THE Water Research FOUNDATION

PROJECT NO.
5048

Integrating Real-Time Collection System Monitoring Approaches into Enhanced Source Control Programs for Potable Reuse

CALIFORNIA Water Boards
STATE WATER RESOURCES CONTROL BOARD
REGIONAL WATER QUALITY CONTROL BOARDS

Integrating Real-Time Collection System Monitoring Approaches into Enhanced Source Control Programs for Potable Reuse

Prepared by:

Andrew Salveson, Amos Branch, Kyle Thompson, Brynne Weeks
Carollo Engineers, Inc.

Scott Mansell
Clean Water Services

Tyler Nading, Ken Thompson, Muhundhan Mohan
Jacobs Engineering Group

Eric Dickenson
Southern Nevada Water Authority

Brendt Thompson
s::can

Phil Ackman, Patricia Hsia
Los Angeles County Sanitation Districts

2023

Co-sponsored by:
California State Water Board

THE Water Research FOUNDATION

The Water Research Foundation (WRF) is the leading research organization advancing the science of all water to meet the evolving needs of its subscribers and the water sector. WRF is a 501(c)(3) nonprofit, educational organization that funds, manages, and publishes research on the technology, operation, and management of drinking water, wastewater, reuse, and stormwater systems—all in pursuit of ensuring water quality and improving water services to the public.

For more information, contact:
The Water Research Foundation

1199 North Fairfax Street, Suite 900	6666 West Quincy Avenue	
Alexandria, VA 22314-1445	Denver, Colorado 80235-3098	www.waterrf.org
P 571-384-2100	P 303-347-6100	info@waterrf.org

WRF ISBN: 978-1-60573-633-4
WRF Project Number: 5048

Prepared by
Carollo Engineers, Inc.
Clean Water Services
Jacobs Engineering Group
Southern Nevada Water Authority
s::can
Los Angeles County Sanitation Districts.

Acknowledgments

Research Team

Principal Investigators:

Andrew Salveson, PE
Carollo Engineers, Inc.

Amos Branch, PhD
Carollo Engineers, Inc.

Scott Mansell, PhD, PE
Clean Water Services

Tyler Nading, PE
Jacobs Engineering Group

Ken Thompson
Jacobs Engineering Group

Eric Dickenson, PhD
Southern Nevada Water Authority

Project Team:

Brynne Weeks, PE
Carollo Engineers, Inc.

Muhundhan Mohan
Jacobs Engineering Group

Erik Desormeaux, PE
Jacobs Engineering Group

Kyle Thompson, PhD, PE
Southern Nevada Water Authority / Carollo Engineers, Inc.

Brendt Thompson
s::can

Phil Ackman, PE
Los Angeles County Sanitation Districts

Patricia Hsia, PE
Los Angeles County Sanitation Districts

Participating Utilities

Hampton Roads Sanitation District (Virginia)
Los Angeles County Sanitation Districts (California)
Morro Bay (California)
Clean Water Services (Hillsboro, OR)
Southern Nevada Water Authority (Henderson, NV)

Project Advisory Committee

Shrirang Golhar, PE, Env. SP
City of Dallas

Charles Bott, PhD, PE
Hampton Roads Sanitation District

Angela Fasnacht, PE, MPH, PhD
Suez Water Management & Services, Inc.

Luis A. Durruty, PE, PMP, BCEE
LA Sanitation & Environment

Daniel Murray, PE, BCEE
USEPA

California State Water Board Liaison

Tricia Lee
California State Water Resources Control Board

Leslie Hart
California State Water Resources Control Board

Cindy Figueroa
California State Water Resources Control Board

WRF Staff

John Albert, MPA
Chief Research Officer

Harry Zhang, PhD, PE
Research Program Manager - Integrated Water and Stormwater

Abstract and Benefits

Abstract:

Real-time monitoring in sewersheds or wastewater influent has the potential to provide a critical early warning system to downstream processes and water quality. However, such monitoring also requires significant ongoing instrument maintenance and calibration. Expenses, such as auxiliary equipment and time for travel, maintenance, troubleshooting, and analysis, usually exceed the cost of the sensors themselves. These challenges can be barriers to the potential major benefits of real-time sewershed monitoring. In addition to an early warning system, real-time sewershed monitoring can detect and track down industrial discharges or other adverse water quality events back to their source. While broadly applicable to wastewater utilities, these benefits are especially important for potable reuse, where the water quality consequences can be significant.

Previous WRF projects, most notably Steinle-Darling et al. 2022, *Demonstrating Real-Time Collection System Monitoring for Potable Reuse*, have explored this and closely related topics. That project found sensors that were resilient and accurate in other contexts required much more maintenance in sewersheds due to collection and buildup of debris, biofilm, etc. on sensors. More innovation was needed for resilient sensors, how to install them to minimizing fouling and ragging, and how to analyze the data to discern errors and real events. Furthermore, a step-by-step guide building upon lessons learned by pioneers in this area could enable utilities to achieve real-time sewershed monitoring with fewer setbacks.

