North Orange County Municipal Separate Storm Sewer System (MS4) Monitoring Evaluation

Kenneth Schiff
Marcus Beck
Elizabeth Fassman-Beck

Southern California Coastal Water Research Project
Technical Report 1221

North Orange County Municipal Separate Storm Sewer System (MS4) Monitoring Evaluation

Kenneth Schiff[1]

Marcus Beck[2]

Elizabeth Fassman-Beck[1]

[1]*Southern California Coastal Water Research Project, Costa Mesa, CA*

[2]*Tampa Bay Estuary Program, Tampa, FL*

October 2021
SCCWRP Technical Report #1221

EXECUTIVE SUMMARY

Monitoring is a key element of any watershed management program. The County of Orange, on behalf of the cities of North Orange County, has been monitoring the North County's watersheds for decades to support the collective Municipal Separate Storm Sewer System (MS4) National Pollutant Discharge Elimination System (NPDES) permit. This monitoring is used for evaluating NPDES compliance, tracking progress towards Total Maximum Daily Loads (TMDL) targets, reporting to the public, and ensuring the County's watersheds' beneficial uses are maintained.

The North Orange County Monitoring Program has many elements encompassing dry weather and wet weather runoff, outfalls and receiving waters, all in a variety of matrices (water, sediment, fish, and bird tissues). The current iteration of the monitoring program was approved by the Santa Ana Regional Water Quality Control Board (Regional Board) in 2009 and partially reviewed and updated for TMDL elements in 2014.

The goal of this project is to conduct a comprehensive review of the North Orange County Monitoring Program to improve both effectiveness and efficiency. This review was accomplished using a five-step process:

1) Identify the monitoring questions of greatest interest to regulated and regulatory agencies,
2) Compile historical data to assess how effectively the current monitoring program answers the priority monitoring questions,
3) Evaluate monitoring design updates and upgrades to increase effectiveness or efficiency,
4) Provide technical guidance and hands-on scientific tools to allow the Regional Board and the North Orange County Permittees define the final details of the future monitoring and reporting program,
5) Develop a reporting template for the County's annual report to the Regional Board.

The five-step process was implemented using a consensus-based process utilizing an Advisory Committee comprised of the principal stormwater regulated party (County of Orange), three Permittees within North Orange County (City of Huntington Beach, City of Newport Beach, and City of Irvine), an environmental advocacy organization (Orange County Coastkeeper), and the state regulatory agency (Regional Board).

This project is intended to create a starting point for updating the MS4 monitoring program. This project does not intend to create a "State of the Watershed" report or a final "Sampling and Analysis Plan" or revisions to the existing approved program. Additional work would be required to complete these goals, including addressing the management context for utilizing such documents. Moreover, this project was not designed to be cost-neutral. Monitoring recommendations are based on increasing effectiveness and efficiency, not cost. Some recommendations may save resources and others may require resources, and the level of resources required will be decided in the next steps of monitoring plan design and implementation. Finally, this project does not address all questions or objectives in the NPDES permit and other steps, activities, and objectives may be required.

Monitoring Questions

Of the 11 monitoring questions identified by the regulated and regulatory agencies in Orange County on the Advisory Committee, four monitoring questions were prioritized:

1. What are the temporal trends in pollutant concentrations and loading? [Mass Emissions and TMDL Monitoring]
2. What are the temporal trends in fish and bird tissue concentrations? [TMDL Monitoring]
3. Are there illegal discharges/illicit connections (ID/IC)? [Dry Weather Monitoring]
4. What are the temporal trends in sediment quality and water column concentrations in estuaries and wetlands? [Estuary/Wetland Monitoring]

Q1: What are the temporal trends in pollutant concentrations and loading?

An inventory of historical monitoring identified sites for the 10-year period 2010 – 2019 across three North Orange County watersheds: San Gabriel River/Coyote Creek, Anaheim Bay/Huntington Harbour, and Newport Bay watersheds. Within these three watersheds, there are multiple sites per watershed, sampled multiple times per year in both wet and dry weather, typically located at the terminus of the watershed and major tributaries.

Historical monitoring results illustrated long-term trends in some pollutant concentrations and loads from a variety of mass emission sites. Concentrations and loads have typically decreased where trends were observed for nutrients, suspended solids, and some trace metals. There were tangible differences in trends between wet and dry weather, between watersheds, and among pollutants. The Advisory Committee agreed to four conclusions after evaluating potential improvements to monitoring efficiency and effectiveness:

- Maintain the ongoing number of monitoring sites, as each generally produces unique information.
- Sampling frequency reductions at many sites and/or parameters will not compromise capability to detect trends.
- When concentrations are approaching thresholds of concern, increasing sampling frequency is an option to increase confidence in trend detection.
- A web application was created with technical tools to optimize monitoring frequency for trend detection by site and parameter for both dry or wet weather when defining the next monitoring program.

Q2: What are the temporal trends in fish and bird tissue concentrations?

An inventory of historical monitoring identified nine bioaccumulation sites for the 10-year period 2010 – 2019 across one Orange County watershed: the Newport Bay Watershed. There are multiple sites per watershed, and multiple species per site, but a limited set of pollutants.

Historical monitoring results illustrated there have been few long-term trends in most pollutant concentrations for either fish tissue or bird egg concentrations. The lack of trend is likely a function of a short time series and insufficient sample size. Significant monitoring design limitations are that bird egg and fish sampling typically occurs only once per year (during the spring/summer nesting season) and bird eggs and fish are present in limited abundance. The Advisory Committee agreed to three conclusions after evaluating potential improvements to monitoring efficiency and effectiveness:

- Since sampling frequency is constrained to annually, the issue of trend detection becomes a function of time to detectable trend. Where sufficient sample size exists for statistical power analysis, optimal trend detection will likely require a decade or longer.
- Trend detection efficiency will improve if monitoring is constrained to a single species and tissue type. Combining across species and tissue types introduces additional variability and reduces power.
- A web application was created with technical tools to estimate optimal time to trend detection by site, species, tissue type, and parameter when defining the next monitoring program.

Q3: Are there illegal discharges/illicit connections (ID/IC)?

An inventory of historical monitoring identified 106 dry weather monitoring sites for the 10-year period 2010 – 2019 across four North Orange County watersheds: Newport Bay, San Gabriel River/Coyote Creek, Santa Ana River, and Anaheim Bay/Huntington Harbour Watersheds. Unlike mass emission sites that occur in receiving waters, dry weather screening occurs in MS4 outfalls. There are hundreds of outfall sites per watershed. Sampling occurs approximately monthly between May and September and analyzed for 71 different pollutants. Frequently, most pollutants are below detection limits. The analysis for effectiveness and efficiency was therefore constrained to the 10 most frequently detected pollutants including bacteria, nutrients, some trace metals, and some pesticides totaling more than 40,000 laboratory analyses.

Historical monitoring results illustrated a number of "hot spots" in every watershed that have frequent exceedances of pollutant thresholds. Tracking these exceedances to sources and remediating them has been problematic. Discharges are intermittent, and laboratory results are not produced for several weeks after the monthly sampling event, well after the pollutant discharge source is gone. The Advisory Committee agreed to four conclusions after evaluating potential improvements to monitoring efficiency and effectiveness:

- The monitoring program should reduce the number of sites sampled and prioritize "hot spot" outfall sites that may have caused or contributed to the high pollutant levels at downstream receiving waters, thereby triggering upstream investigation of sources for remediation.
- This new strategy is best accomplished by replacing the once-a-month periodic dry weather sampling with continuous, near-real-time sensors.
- Continuous data will help focus ID/IC source tracking efforts when pollution events are occurring.

- A web application was created providing a technical tool for prioritizing "hot spot" sites when defining the next monitoring program.

Q4: What are the temporal trends in sediment quality and water column concentrations in estuaries and wetlands?

An inventory of historical monitoring identified 13 sites for the 10-year period 2010 – 2019 across the three North Orange County watersheds with estuaries and wetlands: Anaheim Bay/Huntington Harbour, Santa Ana River (Talbert Marsh), and Newport Bay Watersheds. The sites are sampled multiple times per year for sediment and water column, in dry weather and wet weather. Hundreds of samples were analyzed for a wide variety of parameters including nutrients, trace metals, and organics/pesticides.

Historical monitoring results did not identify many long-term trends in sediment concentrations. Dry weather water column monitoring did observe a limited number of trends including decreasing concentrations of nitrate+nitrite and some metals (copper, mercury, and zinc) in Upper Newport Bay. Similar trends were not observed in Lower Newport Bay or the other waterbodies. Wet weather water column monitoring also observed a limited number of trends including decreasing concentrations of some metals (copper, mercury, selenium, and zinc) in Upper Newport Bay. The Advisory Committee agreed to four conclusions after evaluating potential improvements to monitoring efficiency and effectiveness:

- Continue to monitor both water column and sediment in estuaries and wetlands.

- Sampling frequency reductions at many sites and/or parameters will not compromise capability to detect trends for water column monitoring in both dry and wet weather.

- Modify the existing sediment monitoring question focused on temporal trends in pollutant concentrations at specific locations into temporal trends in percent area above a threshold. To accomplish this, the monitoring design should reduce sampling frequency at existing sites and re-allocate to more sites. This is best accomplished by integrating the design with the Southern California Bight regional monitoring, thereby gaining leverage to bightwide comparisons.

- A web application was created with technical tools to optimize monitoring frequency for trend detection by site and parameter for both dry or wet weather when defining the next monitoring program.

TABLE OF CONTENTS

INTRODUCTION

Monitoring is a key element of any watershed management program (Freedman et al. 2004). Particularly in urban watersheds, a large number of potential stressors – alone or in combination - can impact receiving waters including hydromodification, bacteria, nutrients, trace metals, petroleum hydrocarbons, or pesticides (Walsh et al. 2005). Monitoring helps quantify when and where environmental problems exist, assists in prioritizing which reaches should be remediated first, and ascertains if the management actions taken have been effective at restoring the environmental problems identified and prioritized (PG Environmental 2018, Model Monitoring Technical Committee 2014, National Research Council 1990).

Orange County, California, has been monitoring the local watersheds for decades (Schiff 1995; French et al. 2006). A significant portion of this monitoring is to support two Municipal Separate Storm Sewer System (MS4) National Pollutant Discharge Elimination System (NPDES) permits; one NPDES permit is in north Orange County and the second in south Orange County. A third component of the monitoring is to support Total Maximum Daily Loads (TMDL) implementation. TMDLs for sediment, nutrients, selenium, fecal indicator bacteria, and organochlorine compounds exist in Orange County. Each year, The County of Orange, in cooperation with the Orange County Flood Control District and the cities of Anaheim, Brea, Buena Park, Costa Mesa, Cypress, Fountain Valley, Fullerton, Garden Grove, Huntington Beach, Irvine, La Habra, La Palma, Laguna Hills, Laguna Woods, Lake Forest, Los Alamitos, Newport Beach, Orange, Placentia, Santa Ana, Seal Beach, Stanton, Tustin, Villa Park, Westminster and Yorba Linda (Permittees), reports the results of the watershed monitoring to state and federal regulators, County management, and the public.

The North Orange County Monitoring Program is large and complex, encompassing many different elements (OC Annual Report 2018-19). Orange County monitors water quality in dry weather and wet weather, samples from outfalls and receiving waters, measures a variety of matrices (water, sediment, fish and bird tissues), and analyzes hundreds of different pollutants. The current iteration of the monitoring program was approved by the Regional Board in 2009 and partially reviewed and updated for TMDL elements in 2014 (Schiff et al. 2014).

The goal of this project is to conduct a review of the current North Orange County Monitoring Program to improve effectiveness and efficiency during the next MS4 permit cycle. This review was accomplished using a four-step process:

1) Identify the priority monitoring questions to regulated and regulatory agencies
2) Investigate how well the current monitoring design answers the priority monitoring questions
3) Evaluate monitoring design updates and upgrades to increase effectiveness and efficiency
4) Provide technical guidance and hands-on scientific tools for regulators and regulated parties to negotiate the final details of their future monitoring and reporting program

This project does not intend to create a “State of the Watershed” report (OC Stormwater Program 2014) or a final update of the North Orange County Monitoring Program. A State of the Watershed is a complete waterbody assessment, including recommendations for future management actions, while this project focuses only on monitoring design recommendations. An update of the North Orange County Monitoring Program will require many more sampling details for monitoring implementation (i.e., sampling schedules, final safety plans, monitoring protocols, quality assurance, etc.), while this project focuses on monitoring design attributes (i.e., how many sites, how frequently to sample each site, predicted statistical confidence, etc.). Moreover, this project was not designed to be cost-neutral. Monitoring recommendations are based on increasing effectiveness and efficiency, not cost. Some recommendations may save resources and others may require resources. The level of resources to be expended will be decided in the next steps of monitoring program development. Finally, this project does not address all questions or objectives in the NPDES permit and other steps, activities, and objectives may be required.

METHODS

Advisory Committee

This project relied on an engaged and interactive Advisory Committee (Table 1). The Advisory Committee was composed of regulated, regulatory, and non-governmental agencies, providing a cross-section of interests and goals to maintain a diversity of ideas and perspectives. This committee met eight times between July 2019 and November 2020.

Table 1. Members of the Advisory Committee.

Agency	Representative
Santa Ana RWQCB	Jason Freshwater (representative) Adam Fischer Terri Reeder
OC Coastkeeper	Ray Hiemstra
City of Anaheim	Keith Linker
City of Irvine	Thomas Lo
City of Huntington Beach	Jim Merid
OC Public Works	Stuart Goong Jonathan Humphrey Kelvin Liu Jian Peng (representative)
City of Newport Beach	John Kappeler

Philosophical Approach

The Advisory Committee created a double-edged, self-limiting philosophy as the thematic understory of this project:

1) The Federal Clean Water Act mandates that all pollutant discharges must have an NPDES permit, which are designed to protect the beneficial uses of the public's waters. Thus, as a part of the privilege of discharging, NPDES permittees have an obligation to monitor discharges and/or receiving waters to ensure they are not harming beneficial uses.

2) While permittees have an obligation to monitor the environment, the state and federal regulatory agencies must not make NPDES permittees monitor for monitoring's sake alone. All monitoring must be conducted to answer specific management questions and to assess the need for more (or less) management actions. If monitoring does not answer a management question, then the value of that monitoring should be carefully re-evaluated to protect the public's resources.

Technical Approach

A four-step approach, aligned with the project objectives, was used to conduct this monitoring program review.

1) Identify the monitoring questions of greatest interest

2) Investigate how well the current monitoring design answers the priority monitoring questions

3) Evaluate monitoring design updates and upgrades to increase effectiveness and efficiency

4) Provide technical guidance and hands-on scientific tools

Identify the monitoring questions of greatest interest

Monitoring questions are the most important part of designing a monitoring program. They are also perhaps the most challenging because it is largely non-technical; the monitoring questions should be tied to management concerns and specific management actions.

Monitoring questions are comprised of four parts including:

- Spatial and temporal scale: the where and when to monitor
- Indicators: what to monitor
- Benchmarks: the threshold(s) that define good to bad monitoring results
- Action: what managers, either from regulated or regulatory agencies, are going to do once they have an answer to the monitoring question. Managers are defined as either Permittee or RWQCB staff that are making decisions about monitoring program implementation, or the actions taken based on monitoring results.

Of these four criteria, perhaps the most important is the action(s) to be taken because this adds the element of need.

The Advisory Committee spent three meetings – a total of nearly 15 hours – creating a list of monitoring questions, prioritizing which questions were the most important, and refining the questions to address their most pressing management needs. Then, for each of the prioritized monitoring questions, the Advisory Committee defined the four attributes (see Appendix A).

Identify how well the current monitoring design answers the priority monitoring questions

There were two steps to assess how well the current monitoring design answers the priority monitoring questions. The first step was to compile existing monitoring data. Most often, these data were made available from the County of Orange. These data were then checked for completeness and quality assurance. The output of this step was to create an inventory of available monitoring data.

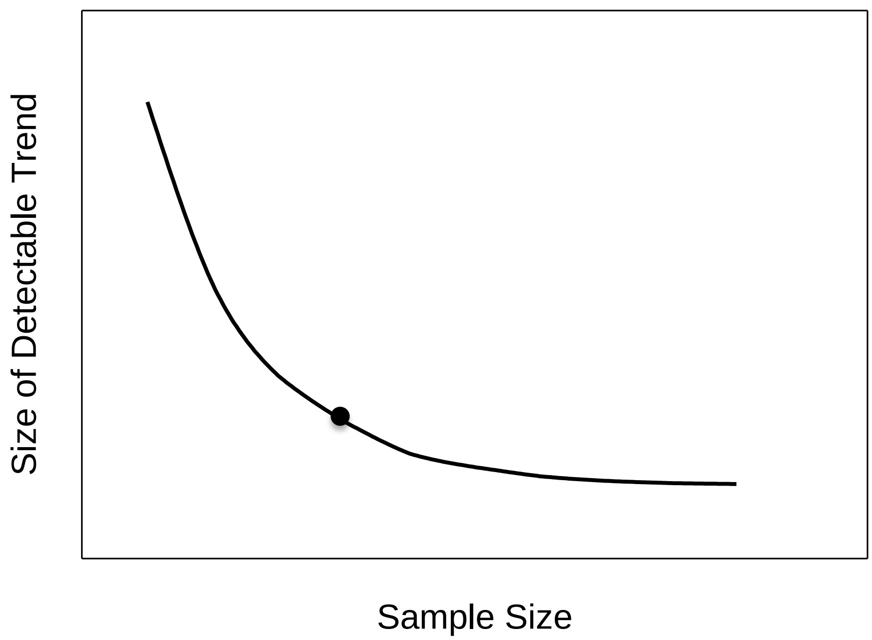
The second step was to conduct data analysis to answer the monitoring questions. This quantitative analysis was conducted independently of previous NPDES reporting. The independence was partly to ensure transparency and objectivity in the monitoring design evaluation. The independence was also partly due to new monitoring questions, which required data analysis that had never been attempted.

Evaluate monitoring design updates and upgrades to increase effectiveness and efficiency

Two steps were used to evaluate monitoring design updates and upgrades. The first step was statistical evaluations of efficiency, which was largely conducted using power analysis. The analysis focuses on the spatial and temporal elements of the monitoring question. Power analysis is a commonly used statistical tool for optimizing sample size, sampling frequency, and site selection with a simple goal of quantifying the confidence in the monitoring question answer (Figure 1). Generally speaking, more samples means greater confidence, but the increase in confidence is not linear. When sample size is small, adding more samples can dramatically increase confidence. At some point, however, more samples do not proportionately increase confidence. In the absence of a specified level of confidence, most managers look to optimize power, which is the inflection point of the sample size-confidence power curve. This inflection point is the level of sampling effort at which reasonable confidence is achieved without expending additional resources in sampling for minimal returns on confidence.

The second step to evaluate monitoring design updates and upgrades was exploring potential uses of new sampling technologies. In this case, the new technologies must address the measurement indicators of the monitoring question. These new indicators might be more accurate, faster to measure, or cheaper than current indicators, while providing the same or improved information.

Figure 1. Example power curve. More samples translate into smaller detectable differences. Optimal sample size is at the inflection point (black circle) of the curve where additional sampling does not yield appreciably greater power to detect smaller trends.



Provide technical guidance and hands-on scientific tools

Several web applications – one for each question – were created and used for technical guidance and providing hands-on scientific tools (see Appendix B for specific web application methods and documentation). Searchable data inventories, on-the-fly data analysis subroutines, and iterative power analysis calculators were created for use by the Advisory Committee during and after project completion. This allows each agency the opportunity to dive into the monitoring design options in detail, and see in real-time, how even small changes in frequency, replication, or sample size can alter confidence and conclusions in monitoring questions. Monitoring questions with larger management actions may require more confidence, while monitoring questions with smaller management consequences can live with less confidence. Links to the web applications are provided in each section below.

MONITORING QUESTIONS

In total, the Advisory Committee generated a list of 11 monitoring questions.

- **What is the temporal trend in pollutant concentrations and loading? [Mass emissions monitoring]**
- **What are the trends in fish and bird tissue concentrations? [Bioaccumulation monitoring for Newport Bay watershed]**
- **Are there illegal discharges/illicit connections (ID/IC)? [Dry weather monitoring]**
- **What are the temporal trends in sediment quality and water column quality in estuaries and wetlands? [Estuary/Wetland Monitoring]**
- Are concentrations below water quality thresholds?
- What are the sources of pollutant loading?
- What are the [natural] background levels of pollutants?
- What is the effectiveness of best management practices (BMPs)?
- What is the status and trends of receiving water biological condition?
- What is the status of program effectiveness?
- How can North Orange County monitoring link/incorporate non-permit monitoring data?

Of these 11 questions, four were ranked and prioritized for evaluation as part of this study (in bold). It is not that the other questions were not important; the Advisory Committee deemed these four questions the most critical in preparing for revision of the North Orange County Monitoring Program. The NPDES permit may still require these additional questions or other activities.

Core monitoring, Regional monitoring, Special studies

Monitoring approaches fall into one of three categories: Core monitoring, Regional monitoring, or Special Studies (Figure 2).

