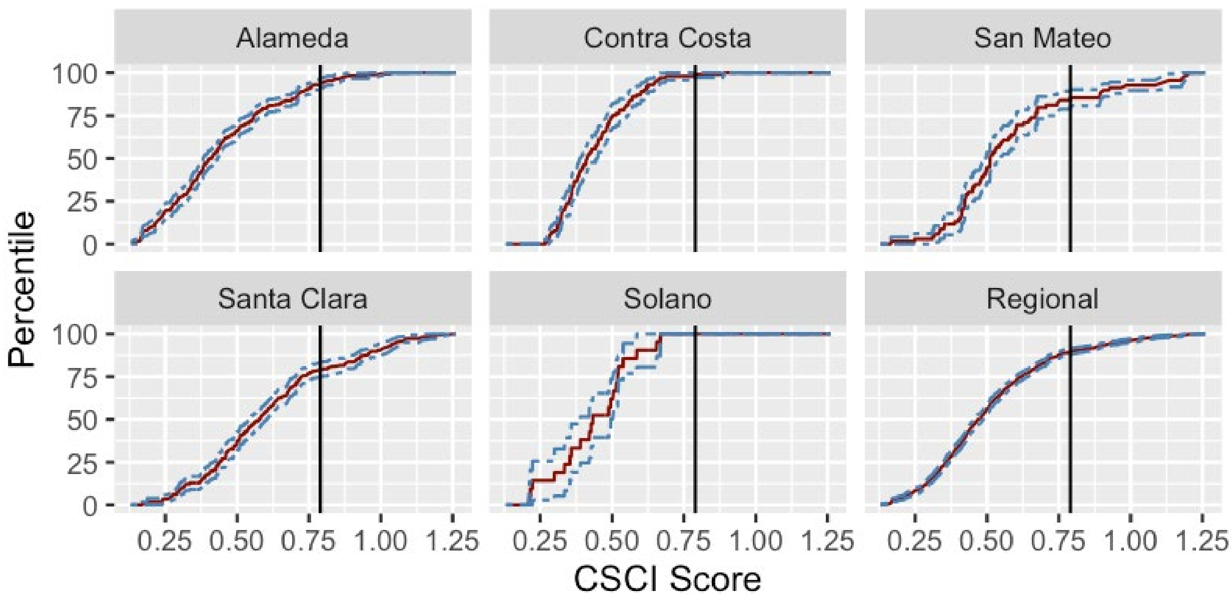

Figure 4. Cumulative distribution function (CDF) of CSCI scores at all urban sites combined and by County (probabilistic sites only; n = 490). Solid line depicts the CSCI threshold for 10th percentile of reference condition (Mazor et al. 2016).

A comparison of biological conditions in streams within each of the four counties is also presented in Figures 4-6. The proportion of urban stream length below the 10th percentile of reference conditions for CSCI was highest for Solano (100%), Contra Costa (99%), followed by Alameda County (94%), San Mateo County (86%), and Santa Clara County (79%) (Figure 4). Both of the algae index scores (D ASCI and H ASCI) also exhibited poor conditions across the five counties, with D ASCI ranging 88%-95% of stream length below the threshold, and H ASCI ranging 82%-100% of stream length below the threshold.

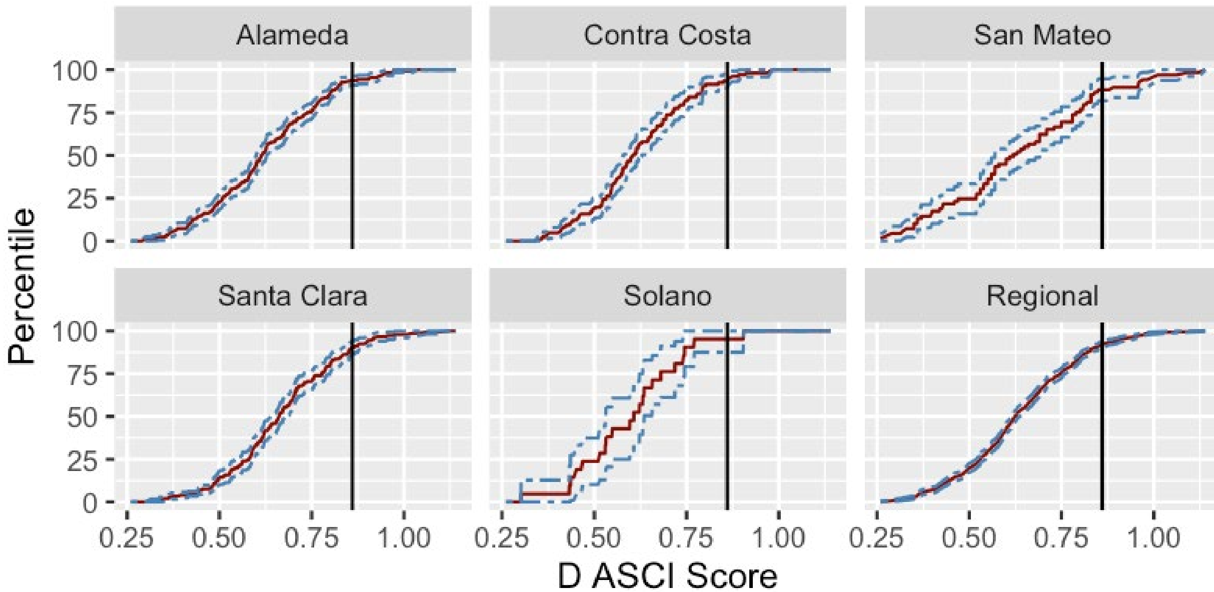

Figure 5. Cumulative distribution function (CDF) of D ASCI scores at all urban sites combined and by County (probabilistic sites only; n = 485). Solid line depicts the D ASCI threshold for 10th percentile of reference condition (Theroux et al. 2020).

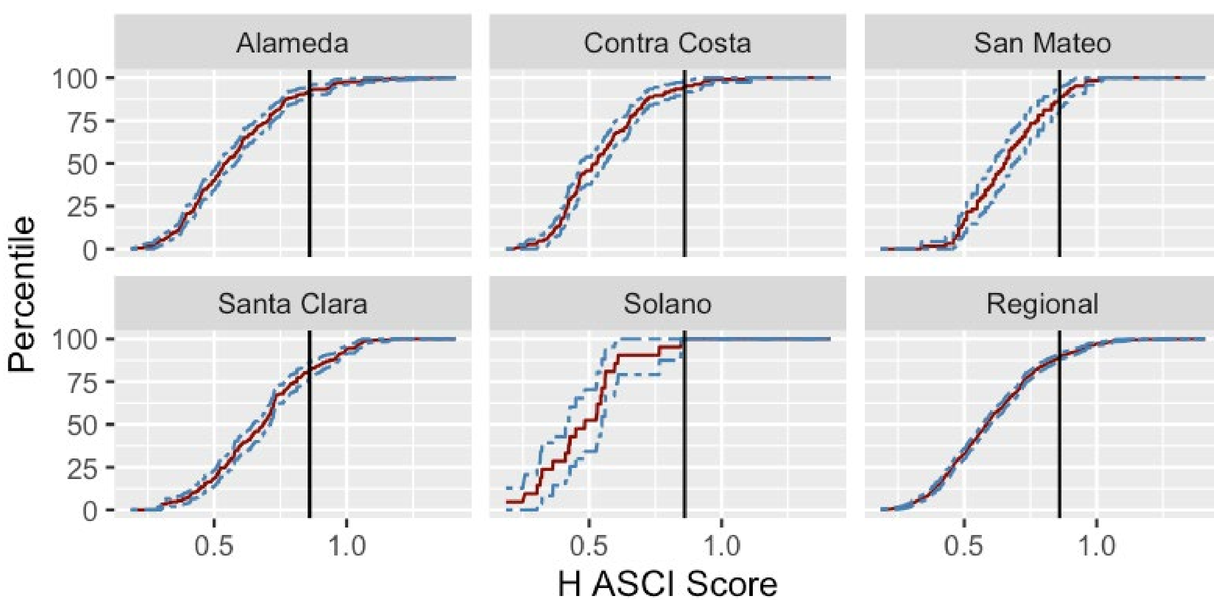

Figure 6. Cumulative distribution function (CDF) of H ASCI scores at all urban sites combined and by County (probabilistic sites only; n = 485). Solid line depicts the D ASCI threshold for 10th percentile of reference condition (Theroux et al. 2020).

3.2.2 Condition Assessment for All Urban and Non-Urban Sites

Regional maps depict the biological condition for urban and non-urban sites in the RMC sample area. Figures 7 through 9 illustrate the biological index scores grouped into four condition categories (Figure 1; Table 3) for every bioassessment site sampled (n = 728) over the past 11 years. County specific maps are provided in Appendix 3. All three sets of index scores (CSCI, D ASCI, H ASCI) indicate most sites corresponded to the Very Likely Altered category (purple symbols). Biological conditions from non-urban areas were predominantly responsible for sites with index scores in the Likely Intact or Possibly Altered condition categories. There was also no apparent regional difference to the distribution of condition score amongst the three indices.

CSCI scores grouped by land use class (urban vs. non-urban) showed that all counties, with the exception of Solano, exhibit higher scores in non-urban areas (Figure 10). The CSCI score for non-urban sites in Santa Clara and San Mateo counties were higher compared to the other counties, with median CSCI score above the level corresponding to 10^{th} percentile of reference conditions (< 0.79). Many sites in these counties had scores at or above 1.0, which is considered reference condition. However, there were also many non-urban sites that had CSCI scores that were below the 10^{th} percentile of reference, indicating degraded conditions.

Generally, algae scores were similar to the CSCI results, with the non-urban area exhibiting higher scores than sites in urban areas within each county (Figures 11 and 12). Although the low sample sizes of the non-urban sites preclude making any definitive comparisons, it was noteworthy that some sites in the urban areas may exhibit similar or higher algae index scores than sites in non-urban areas.

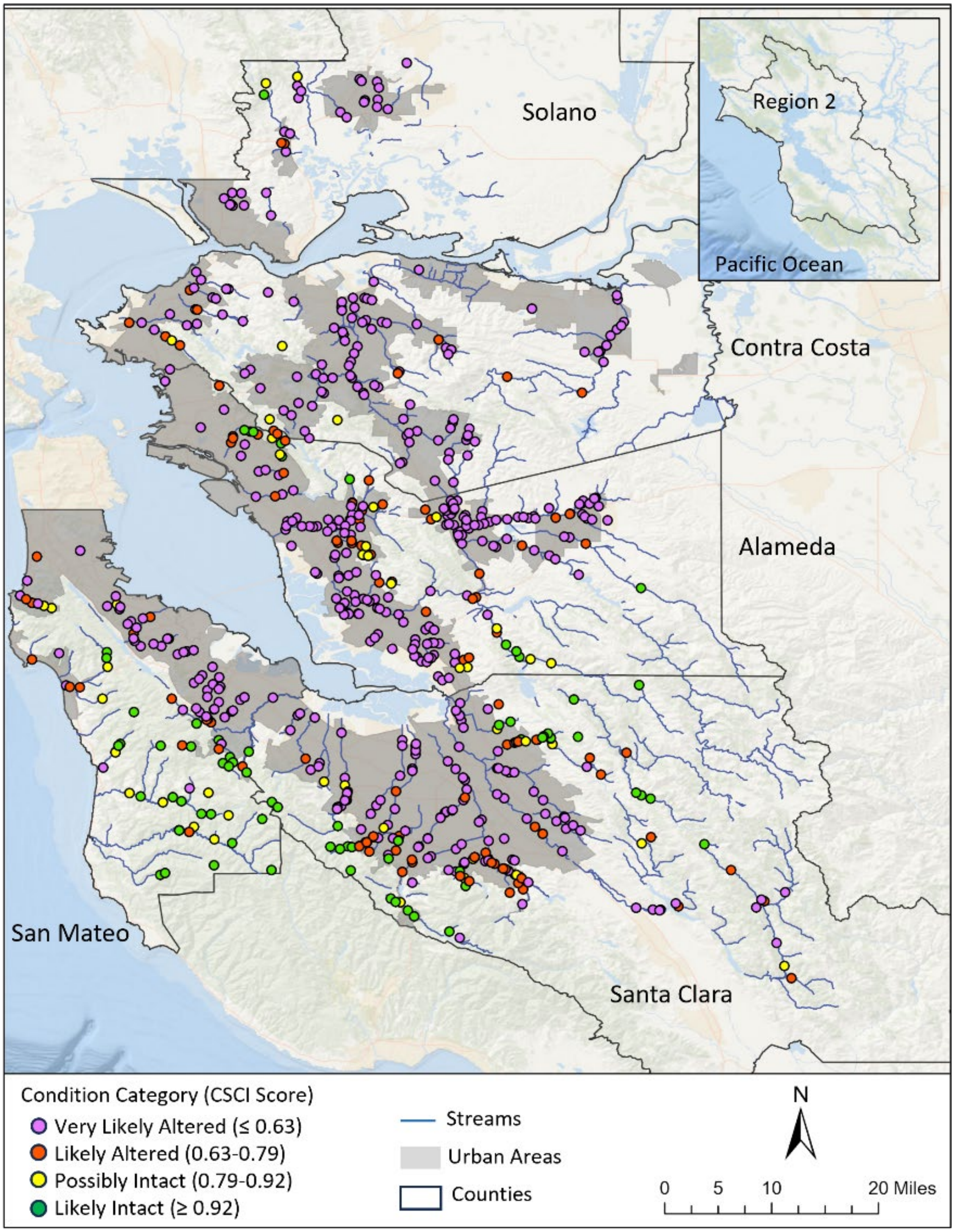

Figure 7. Biological condition of streams in the RMC area based on CSCI scores.

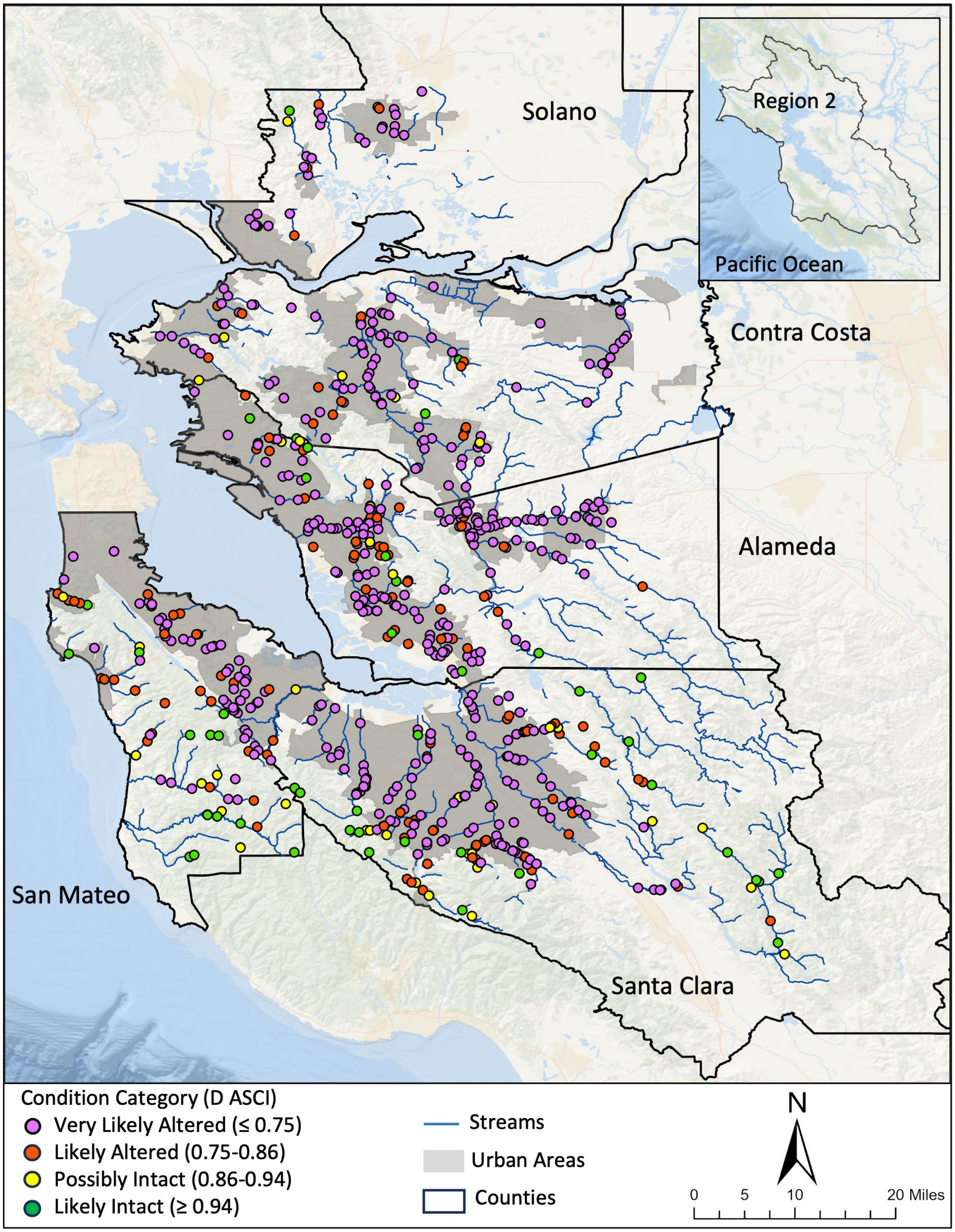

Figure 8. Biological condition of streams in the RMC area based on D ASCI scores.