This new project (1) reviewed available sensors; (2) reviewed past related research reports including Steinle-Darling et al. 2020; (3) conducted new field testing on innovative sensors and sensor-holders; (4) reviewed enhanced source control programs and real-time monitoring programs (ESCPs) and synthesized how they could benefit each other; (5) applied machine learning to real-time wastewater influent and reuse data; and (6) culminated in a step-by-step framework for successful real-time monitoring. The report concluded that the challenges of real-time sewershed monitoring should not be understated but can be overcome with a clear goal and sufficient perseverance and resources, thus realizing the potential benefits of real-time monitoring.

Benefits:

- This report will enable utilities to consider real-time sewershed monitoring with specific examples of what can be achieved but knowing the challenges and ongoing costs.
- Through numerous real-time monitoring case studies, utilities can learn from the experiences of others. This includes both big picture lessons, such as the importance of goal-setting, but also details, such as what maintenance frequency to expect.
- The framework (Chapter 8) serves as a step-by-step guide for utilities seeking to integrate real-time sewershed monitoring into ESCPs.

Contents

Tables

Figures

Acronyms and Abbreviations

AMI	Advanced metering infrastructure
ASR	Aquifer storage and recovery
AWPF	Advanced water purification facility
AWTF	Advanced water treatment facility
BIA	Business intelligent architecture
BOD	Biological oxygen demand
CCP	Critical control points
CIU	Categorical industrial user
CLR	Calcium, lime, and rust remover
COD	Chemical oxygen demand
CWS	Clean Water Services
CWT	Centralized waste treatment
DaaS	Data-as-a-Service
DPR	Direct potable reuse
EC	Electrical conductivity
ECD	Electro-chemical devices
ESCP	Enhanced source control program
FOG	Fats, oils, and grease
ft/s	Feet per second
gpm	Gallons per minute
H_2S	Hydrogen sulfide
HRSD	Hampton Roads Sanitation District
IU	Industrial user
JWPCP	Joint Water Pollution Control Plant
LACSD	Los Angeles County Sanitation District
LANWRP	Lancaster Water Reclamation Plant
LCWRP	Los Coyotes Water Reclamation Plant
µS/cm	Microsiemens per centimeter
MCL	Maximum contaminant level
MES	Microbial electrical signal
mgd	Million gallons per day
mg/L	Milligrams per liter
ML	Mixed liquor
nm	Nanometer
NO_2	Nitrogen dioxide
NO_3	Nitrate
NONC	Notices of non-compliance and correction
NOV	Notice of violation

NPP	National Pretreatment Program
O&M	Operations and maintenance
ORP	Oxidation reduction potential
OWTP	Oxnard Wastewater Treatment Plant
PCA	Principal component analysis
PFAS	Per- and polyfluoroalkyl substances
POTW	Publicly owned treatment works
PVC	Polyvinyl chloride
sBOD	Soluble biochemical oxygen demand
SCADA	Supervisory control and data acquisition
SCWW	Santa Clara Wastewater
sCOD	Soluble chemical oxygen demand
SIU	Significant industrial user
SJCEWRP	San Jose Creek East Water Reclamation Plant
SML	Supervised machine learning
SOP	Standard operating procedure
sTOC	Soluble total organic carbon
TP	Total phosphorus
TSS	Total suspended solids
UV	Ultraviolet
UVA	Ultraviolet absorbance
UVT	Ultraviolet transmittance
VWRF	Ventura Water Reclamation Facility
WRF	The Water Research Foundation
WWTP	Wastewater treatment plant

Executive Summary

ES.1 Key Findings

- Real-time monitoring technology in sewersheds or wastewater influent can now monitor and provide alarms based upon important chemical parameters, such as pH, conductivity, and organic compounds and surrogates. These alarms could trigger investigation, diversion, or treatment adjustment. Nevertheless, stability and accuracy continue to limit the applicability of certain sensors in this context, especially for organic compounds and surrogates.
- Real-time monitoring in sewersheds or wastewater influent can be successfully implemented but often requires significant ongoing instrument maintenance and calibration. Expenses, such as auxiliary equipment and time for travel, maintenance, troubleshooting, and analysis usually exceed the cost of the sensors themselves.
- Utilities should use real-time monitoring with a clear goal and a continuous improvement mindset, using key performance indicators to measure success.
- Careful instrument selection and robust accessories (i.e., power supply, probe holders) are critical to success.
- Real-time monitoring can be integrated into an Enhanced Source Control Program (ESCP) for reuse to detect illicit discharges, identify sources, provide water quality early warning alarms, and communicate with industry.
- Pilot testing is highly recommended to compare sensors and verify site-specific performance and maintenance requirements.
- Machine learning could be applied for early warning systems using data from wastewater influent or within reuse systems, but this requires site-specific modeling.