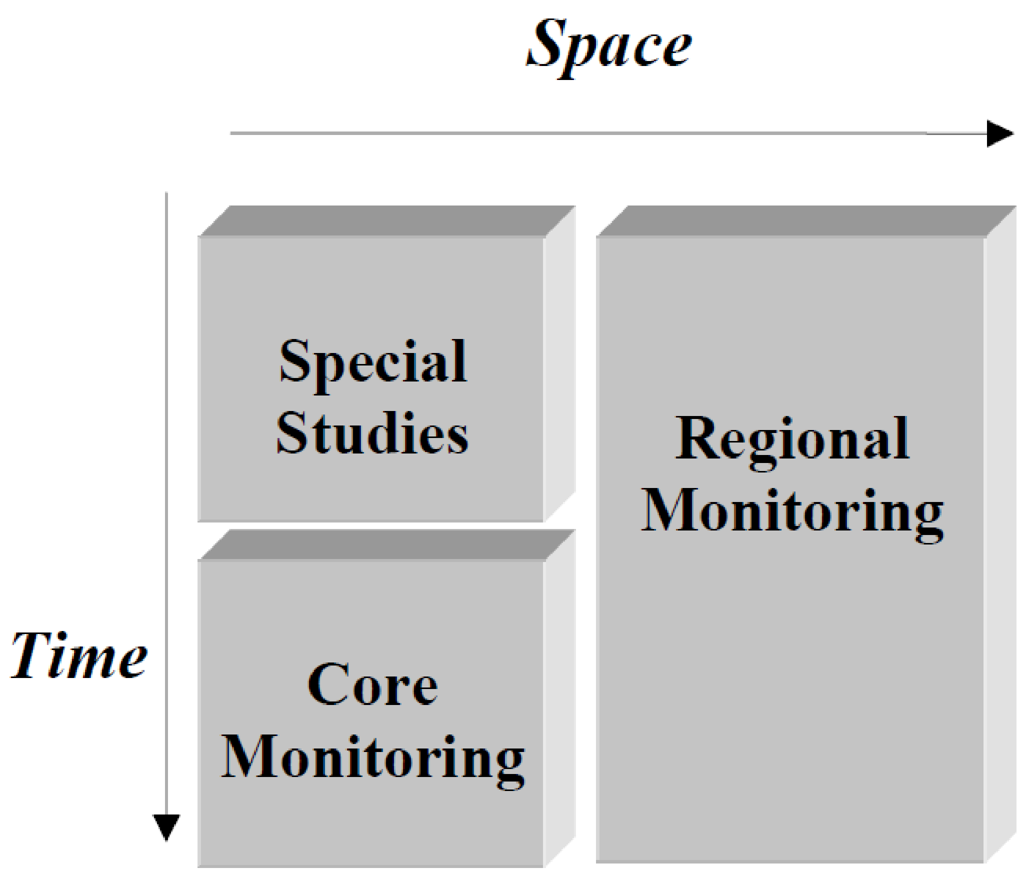
These three categories acknowledge there are different time and space scales required to answer the critical monitoring questions posed by managers.

Core monitoring is often site-specific and long-term, answering monitoring questions that are fundamental to compliance. These measurements are focused on issues such as discharge concentrations or loads; regardless of what today's monitoring tells a manager, these measurements will continue because it is fundamental to compliance requirements.

Regional monitoring can be short- or long-term, but it always large in spatial scale because the monitoring questions require context beyond a single location. Management questions that require regional monitoring typically need the breadth and depth of natural variability to assess reference conditions or need to incorporate multiple stressors from cumulative sources. Good examples include biological assessments, where impacts could be caused by local stressors (i.e., discharges) or by regional stressors (i.e., climate change). Placing a chosen site on the range of impacted sites tells managers the magnitude of change and reinforces the need action (or no action).

Special studies often address small spatial and short-term temporal scales addressing focused questions with a distinct beginning, middle, and end. These studies may last one or two years to address unknown issues associated with core or regional monitoring results. A good example might be following up toxic samples with a toxicity identification evaluation or a source tracking study triggered by a dry weather exceedance.

Figure 2. Schematic of core monitoring, regional monitoring, and special studies.



Q1: WHAT ARE THE TEMPORAL TRENDS IN POLLUTANT CONCENTRATIONS AND LOADING?

Two web applications were created for this question: one for concentration data and the other for pollutant loading data.

- Mass emissions, concentration data: https://sccwrp.shinyapps.io/ocms4review/mass_emissions.Rmd

- Mass emissions, pollutant loading data: https://sccwrp.shinyapps.io/ocms4review/mass_emissions_loads.Rmd

The premise of this question is that receiving water sites located in each of the three North Orange County watersheds – Newport Bay, San Gabriel River/Coyote Creek, Anaheim Bay/Huntington Harbour – are monitored during dry weather and wet weather for concentrations and pollutant loads for up to 42 parameters (inclusive of the hundreds of different isomers and congeners for organic compound groups such as PCBs, DDTs, toxaphene, etc.). Concentrations and loads are compared across years in wet and dry weather, and also compared to TMDL targets, if applicable. Actions managers may take once monitoring questions are answered include:

- Trigger enhanced monitoring as concentrations or loads approach benchmarks (i.e., water quality standard, TMDL target, etc.), to increase confidence that data are above or below the benchmark

- Trigger advanced data analysis for complicating factors that could influence a benchmark exceedance (i.e., precipitation patterns for wet weather)

- Identify trends in concentrations and loads, and factors leading or contributing to the trends

- Determine whether management actions have yielded desired outcome, or design/update future management actions to achieve the desired outcome

- Trigger source identification in order to continually decrease concentrations or loads

The web application content is separated into four main analyses:

- Inventory: Map-based and tabular summaries of monitoring effort and basic characteristics of the data

- Trends and power analyses: Changes over time by selected constituents and locations, including power analyses to help identify optimal sampling effort

- Station differences: A comparison of time series between stations to identify similarities among trends

- Overall trends: Map-based and tabular summary of trend tests for all stations shown together

Each analysis includes sub-tabs or drop-down menus for selecting and viewing different content. Most analyses are also grouped by constituent type such as nutrients, metals, or organics:

- Nutrients: Ammonia, nitrate, nitrite, Total Kjeldahl Nitrogen, orthophosphate, total phosphorus
- Metals/trace elements: Ag, As, Cd, Cr, Cu, Fe, Hg, Ni, Pb, Se, Zn
- Organics: Azinphos methyl (Guthion), Bolstar, Chlorpyrifos, Coumaphos, Demeton-o, Demeton-s, Diazinon, Dichlorvos, Dimethoate, Disulfoton, Ethoprop, Ethyl Parathion, Fensulfothion, Fenthion, Glyphosate, Malathion, Merphos, Mevinphos, Parathion-methyl, Phorate, Ronnel, Tetrachlorovinphos, Tokuthion, Trichloronate, DDTs, PCBs, Pyrethroids

Because concentrations and variance differ in dry or wet weather, each analysis can be performed for either condition.

Inventory

Sixteen stations were monitored in the three OC Watersheds during the 10-year span from January 2010 to December 2019 (Figure 3; Table 2). There were more sites for nutrients (sites=16) compared to metals or organics (sites=12). Over the decade time span, there were roughly 2,500 samples for nutrients, 1,400 samples for organics, and 3,000 samples for metals. The actual number of samples varied by site.

Answer to the question

In general, nutrients were the parameter of greatest interest since nitrate+nitrite was most frequently detected and most frequently exceeded benchmarks. Nearly all of the organic constituents were non-detectable except for malathion and glyphosate. Similarly, mercury was rarely detected.

At some sites, nitrate+nitrite decreased over time. For example, site SDMF05 (San Diego Creek at Campus Dr) significantly decreased between 2010 and 2019 in both dry and wet weather (Figures 4 and 5).

The web application can be used by managers to conduct trend analyses for these and other parameters to decide if progress is being made and whether actions need to be taken.

Updates and Upgrades

Trend monitoring is fundamentally a core component of the North Orange County Monitoring Program. Regardless of an increasing, decreasing, or static trend, the information is critical because there are management actions associated with each answer. Ultimately, trend questions are part of the basic philosophy of monitoring requirements, to ensure that the regulated stormwater discharge(s) is not leading to increased pollution and impacts to beneficial uses.

Given that this trend question is a necessary and ongoing element of the North Orange County's core monitoring, the next step is to ensure efficiency to answer the needs of the regulators most confidently and to address the needs of the regulated parties most effectively. Efficiency of answering trend questions is a function of three factors:

1) Amount of changes in pollutant concentrations or loads

2) Amount of time

3) Amount of variation

As discussed in the technical approach section on increases effectiveness and efficiency of monitoring designs (see page 4), monitoring data with small changes, few samples over a short time, or more variation generally have lower power to detect trends. Monitoring data with large changes, frequent samples over longer time intervals, or smaller variance generally have higher power to detect trends. Statisticians who design monitoring programs use power curves to estimate confidence in answers to trends questions (see Figure 1). Managers interested in the most efficient sampling designs will aim for sampling frequencies (e.g., sample size) at the inflection point of a power curve (i.e., the optimal tradeoff between confidence and sample effort).

For most constituents at most sites, less sampling than is currently being expended can achieve similar levels of statistical confidence. An example using nitrate+nitrite is shown in Figure 6. In this case, regardless of station, trends in concentration can be optimally detected at less than current frequencies (usually at or around 50% of current frequency) regardless of station. Most, but not all, parameters follow this same pattern. The web app can calculate these results for every parameter at every site as regulated and regulatory agencies decide if changes in sampling frequency are appropriate.

The Advisory Committee considered the potential reduction in frequency and decided that frequency reduction may not be appropriate if concentrations or loads were near the benchmark of success or failure. That is, more confidence was necessary when detected values are near a threshold of interest, such as numeric target, and decisions need to be made. Therefore, the web app created a visualization tool for identifying which sites were near the benchmark so that managers may choose greater than optimal sampling frequencies for those sites. In the nitrate+nitrite example from Figure 6, the need for greater confidence at sites near the benchmark would likely still not exceed existing frequencies.

When confronted with the concept of reduced sampling frequencies, the Advisory Committee realized that monitoring logistics also played a role. The rationale was that when sampling, especially wet weather sampling, it is best to collect samples for all analysis at as many sites as possible during the same storm. Therefore, managers should consider all parameters at all sites as expressed in Figure 7. The Advisory Committee discussed several approaches to selecting optimal frequencies with so many different combinations: most sensitive parameter, most sensitive station, median sensitivity, amongst others. The Advisory Committee ultimately decided that, regardless of approach, optimal sampling frequencies will either stay the same or decrease.

Trend monitoring for pollutant loading can be distinctly different than trend monitoring for concentrations. In the example described above, the focus was on trends in concentrations. However, the monitoring data from North Orange County monitoring indicated very similar conclusions for monitoring design upgrades to pollutant load monitoring as the conclusions for concentrations. In general, pollutant loads monitoring can be decreased and still achieve optimal trend detection.

Table 2. Inventory of sample analyses for nutrients by station for mass emission sites in Newport Bay, Anaheim Bay/Huntington Harbour, and San Gabriel River/Coyote Creek between January 2010 and December 2019. Stations are found in Figure 3 and in the web app.

Station Code	Ammonia	Nitrate, Nitrite	Total Kjeldahl Nitrogen	Orthophosphate	Total Phosphorus	Total
ACWF18	132	128	132	131	132	655
BARSED	252	246	252	252	252	1254
BCC02	67	61	67	67	67	329
BCF04	102	100	102	102	102	508
CARB01	50	46	50	47	50	243
CCBA01	51	46	51	50	51	249
CICF25	167	162	167	167	167	830
CMCG02	397	390	397	396	397	1977
EGWC05	70	65	70	70	70	345
FCVA03	60	56	60	58	60	294
LANF08	97	97	97	97	97	485
MIRF07	145	141	145	145	145	721
SADF01	307	300	307	313	307	1534
SDMF05	410	404	410	409	410	2043
SICG03	1	1	1	1	1	5
WYLSED	263	257	263	262	263	1308
Total	2571	2500	2571	2567	2571	12780

Figure 3. Map of mass emission sites in Newport Bay, Anaheim Bay/Huntington Harbour, and San Gabriel River/Coyote Creek between January 2010 to December 2019.

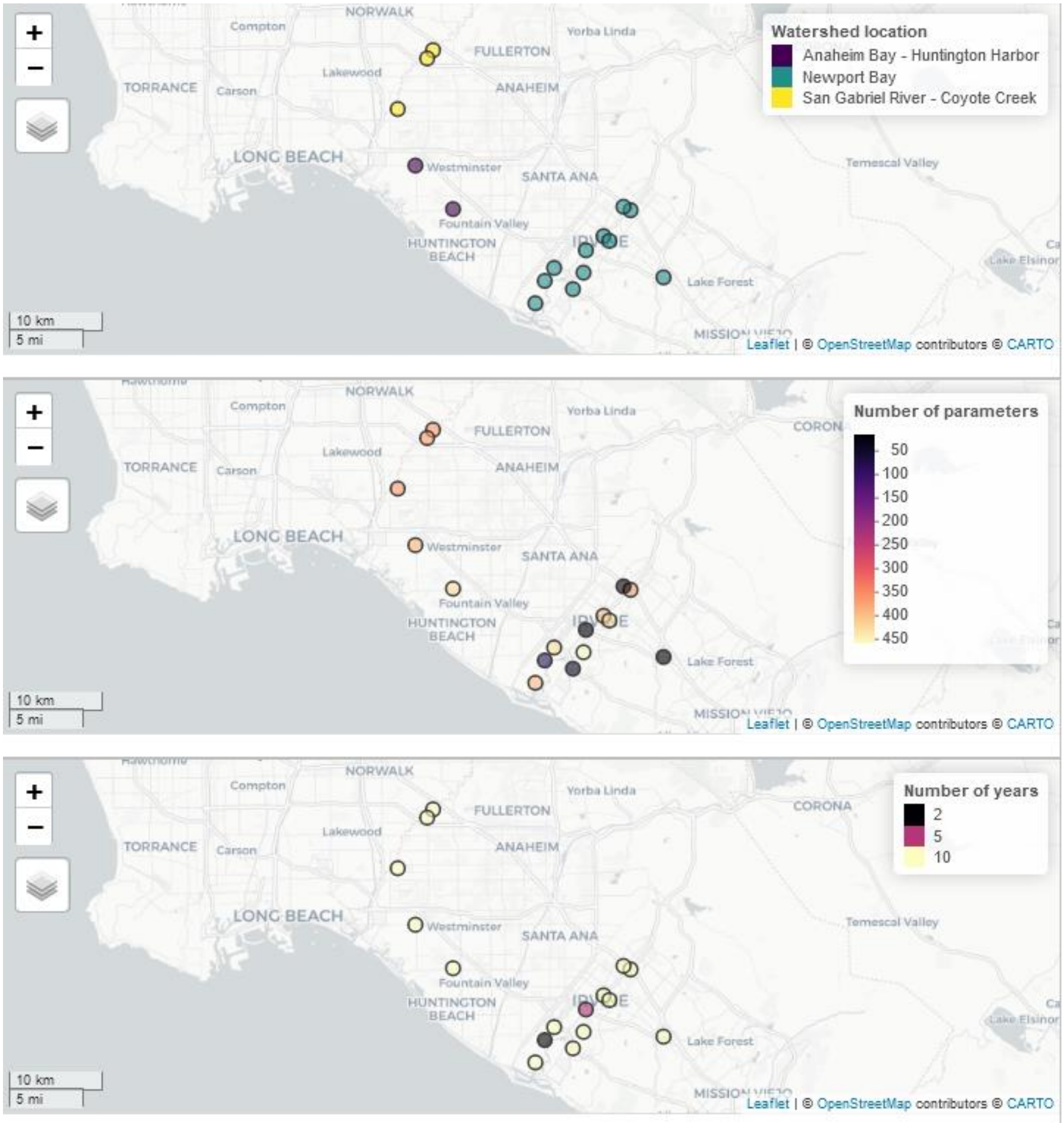
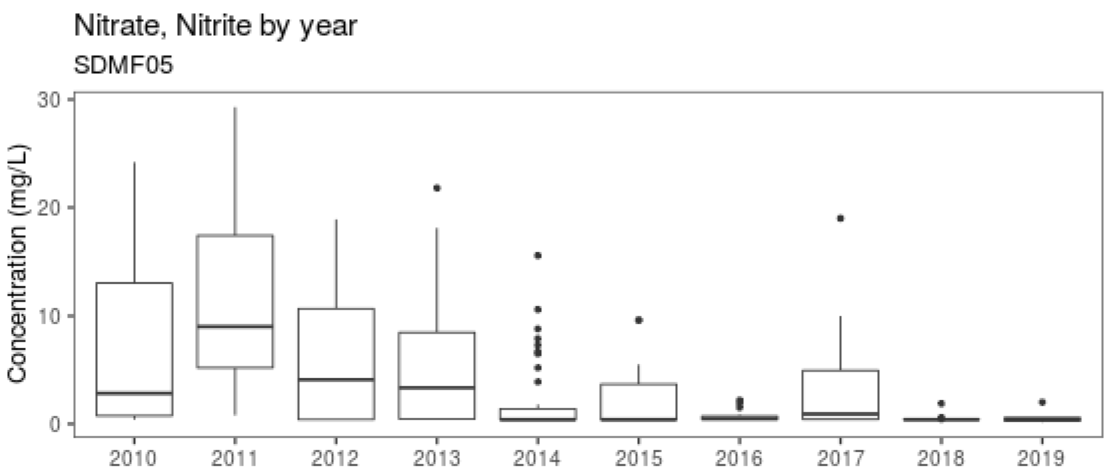
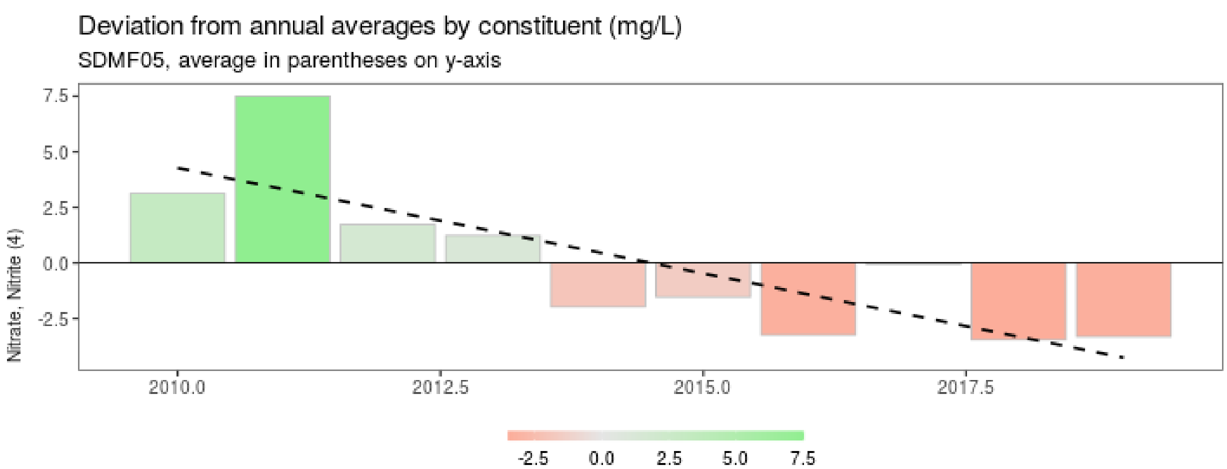


Figure 4. As an example of answers to the monitoring question about trends at mass emission sites, nitrate+nitrite at station SDMF05 (San Diego Creek at Campus Drive) during dry weather between January 2010 and December 2019. A) Box plots of concentration (box hinges are 75, 50, 25 percentiles of the distribution, top and bottom whiskers are 1.5 times the interquartile range); B) Results of trend analysis using residuals from the grand mean (in parenthesis on y-axis).

A)



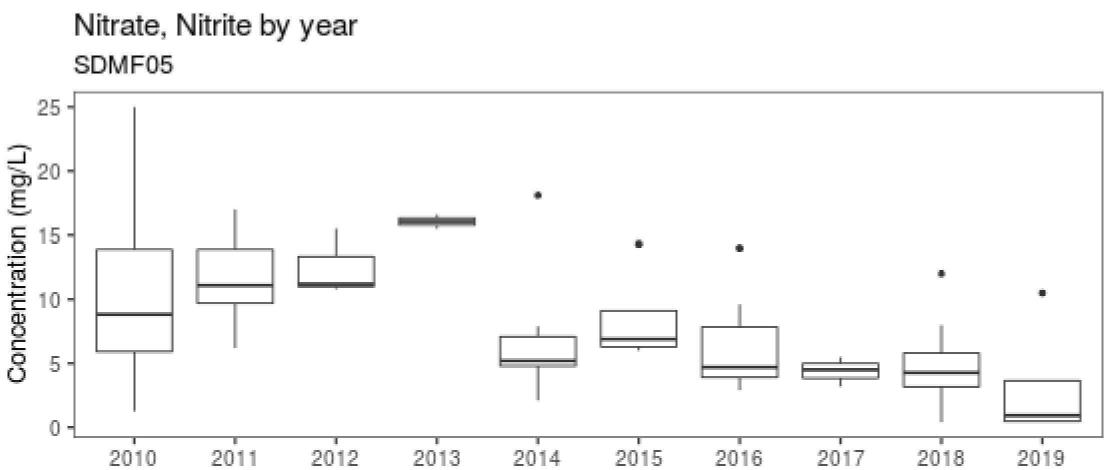
B)



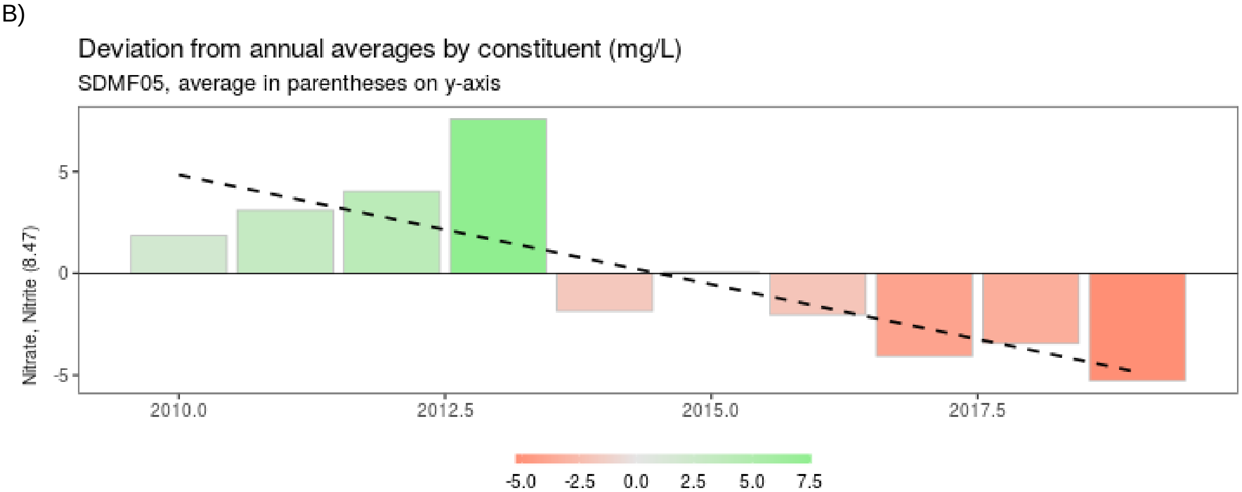
	Dependent variable:
	Nitrate, Nitrite
Year	-0.95***
	(0.23)
Intercept	1,910.31***
	(470.32)
Observations	10
R^2	0.67
F Statistic	16.50*** (df = 1; 8)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 5. As an example of answers to the monitoring question about trends at mass emission sites, nitrate+nitrite at station SDMF05 (San Diego Creek at Campus Drive) during wet weather between January 2010 and December 2019. A) Box plots of concentration (95, 75, 50, 25, 5 percentiles); B) Results of trend analysis using residuals from the grand mean (in parenthesis on y-axis).