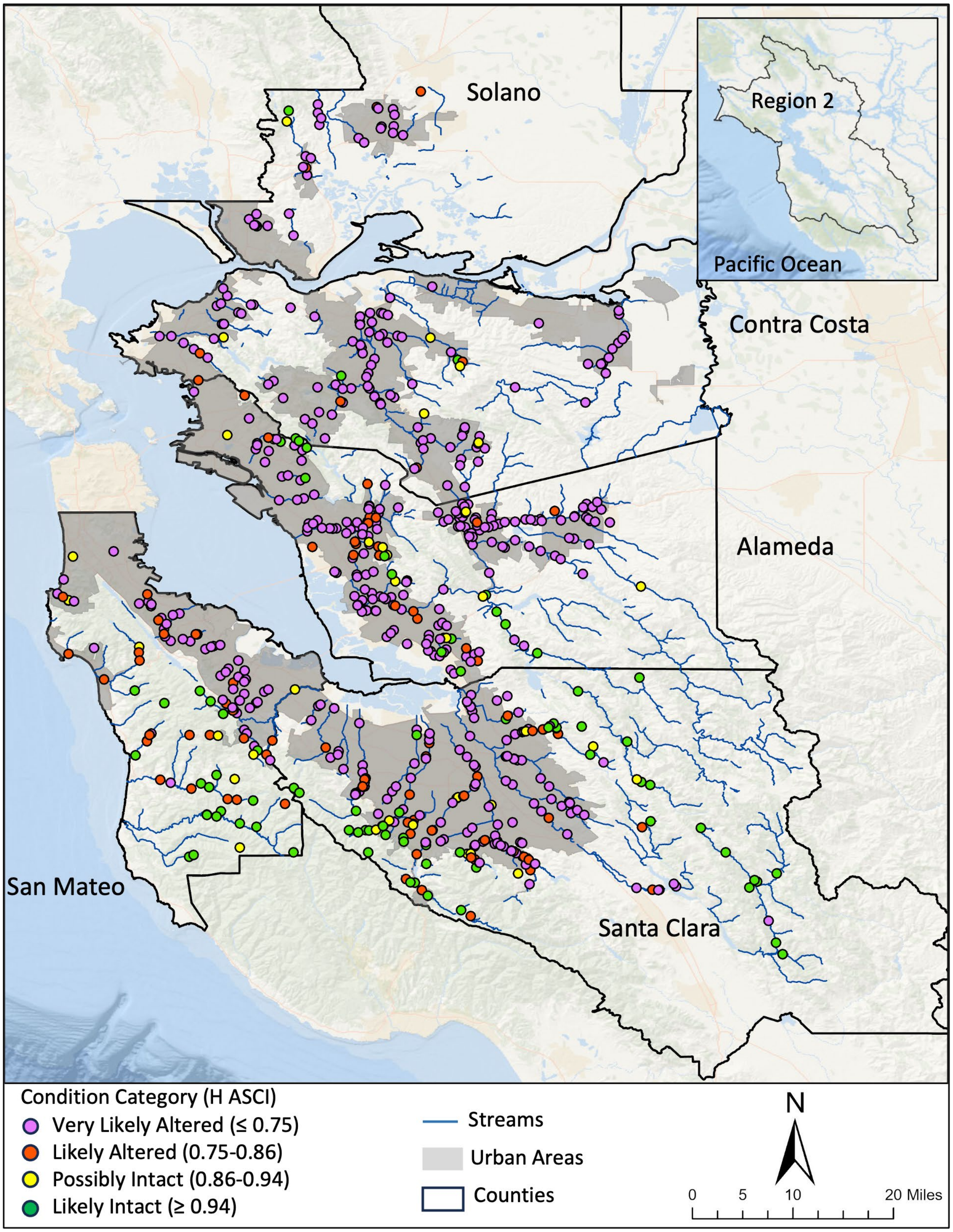

Figure 9. Biological condition of streams in the RMC area based on H ASCI scores.

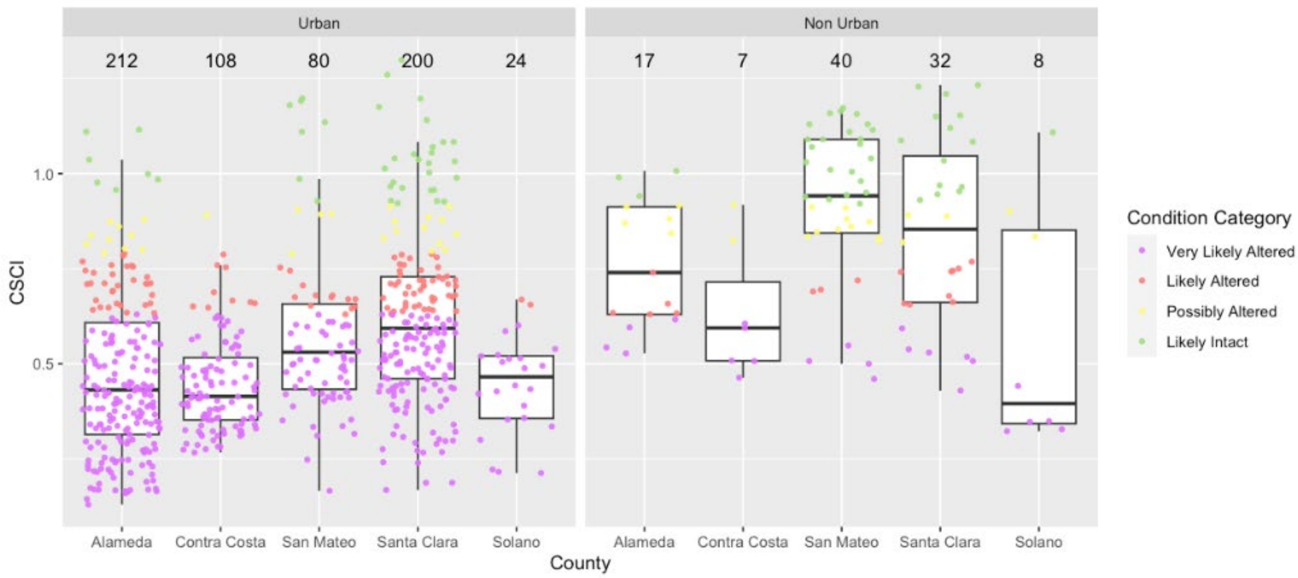

Figure 10. CSCI scores for urban and non-urban sites in each County. Sample size (N) shown above plot by County and land use.

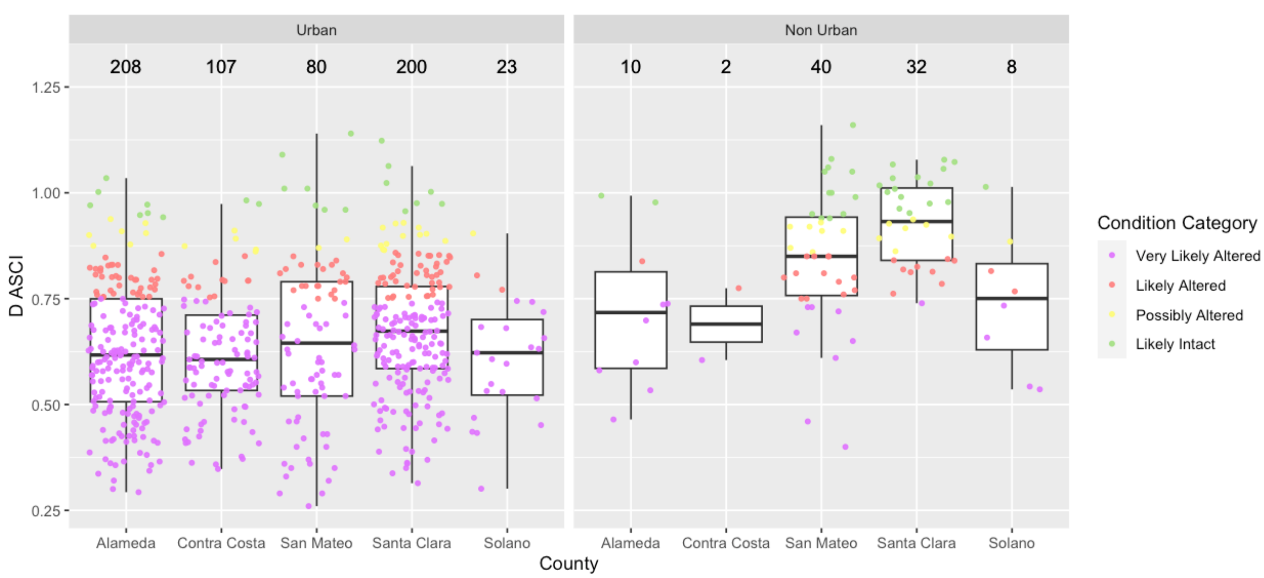

Figure 11. D ASCI scores for urban and non-urban sites in each County. Sample size (N) shown above plot by County and land use.

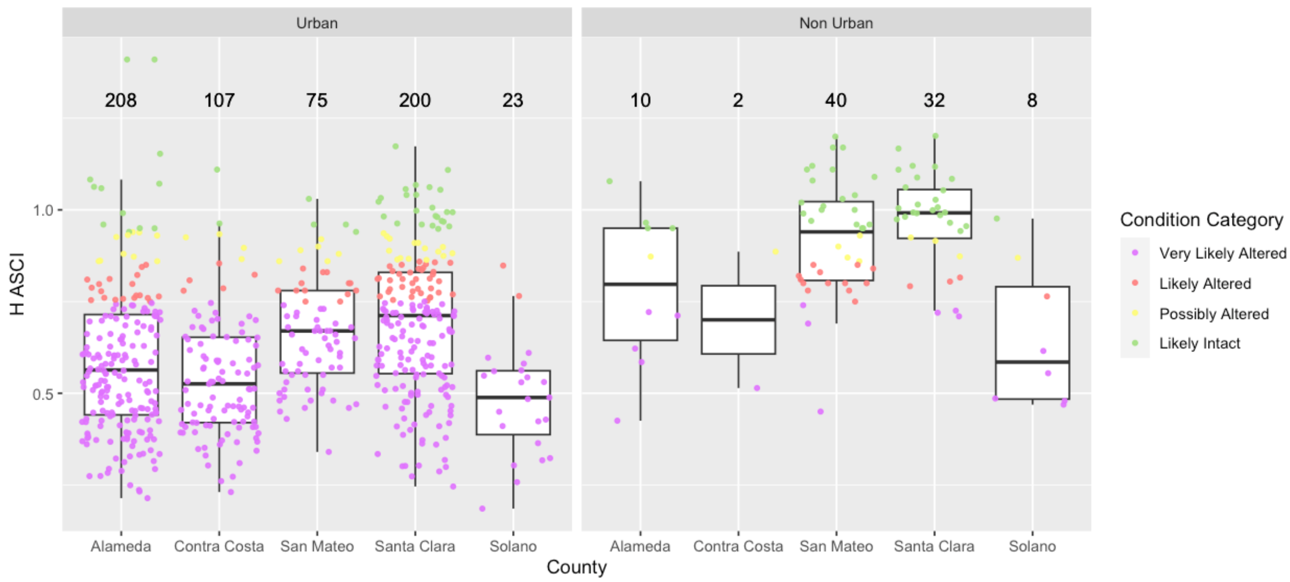

Figure 12. H ASCI scores for urban and non-urban sites in each County. Sample size (N) shown above plot by County and land use.

Biological Conditions by Flow Status

Biological conditions of urban sites, depicted by the four condition categories, were grouped by flow status[4] (perennial vs non-perennial). Condition category is represented by four colors; green represents likely intact conditions, yellow represents possibly altered conditions, red represents likely altered conditions, and purple represents very likely altered conditions. Biological conditions based on CSCI scores for perennial sites were slightly lower compared to non-perennial sites (Figure 13). There was a higher proportion of perennial urban sites with scores in the Very Likely Altered and Likely Altered condition classes; 90%; n = 515) compared to the non-perennial sites (83%; n = 97). This pattern was consistent in Alameda (93% vs. 87%) and Santa Clara County (82% vs. 74%) but less apparent for Contra Costa (99% vs. 100%) and San Mateo (86% vs. 88%), where there was a highly similar proportion of degraded perennial as non-perennial sites.

Overall, there was a similar regional pattern in D ASCI and H ASCI scores between perennial and non-perennial sites. The D ASCI scores showed better scores for non-perennial sites, except for sites in San Mateo County, which had generally higher scores at perennial sites (Figure 14). However, the pool of non-perennial sites was also relatively small for the San Mateo County portion of the RMC sample area (n = 8 sites). The scoring range for H ASCI showed a general pattern of higher scores at non-perennial sites across all counties (Figure 15). It should be noted that many of the urban sites in Santa Clara County (and a small number throughout the RMC) are perennial due to upstream reservoir operations; however, the dataset did not include this classification and no comparisons were made between reservoir-caused vs. natural perenniality.

[4] Flow status data were not available for sites in Solano County.

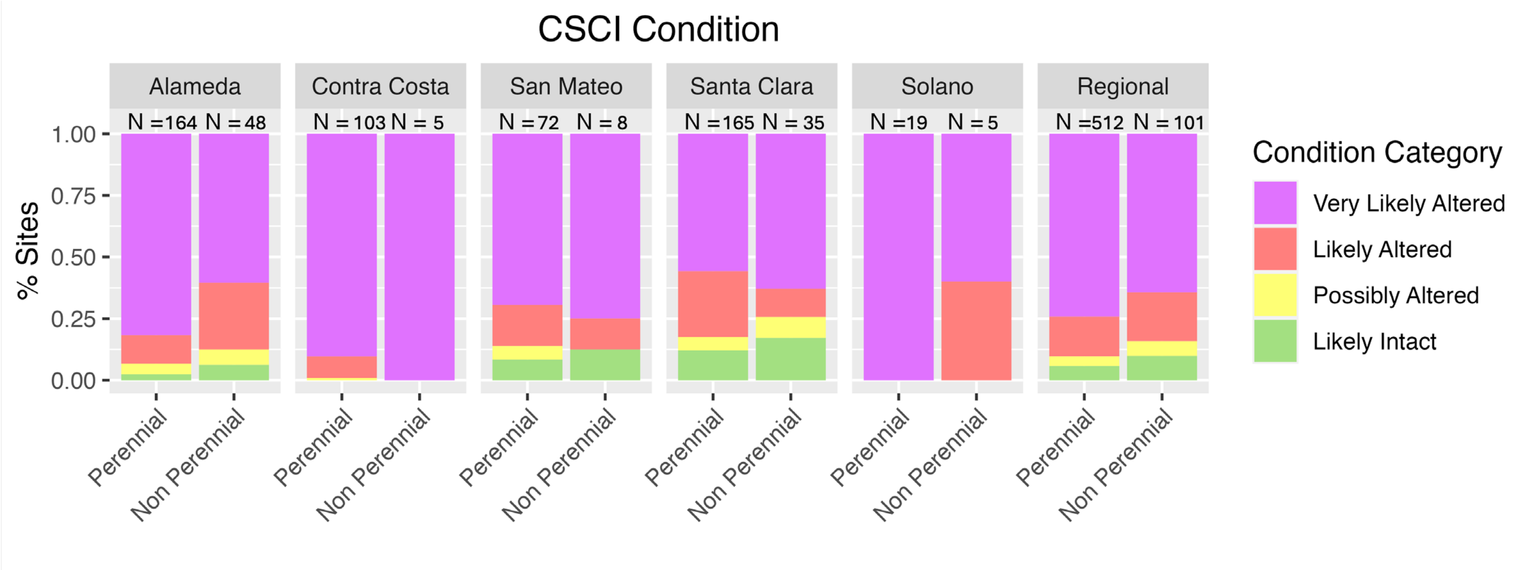

Figure 13. CSCI condition for all urban sites (Regional) and for urban sites by County grouped by flow status. The colors in the plot indicate the proportion of sites in each CSCI condition category. Sample size (N) shown above plot by County and flow status.