ES.2 Background and Objectives

Real-time sewershed monitoring has several benefits for utilities, such as detecting and source-tracking industrial discharges or other adverse water quality events. While broadly applicable to wastewater utilities, these benefits are especially important for potable reuse. Atypical water quality events that could disrupt wastewater treatment plant (WWTP) performance or result in partial pass-through of chemical pollutants could then potentially interfere with the finished water quality from advanced purification. These considerations become even more important as the water sector advances towards direct potable reuse (DPR), which would involve less lead time for response than managed aquifer recharge or augmentation of surface water supplies.

Recent advances in commercially available sensor technology and data connectivity (e.g., internet-of-things, 5G) initially led to high expectations about real-time sewershed monitoring. However, previous projects funded by The Water Research Foundation (WRF) (e.g., Steinle-Darling et al. 2020) revealed a more complex reality. Sensors that were resilient and accurate in other contexts required much more maintenance in sewersheds due to collection and buildup of debris, fats, oils & grease (FOG), etc. on sensors. Utilities are also paying close attention to the cybersecurity considerations of data connectivity systems to prevent intrusions into their supervisory control and data acquisition (SCADA) or business systems. Engineered storage buffers, multibarrier treatment trains, and sensors in reclaimed or purified water can also provide a high degree of protection against atypical water quality events. These challenges and alternative safeguards raised the question of what—if any—level of sewershed monitoring is feasible, appropriate, or necessary for potable reuse. Furthermore, if a utility were to begin sewershed monitoring as part of their ESCP, it would be efficient and expedient to have a step-by-step guide based on the experiences of other utilities in this area.

Therefore, our research objectives were to:

- Determine what level of real-time collection system monitoring is feasible, appropriate, and necessary for protection of downstream potable reuse.
- Develop a framework for integrating real-time monitoring into existing pretreatment program requirements, including data management and security considerations.

ES.3 Project Approach

To fulfill these objectives, this project was organized into four tasks as follows. **Task 1 (Chapter 2)** was a review of currently available real-time monitoring tools. This included a review of currently available sensors (**Task 1A**), a summary of related past WRF sensor efforts (e.g. Steinle-Darling et al. 2020) (**Task 1B**), coordination with the contemporaneous WRF Project 4797 (Thompson *in process*) (**Task 1C**), and summaries of experiences with sensors from project partners (**Task 1D**).

Task 2 (Chapter 3) was new bench and field research to compare the latest sensors and test new methods to reduce fouling and ragging. This included studies at Clean Water Services (CWS) (**Task 2A**), Morro Bay (**Task 2B**), and Los Angeles County Sanitation District (LACSD) (**Task 2C**).

Task 3 was a review of existing ESCPs and real-time monitoring programs and how they might be integrated. This included the subtasks of a review of ESCPs (**Task 3A, Chapter 4**), a review of real-time collection system monitoring results at Ventura and Oxnard (**Task 3B, Chapter 5**), proof-of-concept machine learning analyses for early warning systems using data from CWS and Hampton Roads Sanitation District (HRSD) (**Task 3C, Chapter 6**), and discussion of the potential implementation of real-time monitoring for ESCPs at partner utilities Morro Bay, CWS, and LACSD (**Task 3D, Chapter 7**).

The above gathering, generation, and analyses of data all informed a framework for integrating real-time monitoring into ESCPs (**Task 4, Chapter 8**).

ES.4 Results

Task 1: Many online water quality sensors are now commercially available. Many parameters measurable with these sensors are useful for detecting changes in raw wastewater quality, such as pH, conductivity, and different measurements of organic matter (e.g., chemical oxygen demand [COD]). Optical properties (e.g., absorbance, fluorescence) can be used as indicators directly, or used to estimate other water quality parameters (e.g., COD, total organic carbon, algae). Only a few specific organic chemicals can be measured by online instruments, such as trihalomethanes and compounds with distinct optical signatures (e.g., optical brighteners). It is recommended to select sensors for sewershed monitoring that are least prone to water quality interferences. Also, it is recommended to have sensors without consumables or moving parts. Self-cleaning capabilities (e.g., air blast, ultrasonic cleaning) are especially helpful in this context.

Task 2: Side-by-side trials are an ideal way to compare and select among sensors (Table ES-1). Constructed flumes tested at one of the research sites enabled sensor comparisons in a relatively controlled environment, as well as experiments like intentional ragging and known chemical spikes. CWS developed a custom sensor holder that curves in the direction of flow so that any ragging or debris would be minimized (Figure ES-1). This solution enhanced the accuracy and resilience of the sensors. Testing systems within full-scale plants allowed for detailed research of probe systems without laboratory optimization (at both Morro Bay and LACSD), presenting greater clarity on the performance and maintenance of systems in their current market form. Accuracy and the frequency of cleaning and calibration varied significantly among the sensors. Sensor location was found to be critical in order to optimize the cleaning and data quality.

Table ES-1. Overview of Sensors Tested at Three Utilities for Task 2 Bench and Full-Scale Field Research.
✔ indicates generally successful trial. **X** indicates tested but did not meet the utility's criteria for continued use for real-time monitoring. *Successful when installed within CWS's rag guard sensor holder.