A)



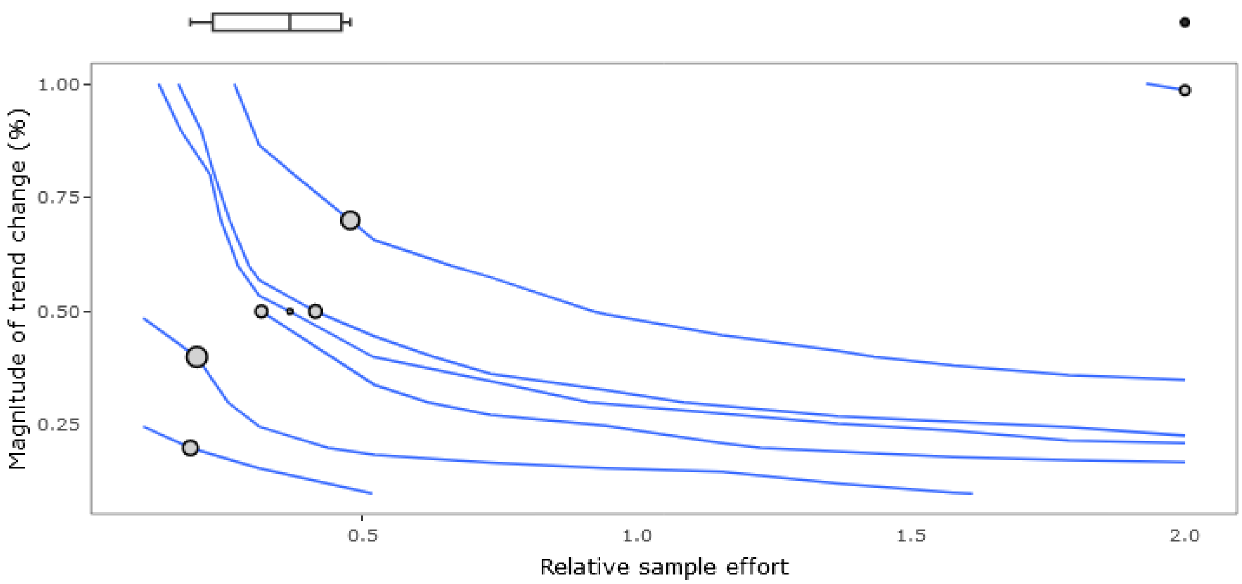
B)



	Dependent variable:
	Nitrate, Nitrite
Year	-1.08***
	(0.29)
Intercept	2,166.37***
	(582.23)
Observations	10
R^2	0.63
F Statistic	13.84*** (df = 1; 8)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 6. Example power curves illustrating the optimal sampling frequency for nitrate+nitrite at SDMF05. Each curve is a different station. Y-axis is the magnitude of change (smaller detectable changes means more samples). X-axis is the sampling frequency relative to existing frequency (< 1.0 means less than current sampling). Gray symbols represent the optimal sampling frequency where smaller detectable differences requires proportionately more samples. Size of gray symbol represents how close the current concentration is to the benchmark (larger symbols are closer, perhaps requiring smaller detectable differences). Box plot above the chart is the distribution of optimal sampling frequencies. All trends assume beta=0.8, alpha=0.05.

A) DRY WEATHER



B) WET WEATHER

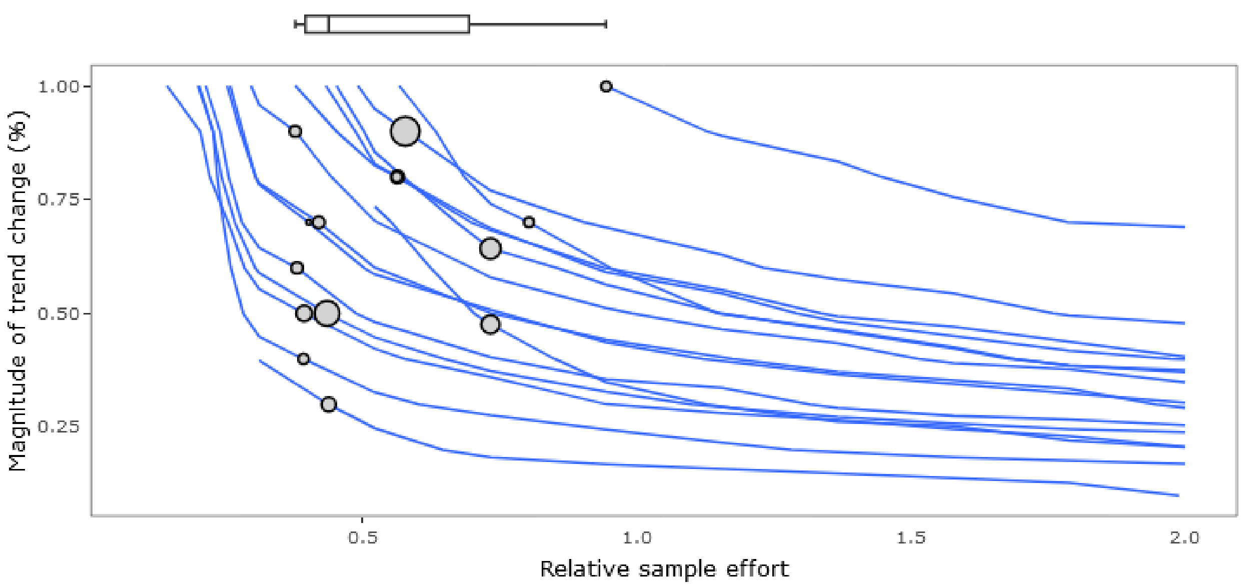
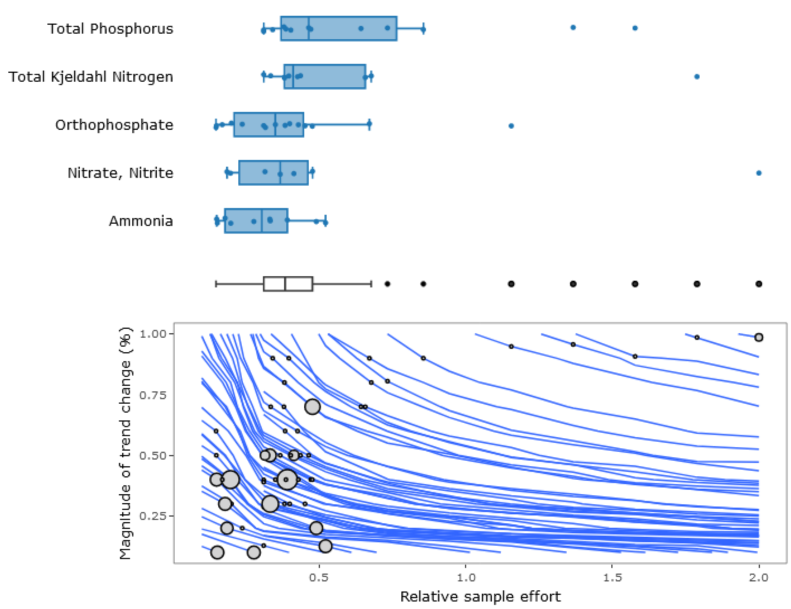
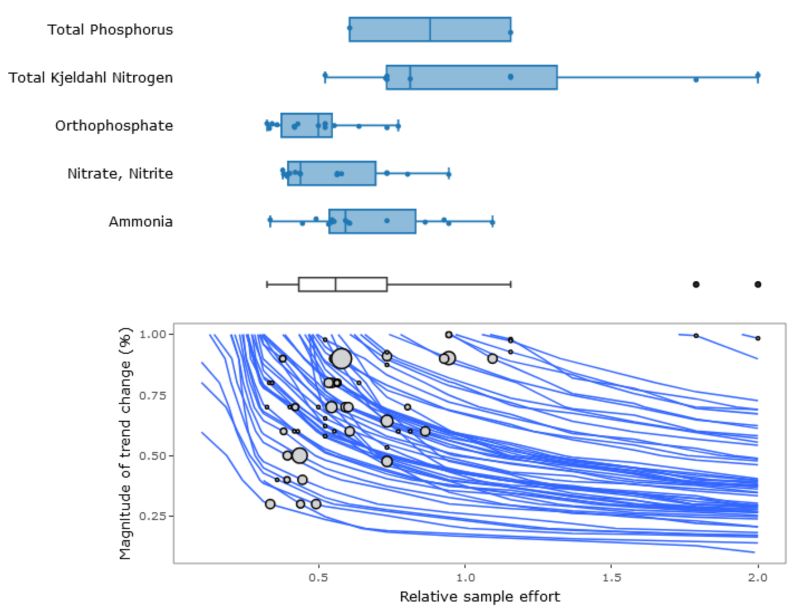


Figure 7. Example power curves illustrating the optimal sampling frequency for all nutrients across all sites. Each curve is a different station-parameter combination. Y-axis is the magnitude of change (smaller detectable changes means more samples). X-axis is the sampling frequency relative to existing frequency (< 1.0 means less than current sampling). Gray symbols represent the optimal sampling frequency where smaller detectable differences requires proportionately more samples. Size of gray symbol represents how close the current concentration is to the benchmark (larger symbols are closer, perhaps requiring smaller detectable differences). Box plots above the chart are the distribution of optimal sampling frequencies. All trends assume beta=0.8, alpha=0.05.

A) DRY WEATHER



B) WET WEATHER



Q2: WHAT ARE THE TEMPORAL TRENDS IN FISH AND BIRD TISSUE CONCENTRATIONS?

A web application was created for this question:

- Tissue contaminant monitoring tool: https://sccwrp.shinyapps.io/ocms4review/tissue.Rmd

The premise of this question is to assess if tissue concentrations in multiple species of fish and bird eggs at sites located across the Newport Bay watershed are increasing or decreasing over time and if the applicable TMDL targets are met (the Anaheim Bay - Huntington Harbour watershed is not currently subject to a TMDL). Actions managers may take once monitoring is conducted include:

- Evaluation of current remediation action effectiveness
- Focusing future remediation planning and implementation
- Listing/delisting decisions
- Additional studies on pollutant transfer/bioaccumulation modeling to better quantify linkage analysis and TMDL targets

The web app content is separated into four main analyses:

- Inventory: Map-based and tabular summaries of monitoring effort and basic characteristics of the data
- Trends and power analyses: Changes over time by select constituents and locations, including power analyses to help identify optimal sampling effort
- Station differences: A comparison of time series between stations to identify similarities among trends
- Overall trends: Map-based and tabular summary of trend tests for all stations shown together

Inventory

In the Newport Bay Watershed, nine stations were monitored for bioaccumulation, eight stations for fish tissues, and seven stations for bird eggs during the 10-year span from January 2010 to December 2019 (Figure 8; Tables 3 and 4). Virtually all samples were collected once per year, in the spring/summer (i.e., nesting season). Samples were analyzed for DDTs, chlordane, PCBs, arsenic, mercury, and selenium, as well as lipids and moisture content. For all analyses, multiple congeners for the same constituent were summed to create the total tissue concentration for DDT, PCB, and chlordane. Over the decade time span, there were over 3,000 analyses conducted on up to 560 samples. The number of samples varied by site, species, and tissue type. Frequently, multiple individuals of the same species were composited to create sufficient tissue mass for sample analysis. Where multiple composite samples were taken of the same species and tissue type on the same day, composite sample results were averaged for this analysis (mathematical compositing).

Answer to the question

In general, there were too few samples to assess trends for almost any parameter (Tables 5 and 6). Although many samples had been collected for bioaccumulation, they often came from a mix of different species and different tissue types. No site had a full complement of data across the 10-year record for a single species, largely based on bird egg availability (i.e., breeding and hatching success). For example, Station IRWD had the most bioaccumulation samples for the Newport Bay watershed. However, four of the six species sampled for bird eggs had less than seven years of data and all 13 of the fish species had less than eight years of data. Maintaining intra-species is critical for trend detection because bioaccumulation can vary dramatically depending on age, sex, prey, life history status, and trophic level (Lavoie et al. 2013; Loganathan et al. 1994).

Given that the number of individual species with sufficient record of data was limited, the web application illustrates what the current trends look like (Figure 9). For example, trends for selenium and total PCBs in bird eggs of the American avocet and composite fish tissues of largemouth bass were examined between 2012 and 2019 at site IRWD. Despite various species-pollutant combinations having increasing trends, another decreasing, and others neither increasing nor decreasing (i.e., near zero slope), none of the trends were statistically significant.

Updates and Upgrades

Similar to the water quality monitoring at mass emission sites, bioaccumulation trend monitoring is fundamentally a core component of the North Orange County Monitoring Program. Regulators, regulated, and the public will always want to know the answer to this question. Likewise, the efficiency of answering bioaccumulation trends questions mirror the mass emission site trends question; trends are a function of three factors:

1) Amount of change
2) Length of time
3) Amount of variation

Unlike water quality at mass emission sites, sampling frequency for bioaccumulation is limited to once-per-year sampling. This is particularly true for bird egg sampling, which is focused entirely around nesting season. Since sampling frequency is fixed, the power analysis switches from how many samples per year to detect a trend based on effort in the existing record of ten years, to how many years will it take to see an optimal trend when sampling once per year.

Consistent with the results to date, power analysis indicates trend detection will not occur rapidly. In nearly all cases, when averaging across sites, bioaccumulation trends may take one or more decades to observe (Figure 10). As an example, web app calculations for time to trend detection for bird eggs (American avocet) will take on average 8 years; time to trend detection for fish tissue composites (largemouth bass) will take on average 15 years.

The Advisory Committee struggled with the challenge of fixed sampling frequencies, but ultimately agreed that bioaccumulation monitoring was important even if significant results could not be observed for many years. The long compliance period of the TMDL also recognized this and compliance dates extend until 2056. The Advisory Committee recognized trends would likely be slow since management actions may take many years to implement. In order to optimize trend detection as much as possible, focusing on fewer, but more consistent species and tissue types is recommended. The criteria for species selection should include:

- Length of time already monitored (longer is preferred)
- Consistency of species availability (consistency is better)
- Tissue mass per individual (more mass is better)
- Existing concentration proximity to wildlife or human health risk thresholds (closer to threshold is preferred)

Monitoring should focus on sites and species that have a longer historical record. This maximizes the efforts to date and the likelihood of detecting trends most expeditiously. Exceptions may be made if there is a specific location of concern, perhaps a site undergoing management actions.

Monitoring should focus on species that are most commonly caught, and likely to be caught in the future. Catch per unit effort is not just a function of resource allocation, but it is also fundamental to detecting trends. Missed sample years because of absent species will reduce the ability to detect trends. Thus, commonly occurring species are preferable to intermittently occurring species.

Monitoring should focus on species with greater tissue mass than species with less mass. Tissue mass is a correlate to species availability. Many instances in the historical data were characterized by imbalanced sample sizes because insufficient tissue mass existed to analyze all pollutants. Therefore, species with more tissue mass are generally preferred over species with less mass, even when compositing, to ensure all lab analyses can be completed.

Monitoring should prioritize site-species-pollutant combinations that are closest to thresholds. While trends are important for all sites, species, and pollutants, there is a clear need to assess trends with greater confidence when the pollutant concentration is about to cross an important management threshold such as a TMDL target. Therefore, managers may want to give preference to those locations and species that are closest to a compliance decision.

Sites with few samples, inconsistent species or tissue types, and further from thresholds may be less important for sampling and managers may consider reducing or eliminating this monitoring effort. However, if there are lingering management concerns about these inconsistent site-species combinations, an optional and less resource intensive study design may focus on pair-wise differences between years. This subtle change in monitoring question assesses differences between two different time periods (i.e., sampling once every 10 years), instead of trying to detect trends over multiple years per decade. The design then switches from annual sampling to replicate sampling in the two years to be compared. For example, the Southern California Bight Regional Monitoring Program and the State Water Board’s Statewide Bioassessment Oversight

Group utilizes a once-per-decade approach for its coastal sportfish bioaccumulation monitoring program design (Davis et al. 2012).

Table 3. Inventory of bird egg tissue samples by pollutant for each monitoring site in the Newport Bay Watershed between January 1, 2010 and December 31, 2019.

StationCode	%Lipid	%Solids	Total chlordane	Se	Hg	DDT	PCB	Toxaphene	Total
BCGC	9	9	15	9	9	18	18	9	96
IRWD	36	36	72	36	36	72	80	42	410
SDC	6	6	13	6	6	12	15	7	71
UCI	22	22	31	22	22	43	44	22	228
UNB	23	23	48	23	23	45	52	32	269
Total	96	96	179	96	96	190	209	112	1074

Table 4. Inventory of fish tissue composites by pollutant for each monitoring site between January 1, 2010 and December 31, 2019.

StationCode	%Lipid	%Solids	Total chlordane	Se	Hg	DDT	PCB	Toxaphene	Total
BCW	12	12	17	12	12	23	25	15	128
IRWD	43	43	72	43	43	80	74	58	456
PCW	39	39	56	39	39	72	67	48	399
SAD	4	4	7	4	4	8	8	4	43
SDC	42	42	71	42	42	78	83	49	449
UCI	9	9	12	9	9	16	15	9	88
UNB	4	4	8	4	4	10	10	4	48
uPCW	1	1	2	1	1	2	1	1	10
Total	154	154	245	154	154	289	283	188	1621

Table 5. Number of sample years with bird egg tissue results for each monitoring site between January 1, 2010 and December 31, 2019.

StationCode	%Lipid	%Solids	DDT	Hg	PCB	Se	Total chlordane	Toxaphene
BCGC	5	5	5	5	5	5	5	5
IRWD	7	7	7	7	7	7	7	7
SDC	2	2	2	2	2	2	2	2
UCI	7	7	7	7	7	7	7	7
UNB	7	7	7	7	7	7	7	7

Table 6. Number of years with fish tissue composite samples for each monitoring site between January 1, 2010 and December 31, 2019.

StationCode	%Lipid	%Solids	DDT	Hg	PCB	Se	Total chlordane	Toxaphene
BCW	8	8	8	8	8	8	8	8
IRWD	8	8	8	8	8	8	8	8
PCW	8	8	8	8	8	8	8	8
SAD	4	4	4	4	4	4	4	4
SDC	8	8	8	8	8	8	8	8
UCI	5	5	5	5	5	5	5	5
UNB	1	1	1	1	1	1	1	1
uPCW	1	1	1	1	1	1	1	1

Figure 8. Tissue station locations in the Newport Bay Watershed.

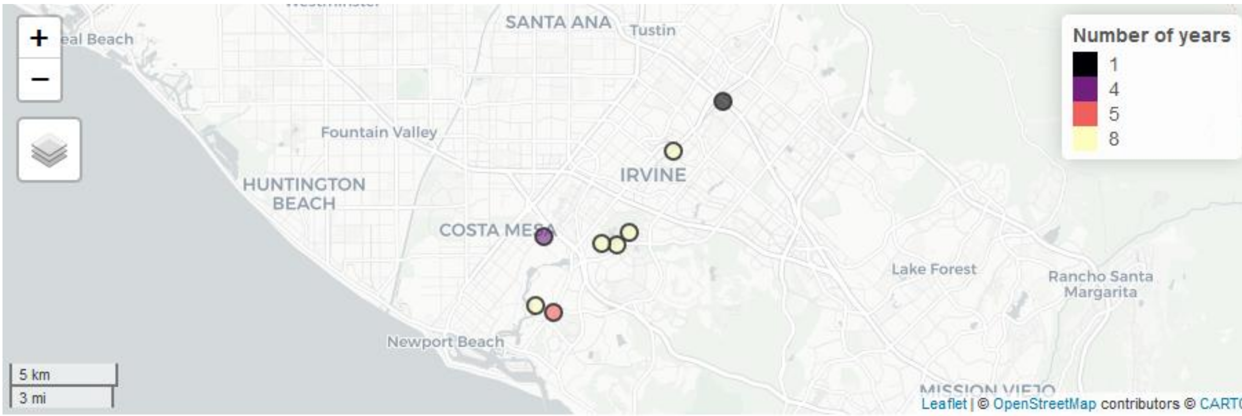


Figure 9. Trends in bird egg and fish tissue. Top = American Avocet, Selenium, PCB. Bottom = largemouth bass, selenium, PCB. None of the trends are statistically significant.