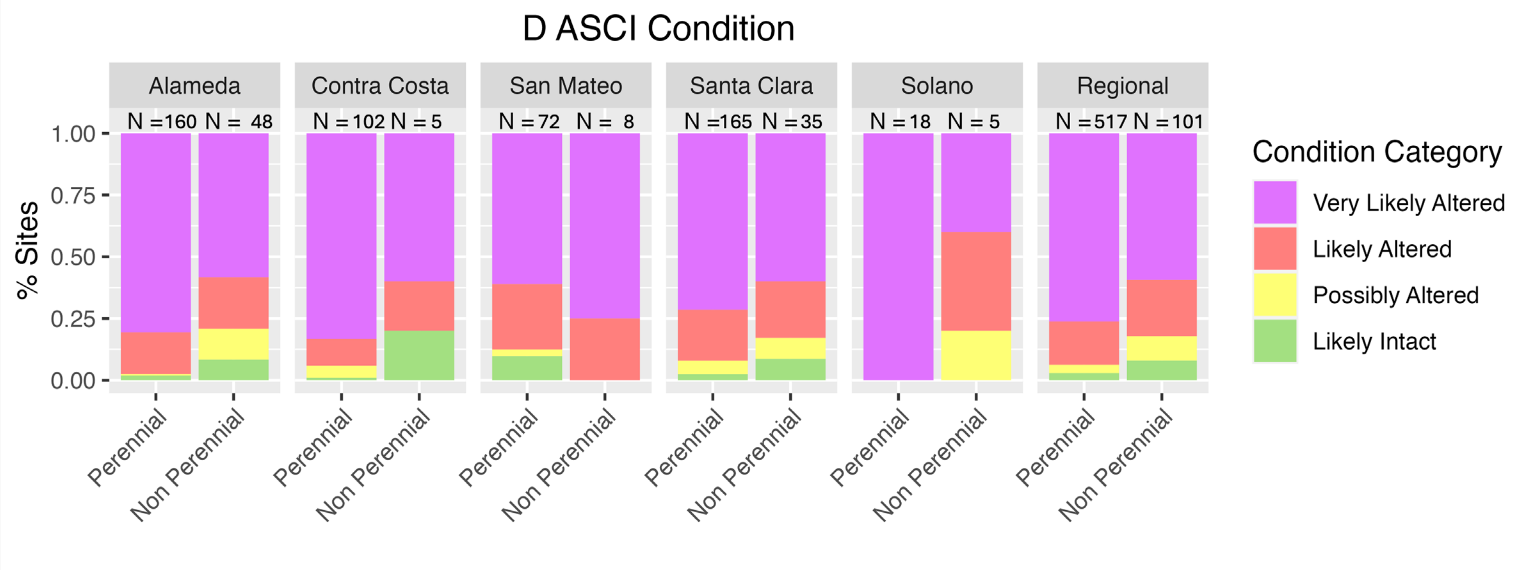

Figure 14. D ASCI condition by county and flow status. The colors in the plot indicate the proportion of sites in each D ASCI condition category. Sample size (N) shown above plot by County and flow status.

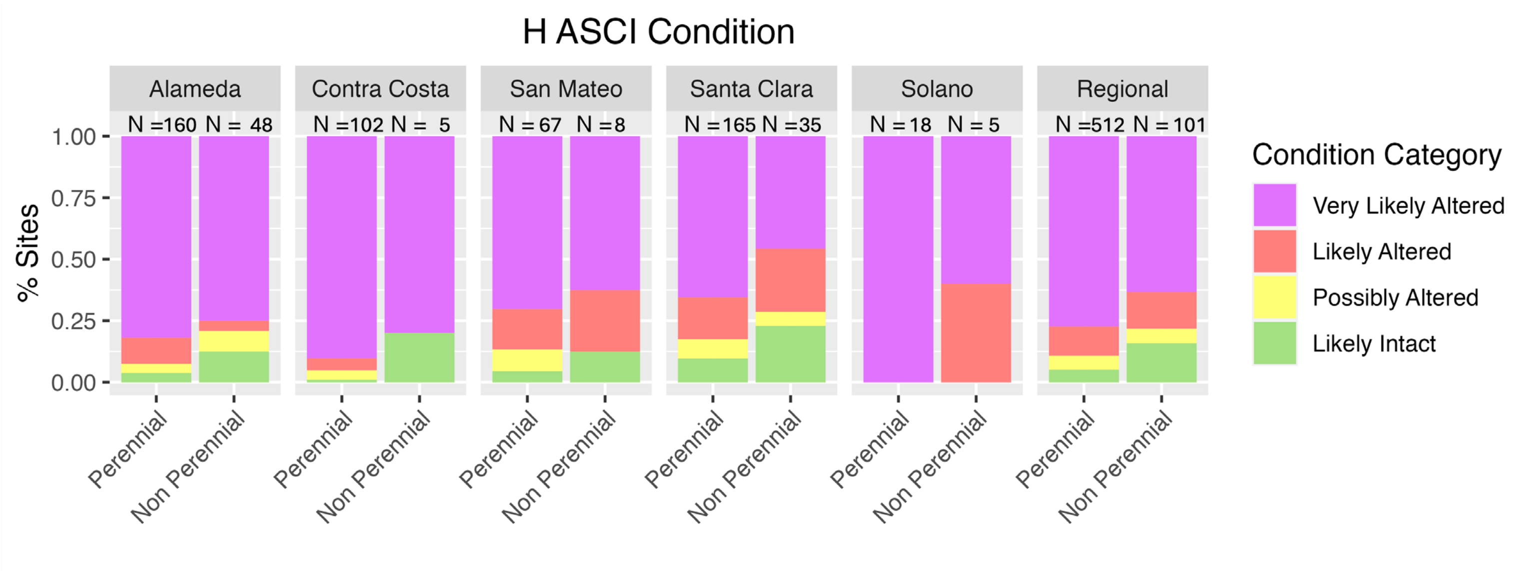

Figure 15. H ASCI condition by county and flow status. The colors in the plot indicate the proportion of sites in each H ASCI condition category. Sample size (N) shown above plot by County and flow status.

3.2.3 Biological Conditions in Modified Channels at Urban Sites

The biological conditions of urban sites, depicted by the four condition categories, were grouped by channel type[5] (hard, soft bed, and natural). The percentage of sites within each condition category for CSCI score is shown in Table 5. All sites with hardened beds had biological conditions in the Very Likely Altered class. Ninety-five percent (95%) of sites that had a modified channel with soft bottom were in the lower two biological condition classes (likely altered and very likely altered). Natural channel had the greatest proportion of sites (15%) in the "likely intact" condition category, however urban sites with natural channels still had over 75% of sites in the two lower condition classes.

Table 5. Summary of CSCI condition scores by channel type for all urban sites (n = 600)

Index	Channel type	N	Likely intact	Possibly altered	Likely altered	Very Likely altered	Mean	Standard Deviation (SD)
CSCI	Hard	47	0%	0%	0%	100%	0.421	0.09
CSCI	Soft Bed	338	2%	3%	11%	84%	0.460	0.18
CSCI	Natural	215	15%	8%	31%	46%	0.665	0.22

Figures 16-17 illustrate the proportion of sites in each for CSCI condition category grouped by channel type.

The percentage of sites within each condition category for D ASCI and H ASCI score is shown in Table 6. Over 90% of sites with hardened beds had biological conditions based on either ASCI score in the Very Likely Altered class. Eighty percent (80%) of sites that had a modified channel with soft bottom were in the lower two biological condition classes (likely altered and very likely altered) for both algae indices. Natural channel had the greatest proportion of sites (13%) in the "likely intact" condition category for H ASCI.

Table 6. Summary of ASCI condition scores by channel type for all urban sites (n = 595)

Index	Channel type	N	Likely intact	Possibly altered	Likely altered	Very Likely altered	Mean	Standard Deviation (SD)
D ASCI	Hard	47	4%	2%	2%	92%	0.582	0.16
D ASCI	Soft Bed	335	1%	2%	16%	80%	0.608	0.15
D ASCI	Natural Channel	213	7%	9%	27%	57%	0.713	0.16
H ASCI	Hard	47	4%	2%	0%	94%	0.545	0.16
H ASCI	Soft Bed	338	4%	5%	10%	81%	0.577	0.18
H ASCI	Natural Channel	208	13%	8%	18%	57%	0.726	0.19

Figures 18 through 21 illustrate the proportion of sites in each for condition category for both D ASCI and H ASCI grouped by channel type.

[5] Channel type data was not available for sites in Solano County.

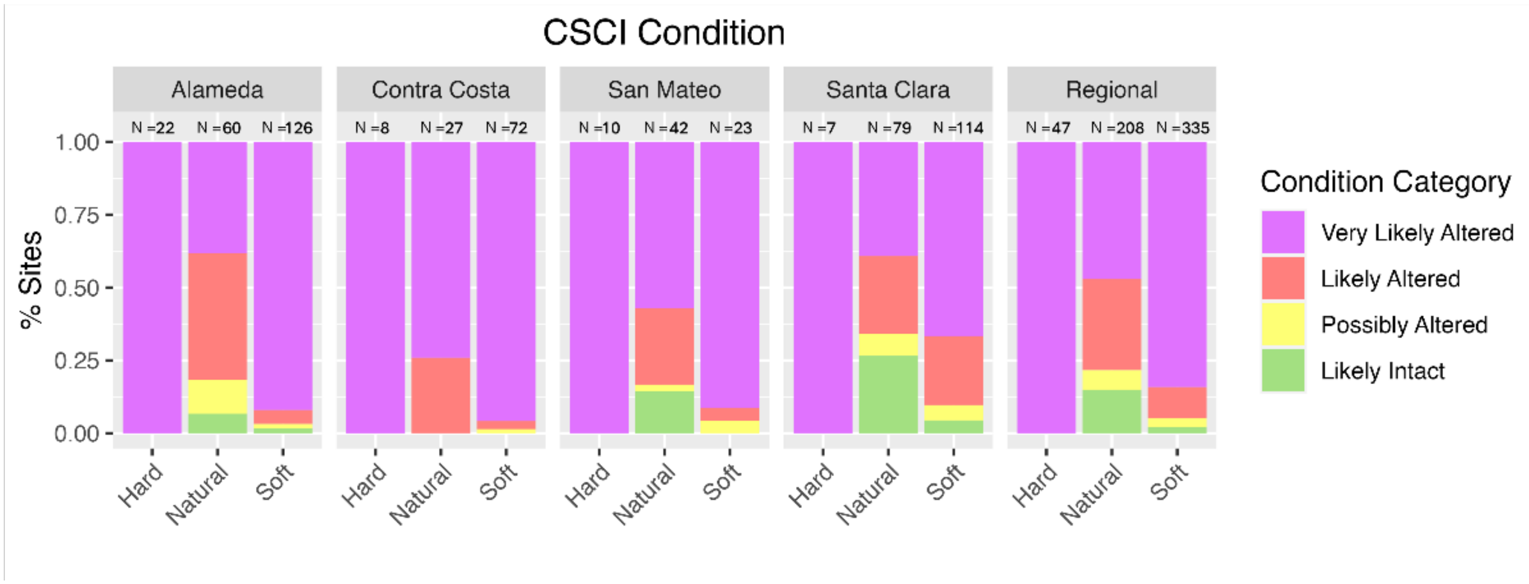

Figure 16. CSCI condition for all urban sites (Regional) and for urban sites by County grouped by channel type. The colors in the plot indicate the proportion of sites in each CSCI condition category. Sample size (N) shown above plot by County and channel type.

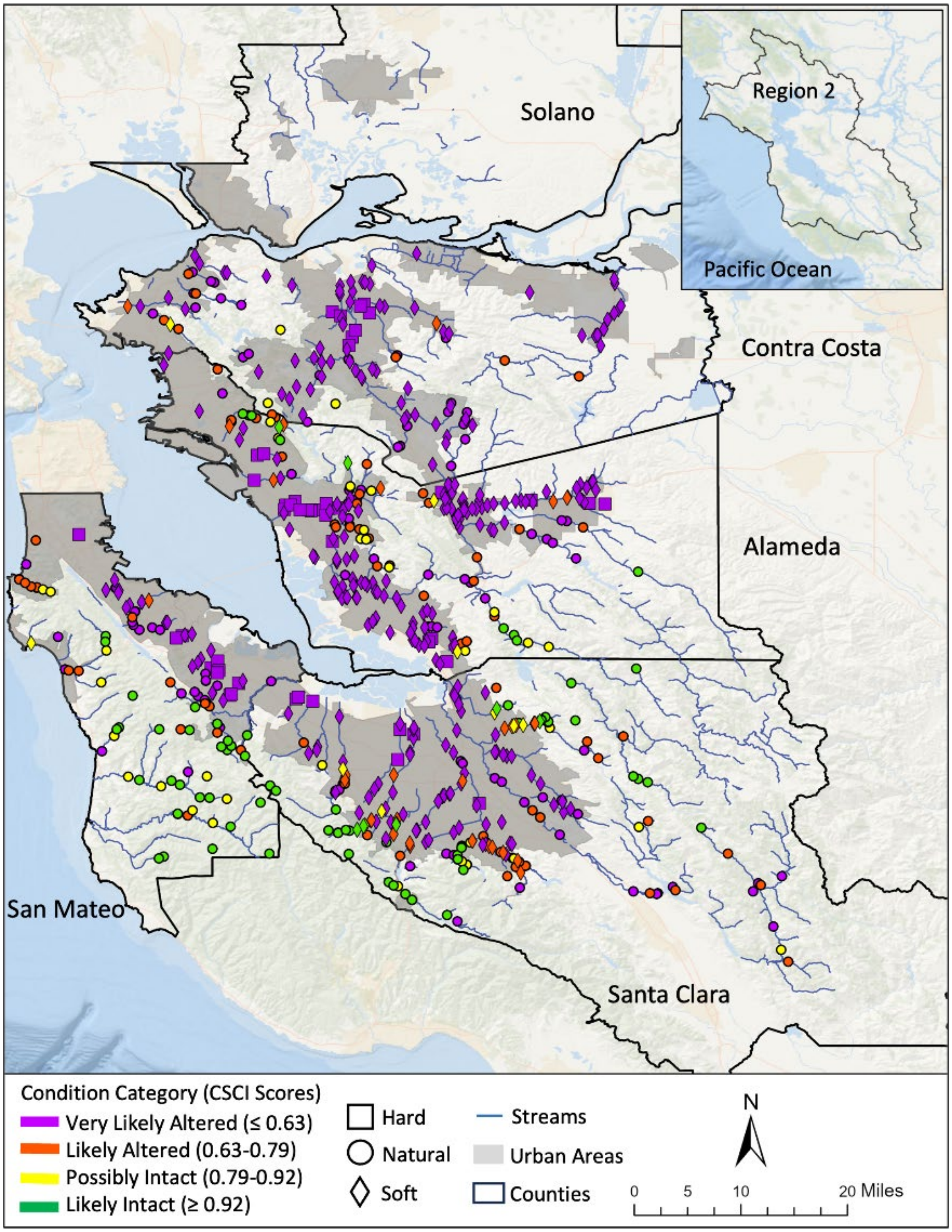

Figure 17. Biological condition of streams in the RMC area based on CSCI scores by channel type category.