Brand	Sensor	CWS	Morro Bay	LACSD
Yosemitech	Y532-A	✔*		
ECD	ORP Pt Cap peek, two-tang probe	X		
ECD	Extended Life pH Electrode RADEL body	✔*		
s::can	Spectro::lyser	✔	✔	
s::can	condu::lyser	✔	✔	
s::can	pH::lyser	✔	✔	
Real Tech	Titanium Ba-X Series SA2010 multi-wave sensor			X
Sentry	Sentry-AD			X

Figure ES-1. Rag Guard Sensor Holder.

Task 3: Utilities practicing or actively planning potable water reuse already have ESCPs in place. Real-time sewershed monitoring often increases the need for grab sampling in the short term for calibration and verification. However, in the long run, real-time monitoring could reduce grab samples if it is deemed sufficiently reliable to replace some existing grab sampling routines. Real-time monitoring could also benefit ESCPs by more closely monitoring industrial discharger compliance, and rapidly notifying industrial dischargers in the event of atypical water quality (Table ES-2). Installing sensors and getting reliable data alone will not make a monitoring program successful. Real-time monitoring requires a connection to the utility's business or SCADA system and dedicated resources to interpret the data and set actionable criteria. Machine learning can expand upon the value of the sensor data through application for early warning systems using data from multiple sensors. However, machine learning can only detect types of events that have occurred before. So, these models should be employed alongside of (not instead of) traditional setpoint-based alerts. The use of spectral alarms and multi-parameter alarms has been shown to be successful at enhancing the detection of an unusual water quality anomaly.

Table ES-2. Aspects of ESCP and Benefit of Effective Online Monitoring.

ESCP Concept	Details	Online Monitoring Benefits Include…
Regulatory Authority	The Sewer Use Ordinance (SUO) provides the authority of a utility to develop and enforce an industrial pretreatment program, including requirements to protect potable water reuse.	Not applicable
Industrial Dischargers	The type and abundance of different industrial dischargers will define the level of effort and cost of a robust ESCP.	Online monitoring systems can provide greater confidence in industrial discharger compliance.
Enforcement Response Plan	The Enforcement Response Plan (ERP) outlines the procedures followed by pretreatment program staff and management to identify, document, and respond to pretreatment violations.	Online monitoring systems can be used to track abnormal discharges in the collection system and to the point of origin.
Monitoring Program	A robust ESCP relies upon a combination of industry-led and utility-led sampling efforts, and a tailored sampling campaign that adjusts with time to minimize laboratory analytical costs while closely monitoring water quality that can be of concern.	Online monitoring systems can reduce the amount of utility-led grab sampling and composite samples through development of a database of “standard” wastewater quality downstream of industrial dischargers.
Outreach Efforts	An effective outreach plan includes: 1. Communication between government departments (e.g., planning department and wastewater department) 2. Engagement of businesses 3. Development and sharing of Best Management Practices (BMPs) 4. Rewarding and acknowledging model industry partners 5. Notifying and enforcing non-compliant industries	Online monitoring results can be used to more rapidly contact industrial dischargers to alert them of water quality changes.
Interagency Agreements	Potable water reuse projects often cross jurisdictional boundaries or require collaboration of water and wastewater utilities. The development of clear roles, responsibilities, and financial commitments from participating parties is central to long term project success.	Online monitoring of wastewater from partner utilities provides for better cost recovery based upon flows and loads.

Task 4: A framework for integrating real-time monitoring into ESCPs was developed based on these results, prior related WRF projects, and the project team’s extensive experience (Figure ES-2). This framework was organized into three phases, as shown below. It is also ordered chronologically, from planning to ongoing improvement. So, this framework serves as a step-by-step guide for utilities seeking to implement real-time sewershed monitoring. Measuring success in Phase 3 of the framework provides the opportunity to continuously improve on system performance.