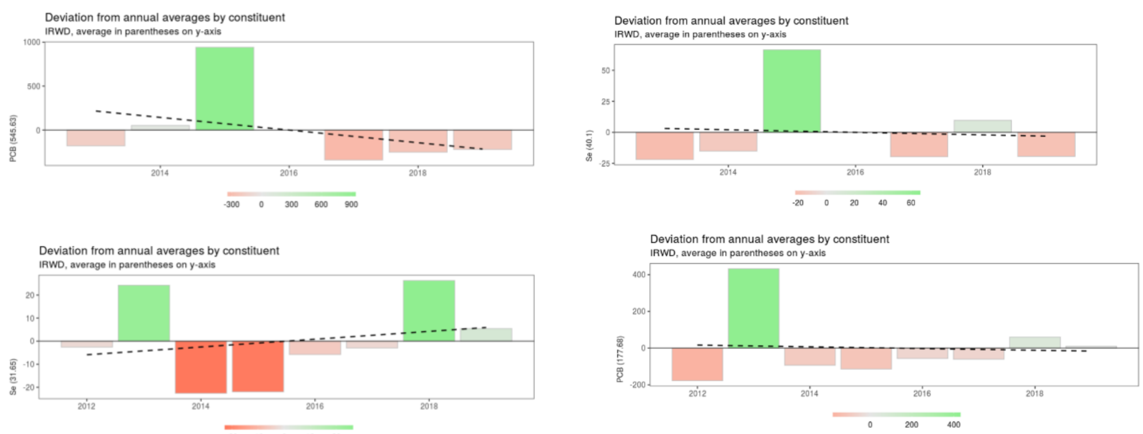
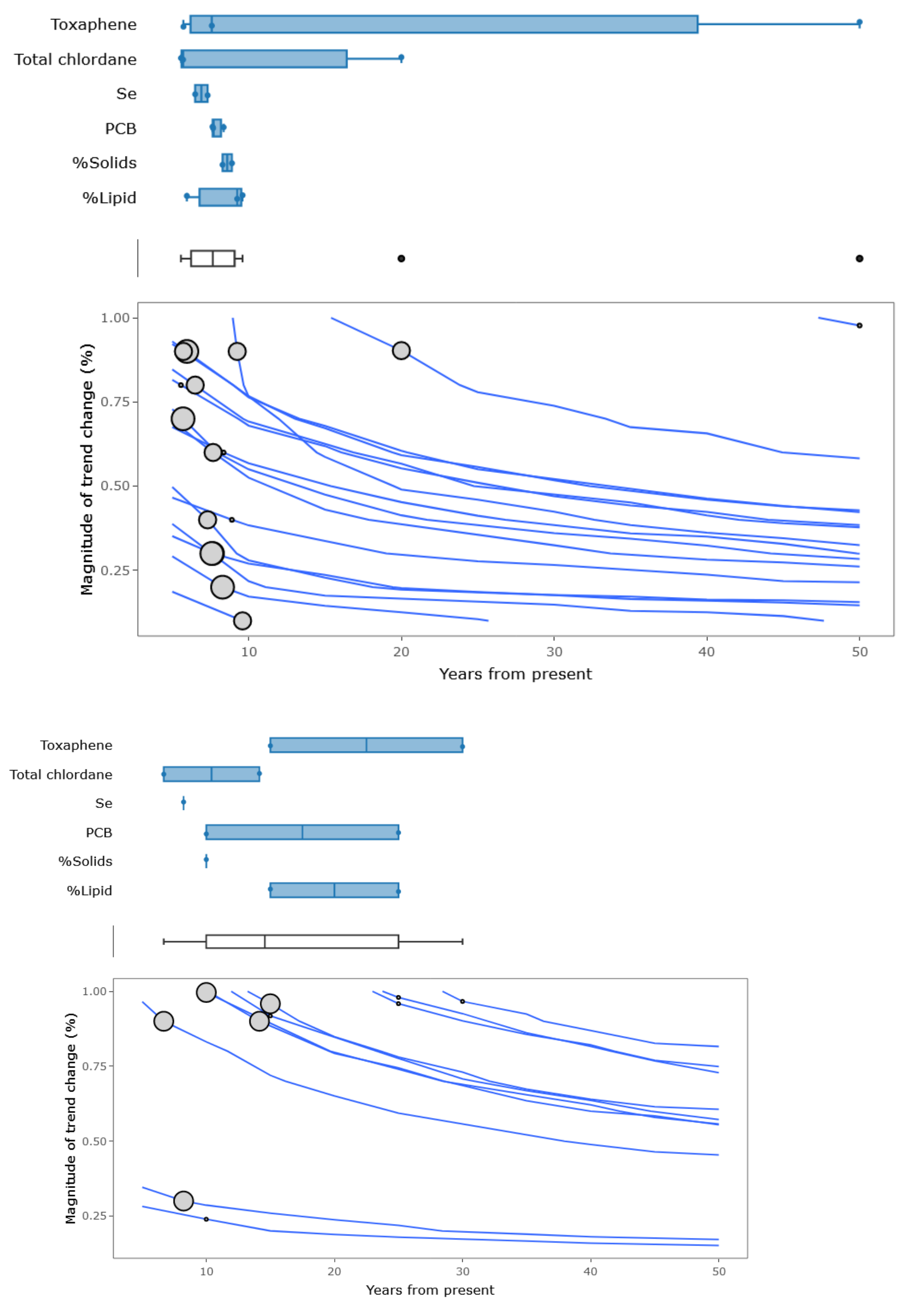


Figure 10. Power analysis estimating the time to trend detection for once-per-year sampling (beta=0.80). Top = American avocet for pollutant-station combinations. Bottom = Largemouth bass pollutant-station combinations.



Q3: ARE THERE ILLEGAL DISCHARGES/ILLICIT CONNECTIONS (ID/IC)?

A web application was created for this question:

- Dry weather monitoring: https://sccwrp.shinyapps.io/ocms4review/dry_weather.Rmd

The premise of this question is to assess the general water quality conditions and trends of discharges from storm drain outfalls and whether the discharges result from illegal discharges/illicit connection (ID/IC) during dry weather. Specifically, managers are looking for ID/IC in the municipal separate storm sewer (MS4) system. The monitoring consists of both field monitoring and laboratory analyses. Field monitoring utilizes a visual evaluation of storm drain discharges and employs water quality probes to measure parameters such as temperature, pH, electric conductivity, and dissolved oxygen. Field kits are used to measure concentrations of chlorine, nitrate, and copper, among others. Water samples are also taken and sent to a laboratory for analysis of a wide range of pollutants. These analyses take some time and results are typically not available for three or more weeks post sampling.

Data analysis to identify "hot spots" utilizes a tolerance interval-based threshold based on the 95^{th} percentile of randomly selected urban storm drain sites; concentrations above the tolerance interval are considered excessive and further site-specific investigations are required to identify, and remove or remediate the ID/IC (Santa Ana Regional Water Quality Control Board 2009, OC Stormwater Program 2003).

Actions managers may take once monitoring is conducted includes:

- Triggers source tracking for identification, confirmation, remediation, enforcement

It is important to note, monitoring is not the only activity used for detecting IC/ID. Regardless of monitoring effort, other activities can and should be used for detecting and eliminating ID/IC such as public complaints, hotlines, and inspection programs, amongst others.

The web app content is separated into three main analyses:

- Inventory: Map-based and tabular summaries of monitoring effort and basic characteristics of the data
- Hot spots: Assessment of hot spot sites based on frequency of threshold exceedances over time
- Analysis by waterbody: A simple analysis of threshold exceedances for select sites shown for the complete time series

Inventory

An inventory of historical monitoring identified 106 dry weather monitoring sites for the 10-year period 2010 – 2019 across four Orange County watersheds: Newport Bay, Anaheim Bay/Huntington Harbour, San Gabriel River/Coyote Creek, and Santa Ana River (Figure 11). Unlike mass emission sites that occur in receiving waters, dry weather screening occurs in MS4 outfalls. There are multiple outfall sites monitored per watershed, ranging from 21 in the Anaheim Bay/Huntington Harbour watershed to 32 in the San Gabriel River/Coyote Creek watershed. The majority of sites are targeted sites that are selected based on their history of ID/IC detection, consistency of flow (i.e., outfalls that infrequently flow are unlikely to be selected), and feasibility of access. The remaining sites are chosen probabilistically (aka “random”), where the data is used to update tolerance interval calculations for different ID/IC pollutants.

Sampling occurs approximately monthly between May and September and samples are analyzed for a long list of pollutants. Seventy-one (71) pollutants have been measured during dry weather monitoring, but frequently most pollutants are not detected. So, the analysis for effectiveness and efficiency for dry weather monitoring was constrained to nutrients and the 10 most frequently detected pollutants including bacteria, some trace metals, and some pesticides (Tables 7 and 8). Nutrients and the most frequently detected pollutants cumulatively amount to 44,072 laboratory analyses.

Answer to the question

Historical monitoring results illustrated there are “hot spots” in every watershed that have frequent exceedances of pollutant tolerance intervals. For example, the tolerance interval for *Enterococcus* is 27,000 cfu/100 mL, and 10 out of the 106 sites in Orange County exceeded this threshold more than 25% of the time, or at least once per summer (Figure 12). Similarly, the tolerance interval for copper is 14 ug/L, and 8 out of the 106 sites in Orange County exceeded this threshold more than 25% of the time (Figure 13).

Updates and Upgrades

Dry weather monitoring for ID/IC is a core monitoring program element. Regardless of how many ID/IC were identified in the previous dry season, dry weather monitoring will occur next dry season, just to ensure that future ID/IC do not occur.

The Advisory Committee discussed several deficiencies in the current dry weather ID/IC monitoring design. First, the sampling design calls for pre-scheduled monthly sampling at fixed stations. This design is not optimal for finding ID/IC, which are often stochastic in time and space. Second, the current sampling design focuses on grab/short-term sampling. This design is not optimal for finding ID/IC, which are typically intermittent and short-lived. Third, the lag time for laboratory analysis can hinder upstream source tracking. Once positive results are returned after 3 weeks, the source has likely moved. Finally, the monitoring design is resource intensive relative to the number of ID/IC identified, enforced, and removed.

After evaluating how dry weather monitoring efficiency and effectiveness could be improved, the Advisory Committee reached the following four conclusions:

- The monitoring program should cut back from the dozens of sites periodically sampled and prioritize “hot spot” outfall sites or subwatershed that may have caused or contributed to the high pollutant levels at downstream receiving waters. These sites or subwatersheds should be targeted for ID/IC identification and elimination
- This new strategy is best accomplished by replacing the approximately once/month periodic dry weather sampling with continuous, near-real-time sensors

- Continuous data will help focus ID/IC source tracking efforts on the times when pollution events are occurring

- A web application was created providing a technical tool for prioritizing "hot spot" sites when defining the next monitoring program

The concept of using continuous real-time or near-real-time sensors for ID/IC is not novel (PG Environmental 2018), but it is a technological advancement from the existing dry weather monitoring. This study design will upgrade the dry weather monitoring from a rote, repetitive core monitoring element to a robust, data-driven monitoring element of the Permittees' program. In this case, placing continuous real-time sensors in hot spot locations will identify the times when ID/IC are occurring, fostering a greater likelihood of catching intermittent and sporadic pollution events. The Advisory Committee felt giving up more sites for more time (via continuous sensor deployments) at the highest priority sites would be most effective for tracking upstream sources. Once the specific time of the ID/IC occurs, the investigation can move upstream in a coordinated, strategic fashion. Once that source is identified and eliminated, the sensors can move onto another site. Ultimately, multiple sensor packages can be deployed throughout the four watersheds.

Use of continuous real-time sensors is far from standardized (PG Environmental 2018). In order to calibrate and validate this new approach as a viable alternative, the revised dry weather screening will need to undertake four steps:

1. Identify key monitoring locations

2. Deploy sensors at strategic locations

3. Establish correlations between measured parameters and permit parameters of concern

4. Develop data processing tools to identify ID/IC based on real-time sensor data

Step 1 is complete utilizing the web application built for this project. However, there is one important factor prior to step 2 that must be considered; the Advisory Committee recommended that hot spot sites should be prioritized for corresponding waterbodies with TMDLs. The goal of dry weather ID/IC monitoring is to focus on reducing pollutants that cause or contribute to impairments. Waterbodies subject to a TMDL are the highest priority locations that need pollutant reductions. Thus, the web app was modified to include a filter for TMDL waterbodies by parameter (Figure 14). In some cases, the sites stayed the same. In other cases, some sites were not prioritized. In the latter case, new sites might be raised to a higher priority

For step 2, there are a wide range of sensors to choose from, ranging from robust and commercially available packages that typically measure surrogates of water quality (i.e., flow, conductivity, pH, temperature, etc.) to research grade prototypes that measure specific pollutants of concern (i.e., metals, nutrients, bacteria, etc.). While the research grade sensors could be evaluated through a special study, the Advisory Committee recommended commercially available sensors for routine use. A list of selection criteria for considering commercially available sensors included:

- Number of concurrent parameters measured

- Where will it be installed (i.e., dry vs wet weather, risk to vandalism, access as a function of ease and frequency)
- Anti-fouling features
- Size
- Calibration requirements (a function of effort and frequency)
- Accuracy and precision
- Sensitivity at low vs. high concentration
- Non-optional "accessories" (i.e., proprietary software, data loggers, communications)
- Up-front cost vs. long-term maintenance
- Power options, requirements, and longevity (battery, solar, direct power)
- Telemetry vs manual download
- Availability of technical support

For step 3, there will be a need to establish correlations between measured parameters and permit parameters of concern (Han et al. 2008; Aumond and Joannis 2006; Muniz et al. 2012). Commercial sensors rarely measure the pollutants of interest, but rather detect surrogates of the water quality pollutants to indicate their presence. The relationship between the sensor and the pollutant(s) of interest needs to be established before managers might be willing to take action. So, some effort for synoptic data measurement of sensor and pollutant(s) will be necessary. For example, Bae et al. (2009) utilized flow, pH, turbidity, and temperature to predict fecal and total coliform bacteria concentrations in the Aliso Creek watershed (Figure 15).

For step 4, data processing tools will need to be created to manage the influx of a significant amount of real-time sensor data, rapidly calculate the relationships with water quality parameters of interest, and then trigger follow-up actions. The true utility of the continuous near-real-time sensor technology is harnessed in this step. Automation will become fundamental, especially if rapid follow-up (i.e., same day or same hour) is necessary to identify ID/IC. This technology, while not standardized, is readily available and could be developed for dry weather ID/IC monitoring using open-source programming, cloud-based technologies, and linked to other data systems within the County or city permittees. While the programming for processing data is site specific, the programming structure will be preserved making the deployment at future sites less challenging than initial deployments.

Considering that real-time sensor applications are not standardized, the Advisory Committee recommended that this core monitoring element initiate with a special study to pilot test the new dry weather monitoring program design. The special study to pilot-test the sensor packages should address the four steps described above, likely requiring 1-2 years.

Table 7. Inventory of dry weather samples by watershed and year for the 10 most commonly detected pollutants, plus nutrients.

Watershed	Ag	As	Cd	Chlorpyrifos	Cr	Cu	Diazinon	Dimethoate	ENT	FC	Fensulfothion	Hg	Malathion	Ni	Pb	Se	TC	TOC	TSS	Zn	Total
Anaheim Bay - Huntington Harbour	325	325	325	341	325	325	341	341	348	348	341	324	341	325	325	325	348	346	342	325	6686
Newport Bay	740	740	740	774	740	740	774	772	780	781	772	740	772	740	740	740	780	787	775	740	15167
San Gabriel River - Coyote Creek	474	474	474	485	474	474	485	485	500	500	485	474	485	474	474	474	500	495	491	474	9651
Santa Ana River	433	433	433	446	433	433	446	446	460	461	446	433	446	433	433	433	460	461	452	433	8854
Total	1972	1972	1972	2046	1972	1972	2046	2044	2088	2090	2044	1971	2044	1972	1972	1972	2088	2089	2060	1972	40358

Table 8. Inventory of dry weather sample size by watershed for nutrients.

Watershed	Ammonia	Nitrate, Nitrite	Total Kjeldahl Nitrogen	Total Nitrogen	Orthophosphate	Total Phosphorus	Total
Anaheim Bay - Huntington Harbour	306	4	4	303	0	4	621
Newport Bay	709	13	13	696	0	12	1443
San Gabriel River - Coyote Creek	402	3	3	399	0	3	810
Santa Ana River	408	10	10	401	1	10	840
Total	1825	30	30	1799	1	29	3714

Figure 11. Map of dry weather monitoring sites, number of parameters measured at each site, and number of years dry weather monitoring had occurred at each site between January 1, 2010 and December 31, 2019.

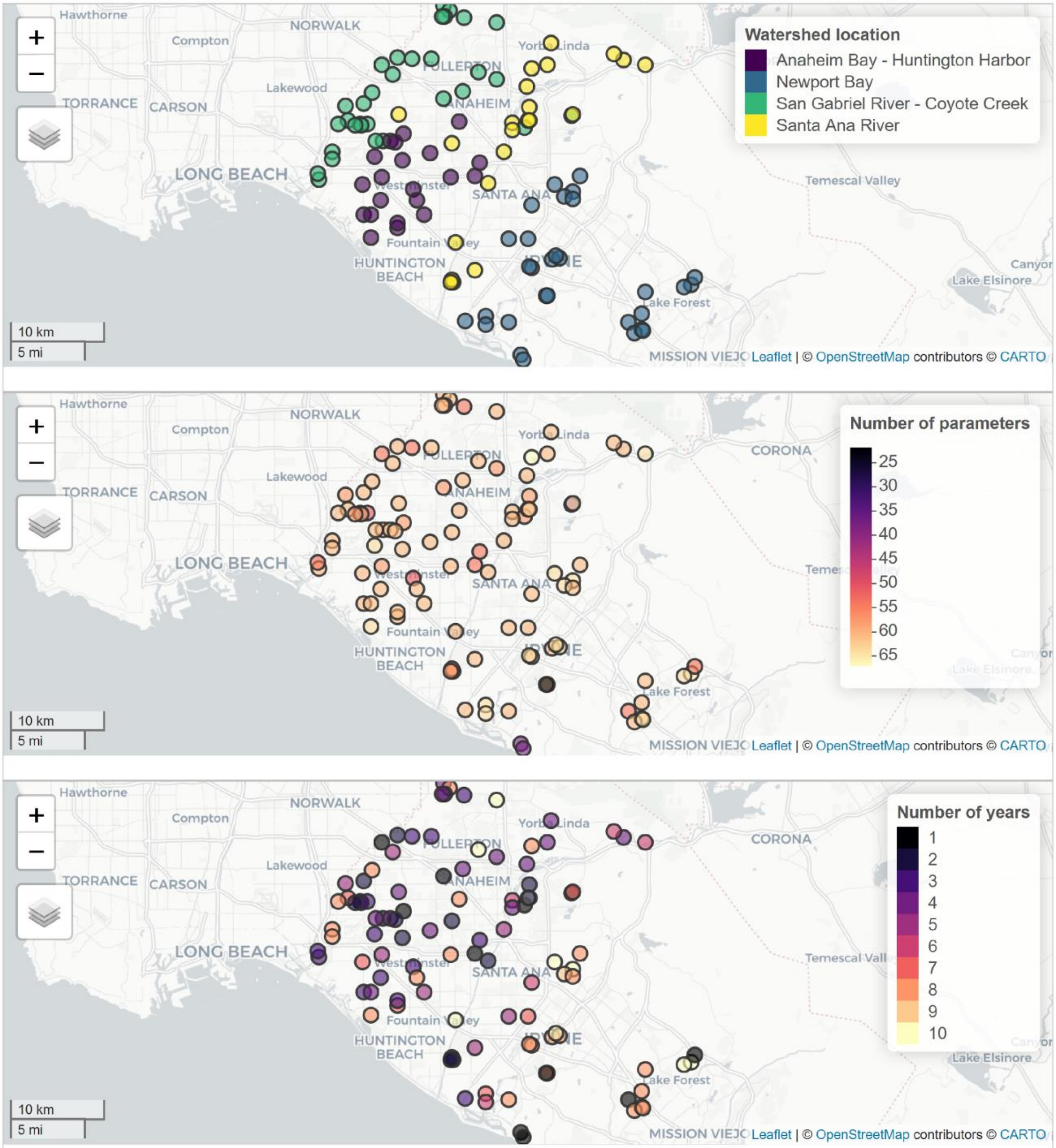


Figure 12. Dry weather hotspot analysis for Enterococcus. Tolerance Interval = 27,000 cfu/100 mL. See web app for station codes in lower figure.

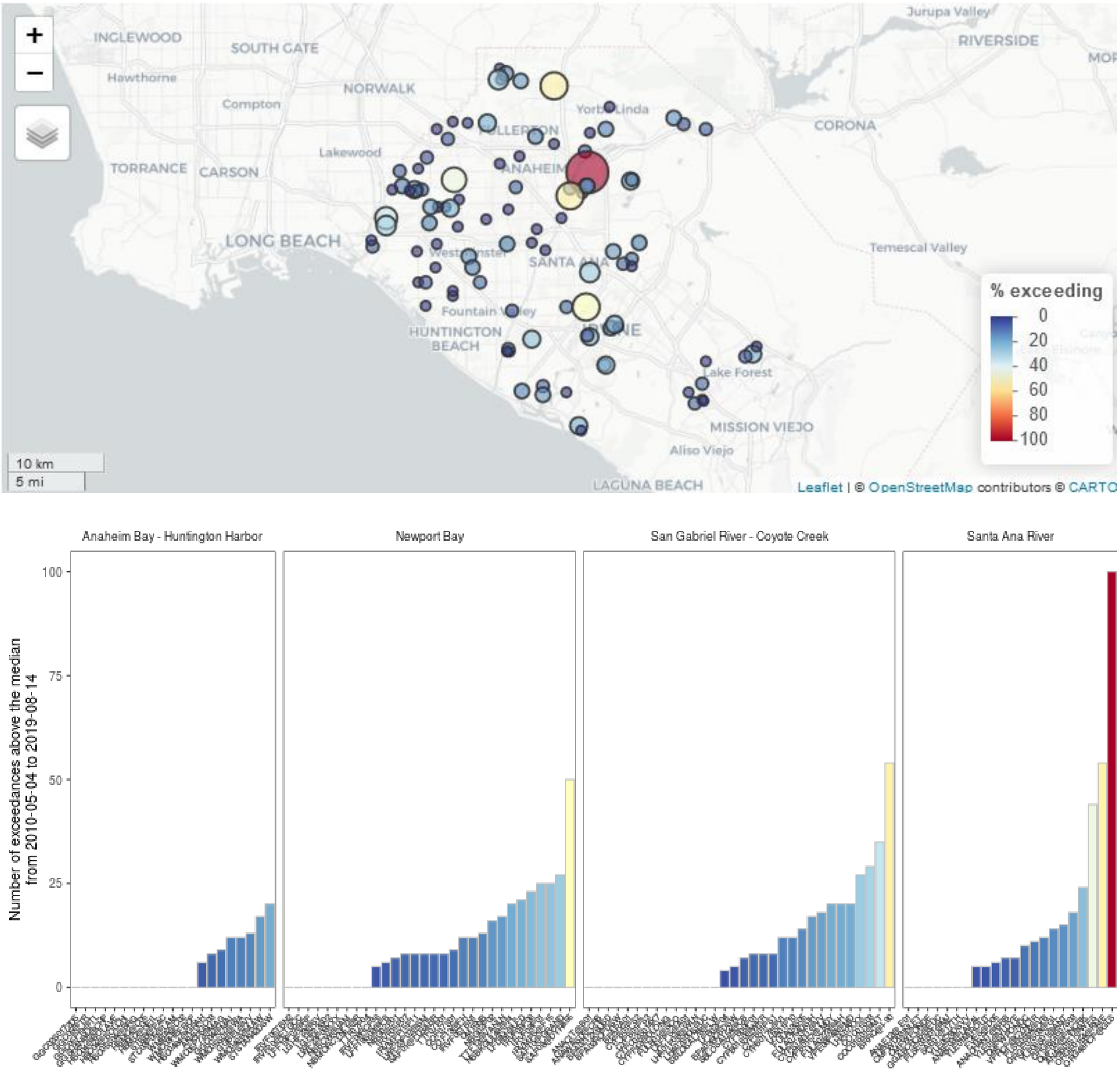


Figure 13. Dry weather hotspot analysis for copper. Tolerance Interval = 14 ug/L. See web app for station codes in lower figure.