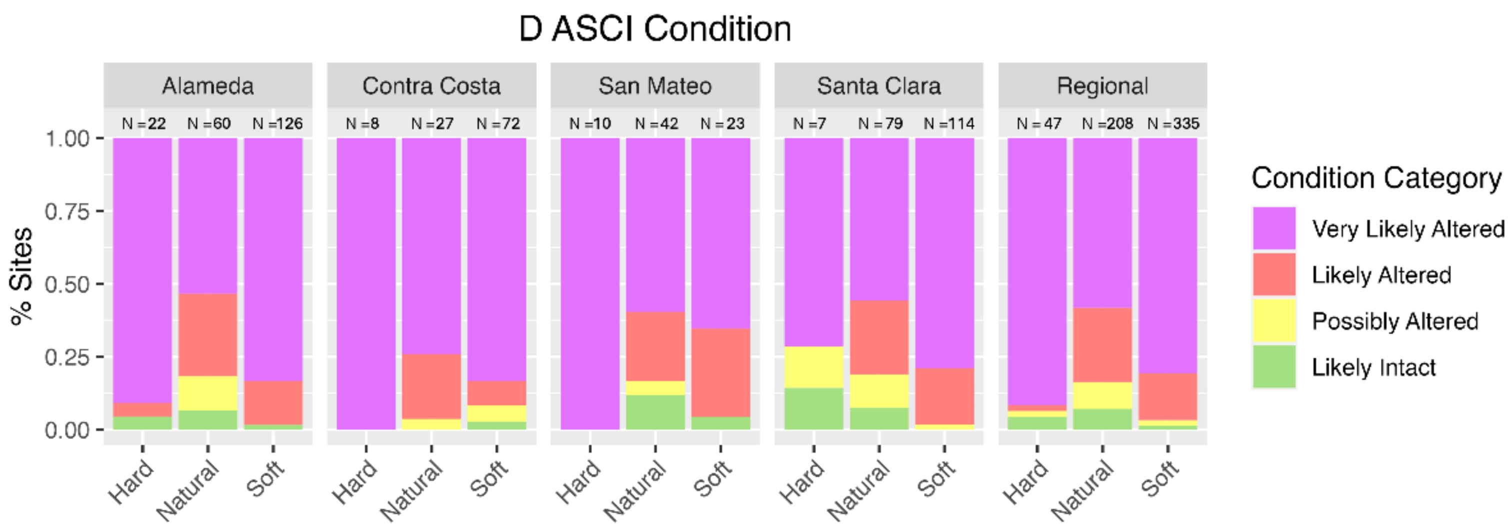

Figure 18. D ASCI condition for all urban sites (Regional) and for urban sites by County grouped by channel type. The colors in the plot indicate the proportion of sites in each D ASCI condition category. Sample size (N) shown above plot by County and channel type.

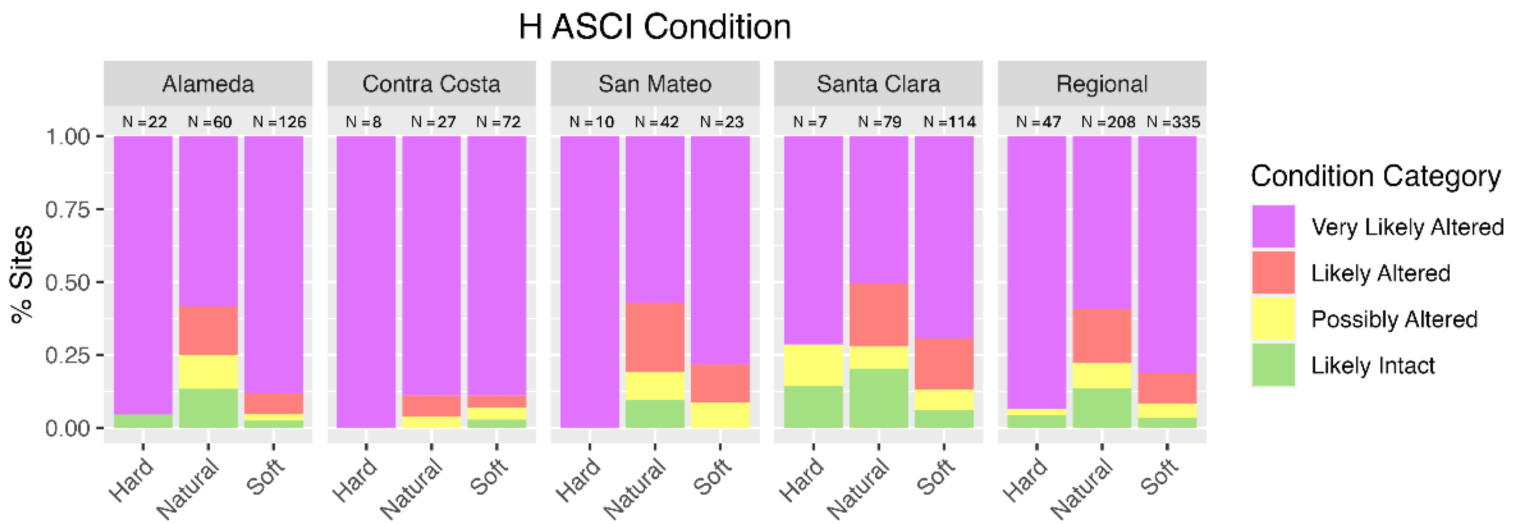

Figure 19. H ASCI condition for all urban sites (Regional) and for urban sites by County grouped by channel type. The colors in the plot indicate the proportion of sites in each H ASCI condition category. Sample size (N) shown above plot by County and channel type.

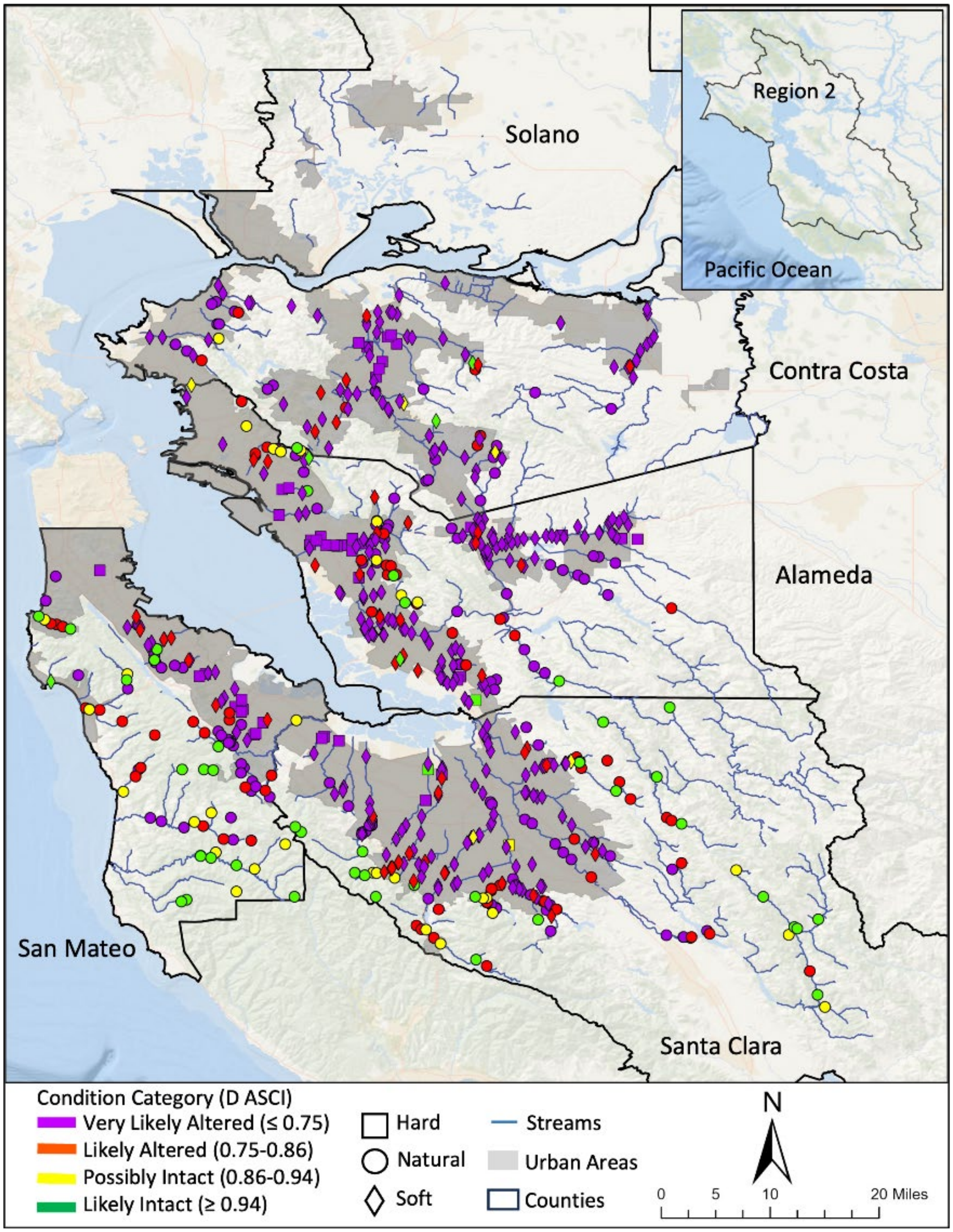

Figure 20. Biological condition of streams in the RMC area based on D ASCI scores by channel type category.

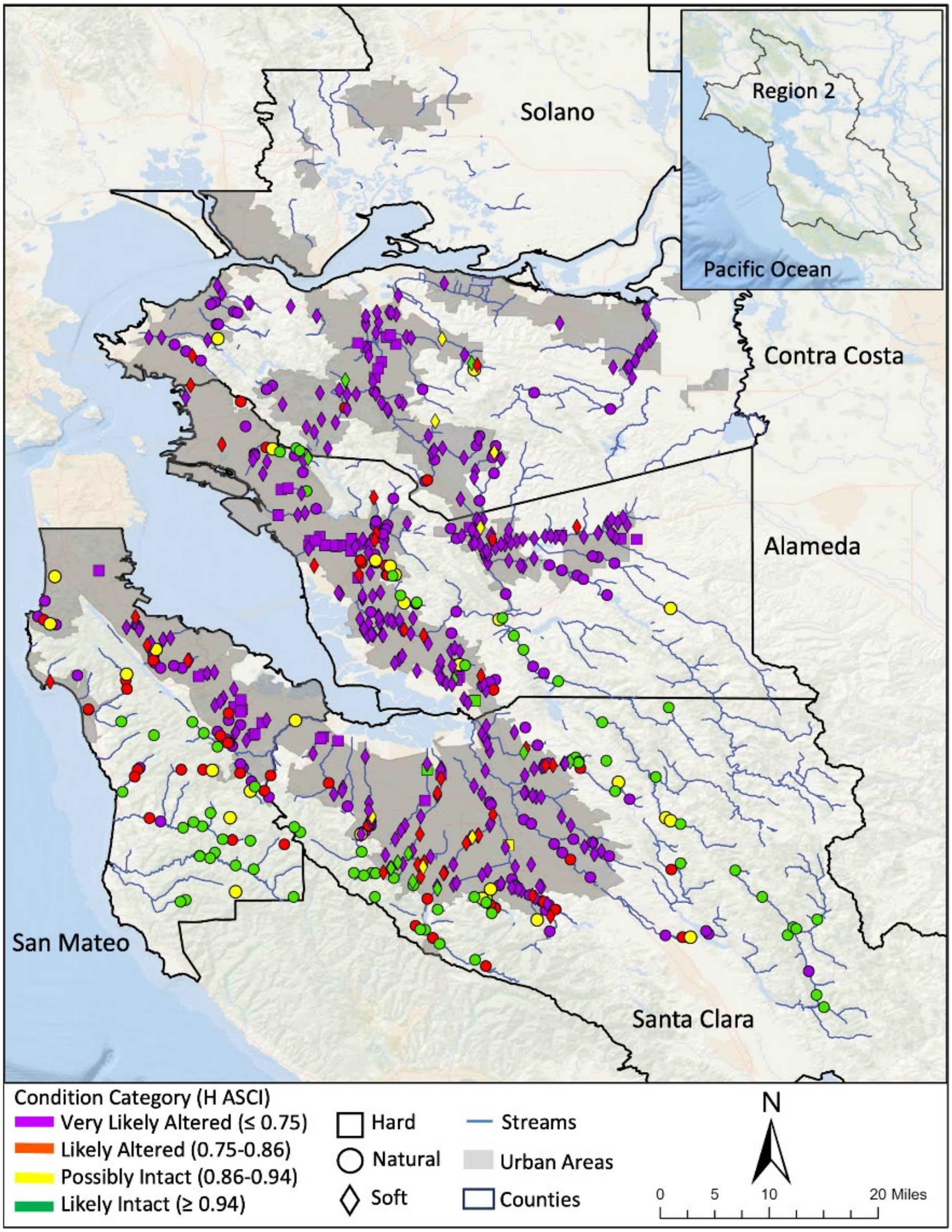

Figure 21. Biological condition of streams in the RMC area based on H ASCI scores by channel type category.

3.2.4 Evaluating the Importance of Channel Classification

An integrated analysis of the channel classification data was conducted to test the significance to the biological condition of urban sites. An ANOVA for CSCI, D ASCI, and H ASCI scores resulted in a model R^2 between 0.15 and 0.32 with statistical significance indicated for variables of County, channel type and the interaction between County and Channel Type (except for H ASCI) (Table 7). The results suggest that there are differences in the biological condition scores amongst Counties, and that condition scores of urban sites also differ by channel type. Channel type effect on CSCI and D ASCI condition scores also differ by County, as indicated by the statistically significant interaction between these two factors. This means that biological condition scores in the region are not consistently found in better or worse condition for the different channel types. Furthermore, the model results with an R^2 of between 0.15 to 0.32, suggest the majority of variability in biological condition scores remain unexplained by these factors alone (see Stressors section below).

Table 7. Results of analysis of variance (ANOVA) testing county and channel type as significant indicators of biological condition scores. * indicates p-value < 0.05. Smaller p-values indicate a higher likelihood that the observed difference represents a significant effect.

Indicator	ANOVA Model R^2	Significance of County (p-value)	Significance of Channel Type (p-value)	Interaction between County and Channel Type (p-value)
CSCI	0.32	< 0.001*	< 0.001*	< 0.001*
D ASCI	0.15	< 0.011*	<0.001*	< 0.001*
H ASCI	0.20	< 0.001*	< 0.001*	0.072

3.3 Stressors Associated with Biological Condition of Urban Sites

3.3.1 Random Forest Model Outputs

To evaluate stressors associated with biological conditions within the RMC sample area, random forest models were developed using the CSCI and D ASCI index results. A parallel analysis was not performed for the H ASCI index due to the lack of soft algae at many of the assessment sites. Stressor data consisted of 26 variables grouped into three types: (1) water quality; (2) physical habitat; and (3) land use (Appendix 1).

Random forest model results indicated better association between stressors and the CSCI index, compared to the model results associated with the D ASCI index. Four of the five selected predictor variables in the CSCI model had a 20% or greater influence on the mean square error (% increase in MSE), while this criterion was only met by one variable in the D ASCI model output, with the top five variables each exhibiting 15-20% influence on the mean square error. Two of the predictor variables selected in the final model outputs overlapped between the CSCI and D ASCI models. Validation of the final random forest models showed that the CSCI model explained 67% of the variance using five predictor (stressor) variables, while the D ASCI model only explained 29% of the variance using five predictors. The CSCI random forest model indicated that physical habitat and land use variables were the most influential variables associated with the biological condition of urban sites (Table 8). Five model predictors exhibited the most influence on CSCI scores (>18% increase in MSE). Two (% fast water and % coarse gravel) physical habitat variables were positively correlated with biological condition, while the other two (% smaller than sand and % fines) were negatively corelated (Spearman's rho: 0.41 – 0.49). Additionally, one land use variable (impervious area within 5 km of the upstream watershed, Figure 22) was also negatively correlated with CSCI scores (Spearman's rho: 0.50).