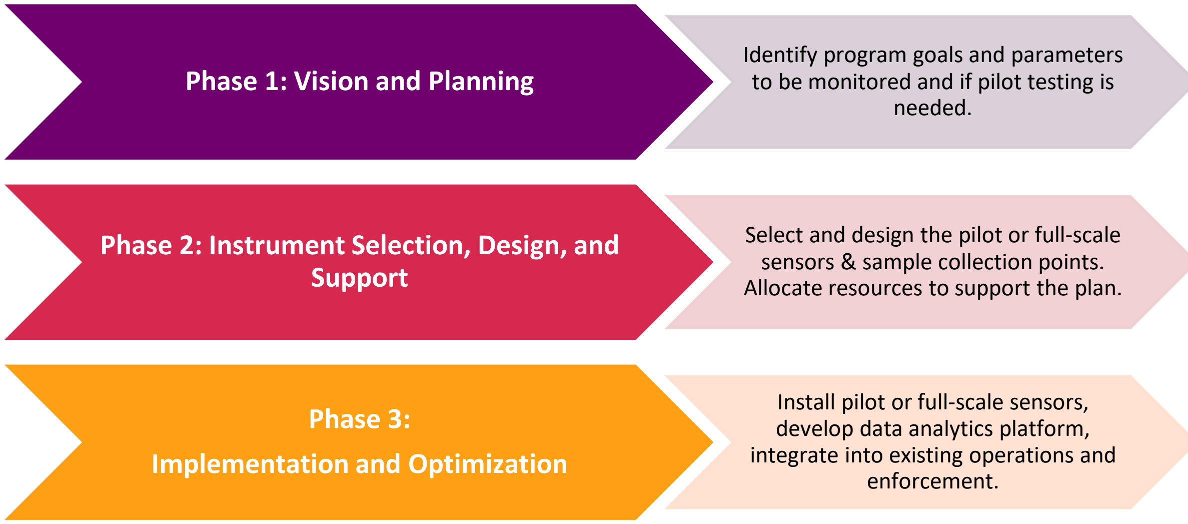


Figure ES-2. Framework for Integrating Real-Time Sewershed Monitoring into ESCPs.

ES.5 Benefits

- Overall, this report will enable utilities to consider real-time sewershed monitoring with specific examples of what can be achieved, knowing the challenges and ongoing costs. Real-time sewershed monitoring is possible, but it is not easy at this point.
- Through numerous new and reviewed real-time monitoring case studies, utilities can learn from the experiences of others. This includes both big picture lessons, such as the importance of goal-setting, but also details, such as what maintenance frequency to expect for different sensors in different waters. The case studies also include two proofs-of-concept for machine learning for improved alerts for water quality events.
- The framework (Chapter 8) serves as a step-by-step guide for utilities seeking to integrate real-time sewershed monitoring into ESCPs.

ES.6 Related WRF Research

- Demonstrating Real-Time Collection System Monitoring for Potable Reuse (4908)
- Designing Sensor Networks and Locations on an Urban Sewershed Scale with Big Data Management and Analytics (4797)
- Integrated Management of Sensor Data for Real Time Decision Making (4759)
- Leveraging Other Industries - Big Data Management: Phase I (4836)
- Compendium of Sensors and Monitors and Their Use in the Global Water Industry (4428)

CHAPTER 1

Introduction

1.1 Background

Industrial and Illicit discharges to municipal sewers are a constant challenge and concern for potable water reuse systems. With a well-regulated and enforced pretreatment program, these industrial dischargers can be well managed to allow for safe and reliable drinking water, protection of water quality, and a thriving industrial community. However, even for well managed systems, periodic accidental or intentional discharges, (e.g., slug discharges) can be problematic to both the downstream WWTP and the potable reuse treatment system if one is in place. For example, oils and grease from food processing can resist biodegradation during secondary treatment due to low solubility and cause unsightly films or ecological harm. Acids or bases can disrupt biological treatment or even damage infrastructure.

Industrial discharges have particularly high human health relevance at WWTPs that are water sources for potable reuse systems. For example, centralized hazardous waste treatment facilities can release extremely toxic radionuclides. Landfills can release bromide, which interferes with treatment by increasing bromate in ozonation or brominated disinfection byproducts in chlorine disinfection (Nading et al. 2022). Industries can release specific chemicals that are challenging to remove such as 1,4-dioxane or per- and polyfluoroalkyl substances (PFAS).

Current regulations under the National Pretreatment Program (NPP) protect against the impact of industrial discharges on WWTPs and on receiving waters. However, existing source control programs under the NPP were designed primarily to protect wastewater infrastructure and the aquatic environment. Potable reuse has greater direct human health relevance and stricter regulatory water quality limits. So, it requires different and more stringent monitoring, i.e., ESCPs (Figure 1-1). Compared to indirect potable reuse like managed aquifer recharge, DPR would potentially have less lead time to respond to adverse water quality changes. This further reinforces the need for tighter monitoring of water quality in the sewer collection system. Accordingly, it is becoming even more important for reuse utilities to promptly detect surges of unwanted industrial effluent.

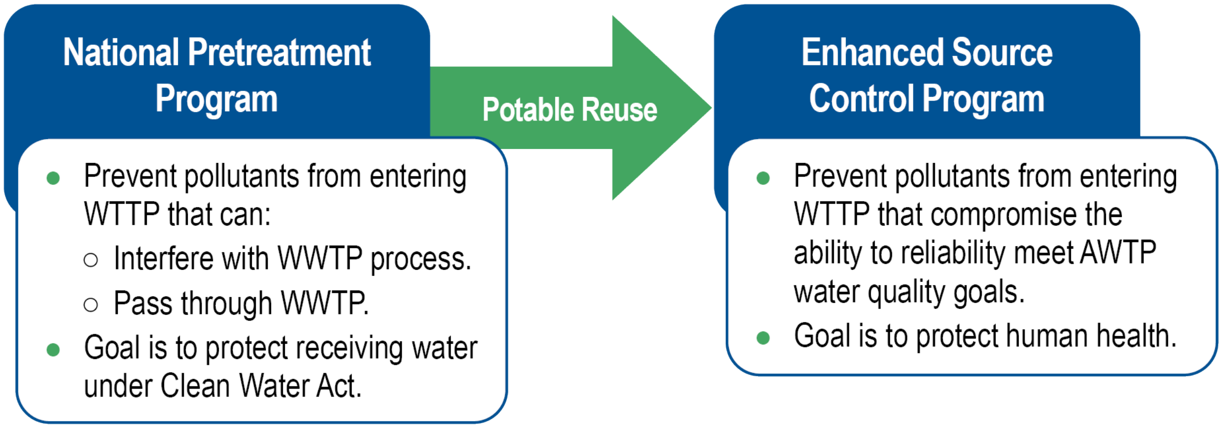


Figure 1-1. Enhancement of Existing Pretreatment Programs for Potable Reuse.
Adapted from Nading et al. 2022.