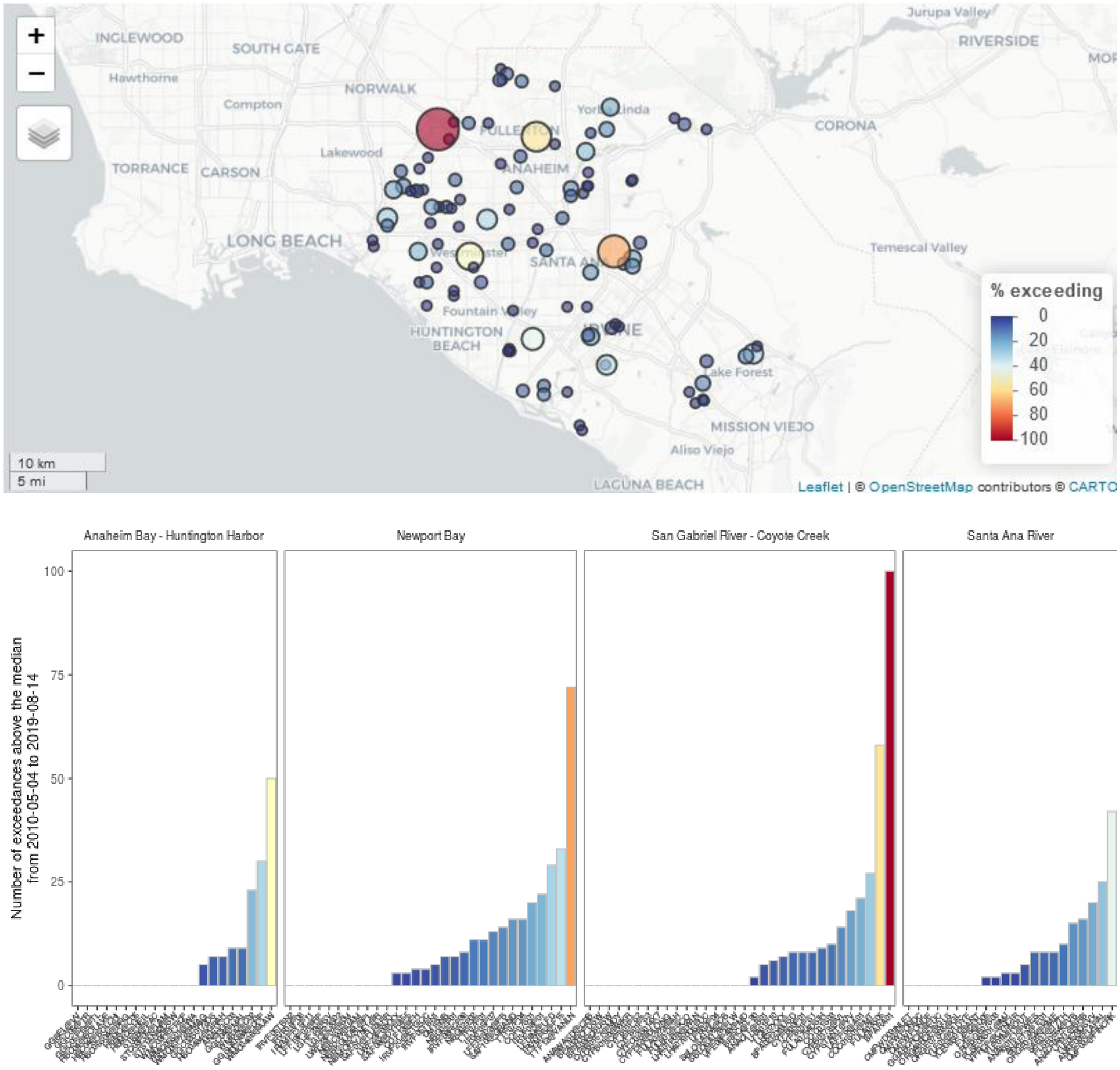


Figure 14. Dry weather MS4 outfall monitoring hot spot prioritization for copper based on waterbodies with TMDLs. See Figure 13 for the relative site ranking of MS4 outfalls independent of TMDL waterbodies. Tolerance interval = 14 ug/L. See web app for station codes in lower figure.

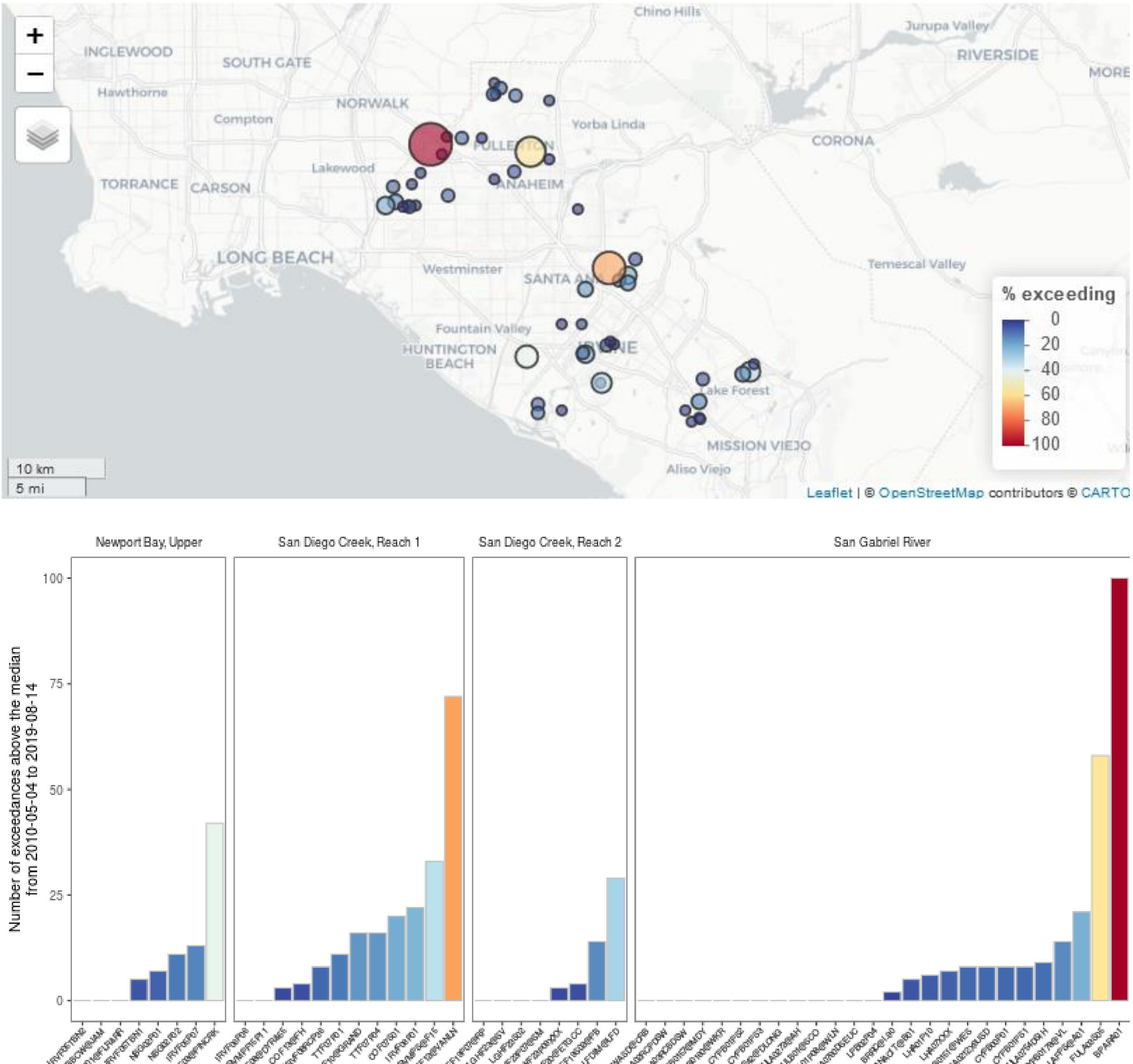
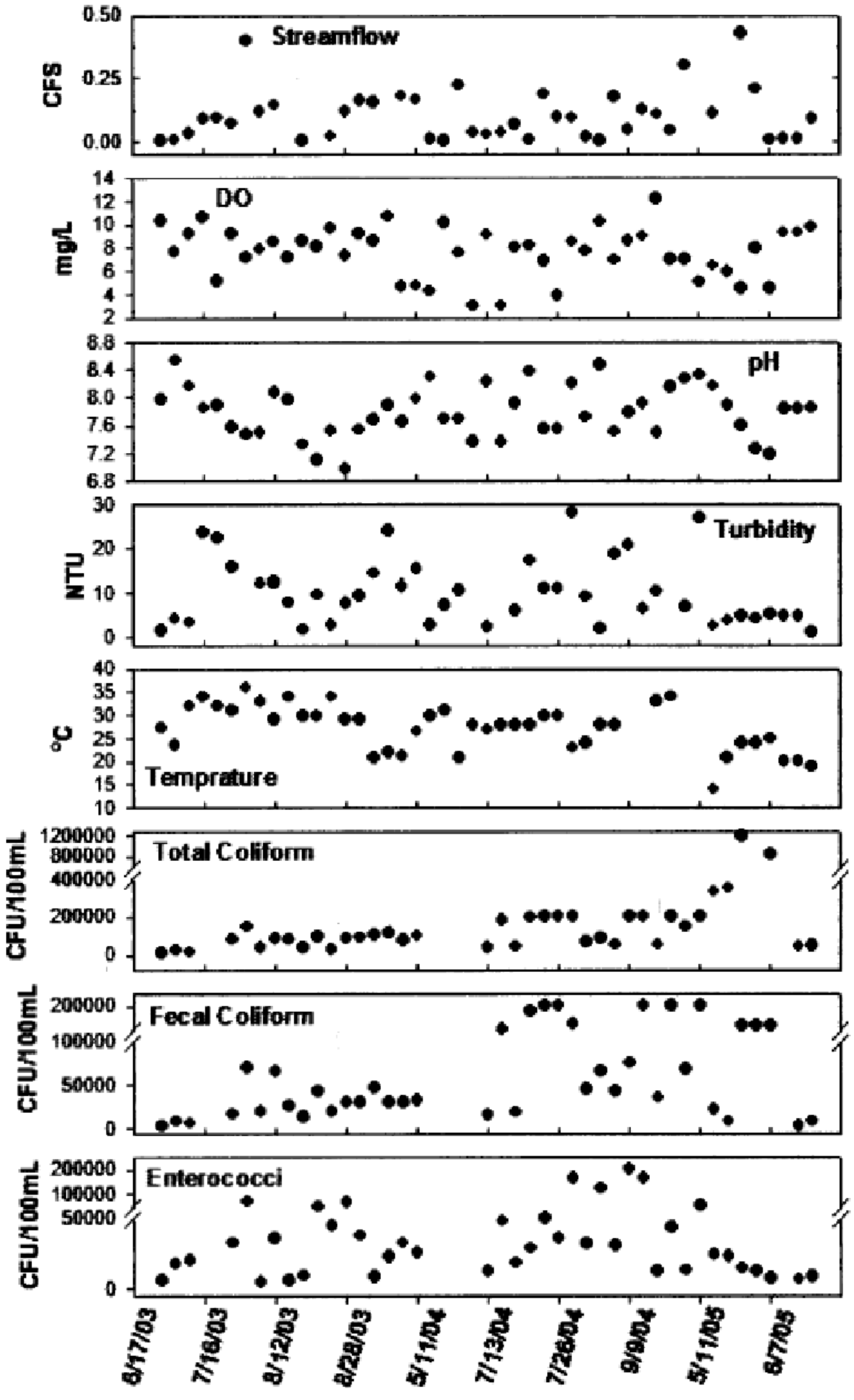


Figure 15. Example of sensor and pollutant synoptic measurements to derive relationships between sensors and pollutants of concern. From Bae et al. 2009.



Q4: What are the temporal trends in sediment quality and water column concentrations in estuaries and wetlands?

A web application was created for this question with filters and toggles based on water column or sediment, and wet weather or dry weather:

- Estuaries and wetlands sediment and water column concentration data: https://sccwrp.shinyapps.io/ocms4review/harbors_estuaries.Rmd

The premise of this question is that receiving water sites located across each of three North Orange County estuaries and wetlands – Newport Bay, Santa Ana River (Talbert Marsh), and Anaheim Bay/Huntington Harbour – are monitored for sediment quality and water column concentrations. Sediment is monitored at targeted sites between two and four times per year depending on waterbody, and then probabilistic sites are sampled once every five years. Water column concentrations are monitored between two and four times per year in dry weather depending on waterbody, then again during or immediately following three different wet weather events. Up to 94 parameters (the number would be hundreds more if individual congeners of compound groups such as PCBs, DDTs, toxaphene, etc. were counted) are measured in each sample. Concentrations are compared across years separately in wet and dry weather, and also compared to TMDL targets, if applicable. Actions managers may take once monitoring question number four are answered include:

- Trigger enhanced monitoring as concentrations or loads approach benchmarks (i.e., water quality standards, TMDL targets, etc.), to increase confidence that data are above or below the benchmark
- Trigger advanced data analysis for complicating factors that could influence a benchmark exceedance (i.e., precipitation patterns for wet weather)
- Identify trends in concentrations and loads from receiving water inputs, and factors leading or contributing to those trends
- Determine whether management actions have yielded desired outcome, or design/update future management actions to achieve the desired outcome
- Trigger source identification in order to continually decrease concentrations.

The web app content is separated into four main analyses:

- Inventory: Map-based and tabular summaries of monitoring effort and basic characteristics of the data
- Trends and power analyses: Changes over time by selected constituents and locations, including power analyses to help identify optimal sampling effort
- Station differences: A comparison of time series between stations to identify similarities among trends
- Overall trends: Map-based and tabular summary of trend tests for all stations shown together

Each analysis includes sub-tabs or drop-down menus for selecting and viewing different content. Most analyses are also grouped by constituent type such as nutrients, metals, or organics:

- Nutrients: Ammonia, nitrate, nitrite, Total Kjeldahl Nitrogen, orthophosphate, total phosphorus
- Metals/trace elements: Ag, As, Cd, Cr, Cu, Fe, Hg, Ni, Pb, Se, Zn
- Organics: 2,4'-D, 2,4-DB, 2,4,5 TP-Silvex, 2,4,5,-T, Aldrin, Allethrin, Alpha-BHC, Atrazine, Azinphos methyl (Guthion), Be, Beta-BHC, Bifenthrin, Bolstar, Chlordane, Chlorpyrifos, Cis-Permethrin, Coumaphos, Cyfluthrin, Cypermethrin, Dalapon, DDT, Delta-BHC, Deltamethrin, Demeton-o, Demeton-s, Diazinon, Dicamba, Dichlorprop, Dichlorvos, Dieldrin, Dimethoate, Dinoseb, Disulfoton, Endosulfan I, Endosulfan II, Endosulfan Sulfate, Endrin, Endrin Aldehyde, Endrin Ketone, Ethoprop, Ethyl Parathion, Fensulfothion, Fenthion, Gamma-BHC, GLYP, Heptachlor Epoxide, HPAH, L-Cyhalothrin, LPAH, Malathion, MCPA, MCPP, Merphos, Mevinphos, Mirex, OxyChlordane, Parathion-methyl, PCB, Permethrin, Perthane, pH, Phorate, Prallethrin, Prometon, Prometryn, Ronnel, Sb, Simazine, Tetrachlorovinphos, Tl, TOC-S, Tokuthion, Total chlordane, Toxaphene, Trans-Permethrin, Trichloronate

Because concentrations and variance differ in dry or wet weather, each analysis can be performed for either condition.

Inventory

Thirteen targeted stations were routinely monitored in the three watersheds during the 10-year span from January 2010 to December 2019 (Figure 16). An additional seven sites were monitored once as part of the probabilistic design. Over the decade time span, there were up to 332 sediment samples analyzed for nutrients, 410 samples analyzed for organics, and 357 samples analyzed for metals. Over the decade time span, there were up to 742 sediment samples analyzed for nutrients, 718 samples analyzed for organics, and 1,153 samples analyzed for metals. The actual number of samples varied by site, matrix, and parameter.

Answer to the question

Sediment

In general, organic contaminants were largely not detected. In contrast, most metals were routinely quantified, but there were almost no trends detected at statistically significant levels. As an example, Figure 17 illustrates a common pattern at two sites in Newport Bay, one in upper Newport Bay (UNBJAM) and one in lower Newport Bay (LNBTUB). After an initial drop in concentration at the beginning of the monitoring period, concentrations did not change with time. The Advisory Committee commented that much of the trend occurred prior to 2010, the beginning of the monitoring period.

Water Column

Like sediment, organic contaminants were rarely detected in water column samples. However, some significantly decreasing trends were observed in other contaminants from Upper Newport Bay, including concentration decreases of nitrate+nitrite and some metals (Cu, Hg, Zn). As an example, Figure 18 illustrates significant concentration decreases in copper from Upper Newport Bay (UNBJAM), with decreasing but non-significant decreases in copper concentrations observed from Lower Newport Bay (LNBHIR) and Huntington Harbour (HUNCRB).

Wet weather mirrored the trends observed in dry weather. Organic contaminants were rarely detected in water column samples during wet weather. However, some significantly decreasing trends ($p < 0.05$) were observed in other contaminants from Upper Newport Bay, including concentration decreases of some metals (Cu, Hg, Se, Zn). As an example, Figure 19 illustrates significant concentration decreases in copper from Upper Newport Bay (UNBJAM), with decreasing but non-significant decreases in copper concentrations observed from Lower Newport Bay (LNBHIR) and Huntington Harbour (HUNCRB). Similar to temporal trends in sediment concentrations, the Advisory Committee commented that much of the decreasing trend in nutrient concentrations occurred prior to 2010.

Updates and Upgrades

Trend monitoring of estuaries and wetlands is fundamentally a core component of the North Orange County Monitoring Program. Regardless of an increasing, decreasing, or static trend, the information is critical because there are management actions associated with each answer. In this case, continual sampling of these receiving waters is necessary to ensure impacts to beneficial uses are not occurring from MS4 discharges, or that existing impacts are improving.

Given that this trend question is a necessary and ongoing element of the North Orange County Monitoring Program, the next step is to ensure efficiency and effectiveness of the monitoring, to confidently address the needs of both regulators and regulated parties.

For most constituents at most sites, less sampling than is currently being expended can achieve similar levels of statistical confidence in temporal trend detection. This applies to both sediment and water column sampling. An example using trace metals is shown in Figure 20. In this case, regardless of station, trends in concentration can be optimally detected at less than current frequencies. The median reduction was approximately 50% across all stations and all metals for dry weather. Similar sampling frequency reductions for trace metals were observed for wet weather. For organic contaminants, the reduction might even be greater due to the lack of quantifiable concentrations. The actual confidence in sampling frequency may be altered if a parameter-site combination is approaching a critical threshold of concern. However, it should be recognized that just as in questions 1 and 2, there are also sampling logistics considerations. The web app can calculate these results for every parameter at every site as regulated and regulatory agencies decide if changes in sampling frequency are appropriate.

The Advisory Committee recommended evolving the sediment trends question. Currently, the monitoring question focuses on temporal trends in sediment concentrations at a small number of sites per waterbody. Instead, the monitoring question should refocus on spatial trends in sediment quality. Hence, the monitoring design should pivot from many sampling events at a small number of sites, to fewer samples at a larger number of sites. Ultimately, this refocused monitoring design is trying to achieve more information about space than time.

Since there were almost no temporal trends in sediment concentrations, managers give up virtually no information by refocusing monitoring on spatial trends. Temporal trends can still be gleaned at a much-reduced frequency. Gaining the spatial information will help substantially, however, since the SWRCB's new 303(d) listing policy for sediment quality is based on percent area rather than number of samples. The 303(d) listing policy states impacted sediment quality that exceeds 15% of the area in an enclosed bay or estuary will result in an impaired waterbody listing. Thus, both regulated and regulatory agencies want greater precision in percent area estimates. The new monitoring question could be restated as, "What is the trend in percent area that exceeds sediment quality objective thresholds and is it related to MS4 discharges?"

One way to efficiently achieve an answer to a spatial extent question about sediment quality is to leverage with the Southern California Bight Regional Marine Monitoring Program (Bight Program). This long-standing regional monitoring program collects about 400 samples every five years to assess sediment quality, including every estuary and wetland from San Diego to Santa Barbara. Utilizing standardized protocols and robust quality assurance, over 100 organizations (including regulators and regulated agencies) participate to answer the same question proposed by the Advisory Committee. A good example of the Bight Program output can be found in Figure 22.

If the percent of area that exceeds sediment quality objectives exceeds allowable spatial extent, then subsequent actions should be taken to ensure the exceedances are attributable to MS4 discharges.

Integrating and leveraging the North Orange County Monitoring Program sediment monitoring with the Bight Program will be relatively seamless since there has already been coordination on probabilistic sediment quality sites in the last two Bight surveys. The primary action will be to increase the density (number of sites) in the waterbodies of interest. Since the monitoring methods are comparable among participating agencies, the spatial extent in Orange County estuaries and wetlands can be compared across time, as well as across estuaries outside of Orange County.