The results of the random forest model for D ASCI indicated some overlap in the stressor list that explained biological condition relative to the CSCI model results. Two land use variables, two water quality variables (specific conductivity and temperature), and one physical habitat variable were identified as the most influential predictors on D ASCI scores (Table 9). The two variables (% smaller than sand, impervious area 5km) were the same as in the CSCI model results, though with lower explanatory variance (15-20%) and correlation (0.24-0.26) to index scores.

Based upon the random forest model outputs, plots of individual variables versus observed BMI response values (i.e., CSCI scores) were developed to illustrate relationships between stressors and biological condition (Figures 22 to 24). Sites were grouped by condition category and channel type to further explore the overall patterns to the model results. The plots of CSCI scores as a function of fast water indicate that good condition scores (> 10th percentile of reference) are observed predominantly in natural channels where 10% or more of the stream reach is comprised of fast water habitat (Figure 24). Similar plots of CSCI scores for coarse gravel (Figure 25) suggest the best condition scores were also apparent for natural channels with >10% coarse gravel. These results may be suggestive of the potential to improve condition scores through upstream habitat improvements.

Table 8. Summary statistics for the CSCI random forest model.

Stressor Variable	% Increase MSE	Increase Node Purity	Rank Correlation Coefficient (Rho)
Percent Fast Water of Reach	25.4	2.23	0.41
Percent Impervious 5K	24.7	2.60	-0.50
Percent Smaller than Sand	21.8	1.40	-0.46
Percent Coarse Gravel	20.0	1.85	0.49
Percent Fines	18.6	0.77	-0.41

Note: Rank of importance of selected stressor variables are colored according to categories: physical habitat (green), land use (brown), and water quality (blue), if applicable. The correlation coefficient (rho) for each stressor variable is also presented.

Table 9. Summary statistics for the D ASCI random forest model.

Stressor Variable	**% Increase MSE**	**Increase Node Purity**	**Rank Correlation Coefficient (Rho)**
Percent Impervious 5K	20.4	0.91	-0.26
Specific Conductivity	19.4	1.22	-0.34
Temperature	18.2	0.91	-0.33
Road Crossings 5K	15.8	0.66	-0.29
Percent Smaller than Sand (<2mm)	15.2	0.47	-0.24

Note: Rank of importance of selected stressor variables are colored according to categories: physical habitat (green), land use (brown), and water quality (blue), if applicable. The correlation coefficient (rho) for each stressor variable is also presented.

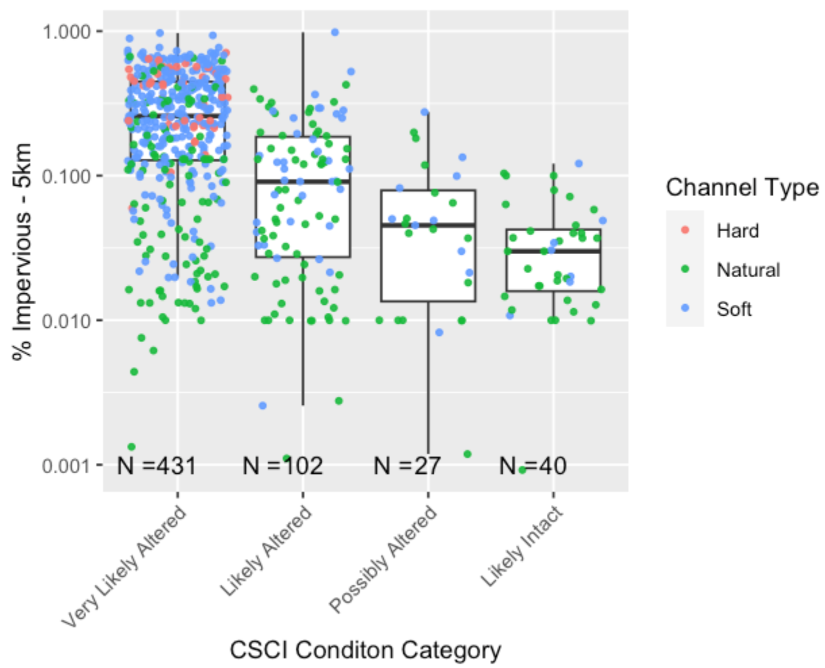

Figure 22. Boxplot of the Percent Impervious in 5 km radius around the site, where sites are categorized as either hard, soft, or natural channels, and grouped by CSCI Condition Category. Urban sites only. Note a logarithmic scale has been applied to the y-axis to spread out the data. Sample size (N) shown below plot for each CSCI Condition Category.

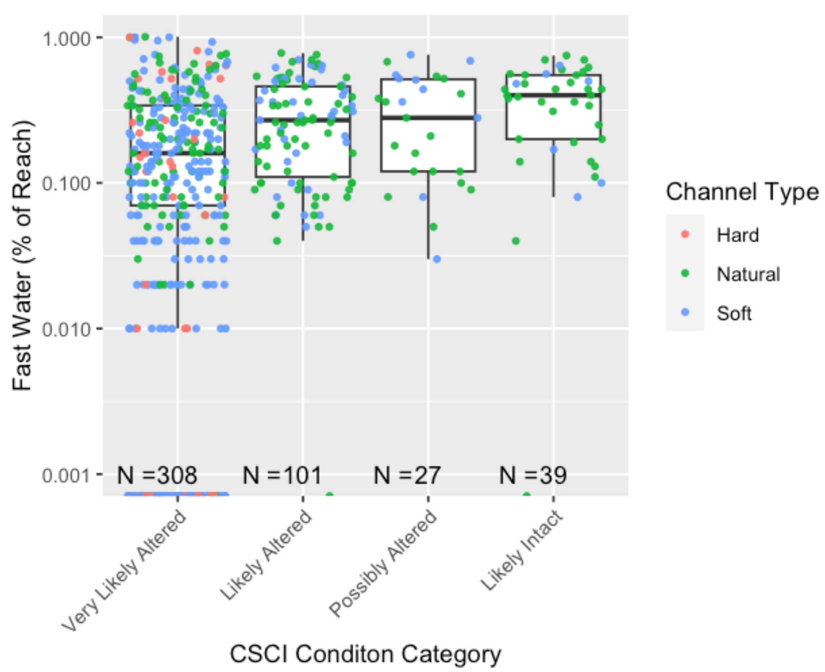

Figure 23. Boxplot of the percentage of fast water in the reach around the site, where sites are categorized as either hard, soft, or natural channels, and grouped by CSCI Condition Category. Urban sites only. Note a logarithmic scale has been applied to the y-axis to spread out the data. Sample size (N) shown below plot for each CSCI Condition Category.

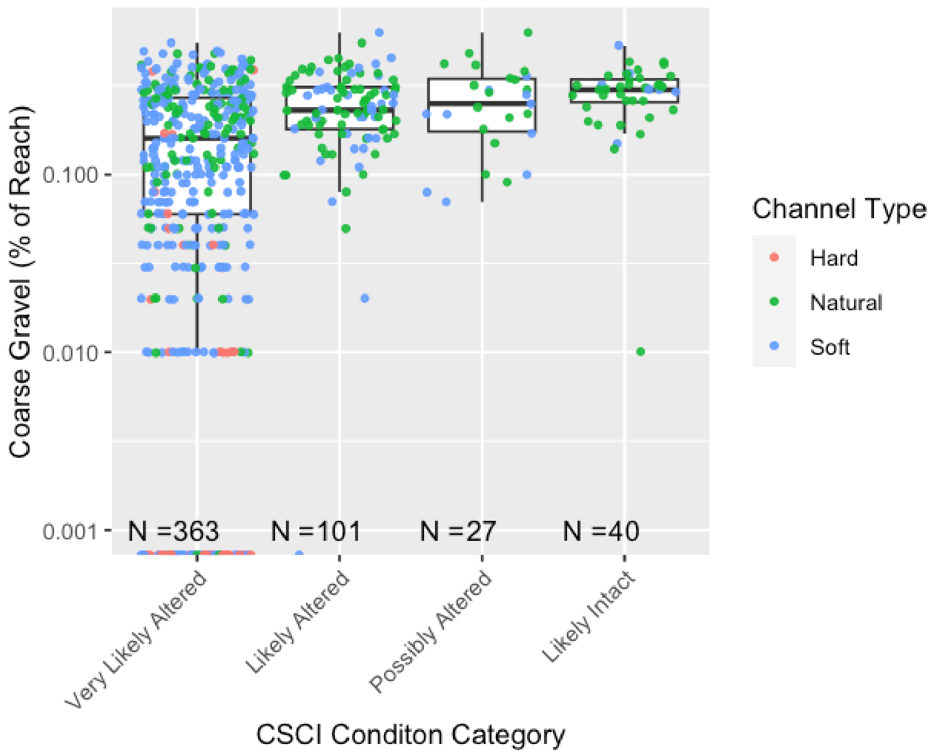

Figure 24. Boxplot of the percentage of coarse gravel in the reach around the site, where sites are categorized as either hard, soft, or natural channels, and grouped by CSCI Condition Category. Urban sites only. Note a logarithmic scale has been applied to the y-axis to spread out the data. Sample size (N) shown below plot for each CSCI Condition Category.

4. Conclusions

The results and conclusions of the BAMSC regional bioassessment data evaluation are discussed below as they relate to the two goals identified for the project.

4.1 What are the biological conditions of Urban streams in the RMC Area?

Biological Conditions

The biological conditions of urban streams in the RMC area were assessed using two ecological indicators: benthic macro-invertebrates (BMIs) and algae. A probabilistic survey design was developed to provide an objective estimate of biological conditions of sampleable streams (i.e., accessible streams with suitable flow conditions) at both the RMC area and countywide scale. RMC Programs focused their monitoring efforts at probabilistic sites in urban streams, sampling 527 of the total 617 (85%) probabilistic sites within the urban area.

Results of the survey indicate that urban streams in the RMC area are generally in poor biological condition:

- The Cumulative Distribution Function (CDF) of CSCI scores for benthic macroinvertebrates (BMIs) indicates that 89% of urban stream length in the region was below the 10^{th} percentile of reference conditions.
- The Diatom and Hybrid Algae Indices (D ASCI and H ASCI) exhibited a very similar pattern, with 92% and 90%, respectively, of stream length in the region below the 10^{th} percentile of reference conditions for each index.

The biological conditions of streams were variable across the five counties. Stream health based on CSCI scores was generally lower in Alameda, Contra Costa and Solano Counties, ranging from 94% to 100% of urban stream length below the 10^{th} percentile of reference conditions. Conditions were slightly better in San Mateo and Santa Clara Counties, with 79% to 86% of stream length below the 10^{th} percentile of reference condition. A similar pattern was observed with biological conditions based on D ASCI and H ASCI scores.

The number of non-urban probabilistic sites (n=90) sampled during the study period was insufficient to confidently assess the overall condition of non-urban streams at the RMC area or countywide scale. However, a comparison of biological conditions by land use (urban vs non-urban) showed higher scores at non-urban sites across all counties, except for Solano County. The CSCI scores for non-urban sites in Santa Clara and San Mateo counties were higher compared to the other counties, with scores at many sites at or above 1.0, which is considered reference condition. However, there were also many non-urban sites with CSCI scores below the 10^{th} percentile of reference, indicating biological conditions were degraded from non-urban types of land use impacts (e.g., grazing, logging, road crossings).

Higher overall conditions, based on CSCI scores, in Santa Clara and San Mateo may be associated with regional differences in rainfall and flow duration. For example, San Mateo County and western Santa Clara County watersheds drain the Santa Cruz mountains, which typically receive higher rainfall, in contrast to Alameda and Contra Costa counties, which primarily contain watersheds that drain the western slopes of the drier Diablo range.

Comparison of biological condition at urban sites showed that index scores differed by flow regime. In general, all three biological indices had lower scores at perennial sites compared to non-perennial sites. However, this pattern was not always consistent among Counties and likely affected by low sample sizes. It would be expected that perennial streams would provide a greater range of habitat types over longer periods of time to support a higher number of BMI and algae taxa. It should be noted however, that defining flow status comes with uncertainty. Many sites were observed to be wet one year and dry the next due to inter-annual variability in precipitation. There are likely many factors related to flow influencing biological conditions, including the timing, duration and extent of dry channel conditions. As a result, assessing flow at one place and time to define two classes representing perennial and non-perennial flow may not be the best approach to evaluate differences in sites, which experience a wide range of flow conditions during the dry season.

Channel type appears to be an important factor that influences biological condition. Unsurprisingly, all the sampling locations in channels with hardened beds had CSCI scores in the lowest condition category, very likely altered. Conversely, urban sites with a natural channel, had generally much higher CSCI scores, with approximately 50 of the 215 sites (23%) above the 10^{th} percentile of reference. However, the remaining 77% of urban sites within natural channels were below the threshold, indicating that factors other than physical habitat, such as water quality, quantity/flow, may be affecting biological conditions.

Soft bed channels were the largest category of sites (n=338) with 95% of sites below the 10^{th} percentile of reference. Although these sites were determined to not have hardened beds, other impacts associated with channel modification (e.g., flashier flows, low habitat quality) may have negative impacts on biological condition. Habitat restoration efforts at these sites may have limited success.

Channel type information may be useful to watershed managers to identify and prioritize reaches for future stressor source identification studies. In particular, urban sites with natural channels may provide better opportunities to implement management actions to improve stream health.

4.2 What Stressors Are Associated with Biological Conditions?

This question was addressed by evaluating the relationships between biological indicators (CSCI and D ASCI) and stressor data through random forest analysis. The study results indicate that each of the biological indices responded to different types of stressors and therefore the two may be best used in combination to assess potential causes of poor (or good) biological conditions in streams:

- Biological condition, based on CSCI scores, was strongly influenced by physical habitat variables and land use within the vicinity of the site. Habitat variables associated with substrate size and water velocity appear to have the largest influence on CSCI scores based on the random forest model results. The percentage of the land area within a 5 km radius that has impervious urban development and road crossings within a 5 km radius were also important variables.
- Biological condition, based on D ASCI scores, was moderately correlated with land use development and water quality variables (conductivity and water temperature,) and to a lesser degree, variables associated with habitat.

In general, CSCI scores at urban sites were consistently low in all RMC counties, indicating that degraded physical habitat conditions in and around streams do not support healthy in-stream biological communities. D ASCI scores at urban sites were more variable, but overall indicated that degraded diatom assemblages occur at sites with adverse water quality conditions and to a lesser extent poor physical habitat and land use development. However, nutrient-related stressors included in the analysis (total nitrogen, orthophosphate, ammonia, unionized ammonia, AFDM, chlorophyll a) were not found to be important in explaining variability in either CSCI or D ASCI scores.

Although results show associations between some stressors and biological condition, they do not establish causation. There are several factors that may affect the strength of the association between stressors and biological condition:

- Stressors are not independent of one another and may have synergistic or mediating effects on condition. For example, elevated temperatures reduce the amount of oxygen that can be dissolved in the water column and both stressors may result in adverse effects to aquatic biota.
- Potential variability of stressor concentrations over time may not be represented in a single grab sample. For example, temperature can have a wide range over a 24-hour period, and several anomalous observations of water quality are evident in the time-series.
- Many of the physical habitat variables can be highly variable throughout the sample reach. For example, a wide range of substrate grain sizes can occur within a single transect. Thus, degraded habitat conditions that may exist at selected transect(s) of the assessment reach may not be well represented in reach-wide averages used as endpoints for the stressor analysis.
- Stressor impacts may be dependent on other factors (possibly not measured) for effects to occur. For example, favorable habitat conditions do not necessarily result in healthy biological condition. Stream locations that have minimal exposure to sunlight, cooler water and higher flow rates may still develop degraded conditions, if co-existing with the presence of poor habitat complexity or significant urban development.

5. References

Bay Area Stormwater Management Agencies Association (BASMAA). 2016. Regional Monitoring Coalition Creek Status Monitoring Standard Operating Procedures. Version 3, March 2016.

Bay Area Stormwater Management Agencies Association (BASMAA). 2020. Regional Monitoring Coalition Creek Status Monitoring Program Quality Assurance Project Plan. Version 3, March 2016.

Bay Area Stormwater Management Agencies Association (BASMAA). 2019. BASMA Regional Monitoring Coalition Five-Year Bioassessment Report, Water Years 2012 – 2016. Consulting report prepared by EOA, Inc. and Applied Marine Sciences for BASMAA. March 2019.