Real-time monitoring is the use of sensors that produce data continuously or frequently, i.e., at least hourly. If sufficiently precise and reliable, real-time monitoring in the sewershed could enable much more rapid and consistent detection of water quality excursions such as industrial discharges (Figure 1-2). Specific strategies include continuously monitoring at the WWTP raw influent, key nodes of the collection system, or the discharge of point sources of concern. It could also include mobile units to track contaminants or surrogates upstream towards their source. The potential benefits of real-time monitoring are great enough that certain states are considering requiring it. For example, the draft California DPR regulations would require "…a sewershed surveillance program to receive early warning of a potential occurrence that could adversely affect the DPR treatment and … on-line monitoring instrumentation at critical locations that measure surrogate(s) that may indicate a chemical peak."

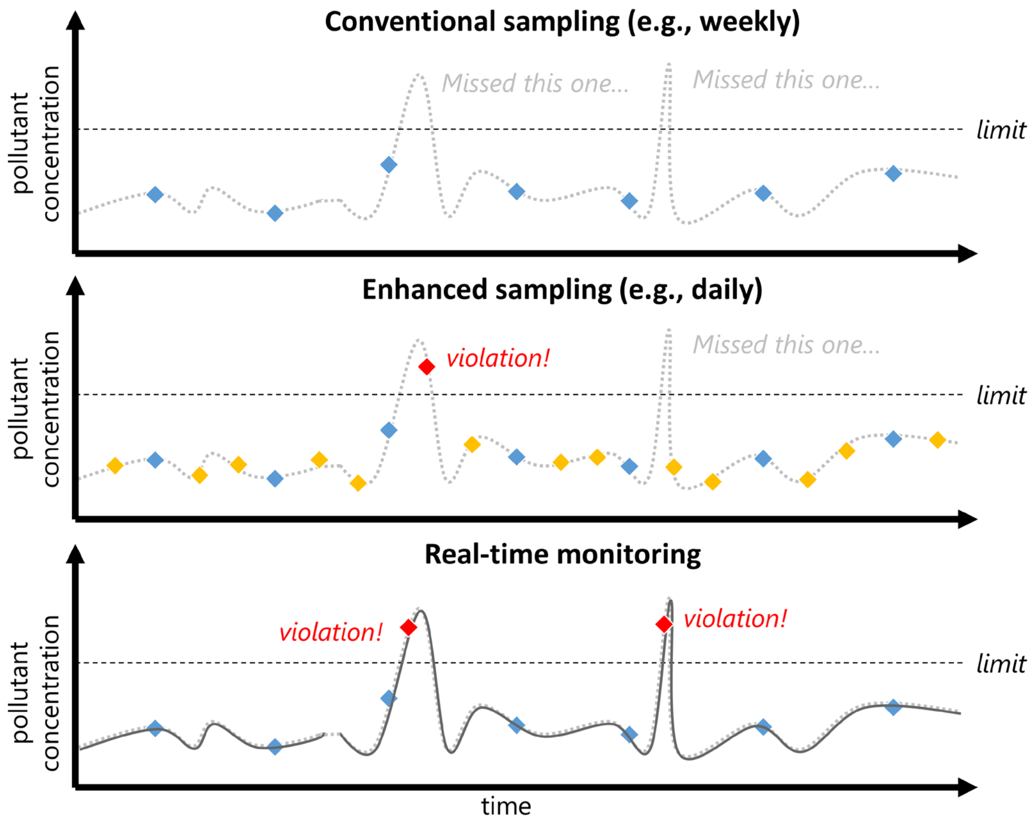


Figure 1-2. Theoretical Faster, More Consistent Excursion Detection with Real-Time Monitoring.
The dashed line represents the true pollutant concentration. Diamonds represent grab samples. The solid dark gray line represents the real-time monitoring signal.

Advances in commercially available online instrumentation—such as affordable, self-cleaning optical sensors—initially led to much optimism and enthusiasm about real-time sewershed monitoring (Figure 1-3). The WRF identified that less research had been done on real-time sewershed monitoring compared to advanced treatment, and funded several related projects, including:

- SENG6R16 Phase 1/WRF 4835 – Designing Sensor Networks and Locations on an Urban Sewershed Scale – (Liggett et al. 2018).
- SENG6R16 Phase 2/WRF 4797 – Designing Sensor Networks and Locations on an Urban Sewershed Scale with Big Data Management and Analytics – (Thompson *in process*).
- Reuse 17-30/WRF 4908 – Demonstrating Real-Time Collection System Monitoring for Potable Reuse (Steinle-Darling et al. 2020).
- WRF 4759 – Integrating Management of Sensor Data for a Real Time Decision Making and Response System (Neemann et al. 2019).