Figure 16. Map of wetland and estuary monitoring sites in Newport Bay, Talbert Marsh, and Anaheim Bay/Huntington Harbour between January 2010 to December 2019.

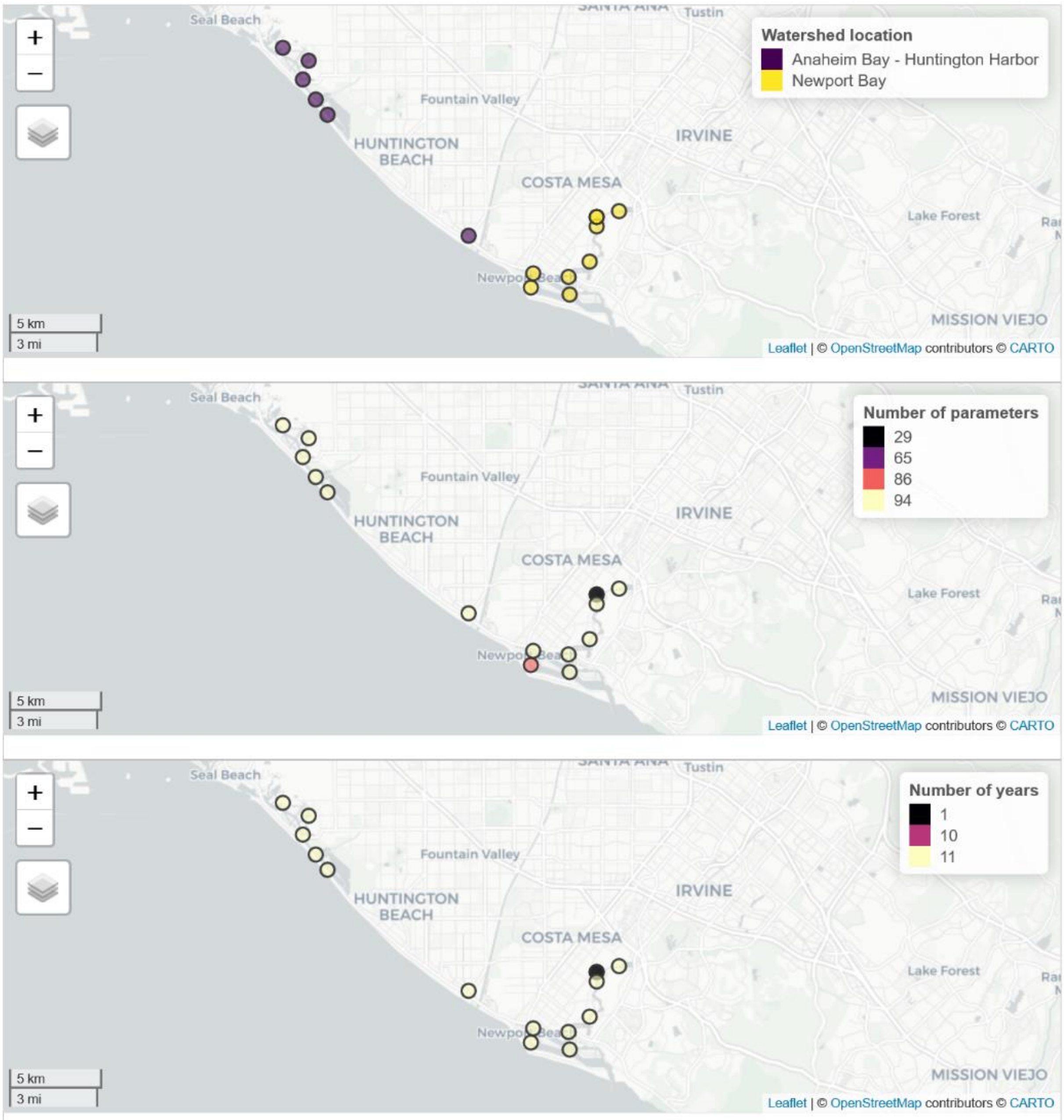
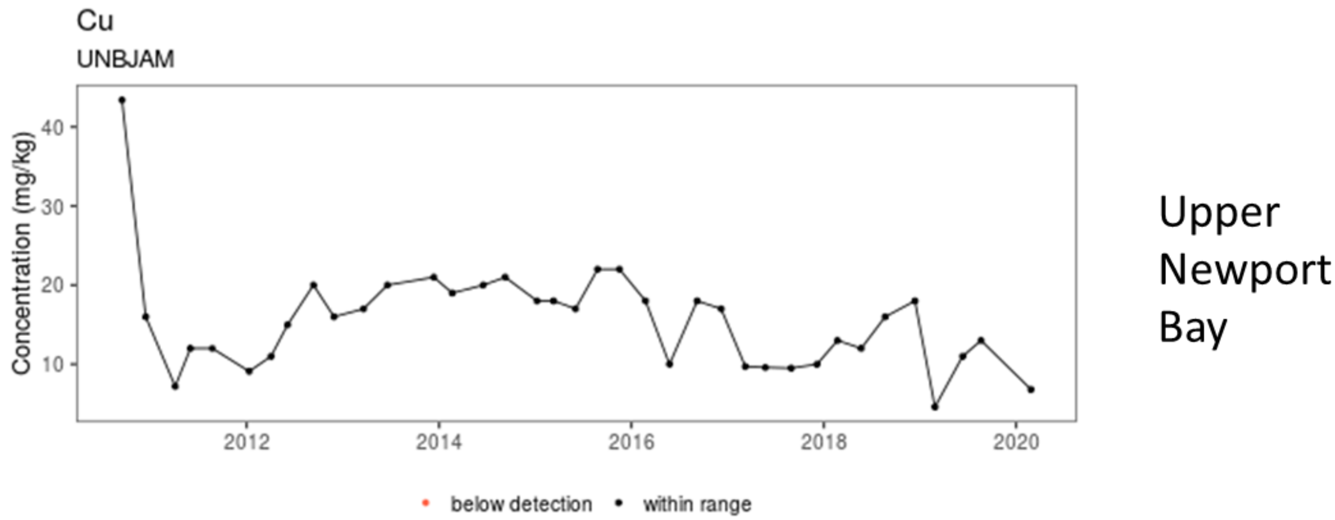


Figure 17. Sediment sampling for copper (Cu) in Upper and Lower Newport Bay during the monitoring period 2010-2020.



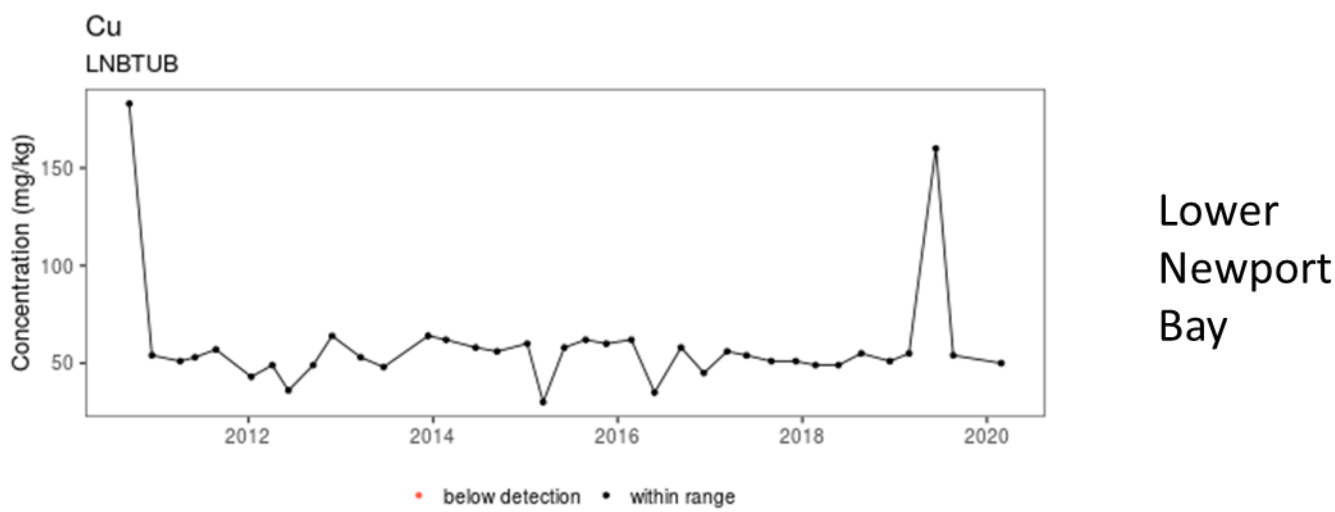


Figure 18. Dry weather water column monitoring in Upper Newport Bay, Lower Newport Bay, and Huntington Harbour during the monitoring period 2010-2020.

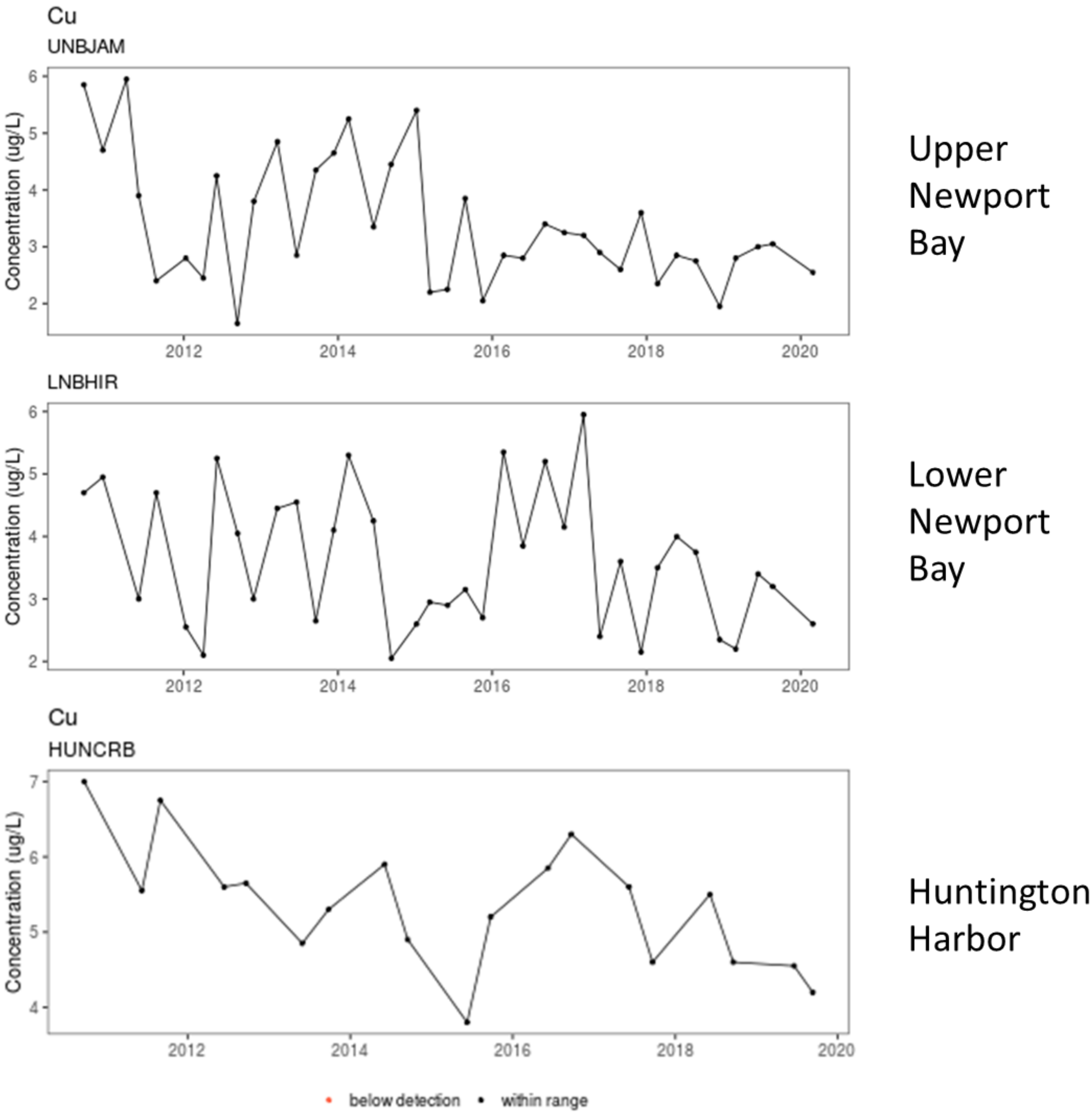


Figure 19. Wet weather water column monitoring in Upper Newport Bay, Lower Newport Bay, and Huntington Harbour during the monitoring period 2010-2020.

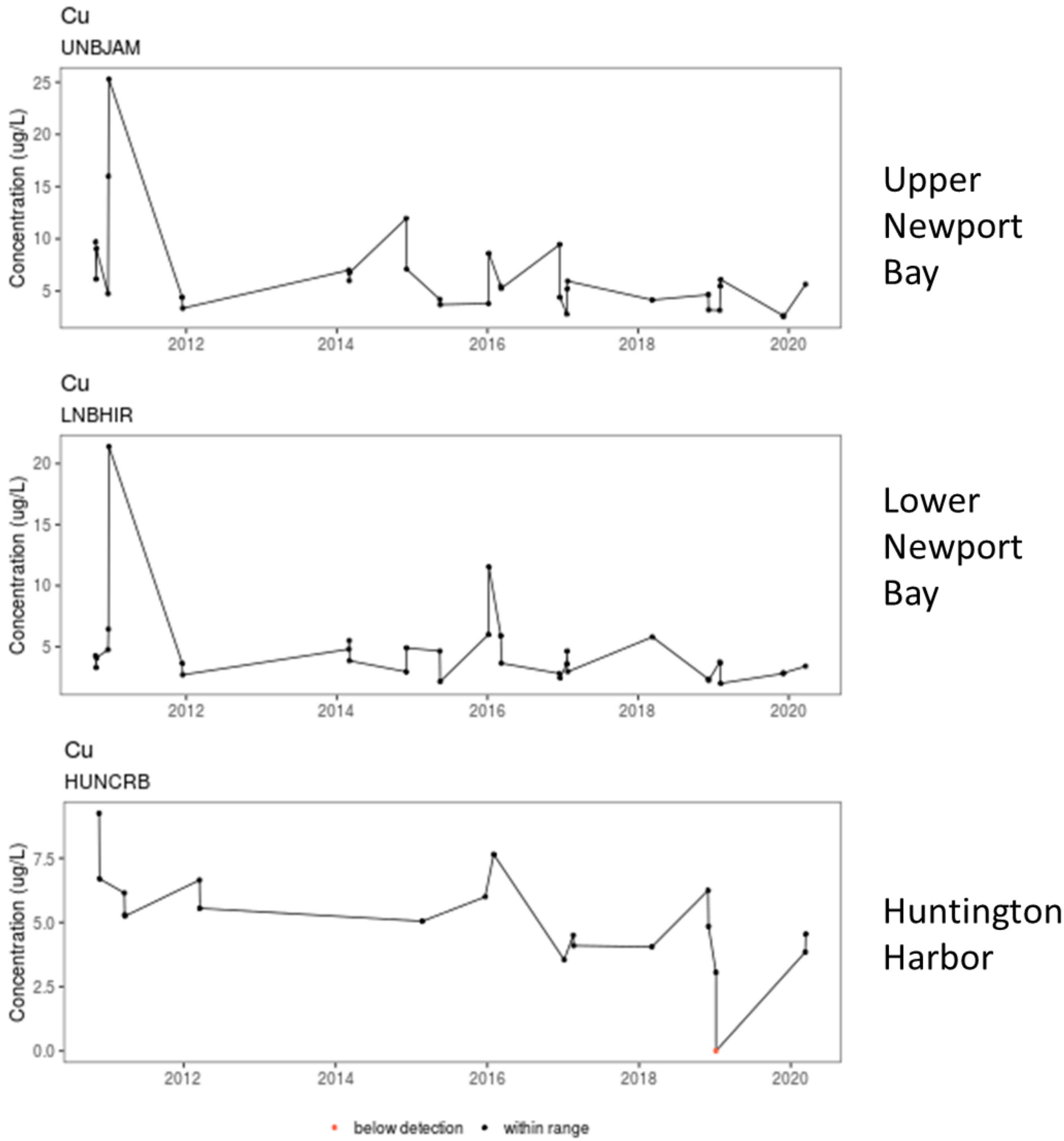


Figure 20. Example power curves illustrating the optimal sampling frequency for all trace metals across all water column sites during wet weather (left) and dry weather (right). Each curve is a different station-parameter combination. Y-axis is the magnitude of change (smaller detectable changes means more samples). X-axis is the sampling frequency relative to existing frequency (< 1.0 means less than current sampling). Gray symbols represent the optimal sampling frequency where smaller detectable differences requires proportionately more samples. Box plots above the chart are the distribution of optimal sampling frequencies. All trends assume beta=0.8, alpha=0.05.

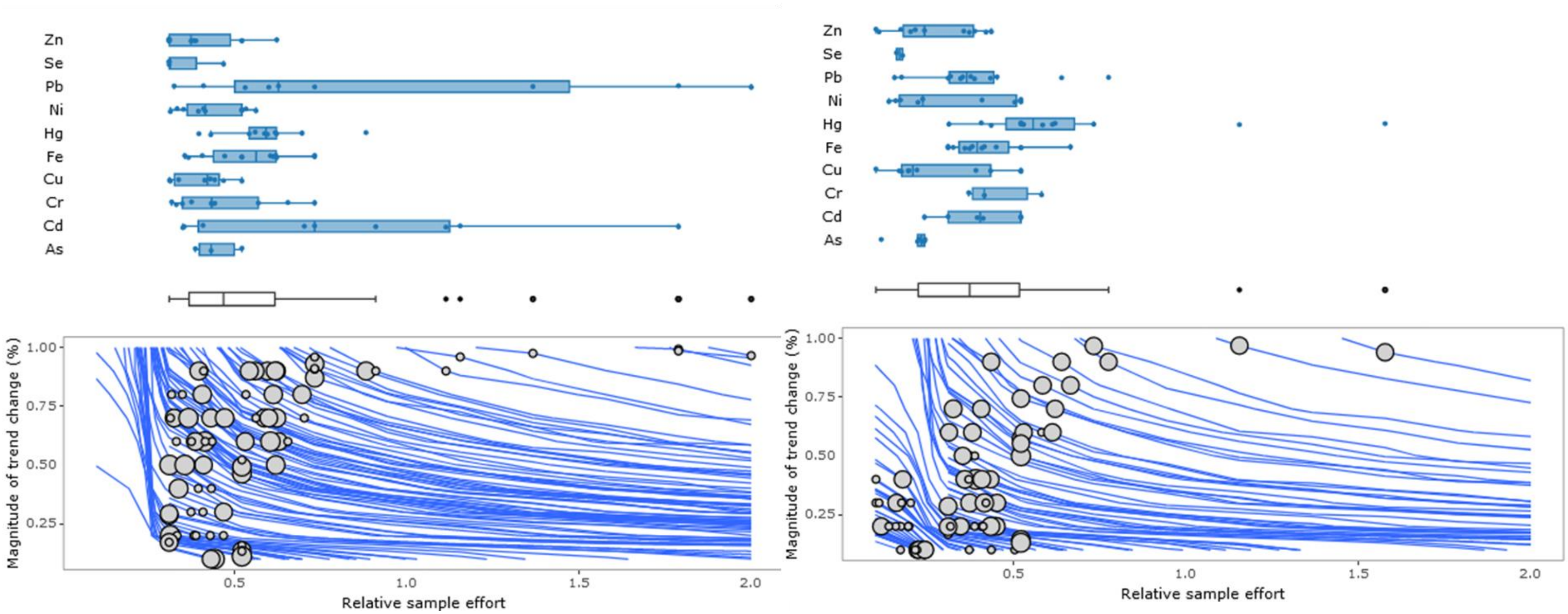


Figure 21. Example power curves illustrating the optimal sampling frequency for all trace metals across all water column stations during wet weather (left) vs sediment stations (right). Each curve is a different station-parameter combination. Y-axis is the magnitude of change (smaller detectable changes means more samples). X-axis is the sampling frequency relative to existing frequency (< 1.0 means less than current sampling). Gray symbols represent the optimal sampling frequency where smaller detectable differences requires proportionately more samples. Box plots above the chart are the distribution of optimal sampling frequencies. All trends assume beta=0.8, alpha=0.05.