Brown, J.S., and R.D. Mazor. 2023. An assessment of the biological condition of streams in the San Francisco Bay. Southern California Coastal Water Research Project Technical Report 1340. December 2023.

Elith, J., Leathwick, J. R. and Hastie, T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77.4: 802-813.

Fetscher, A. E., Stancheva, R., Kociolek, J. P., Sheath, R. G., Stein, E. D., Mazor, R. D., & Busse, L. B. 2014. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. Journal of applied phycology, 26(1), 433-450.

Kincaid, T. M. and Olsen, A. R. 2016. spsurvey: Spatial Survey Design and Analysis. R package version 3.3.

Maloney, K., Weller, D., Russell, M., Hothorn, T. 2009. Classifying the biological condition of small streams: an example using benthic macroinvertebrates. J North Am Benthol Soc 28(4): 869–884.

Mazor, R.D. 2015a. Bioassessment of Perennial Streams in Southern California: A Report on the First Five Years of the Stormwater Monitoring Coalition's Regional Stream Survey. SCCWRP Technical Report #844. May 2015.

Mazor, R.D. 2015b. Bioassessment Survey of the Stormwater Monitoring Coalition. Workplan for Years 2015 through 2019. Version 1.0. SCCWRP Technical Report #849. February 2015.

Mazor R.D., Rehn A.C., Ode P.R., Engeln M., Schiff K.C., Stein E.D., Gillett DJ, Herbst D.B., Hawkins C.P. 2016. Bioassessment in complex environments: designing an index for consistent meaning in different settings. Freshwater Science 35(1):249-71.

Mazor, R., Ode, P.R., Rehn, A.C., Engeln, M., Boyle, T., Fintel, E., Verbrugge, S., and Yang, C. 2016. The California Stream Condition Index (CSCI): Interim instructions for calculating scores using GIS and R. SWAMP-SOP-2015-0004. Revision Date: August 5, 2016.

Mazor, R., M. Beck, and J. Brown. 2018. 2017 Report on the Stormwater Monitoring Coalition Regional Stream Survey. SCCWRP Technical Report #1029. Southern California Coastal Water Research Project. Costa Mesa, CA.

Ode, P.R., Fltscher, A.E. and Busse, L.B. 2016. Standard Operating Procedures for the Collection of Field Data for Bioassessments of California Wadeable Streams: Benthic Macroinvertebrates, Algae, and Physical Habitat. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 004.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (https://www.R-project.org/).

Rehn, A.C., Mazor, R.D. and Ode, P.R. 2015. The California Stream Condition Index (CSCI): A new statewide biological scoring tool for assessing the health of freshwater streams. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) TM-2015-0002. September 2015.

State Water Resources Control Board (SWRCB). 2015. Surface Water Ambient Monitoring Program (SWAMP) Perennial Stream Assessment Management Memo. SWAMP-MM-2015-0001. June 2015.

Stevens, D.L., Jr., and Olsen, A.R. 2004. Spatially-balanced sampling of natural resources. Journal of the American Statistical Association 99: 262-278.

San Francisco Bay Regional Water Quality Control Board (SFBRWQCB). 2017. San Francisco Bay Basin (Region 2) Water Quality Control Plan (Basin Plan). Incorporating all amendments approved by the OAL as of May 4, 2017.

San Francisco Bay Regional Water Quality Control Board (SFBRWQCB). 2015. California Regional Water Quality Control Board San Francisco Bay Region Municipal Regional Stormwater NPDES Permit (MRP 2.0). Order No. R2-2015-0049. NPDES Permit No. CAS612008. November 19, 2015.

San Francisco Bay Regional Water Quality Control Board (SFBRWQCB). 2009. California Regional Water Quality Control Board San Francisco Bay Region Municipal Regional Stormwater NPDES Permit (MRP 1.0). Order No. R2-2009-0049. NPDES Permit No. CAS612008. October 14, 2009.

van Buuren, S. and Groothuis-Oudshoorn, K. 2011. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1–67.

Waite, I. R., Kennen, J. G., May, J. T., Brown, L. R., Cuffney, T. F., Jones, K. A. and Orlando, J. L. 2012. Comparison of Stream Invertebrate Response Models for Bioassessment Metrics. JAWRA Journal of the American Water Resources Association, 48: 570-583.

APPENDICES

1. Stressor Variables in RMC Dataset
2. Countywide Maps of Biological Condition
3. Summary of Index Scores by Channel Type

APPENDIX 1 STRESSOR VARIABLES

Table 1-A. Variable group, description, and summary statistics of response variables (condition indices) and explanatory environmental variables (landscape, habitat, and water quality) used for random forest model development (data represent urban sites only).

Variable Group	Description	Units	Min Value	Max Value	Mean Value
Response	California Stream Condition Index (CSCI)	Unitless	0.13	1.30	0.53
Response	Diatom Algae Stream Condition Index (D ASCI)	Unitless	0.26	1.14	0.64
Habitat	Combined Riparian Human Disturbance Index	Unitless	0	7.23	3.06
Habitat	Evenness Flow Habitat	Unitless	0	1	0.538
Habitat	Evenness Natural Substrate	Unitless	0	1	0.743
Habitat	Natural Shelter Cover	Unitless	0	250	45.9
Habitat	Shannon Diversity Natural Substrate Types	Unitless	0	2.01	1.37
Habitat	% Fast Water	%	0	1.00	0.21
Habitat	% Coarse Gravel	%	0	0.63	0.18
Habitat	% Fines	%	0	1.00	0.20
Habitat	% Smaller than Sand (<2mm)	%	0	1	0.41
Water Quality	Dissolved Oxygen	mg/L	0.57	96.7	9.63
Water Quality	Temperature	˚C	9.3	31.6	16.5
Water Quality	Specific Conductivity	psu	2.0	8197	964
Water Quality	Ash Free Dry Mass	g/m2	0.10	30500	660
Water Quality	Chlorophyll a	mg/m2	0.79	8000	189
Water Quality	Ammonia	mg/L	0.008	1.70	0.17
Water Quality	Total Nitrogen	mg/L	0.015	50.4	1.83
Water Quality	Orthophosphate	mg/L	0.002	4.3	0.10

Variable Group	Description	Units	Min Value	Max Value	Mean Value
Water Quality	Total Phosphorus	mg/L	0.004	3.50	0.12
Landscape	Impervious Area within 1km upstream of site	%	0	0.99	0.35
Landscape	Impervious Area within 5km upstream of site	%	0	0.98	0.24
Landscape	Urban Area within 1km upstream of site	%	0	1.00	0.61
Landscape	Urban Area within 5km upstream of site	%	0	1.00	0.42
Landscape	Road Crossings within 1km upstream of site	# per 1km	0	19.0	3.17
Landscape	Road Crossings within 5km upstream of site	# per 5km	0	182.0	19.3
Landscape	Road Density within 1km upstream of site	# per 1km	0	22.0	8.84
Landscape	Road Density within 5km upstream of site	# per 5km	0	17.6	6.28

APPENDIX 2 COUNTYWIDE MAPS OF BIOLOGICAL CONDITION

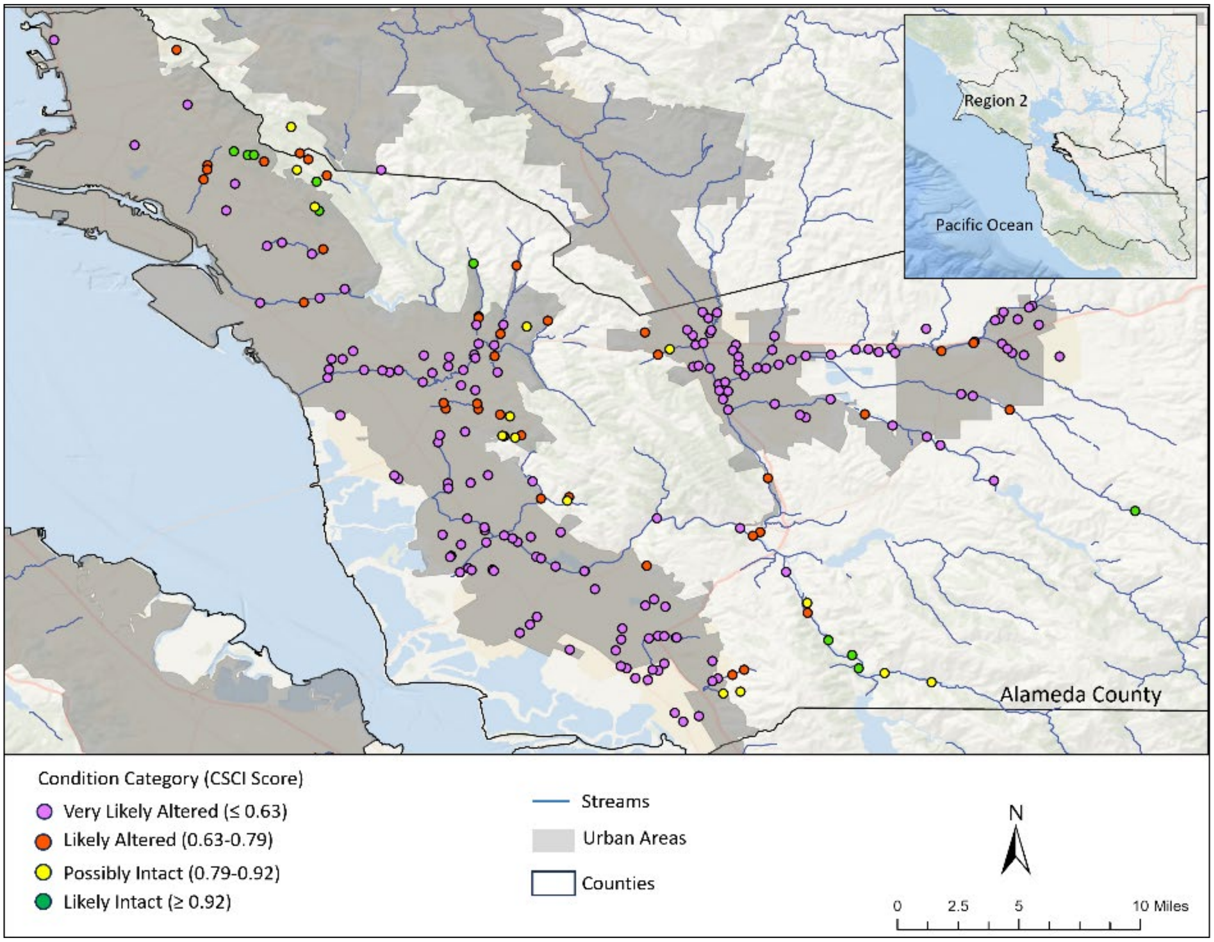

Figure 2-A. Biological condition based on CSCI scores in Alameda County.

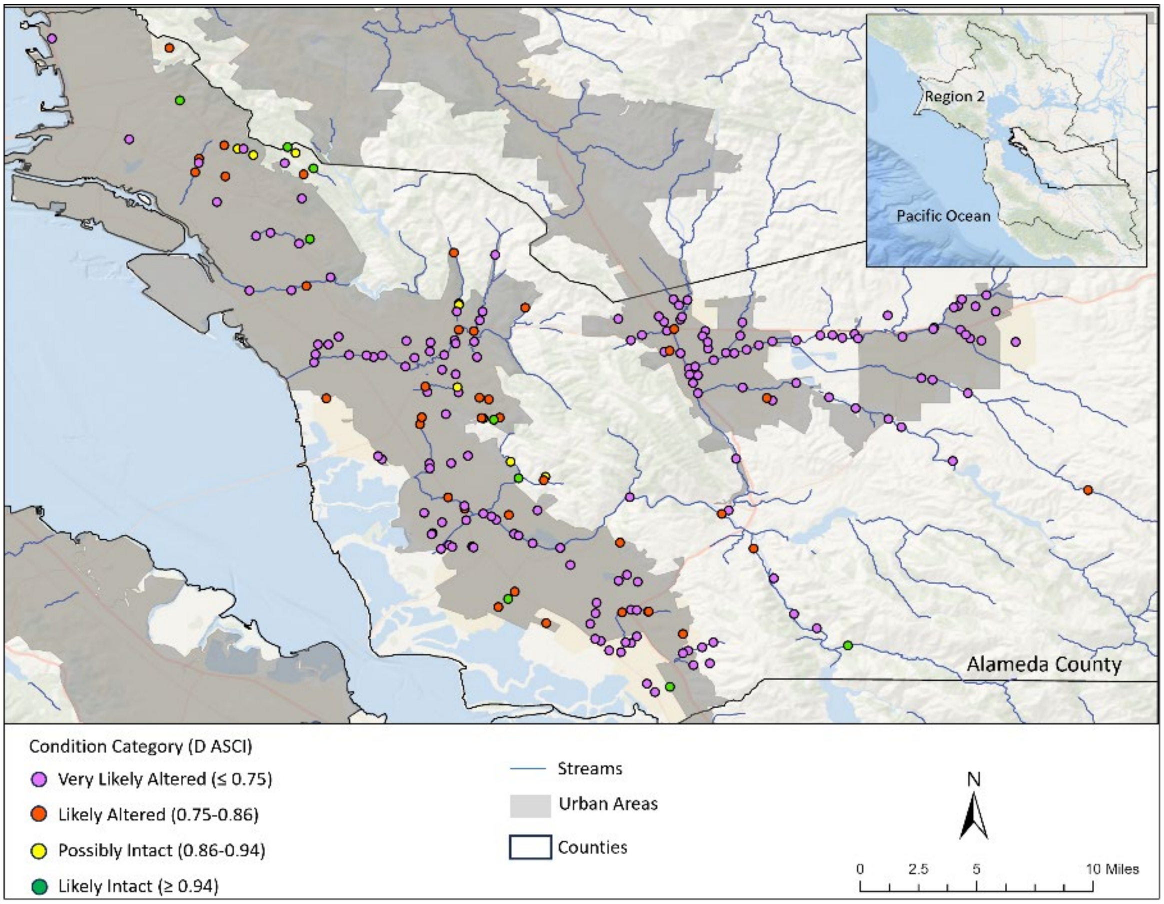

Figure 2-B. Biological condition based on D ASCI scores in Alameda County.

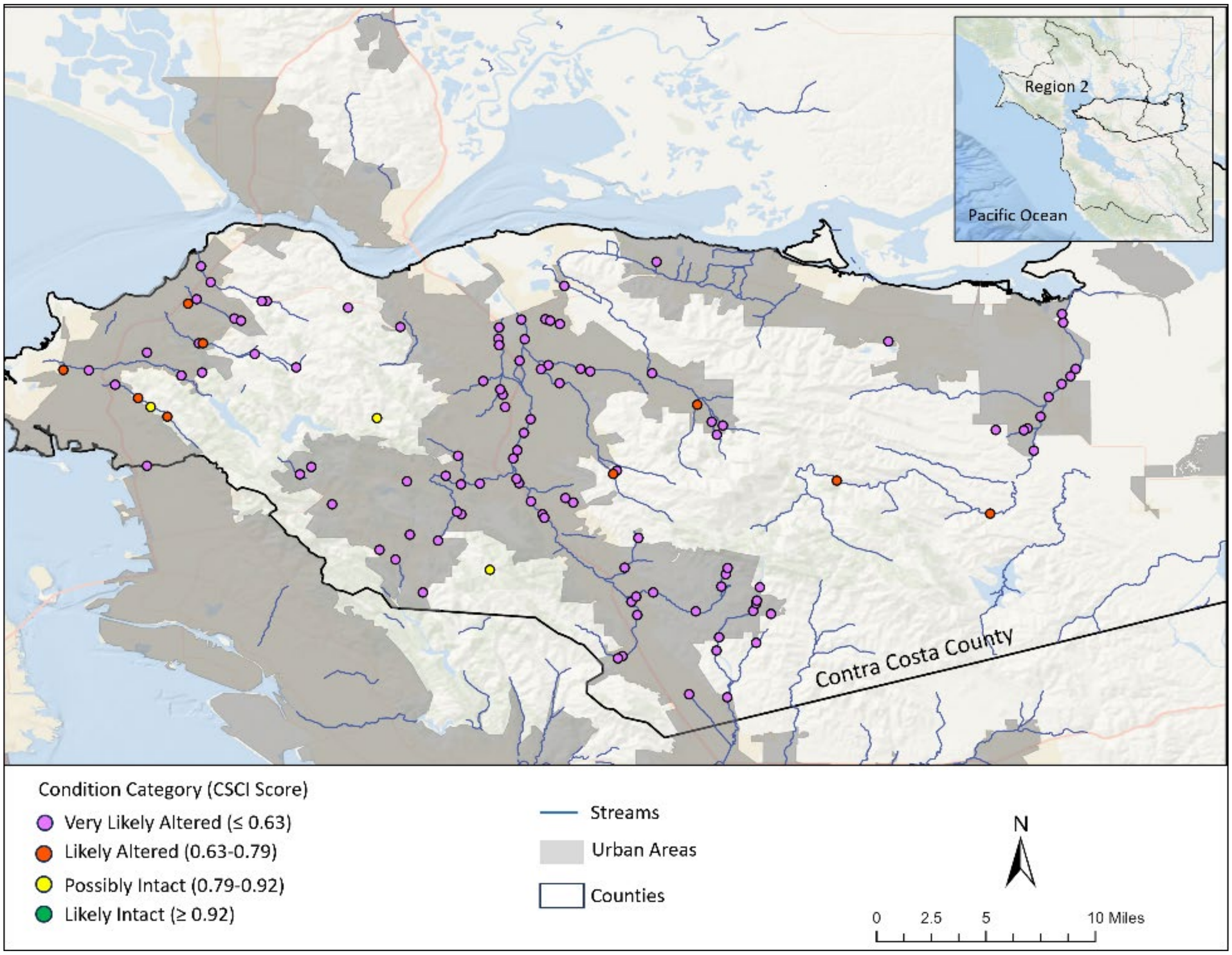

Figure 2-C. Biological condition based on CSCI scores in Contra Costa County.