However, practical limitations soon dampened these expectations. The unfavorable and highly variable water quality in raw wastewater caused instruments to require frequent maintenance or even replacement. Power supply, data connectivity (e.g., automatically, reliably transferring data to a platform where it can be readily visualized and used), and physical security were all major challenges in sewers. Growing cybersecurity concerns raised questions about the wisdom of and best practices for transferring data from these widespread sensors to facility networks. Multibarrier treatment trains, other ESCP measures, and sensors in treated water at WWTPs or advanced treatment facilities already provide a high degree of protection against industrial discharges. So, the practicality, necessity, and cost-effectiveness of real-time monitoring came into question.

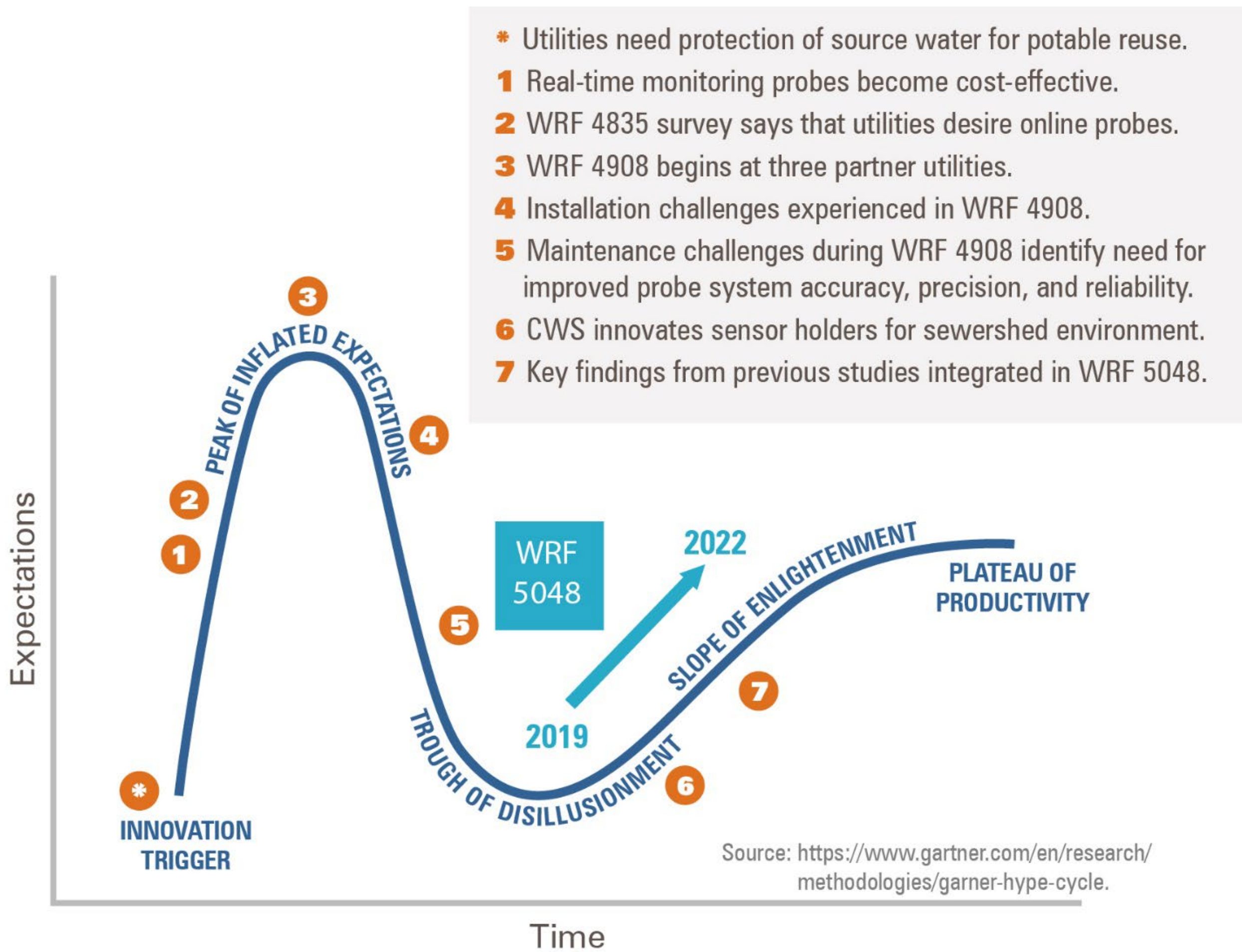


Figure 1-3. Gartner Hype Cycle as Applied to Real-Time Sewershed Monitoring.