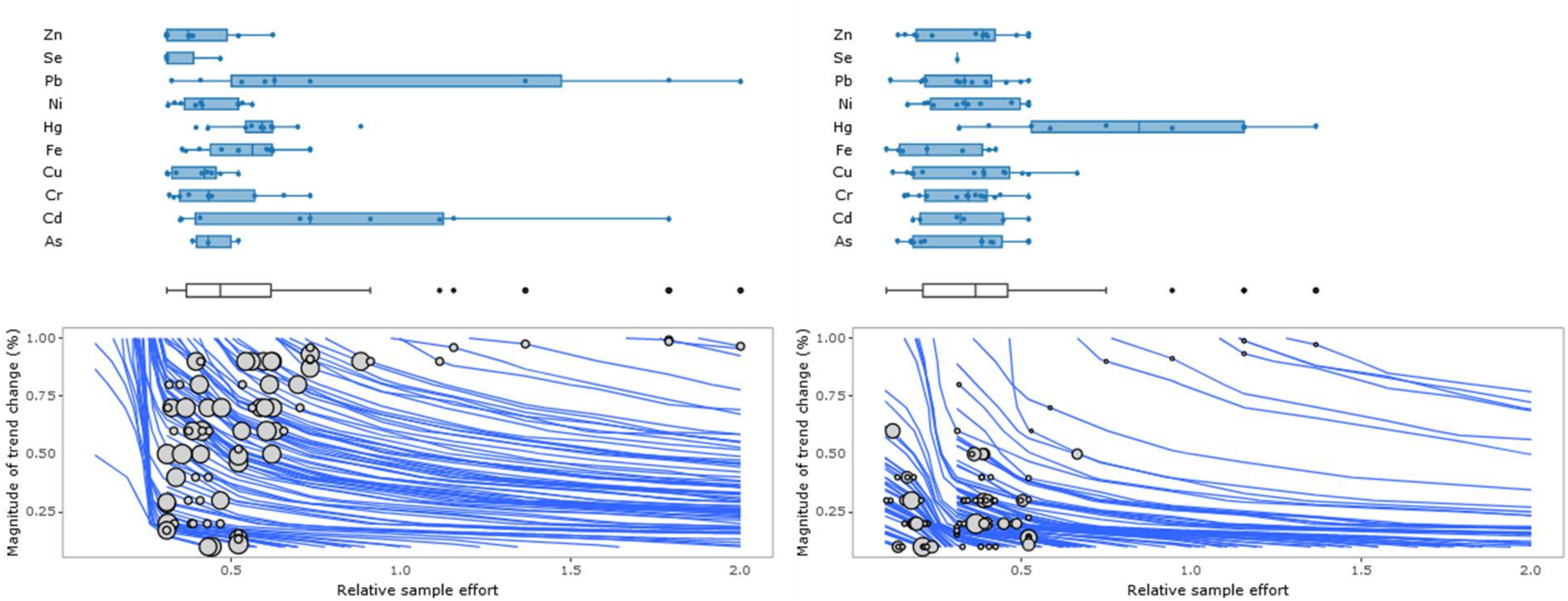
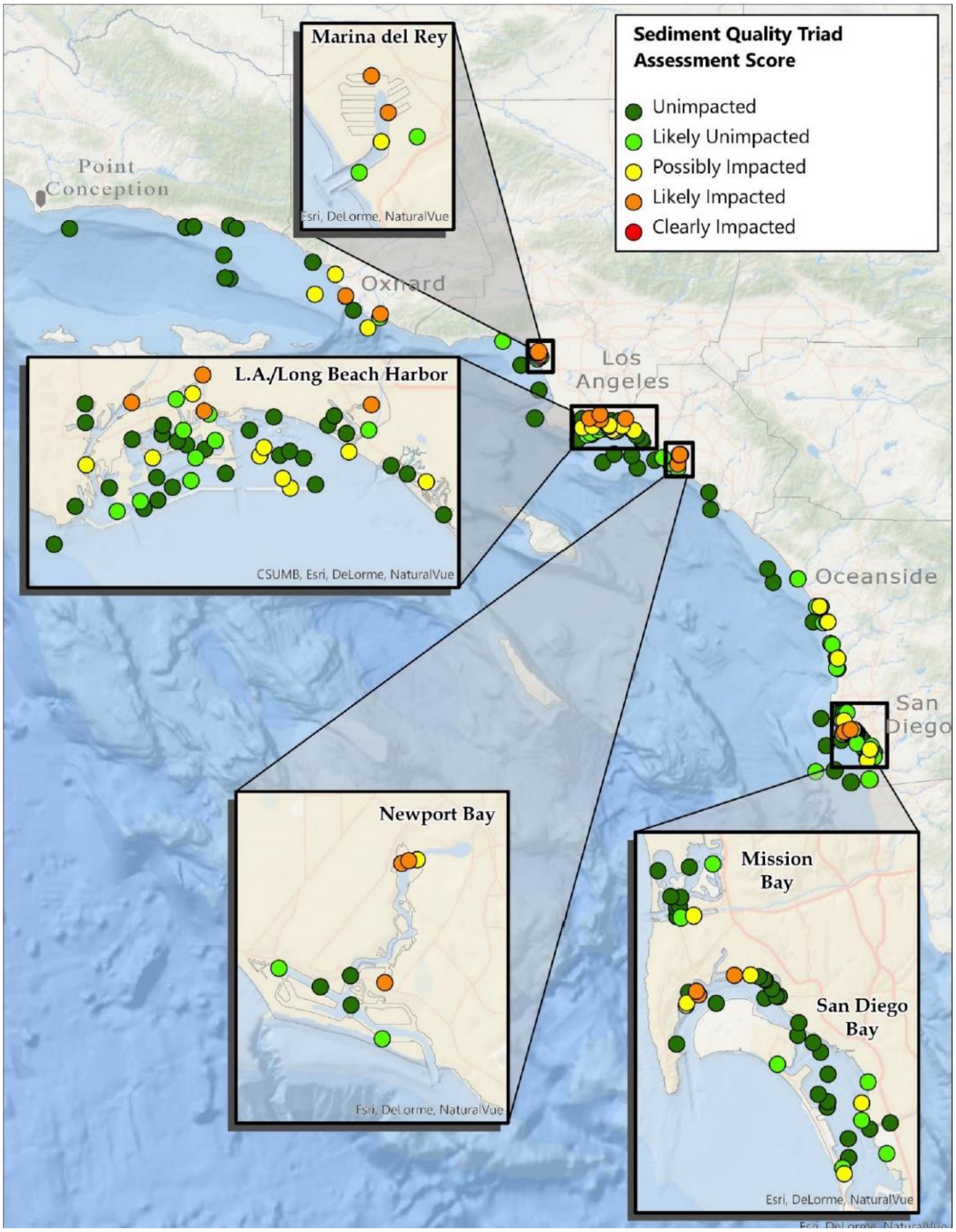


Figure 22. Example of sediment quality monitoring from the Southern California Bight Regional Marine Monitoring Program.



UPDATED REPORTING

Once the study designs were evaluated, and recommendations for study designs were provided, the North Orange County Permittees wanted a review of reporting. Reporting is closely linked to, but independent of, monitoring study design. Instead, reporting is important for communicating the answers to monitoring questions. Thus, this aspect of the monitoring program evaluation is targeted at assessing effectiveness and efficiency of the monitoring program Annual Reporting.

Inventory

The North Orange County permittees currently submits two types of reports to the Regional Board. The first is a multi-year assessment as part of their Report of Waste Discharge (ROWD), which is due prior to the NPDES permit expiration. The second is an Annual Report, which incorporates monitoring as one aspect of their overall multi-faceted program and includes much more than environmental monitoring information. Last year's Annual Report (2019-20) included a single chapter summarizing their monitoring activities, then dedicated an appendix for each of the elements of the North Orange County Monitoring Program. The Annual Report is the focus of the efficiency and effectiveness evaluation in this section.

The summary section from the Annual Report utilized management friendly indices of condition that translate detailed indicator measurements into easily understood, non-scientific descriptors of condition. Examples included:

- The Canadian Council of Ministers of the Environment (CCME) for combining the magnitude and frequency of chemical contaminants that exceed water quality thresholds.
- The California Stream Condition Index (CSCI) for assessing the health of stream microbenthic invertebrate communities.
- Heal the Bay's Beach Report Card grades for frequency and magnitude of fecal indicator bacteria that exceed water quality thresholds.

In contrast, the appendices are dense technical documents. They include a short background, then focus on methods, quality assurance, and detailed results. The dense appendices are a regulatory necessity to transmit all of the monitoring data to the RWQCB as part of the NPDES permit.

Updates and Upgrades

The primary challenge for the monitoring section of the Annual Report is its sole focus on compliance. Based on the sections in this monitoring evaluation report, the answers to monitoring questions were not solely focused compliance, but were related to management decisions to achieve compliance. If the Annual Report is refocused around the monitoring questions, and the resulting management actions, this will enhance the Annual Report because there will be new and expanded purpose to the data interpretation.

To achieve this focus on management questions, a new Annual Report format is recommended. The summary document should continue, with upgrades to the key graphics. There should be one key graphic associated with each monitoring question, and the text should focus on what the key graphic means rather than on the methods used to produce it (as it is now).

In addition to the Annual Report summary, the North Orange County permittees should take an annual deep dive on one or two monitoring questions. The deeper dive can provide some detail about the background, methods and quality assurance, lifting some abbreviated material from the 2018-19 document appendices. The deeper dive has the added benefit of highlighting any special studies to be planned and implemented to answer site-specific or emerging water quality issues. However, the deeper dive section should also include an extended discussion section about the meaning and context of the data. For example, the trends questions may discuss magnitude of trend, trend confounders, or when the trend will approach thresholds. The deeper dive should also have a section focused on decision making and adaptive management - what are the MS4 Permittees going to do now that they have an answer to the monitoring question? This adaptive management strategy is highlighted in every chapter of this monitoring evaluation report. The North Orange County Permittee Annual Report is the time and place to respond to the monitoring question.

Creating a deeper dive on every question every year is not necessary. In many cases, such as the fish tissue and bird egg bioaccumulation monitoring, answers to question will take many years. Therefore, a detailed section every year makes little sense. This approach of not critically evaluating every question every year will save effort, while simultaneously providing greater meaningfulness.

While not every monitoring question is recommended to be answered every year, every monitoring question should be addressed at least once every permit cycle. The ROWD is the opportunity to summarize the answers to all monitoring questions over the course of the entire NPDES permit cycle, which typically occurs every five years.

Finally, the Annual Report should avoid large data tables by making the monitoring data available online. Data submittal is a condition of compliance and an important element of public accessibility and transparency. Moving the data online will save tremendous effort by not producing hardcopy data, and it will also automate time-consuming data requests from the public after the Annual Report and ROWD are released. Data portals are becoming commonplace, and this technology is easily within reach. Finally, the new Annual Report and ROWD format can serve as the metadata documentation to accompany the online data. If the North Orange County Permittees want to increase data usability, new user interfaces allow for calculations “on-the-fly” of summary indices, which will help transmit the primary messages (i.e., key graphics) they want to communicate. The examples used in the web applications used for this report are a good example of what the North Orange County Permittees may wish to pursue.

CONCLUSIONS AND RECOMMENDATIONS

Based on the review of the MS4 Permittee's North Orange County Monitoring Program over the last 10 years, there are a number of over-arching themes from the Advisory Committee assessment of effectiveness and recommendations for increased efficiency.

- **The Advisory Committee prioritized four monitoring questions**

  After spending many hours creating lists of monitoring questions and how they would impact their management decisions, the Advisory Committee prioritized four questions.

  What is the temporal trend in pollutant concentrations and loading? [Mass emissions monitoring]

  What are the trends in fish and bird tissue concentrations? [Bioaccumulation]

  Are there illegal discharges/illicit connections (ID/IC)? [Dry weather monitoring]

  What are the temporal trends in sediment quality and water column quality in estuaries and wetlands? [Estuary/Wetland Monitoring]

- **Overall, the MS4 Permittee North Orange County Monitoring Program is answering the monitoring questions**

  Across the four monitoring questions, existing monitoring designs are effectively answering the questions. Many pollutants monitored at wet and dry weather mass emission sites are detecting trends, many of them downward trends, over the last 10 years. Pollutants are remaining steady in tissues of fish and birds in Upper Newport Bay. Dry weather IDDE are detecting and defining hot spots in the three North Orange County watersheds. This does not imply that the environment is healthy everywhere. There are clearly places with exceedances of water quality standards and TMDL targets. However, the monitoring program is answering questions about trends, which is a primary objective of this program review.

- **Increased monitoring efficiency can be achieved**

  In most cases, statistical power analysis determined that reduced sampling frequencies could achieve similar levels of confidence in detecting trends. In some cases, such as mass emission monitoring or water column monitoring in estuaries and wetlands, monitoring frequency could be reduced by up to 50% and still maintain an optimal statistical ability to detect temporal trends. However, statistical power is not the only factor to be considered. Other factors, such as sampling logistics can influence sampling frequency. Likewise, the Advisory Committee described situations where increased confidence is needed, such as when concentrations are near thresholds of concern and the management reaction is particularly expensive or time-sensitive.

- **Some monitoring questions or designs need to evolve**

  When a question is sufficiently answered, new questions can arise. Two examples occurred during this monitoring program review. The first is the dry weather monitoring program where the Advisory Committee agreed that the contaminant hotspots were generally known, but once per month sampling did not provide sufficient information for taking the next step towards source identification. Therefore, the Advisory Committee recommended monitoring the worst locations, but using new technology for continuous monitoring to address when and where the IDDE is occurring. This information should dramatically enhance the labor-intensive management action of source tracking to be successful. The second example is the estuary and wetland sediment monitoring. In this case, the pertinent monitoring question needed to evolve from temporal trends to spatial trends. The result is monitoring more sites, but at reduced frequency. Since temporal trends for sediment concentration were minimal, little information is lost by reducing frequency, but adding more sites enables enhanced information on spatial extent of contamination, which is consistent with the new management decision making criteria regarding 303(d) listing.

REFERENCES

Aumond, M. and C. Joannis. 2006. Mesure en continu de la turbidité sur un réseau séparatif eaux usées : mise en œuvre et premiers résultats. *La Houille Blanche*(4): 121-128.

Bae, H.K., B.H. Olsen, K.L. Hsu, S. Sorooshian. 2009. Identification and application of physical and chemical parameters to predict indicator bacterial concentration in a small Californian creek. *Water Environment Research* 81:633-640

Davis, J.A., J.R.M. Ross, S.N. Bezalel, J.A. Hunt, A.R. Melwani, R.M. Allen, G. Ichikawa, A. Bonnema, W.A. Heim, D. Crane, S. Swenson, C. Lamerdin, M. Stephenson, K. Schiff. 2012. Contaminants in fish from the California coast, 2009-2010: summary report on a two-year screening survey. Report to the State Water Resources Control Board. Sacramento, CA. 87 pp.

Díaz-Muñiz, C., P.J. García Nietro, J.R. Alonso Fernández, J. Martínez Torres, J. Taboada. 2012. Detection of outliers in water quality monitoring samples using functional data analysis in San Esteban estuary (Northern Spain). *Science of the Total Environment* 439: 54-61.

Han, Y. H., S.L. Lau, M. Kayhanian, M.K. Stenstrom. 2008. Correlation analysis among highway stormwater pollutants and characteristics. *Water Science and Technology* 53(2): 235-243.

Loganathan, B.G. and K. Kannan. 1994. Global organochlorine contamination trends: an overview. *Ambio*, pp. 187-191.

Lavoie, R.A., T.D. Jardine, M.M. Chumchal, K.A. Kidd, L.M. Campbell. 2013. Biomagnification of Mercury in Aquatic Food Webs: A Worldwide Meta-Analysis. *Environmental Science & Technology*. 47 (23): 13385-13394. DOI: 10.1021/es403103t

Freedman, P.L., A.D. Nemura, and D.W. Dilks. 2004. Viewing Total Maximum Daily Loads as a Process, Not a Singular Value: Adaptive Watershed Management, *Journal of Environmental Engineering* 130(6): 695-702

French, C., L. Wu, T. Meixner, D. Haver, J. Kabashima, and W.A. Jury. 2006. Modeling nitrogen transport in the Newport Bay/San Diego Creek watershed of southern California. *Agricultural Water Management*, 81(1-2), pp.199-215.

Model Monitoring Technical Committee. 2004. Model monitoring program for municipal separate storm sewer systems in southern California. Technical Report 419. Southern California Coastal Water Research Project. Westminster, CA.

National Research Council. 1990. *Managing Troubled Waters: The Role of Marine Environmental Monitoring*. Washington, DC: The National Academies Press. https://doi.org/10.17226/1439.

OC Stormwater Program. 2014. State of The Environment. 2013 Report of Waste Discharge, Santa Ana Region. Orange County Public Works, Orange CA. 38 pp.

OC Stormwater Program. 2003. Drainage Area Management Plan. Orange County Public Works, CA. https://ocerws.ocpublicworks.com/service-areas/oc-environmental-resources/oc-watersheds/documents/drainage-area-management-plan-7

Santa Ana Regional Water Quality Control Board. 2009. Santa Ana Region Water Quality Monitoring Program. NPDES Permit No. CAS618030, Order No. R2009-030

Schiff, K. 1995. Review of existing stormwater monitoring programs for estimating bight-wide mass emissions from urban runoff. In, Weisberg and Elmore (eds): Southern California Coastal Water Research Project Annual Report, pp.44-55.

Schiff, K.C., J.S. Brown, S.B. Weisberg. 2001. Model monitoring program for large ocean discharges in southern California. Technical Report 357. Southern California Coastal Water Research Project. Westminster, CA.

Schiff, K.C., A.E. Fetscher, M.M. Hanken. 2014. Newport Bay Watershed Monitoring Evaluation. Technical Report 815. Southern California Coastal Water Research Project. Costa Mesa, CA.

PG Environmental, 2018. Improving Stormwater Program Monitoring, Evaluation, Tracking, and Reporting: Workshop Report and Recommendations. Prepared for U.S. Environmental Protection Agency Region 9, October 12, 2018. https://www.epa.gov/npdes-permits/workshop-reports-improving-municipal-stormwater-program-monitoring-evaluation

Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M. and Morgan, R.P. 2005. The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society* 24(3), pp.706-723.

APPENDIX A: LIST OF MONITORING QUESTION ATTRIBUTES

Question #1: What is the temporal trend in pollutant concentrations and loading? [Mass Emissions Monitoring]	
Space and time scale:	Mass Emission Stations at Anaheim Bay/Huntington Harbour and Newport Bay watersheds
Indicator(s) to be measured:	Bacteria, total and dissolved metals, nutrients, sediment, pesticides (CHCs, PCBs, OP, pyrethroids)
Benchmark for comparison:	TMDL Target(s)
Action to be taken once you have an answer:	Triggers enhanced monitoring as you approach the threshold Triggers advanced data analysis for complicating factors Triggers source identification

Question #2: What are the trends in fish and bird tissue concentrations? [TMDL Monitoring]	
Space and time scale:	Tissue Stations in Newport Bay Watershed, Annual (Summer)
Indicator(s) to be measured:	Organics (DDTs, PCBs), selenium (Se)
Benchmark for comparison:	TMDL Target(s)
Action to be taken once you have an answer:	Evaluation of current remediation action effectiveness, Focusing future remediation planning and implementation, Listing/delisting decisions, Additional studies on pollutant transfer/bioaccumulation modeling to better quantify linkage analysis and TMDL targets

Question #3: Are there illegal discharges/illicit connections (ID/IC)? [Dry Weather Monitoring]	
Space and time scale:	Probability & Revisit sites in Anaheim Bay/Huntington Harbour, and Newport Bay Watersheds, Dry weather only (May – Sept)
Indicator(s) to be measured:	Focused on Top 10 parameter detection frequencies- Ag, As, Cd, Chlorpyrifos, Cr, Cu, diazinon, dimethoate, ENT, FC, Fensulfothion, Hg, malathion, Ni, Pb, Se, TC, TOC, TSS, Zn
Benchmark for comparison:	Dry Weather Performance Limit, Water Quality Standard, TMDL Target(s)
Action to be taken once you have an answer:	Triggers source tracking for identification, confirmation, remediation, enforcement

Question #4: What are the temporal trends in sediment quality and water column concentrations of estuaries and wetlands?	
Space and time scale:	Probabilistically selected stations in Newport Bay Watershed, Annual (Summer)
Indicator(s) to be measured:	Sediment chemistry, sediment toxicity, benthic infauna, commensurate with reginal monitoring protocols
Benchmark for comparison:	303(d) listing guidance (< 15% of area)
Action to be taken once you have an answer:	Trigger enhanced monitoring as concentrations or loads approach benchmarks (i.e., water quality standards, TMDL targets, etc.), to increase confidence that data are above or below the benchmark, Trigger advanced data analysis for complicating factors that could influence a benchmark exceedance (i.e., precipitation patterns for wet weather), Identify trends in concentrations and loads from receiving water inputs, and factors leading or contributing to those trends, Determine whether management actions have yielded desired outcome, or design/update future management actions to achieve the desired outcome, Trigger source identification to ensure the exceedance is coming from MS4 discharges.

APPENDIX B: WEB APPLICATION DESCRIPTIONS

Web Application descriptions

Five primary datasets were evaluated to answer key management questions for review of the North Orange County Monitoring Program. For each dataset, an R Shiny website (Chang et al. 2020; RDCT (R Development Core Team) 2020) was created to provide online access to analyses that addressed key management questions.

- Mass emissions, concentration data: https://sccwrp.shinyapps.io/ocms4review/mass_emissions.Rmd
- Mass emissions, load data: https://sccwrp.shinyapps.io/ocms4review/mass_emissions_loads.Rmd
- Dry weather or illicit discharge/illicit contaminant data: https://sccwrp.shinyapps.io/ocms4review/dry_weather.Rmd
- Tissue contaminant monitoring: https://sccwrp.shinyapps.io/ocms4review/tissue.Rmd
- Estuary/Wetlands water column and sediment monitoring: https://sccwrp.shinyapps.io/ocms4review/harbors_estuaries.Rmd

The structure of each website was similar, with some variation depending on specific questions relevant to each dataset. In general, each website included the following tabs to access and view results for information about each dataset:

1. Inventory: Map-based and tabular summaries of monitoring effort and basic characteristics of the data
2. Trends and power analyses: Changes over time by select constituents and locations, including power analyses to help identify optimal sampling effort
3. Station differences: A comparison of time series between stations to identify similarities among trends
4. Overall trends: Map-based and tabular summary of trend tests for all stations shown together

The website for the dry weather monitoring dataset also included the following tabs:

- Hotspots: Assessment of hotspot sites based on threshold exceedances over time
- Analyses by waterbody: A simple analysis of threshold exceedances for select sites shown for complete time series

Each main tab on the website (for the numbered items above) included additional sub-tabs or drop-down menus for selecting and viewing different content relevant to each dataset. For datasets where parameters lists were large (i.e., including all parameters on the website was impractical), select constituents were often included that were the most often observed or otherwise important for regulatory or other management needs. Decisions on specific constituents to include were vetted by the advisory committee.

Trend Analyses

All websites, excluding that for the dry weather monitoring program, included an assessment of trends over time to understand 1) long-term changes at specific locations for parameters of interest, 2) whether any changes were above or below thresholds of management or regulatory concern, and 3) if management actions to reduce pollutant concentrations or loads have been successful. Trends were evaluated visually using plots of concentration or loading of observed data over time, including an evaluation of monthly and annual trends using boxplots (McGill et al. 1978). Formal hypothesis tests for trends also included linear regression analyses and non-parametric Kendall tests of annual averages. For both linear regression and Kendall tests, trends were evaluated based on deviations of the annual average of a parameter from the grand mean at an individual monitoring station. The regression analysis reported an estimate of slope (change in concentration or load per year) and overall significance of the regressions. Kendall tests provided an alternative indication of trend by evaluating magnitude, direction, and significance of a change over time using a non-parametric approach. For the latter, the kendallTrendTest function from the EnvStats package for R was used (Millard 2013).