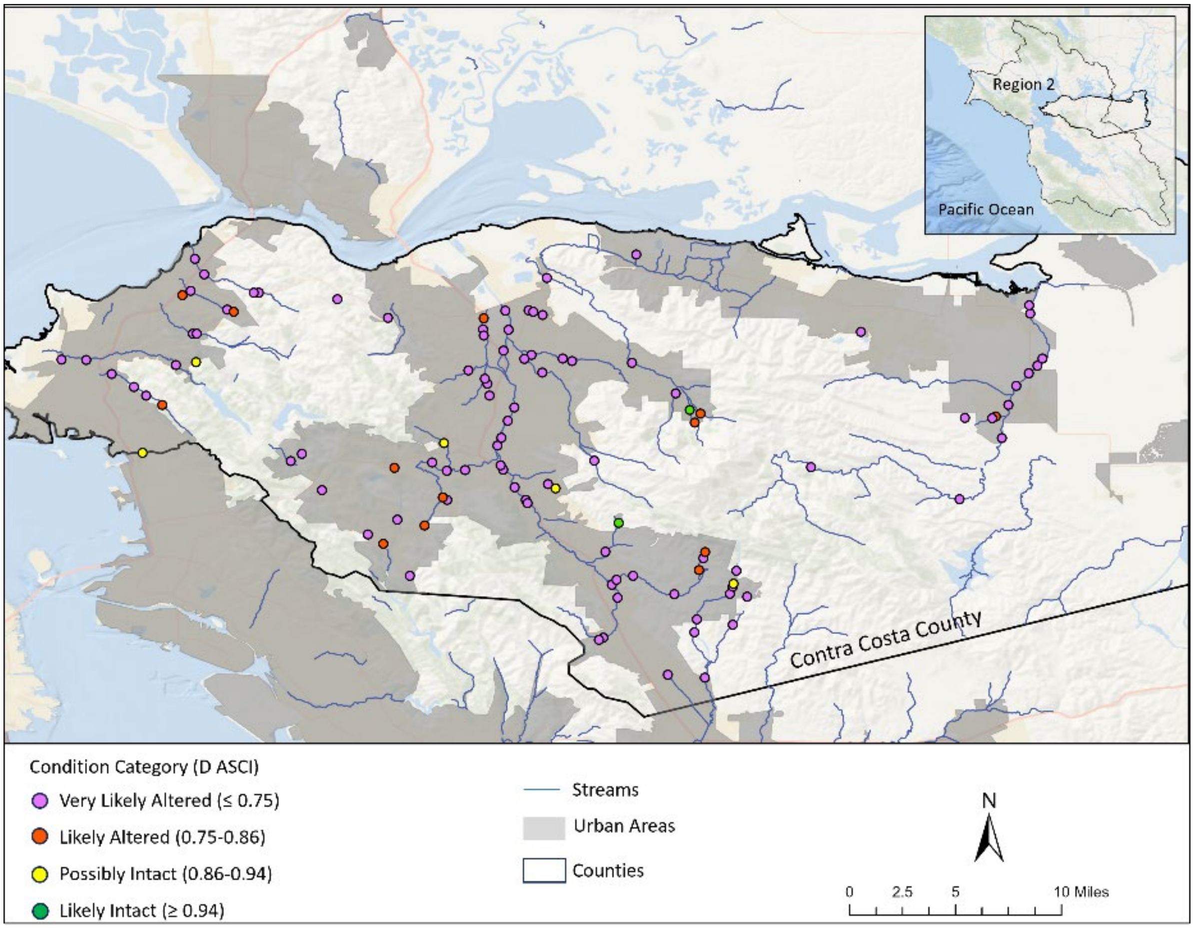

Figure 2-D. Biological condition based on D ASCI scores in Contra Costa County.

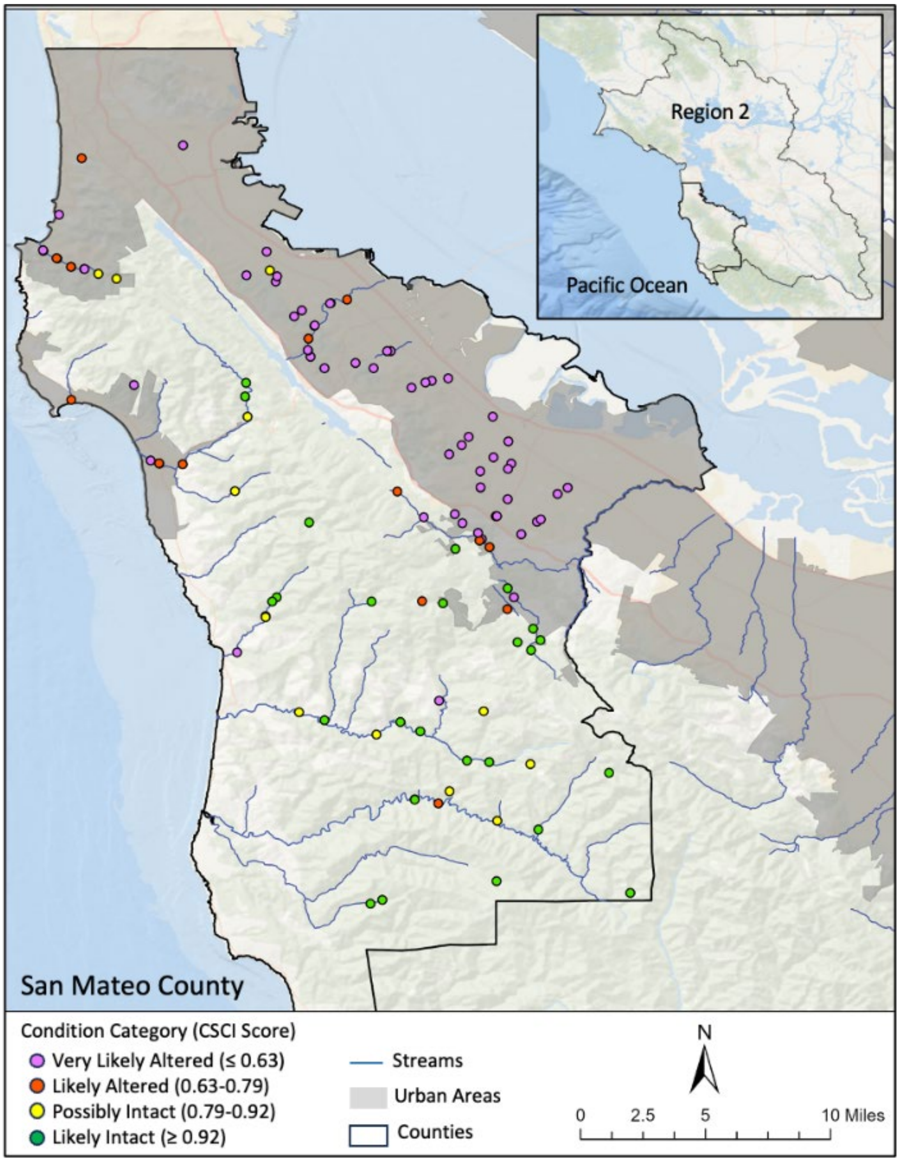

Figure 2-E. Biological condition based on CSCI scores in San Mateo County.

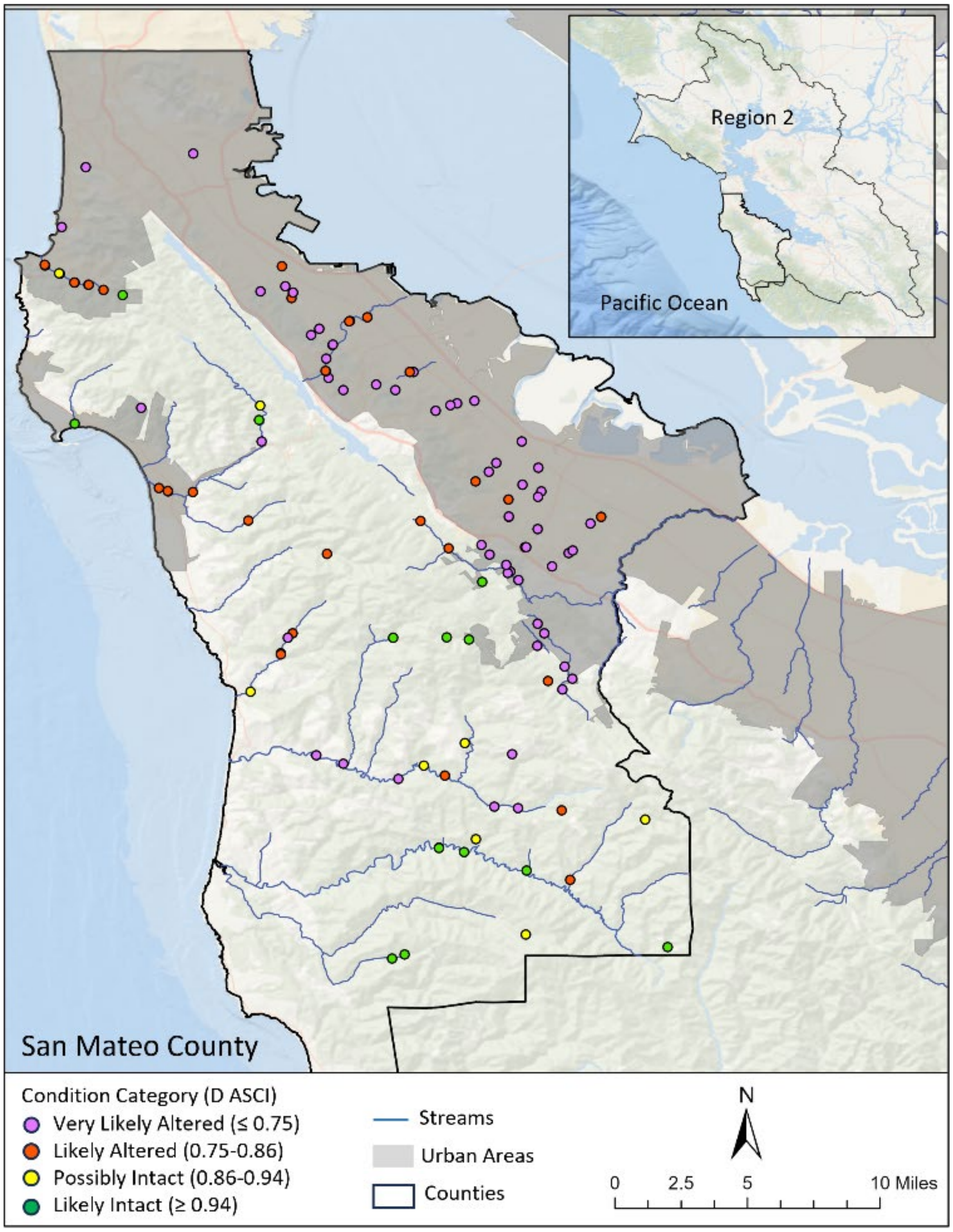

Figure 2-F. Biological condition based on D ASCI scores in San Mateo County.

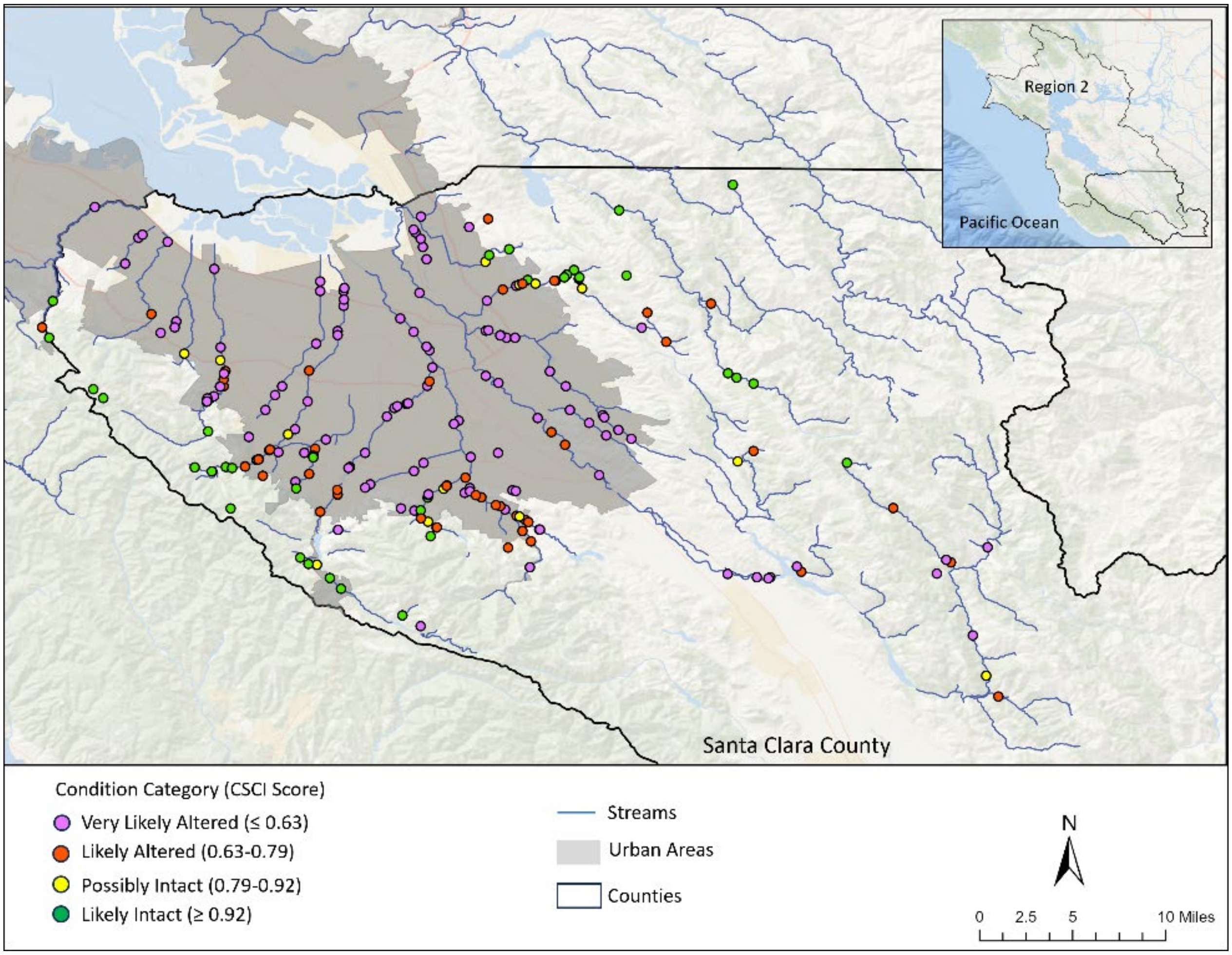

Figure 2-G. Biological condition based on CSCI scores in Santa Clara County.

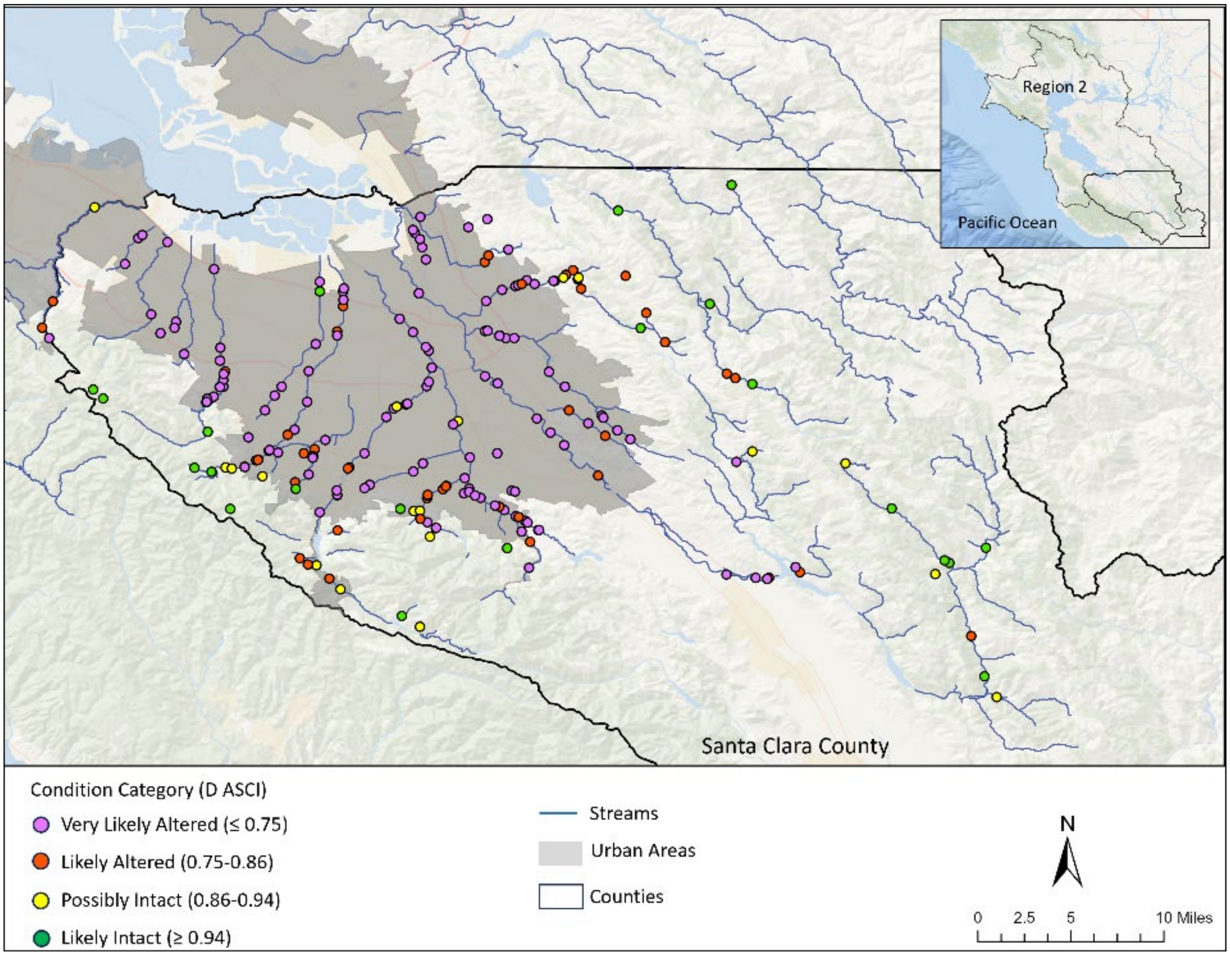

Figure 2-H. Biological condition based on D ASCI scores in Santa Clara County.

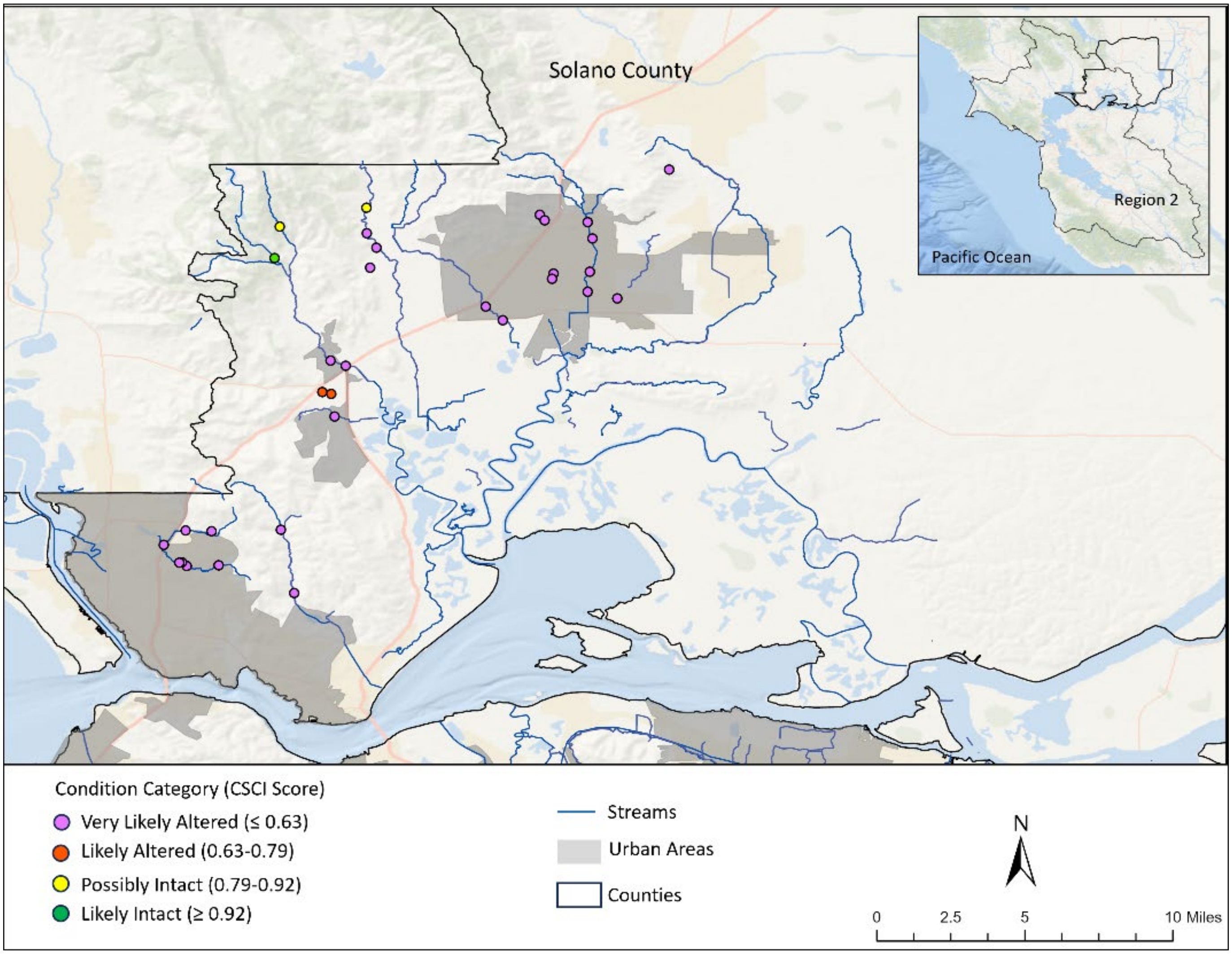

Figure 2-I. Biological condition based on CSCI scores in Solano County.

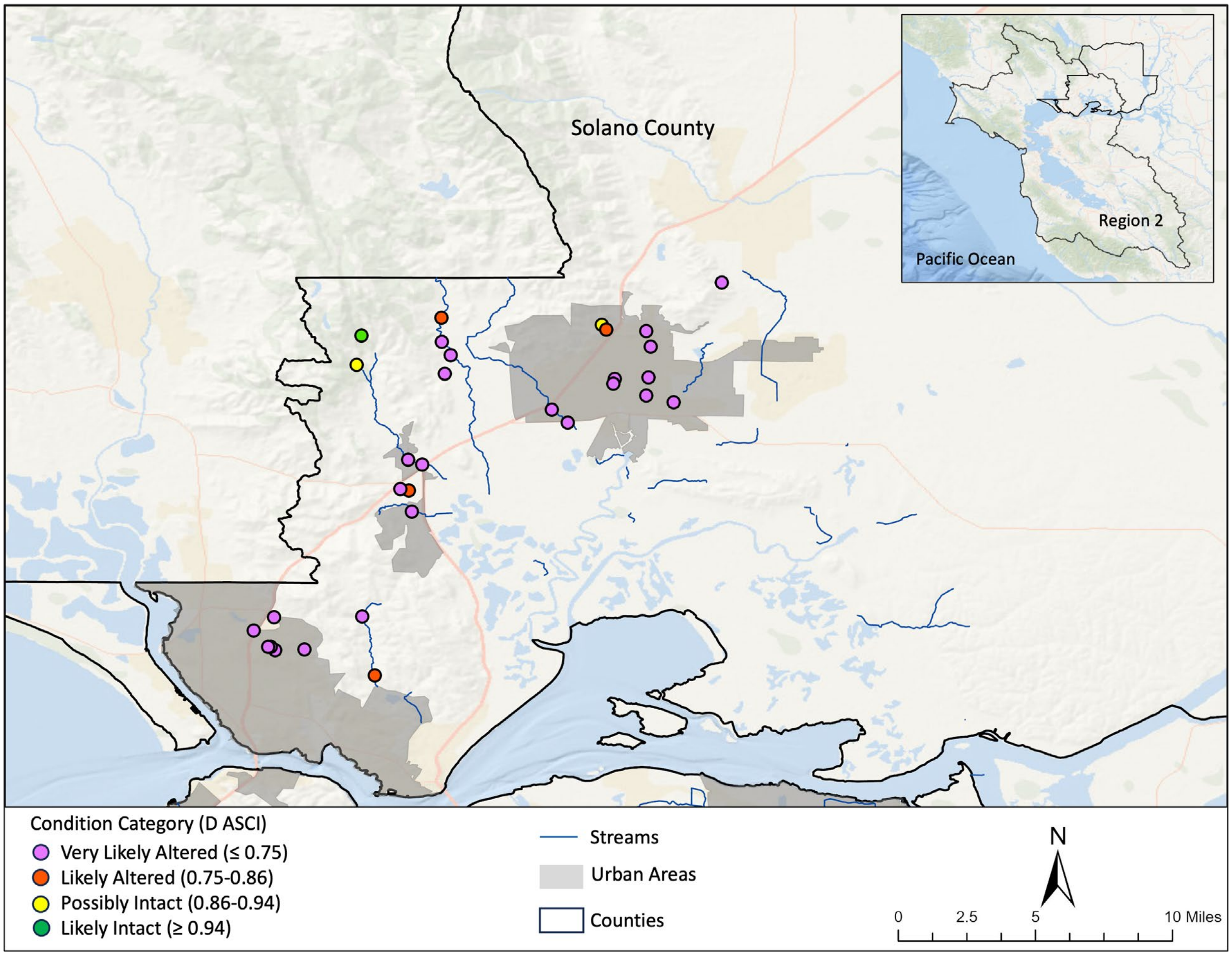

Figure 2-J. Biological condition based on D ASCI scores in Solano County.

Appendix 3 Summary Of Biological Condition At Urban Sites by Channel Type

Table 3-A. Summary of CSCI condition scores by channel type for all urban sites in each County.

Index	County	Channel type	N	Likely intact	Possibly altered	Likely altered	Very Likely altered	Mean	SD
CSCI	Alameda	Hard	22	0%	0%	0%	100%	0.402	0.08
CSCI	Alameda	Soft Beds	128	2%	2%	5%	92%	0.381	0.17
CSCI	Alameda	Natural Channel	62	8%	13%	42%	37%	0.675	0.18
CSCI	Contra Costa	Hard	8	0%	0%	0%	100%	0.487	0.06
CSCI	Contra Costa	Soft Beds	73	0%	1%	3%	96%	0.421	0.12
CSCI	Contra Costa	Natural Channel	27	0%	0%	26%	74%	0.497	0.15
CSCI	San Mateo	Hard	10	0%	0%	0%	100%	0.445	0.09
CSCI	San Mateo	Soft Beds	23	0%	4%	4%	91%	0.497	0.10
CSCI	San Mateo	Natural Channel	47	15%	6%	26%	53%	0.650	0.25
CSCI	Santa Clara	Hard	7	0%	0%	0%	100%	0.373	0.10
CSCI	Santa Clara	Soft Beds	114	4%	5%	24%	67%	0.567	0.19
CSCI	Santa Clara	Natural Channel	79	27%	8%	27%	39%	0.722	0.23

Table 3-B. Summary of D ASCI condition scores by channel type for all urban sites in each County.

Index	County	Channel type	N	Likely intact	Possibly altered	Likely altered	Very Likely altered	Mean	SD
D ASCI	Alameda	Hard	22	5%	0%	5%	90%	0.634	0.11
D ASCI	Alameda	Soft Beds	126	2%	0%	15%	83%	0.575	0.15
D ASCI	Alameda	Natural Channel	60	7%	12%	28%	53%	0.729	0.14
D ASCI	Contra Costa	Hard	8	0%	0%	0%	100%	0.548	0.13
D ASCI	Contra Costa	Soft Beds	72	3%	6%	8%	83%	0.615	0.15
D ASCI	Contra Costa	Natural Channel	27	0%	4%	22%	74%	0.652	0.11
D ASCI	San Mateo	Hard	10	0%	0%	0%	100%	0.445	0.09
D ASCI	San Mateo	Soft Beds	23	4%	0%	31%	65%	0.497	0.10
D ASCI	San Mateo	Natural Channel	47	13%	4%	30%	53%	0.650	0.25
D ASCI	Santa Clara	Hard	7	14%	14%	0%	71%	0.373	0.10
D ASCI	Santa Clara	Soft Beds	114	0%	2%	19%	79%	0.567	0.19
D ASCI	Santa Clara	Natural Channel	79	8%	11%	25%	56%	0.722	0.23

Table 3-C. Summary of H ASCI condition scores by channel type for all urban sites in each County (N=590).

Index	County	Channel type	N	Likely intact	Possibly altered	Likely altered	Very Likely altered	Mean	SD
H ASCI	Alameda	Hard	22	5%	0%	0%	95%	0.540	0.170
H ASCI	Alameda	Soft Beds	126	2%	2%	8%	88%	0.529	0.174
H ASCI	Alameda	Natural Channel	60	13%	12%	17%	58%	0.737	0.212
H ASCI	Contra Costa	Hard	8	0%	0%	0%	100%	0.469	0.105
H ASCI	Contra Costa	Soft Beds	72	3%	4%	4%	89%	0.533	0.175
H ASCI	Contra Costa	Natural Channel	27	0%	4%	7%	89%	0.603	0.146
H ASCI	San Mateo	Hard	10	0%	0%	0%	100%	0.552	0.075
H ASCI	San Mateo	Soft Beds	23	0%	9%	13%	78%	0.638	0.140
H ASCI	San Mateo	Natural Channel	42	9.5%	9.5%	24%	57%	0.713	0.149
H ASCI	Santa Clara	Hard	7	14%	14%	0%	72%	0.639	0.244
H ASCI	Santa Clara	Soft Beds	114	6%	7%	18%	69%	0.644	0.184
H ASCI	Santa Clara	Natural Channel	79	20%	8%	21%	51%	0.768	0.184