Power analysis

A critical question addressed during the evaluation of the North Orange County Monitoring Program was how well the current sampling design was able to detect trends of interest. In particular, questions were evaluated regarding the ability to detect a specified magnitude of change (e.g., 30% decrease over ten years) and the likelihood of observing an exceedance of a concentration (e.g., is the true mean above a regulatory threshold) for a given parameter over a period of time. For both questions, power analyses were conducted for specific parameters and locations where sufficient data were available. In essence, power describes the probability of observing a true event in a population, based on a sample of the population and if the true event actually occurred. This is analogous to observing an actual change in water quality conditions for a given sample design with the knowledge that sampling is discontinuous over time and at varying time intervals depending on location.

The first power analysis estimated the ability to detect a specific trend for a desired sampling frequency. For a chosen parameter and location, the observed time series was first detrended by taking the residuals of a regression of concentration or load vs time. From the residuals, the variance of the dataset around the mean was estimated and used to simulate new time series from which power was evaluated. For example, if a 50% change (increase or decrease) was considered the true change, a simulated time series was created by first estimating the linear change over the length of time that the true time series was observed (e.g., ten years) and then imposing uncertainty in the linear estimate by adding variance from the residuals to the linear trend. The observed level of sample effort was considered 100% of the current effort if the number of observations in the simulated time series was the same as the observed. Evaluating power at different levels of effort required subsampling of each simulated time series for the selected level of effort. For a large number of simulated time series (n = 1000), power was estimated as the percentage of simulations where the change was significant based on linear regression. This was repeated for varying sample effort from 10% to 200% of the current for a given time series.

A second power analysis was conducted to quantify the likelihood of observing an exceedance of a concentration or load of a parameter at a given sample density. Similar methods as the first analysis were used such that power was estimated by repeated sampling of a simulated time series using an estimate of variance for each parameter. Rather than estimating significance of a trend at a simulated magnitude of change, power was defined as the percentage of simulations where a simple t-test identified a significant difference of the simulated values above a threshold. The evaluated thresholds for each parameter were chosen across the range of observations of a parameter from the mean value to the 95th percentile. The simulated values were created as before, except that zero change in the linear difference from the beginning of each time series was assumed (i.e., time series had 0% magnitude change across the period of record). As for the first power analysis, power was evaluated for varying levels of sample effort from 10% to 200% of the observed. Importantly, the interpretation of the power estimates were slightly different than the first analysis. Rather than showing the percentage of time for which a trend would be detected, the power values shows the ability to determine that the true mean of a time series is equal to the threshold as an indication of confidence in exceeding a value of interest for a given sample design.

Optimal sample effort

An optimal level of sample effort was derived from the power analyses to describe the balance between over- and under-sampling. From a programmatic perspective, the optimal level of effort minimizes sampling cost by identifying the level of effort where any additional samples do not substantially increase the ability to detect a trend, whereas reductions in sample effort cause a disproportionate increase in the magnitude of the trend to be detected for a desired level of power. Graphically, the optimal level of effort is the inflection point on a power curve where the y-axis shows the magnitude of the trend to detect, and the x-axis is the level of sampling effort. This inflection point was determined quantitatively for each water quality parameter as the point in a monotonic power curve where the slope of y versus x (i.e., trend to detect vs sample effort) exceeded that of x versus y (i.e., sample effort vs trend to detect). Given that sample effort and variance of each time series differed considerably among the observed time series, optimal effort was identified for power curves only where sufficient data were available. A power estimate of 80% was considered a sufficient target for optimal effort.

Optimal effort was estimated for each parameter at each monitoring station and based on aggregates of optimal effort across parameters and stations. For aggregate estimates, optimal effort was identified for one parameter across all stations as the median optimal sample effort across stations. Similarly, optimal effort for multiple parameters and multiple stations was simply the median across all estimates. In cases, where the aggregate optimal effort varied considerably from the median, boxplot summaries were used to characterize the spread. Separate boxplots were also retained for parameters across stations so that variation across parameters could be determined, i.e., optimal effort for one parameter may be larger than another, and an *a priori* preference for one parameter could determine future sampling design. A final management question relative to optimal effort addressed the need to focus attention at locations where observed values were close to important regulatory thresholds. For the plots of optimal effort, points showing the inflection point were sized based on proximity of the average value for a parameter at a station to an appropriate threshold. Larger points indicated a station had a parameter close to a threshold, whereas smaller points indicated the average value was either much less or much greater than the threshold. Proximity to the threshold provided added context that additional attention could be focused on sites where management intervention may be needed.

Station similarity

The interactive websites also included a tab for identifying similarities among stations within each monitoring dataset. These analyses were provided to support decisions where potentially redundant sites with similar characteristics may be dropped or one site may be preferred over another for continued sampling if no additional information is gained by sampling multiple sites. Conversely, future sampling designs could be focused on locations where parameters or groups of parameters had maximum differences, thereby focusing efforts at locations with the highest variation among all sites. For individual parameters, dissimilarity measures across all sites were estimated by calculating the Euclidean distances of the standardized (zero mean, unit variance) values between each time series (Oksanen et al. 2018). This resulted in pairwise estimates of dissimilarity between all sites as a single relative number to quantitatively evaluate which sites had time series with similar characteristics. Further, a principal components analysis (PCA) was used on the average values of each time series across all parameters and sites to group similar sites using biplots (Venables and Ripley 2002). Each biplot showed site groupings relative to the dominant principal components and vectors for each parameter used in the PCA. Pairs of sites could also be selected on the website to view relative overlap for assessing similarity and which vectors (i.e., parameters) explained the groupings.

Hotspots

The dry weather monitoring dataset (illicit discharge, illicit contaminants) included an analysis of "hotspots" that was not provided for the other datasets given the priority management questions of the review committee. This analysis was motivated by the monitoring question of which sites were more likely to have exceedances for important thresholds relative to those that were stable and of low concentration over time. Hotspots were identified for individual parameters across all stations as the proportion of observations that were above a threshold of interest, defined as a threshold of regulatory concern (e.g., for TMDL compliance) or as the median value across all stations if a threshold was not applicable. These estimates produced a map where site points were sized and colored relative to the number of exceedances to identify groups of sites where exceedances were more common. A similar analysis was conducted by grouping all parameters across all sites, where the percentage of exceedances were estimated as all instances of an exceedance across parameters divided by all observations at a site. This provided a similar map where hotspots could be evaluated relative to multiple parameters. For both analyses, sites could be filtered by relative sample effort (i.e., total number of observations) and by date ranges to subset sites with similar sampling characteristics. Options to filter sites by receiving waterbodies where TMDLs are currently in place was also provided.

References

Chang, W., J. Cheng, J.J. Allaire, Y. Xie, and J. McPherson. 2020. *Shiny: Web Application Framework for R*. https://CRAN.R-project.org/package=shiny.

McGill, R., J.W. Tukey, and W.A. Larsen. 1978. "Variations of Boxplots." *The American Statistician* 32 (1): 12–16.

Millard, S.P. 2013. *EnvStats: An R Package for Environmental Statistics*. New York: Springer. http://www.springer.com.

Oksanen, J., F. Guillaume Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.H. Stevens, E. Szoecs. And H. Wagner. 2018. *Vegan: Community Ecology Package*. https://CRAN.R-project.org/package=vegan.

RDCT (R Development Core Team). 2020. "R: A language and environment for statistical computing, v4.0.1. R Foundation for Statistical Computing, Vienna, Austria."

Venables, W.N., and B.D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth. New York, New York: Springer-Verlag. http://www.stats.ox.ac.uk/pub/MASS4.

APPENDIX C: CONTINUOUS SENSOR SELECTION

Introduction

The objective of introducing continuous, near-real-time water quality monitoring using automated sensors is to get more data with less effort to help North Orange County Permittees identify long-term trends in water quality and/or isolate occurrences of unusual water quality events. The high-level strategy proposed herein is to deploy commercially available sensors at strategic locations in the network, measuring water quality continuously, in near-real time. Sensors are deployed for long enough to reliably detect trends in dry weather water quality, and/or identify potential problems at a given location before being relocated. The strategy is implemented in three major tasks:

1. Identifying key monitoring locations
2. Establishing correlations between measured parameters and permit parameters of concern
3. Developing data processing tools

The intent of this section is to outline a process by which Permittees can proceed to develop an automated monitoring program. Available sensors and a process for sensor selection are suggested, followed by the overview of processes required for tasks 2 and 3.

Sensor Selection

Implementing a strategy to obtain more data with less effort poses a series of trade-offs in terms of monitoring needs vs. existing sensor capabilities, logistics of sensor deployment, operation, and maintenance, and capital and operating cost. Key considerations for selecting sensors for continuous and/or near-real-time in-line water quality measurement are summarized in Table 1.

Using commercially available/off-the-shelf sensors is suggested for the purposes of identifying occurrences of dry-weather water quality deviations in the current North Orange County Monitoring Program. Various alternative products offer single- to- multi-parameter instruments plus data loggers and telemetry which are intended to be deployed by municipal crews, among others. Most vendors offer capability to measure physico-chemical parameters such as pH/ORP, temperature, dissolved oxygen (DO), turbidity, electrical conductivity (EC), water level (by pressure transducer). More advanced instruments might also offer capability to measure nitrogen species including nitrate and ammonia, algae, and chlorophyll, among others, as listed in Table 2. Parameters like salinity, total suspended solids (TSS), or biochemical oxygen demand (BOD), are typically calculated based on a site-specific calibration activity and/or combinations of multiple other directly measured parameters. Single-parameter instruments may be complexed for multiple parameter measurements in a user-configured data logger, requiring substantial expertise. Multi-parameter instruments (sondes or spectrometers) allow the user to customize the suite of parameters, but the number of parameters concurrently measured is uniquely limited by the instrument itself.

Four vendors offering readily-available, in-line water quality measurement products were identified. The systems shown in Figure 1 represent the top-of-the-line products for each vendor, and all vendors offer reduced capability models. For reference, the s::can spectro::lyser, YSI EXO2 and In-Situ AquaTroll 600 are a few inches in diameter, and 2-3 feet in length. The parameters for which each system can be customized are summarized in

Table 2. The s::can spectro::lyser is the only device amongst them that can measure its full suite of potential parameters in a single instrument. On the other hand, it does not measure water level while others can integrate a pressure transducer in addition to the water quality sensors. ATI sells four types of single sensor + loggers, and it is included here because they are the only vendor identified that appears to specialize in the specific application of in-line water quality monitoring. Multiple vendors can configure user-specified data logger + sensor arrays (e.g., Campbell Scientific or Hach).

Selecting a sensor to suit project needs is complicated. Trade-offs are not necessarily equal amongst products. While it is not the intent of this report to recommend a specific instrument or product, an example of the beginnings of a selection matrix is illustrated in Table 3. It is important to note that the ballpark base prices listed are presented as a general guide. Each instrument is highly customized, and prices vary according to the parameter suite chosen by the user. A decision should not be made based on the base price in Table 3. It is strongly recommended that thorough attention be given to the range of considerations in Table 1.

Unfortunately, a major drawback of using existing commercially available sensors is the mis-match between parameter suite in North Orange County's current ambient monitoring program (nutrients [nitrogen and phosphorus species], heavy metals, and pesticides) versus parameters currently available in commercial sensors. Research-grade and prototype sensors can be found for continuous and/or near-real-time measurement of several these parameters; for example, many can be found in the Open Access Journal, *Sensors*. However, practical implementation is challenging. Extensive site-specific calibration with manual samples and advanced training are often required in order to deploy and operate them.

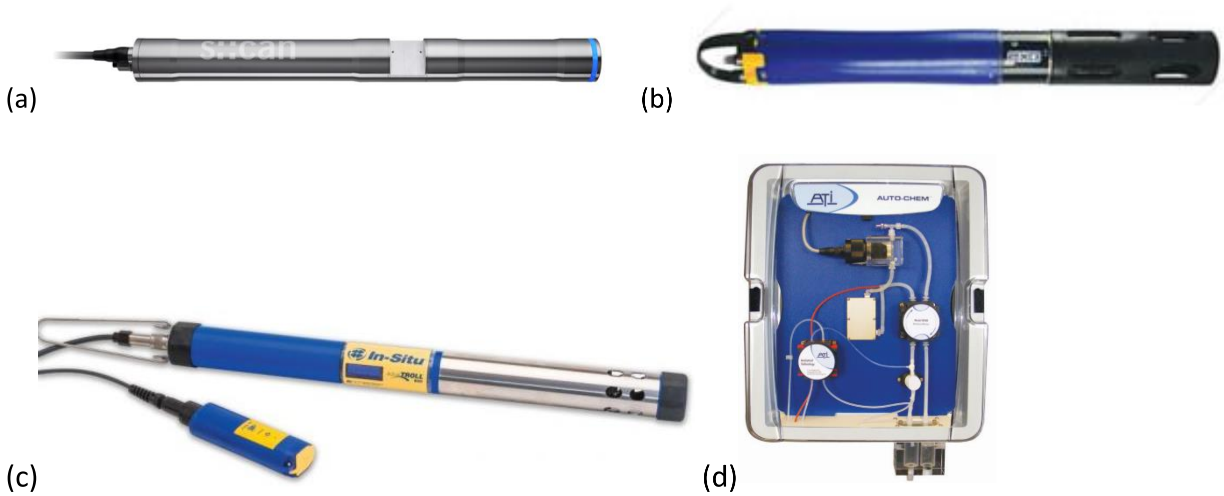


Figure 1. Top-of-the-line multi-parameter capable, commercially available sensors and loggers: (a) s::can spectro::lyser spectrometer (https://www.s-can.at/products/spectrometer-probes); (b) YSI EXO2 multi-parameter sonde (https://www.ysi.com/exo2); (c) In-Situ AquaTroll 600 multi-parameter sonde (https://in-situ.com/us/aqua-troll-600-multiparameter-sonde); (d) ATI Technologies data logger https://www.analyticaltechnology.com/analyticaltechnology/gas-water-monitors/product.aspx?ProductID=1038.

Table 1. Sensor selection considerations.

Measurement parameters	Number of concurrent parameters measured?
Logistics	Where will it be installed? Dry vs wet Safety/vandalism Access
	Anti-fouling features
	Size (how big/small is it?)
	Calibration requirements (activity & frequency)
	Accuracy, precision
	Sensitivity at low vs high concentration
Transparency/ Real cost of operation	Non-optional “accessories” Proprietary software Data loggers Communications
	Purchase cost vs long term maintenance cost
Power	Requirements, longevity Battery Solar Direct
Data & tech transfer	Telemetry vs manual download
	Availability of technical support

Table 2. Parameters potentially measured or calculated by commercially available instruments.

	S::can Spectro::lyser	YSI EXO2	In-Situ AquaTroll 600	ATI
pH/ORP		√	√	√
Temperature	√	√	√	√
Conductivity		√	√	√
Dissolved oxygen		√	√	√
Turbidity	√	√	√	
TSS	√	√	√	
TDS		√	√	
Chlorophyll a	√	√	√	
Algae • Total algae or • Phycocyanin (BGA-PC) or • Phycoerythrin (BGA-PE)		√	√	
Ammonium		√	√	
Chloride		√	√	
Nitrate	√	√	√	
Rhodamine		√	√	
Total organic carbon (TOC)	√			
Dissolved organic carbon (DOC)	√		√	
Organic matter • fDOM or • UV254	√	√		
Biochemical oxygen demand (BOD)	√			
Chemical oxygen demand (COD)	√			
Chloramine	√			
Ozone	√			
Hydrogen sulfide	√			
BTX	√			
Color	√			
Others	√	√	√	

Table 3. Example sensor selection matrix.

Product	Max. # Water Quality Parameters per Device	OK in a dry channel?	Configuration	Ballpark Base Purchase Price with Max # of Sensors and Anti-fouling*	Regular Replacement Parts Necessary?
In-Situ AquaTroll 600	4 + pressure	Mostly (not pH)	Multi-parameter Sonde	$8K-$9K	Yes
YSI EXO2	5 + Temp & Conductivity (required) + pressure	Some parameters (not pH or DO)	Multi-parameter Sonde	$8K-$10K	Yes
s::can Spectro::lyser	17	Yes	Spectrometer	$18K-$20K	No
ATI Technologies	4	No	Individual/single sensors connected to data logger	$3K per analyte	Yes

* Presented as an initial indicator only. Actual purchase price depends primarily on the specific suite of parameters that will be measured and other options such as power, telemetry, etc.

Deployment to Establish Correlations

Since specific parameters of interest cannot be easily measured with commercially available sensors, surrogate indicators must be identified for trend monitoring. The behavior of surrogate indicators must be relatable in a quantitative manner to behavior of the constituents of concern, for example via statistical techniques such as correlation or machine learning. In essence, a sensor or suite of sensors is deployed at a given location to continuously measure candidate indicators. Periodically, manual samples are collected and analyzed for the actual constituents of concern. Manual samples must be collected often enough to represent a range of environmental conditions and generate confidence in subsequent statistical analysis. Finally, data are investigated for statistically significant relationships between candidate indicators and contaminants of concern.

The approach is quite common in the scientific literature. It is clear that correlations are site- or waterbody- specific and must be uniquely developed for locations of interest. For example, monitoring of highway runoff in West Los Angeles found correlations between dissolved organic carbon, chemical oxygen demand (COD), and TKN, but none of these parameters were correlated to total suspended solids (TSS). TSS was correlated to particulate metals, but not cadmium, while COD was correlated to oil and grease (Han et al. 2008). A different study of an urban catchment found that turbidity was a good indicator of TSS and COD, while conductivity was correlated to total nitrogen (Aumond and Joannis 2006). Elsewhere again in an urban setting, turbidity was identified as an indicator of COD, ammonium, sulfate, or nitrate, while generally conductivity is a parameter indicative of human activity in the watershed (Díaz-Muñiz et al. 2012).

As with any water quality investigation based on indicators, since the most useful indicator is unknown at the start of the study, it is worthwhile to measure multiple parameters during the correlation phase. For example, in a study of Aliso Creek, researchers measured five common parameters (flow, DO, pH, temperature, turbidity) in order to find an indicator for bacteria (fecal coliform, total coliform, or enterococci). Individually, none of the candidate parameters was a suitable indicator. Statistically combining all five candidate parameters did yield an acceptable predictor for fecal coli. and total coli. (Bae et al. 2009). Concurrent monitoring of COD and TSS revealed that COD was a preferred indicator of wastewater discharge into surface water (Bareš et al. 2009). The U.S. Environmental Protection Agency (2016) offers guidance on core water quality parameters and what they might indicate in a potable water distribution system. This guidance could offer useful starting points for indicators in surface waters, with additional interpretation.

In at least one location in each of the four watersheds of North Orange County, it is recommended to deploy an array of continuous measurement sensors. To investigate viable indicators, manual samples for parameters of interest should be collected weekly- to monthly, and during or after known events such as spills or storms. Manual sampling is considered sufficient when statistically valid relationships between indicators and parameters of interest are identified.

Develop Data Processing Tools

Environmental data is highly variable. All sensors have limitations on accuracy and precision, which often varies with relative concentration (e.g., less sensitivity or compromised accuracy at upper or lower ends of detection ranges). Altogether, the continuous data to be collected from the North Orange County network is anticipated to be very noisy, complicating the interpretation of trends and/or identification of outliers.

Three categories of data processing are envisioned to convert continuous water quality data into a tool for solving dry weather water quality violations in North Orange County:

1. Data quality checks to validate logged values. Processing large, in-line data sets introduces demand for advanced, and/or new computational procedures.
2. Trend analysis, which may require data smoothing functions and other pre-processing to reduce noise.
3. Validation of outliers with upstream source tracking.

Many studies on data checkers, smoothing functions, and trend analysis can be found in the literature (Di Blasi et al. 2013; Bridle et al. 2014; Alferes and Vanrolleghem 2016). These studies serve as useful examples, but research will be required to calibrate site-specific procedures. To actually solve water quality problems, the upstream source tracking completes a comprehensive approach.

References

Alferes, J. and P.A. Vanrolleghem. 2016. "Efficient automated quality assessment: Dealing with faulty on-line water quality sensors." AI Communications **29**(6): 701-709.

Aumond, M. and C. Joannis. 2006. "Mesure en continu de la turbidité sur un réseau séparatif eaux usées : mise en œuvre et premiers résultats." La Houille Blanche(4): 121-128.

Bae, H.-K., B.H. Olson, K.-L. Hsu, and S. Sorooshian. 2009. "Identification and Application of Physical and Chemical Parameters to Predict Indicator Bacterial Concentration in a Small Californian Creek." Water Environment Research **81**(6): 633-640.

Bareš, V., D Stránský, and P. Sýkora. 2009. "Sewer infiltration/inflow: long-term monitoring based on diurnal variation of pollutant mass flux." Water Science and Technology **60**(1): 1-7.

Bridle, H., B. Miller, and M.P.Y. Desmulliez. 2014. "Application of microfluidics in waterborne pathogen monitoring: A review." Water Research **55**: 256-271.

Di Blasi, J.I.P., J. Martínez Torres, P.J. García Nietro, J.R. Alonso Fernández, C. Díaz Muñiz, J. Taboada. 2013. "Analysis and detection of outliers in water quality parameters from different automated monitoring stations in the Miño river basin (NW Spain)." Ecological Engineering **60**: 60-66.

Díaz-Muñiz, C., P.J. García Nietro, J.R. Alonso Fernández, J. Martínez Torres, and J. Taboada. 2012. "Detection of outliers in water quality monitoring samples using functional data analysis in San Esteban estuary (Northern Spain)." Science of the Total Environment **439**: 54-61.

Han, Y. H., S.L. Lau, M. Kayhanian, and M.K. Stenstrom. 2008. "Correlation analysis among highway stormwater pollutants and characteristics." Water Science and Technology **53**(2): 235-243.

U.S. Environmental Protection Agency. 2016. Online Source Water Quality Monitoring For Water Quality Surveillance and Response Systems, Office of Water. **EPA 817-B-16-003**.