1.2 Objectives

Our goal was to make recommendations on whether and how to integrate real-time sewershed monitoring into ESCPs, considering both the benefits and the practical challenges. To this end, our research objectives were to:

- Determine what level of real-time collection system monitoring is feasible, appropriate, and necessary to protect potable reuse.
- Develop a framework for integrating real-time monitoring into existing pretreatment programs, including data management and security considerations.

1.3 Report Contents and Technical Approach

To fulfill these objectives, this project was organized into the four tasks below. The chapter numbers for these tasks are shown in parenthesis. The team conducted state-of-the-industry reviews of instruments available, real-time monitoring projects, and ESCPs. New field research was conducted to test the accuracy and maintenance requirement of cutting-edge instruments under different conditions (water velocity, ragging) and in different locations within sewersheds or WWTPs. The team conducted two desktop proof-of-concepts for machine-learning-based alert systems using data from WWTPs or advanced treatment trains. These activities, as well as the project team's experience and prior related WRF-funded projects, then informed a step-by-step framework for integrating real-time monitoring in ESCPs.

Task 1 – Review Current Available Real-Time Monitoring Tools (Chapter 2)

- 1A – Review of Currently Available Sensors
- 1B – Summary of Related WRF Sensor Efforts
- 1C – Coordination with WRF 4797
- 1D – Experience with Sensors from Project Partners

Task 2 – New Bench and Field Research (Chapter 3)

- 2A - Clean Water Services Field Research (Appendix A)
- 2B - Morro Bay Field Research
- 2C - Los Angeles County Sanitation Districts Field Research

Task 3 – ESCP and Real-Time Monitoring Review

- 3A - Review of ESCPs (Chapter 4)
- 3B - Review of Real-Time Collection System Monitoring Results (Chapter 5)
- 3C – Data Analysis for Online Monitoring Systems within WWTPs or Advanced Water Treatment Facilities (Chapter 6, Appendix B, Appendix C)
- 3D – Potential Implementation of Real-Time Monitoring for ESCPs (Chapter 7)

Task 4 – Framework for Real-Time Monitoring (Chapter 8)

CHAPTER 2

Review of Currently Available Real-Time Monitoring Tools

The goal of this chapter was to provide a comprehensive look at commercially viable and proven sensor technologies as well as some emerging sensor technologies for biologic monitoring that can potentially be used within the sewer collection system and/or at various locations at a WWTP. The information presented within this chapter is organized as follows:

- A review of currently available sensors from different suppliers, which documents a large range of sensor technologies that can be deployed;
- A summary of experience from different WRF grants on similar sensor-related topics, which documents both the potential and the challenges of existing sensor systems;
- Coordination and information sharing with ongoing Thompson *in process* (Designing Sensor Networks in an Urban Sewershed); and
- Experience with sensors from project partners.

2.1 Review of Currently Available Sensors

The project team reviewed information from both internal and external resources to assemble the information for available physical and inorganic parameters, organic parameters, radionuclides, biological parameters, and integrated systems. Key resources included the following:

- SENG6R16 Phase 1/WRF 4835 – Designing Sensor Networks and Locations on an Urban Sewershed Scale – (Liggett et al. 2018).
- SENG6R16 Phase 2/WRF 4797– Designing Sensor Networks and Locations on an Urban Sewershed Scale with Big Data Management and Analytics – *(*Thompson *in process*).
- SENG7R16/WRF 4836 – Big Data Management - Phase I (Kadiyala and Macintosh 2018).
- WRF 17-30/4908 – Demonstrating Real-Time Collection System Monitoring for Potable Reuse (Steinle-Darling et al. 2020).
- WRF 4759 – Integrating Management of Sensor Data for a Real Time Decision Making and Response System (Neemann et al. 2019).
- Water Quality Sensors – Global Horizon Scan (Livaniou et al. 2020).
- SENG1C11– Compendium of Sensors and Monitors and Their Use in the Global Water Industry (GWRI & WERF) (van den Broeke 2014).
- WE&RF 11-01/WRF 1688 – Monitoring for Reliability and Process Control of Potable Reuse Applications (Pepper and Snyder 2016).
- Prior evaluations conducted for the USEPA Water Security Division Surveillance and Response Program (USEPA 2021b).

The information includes provider, sensors, parameters measured, and maturity level, as listed in Table 2-1.

This project focused on real-time monitoring systems that can track and detect challenging water quality from upstream industrial dischargers. This could include physico-chemical properties (e.g., pH), bulk organics, optical signatures, or specific inorganics that could be problematic themselves or surrogates correlated with specific chemicals. A small number of online biological instruments (i.e., measuring pathogens or measuring toxicity using microbes) are commercially available or may become available soon (Table 2-1). If sufficiently accurate and reliable, online pathogen sensors could be used for an early warning system or even adaptive control. However, with the possible exception of hospitals, pathogens would not come from industries or point sources. Rather, pathogens would inevitably come from domestic wastewater with no intention or fault from any person or organization. Source control is designed to "reduce conventional and toxic pollutant levels discharged by industries and other nondomestic wastewater sources into municipal sewer systems" (USEPA 2021a). So, online pathogen sensors could not be used to identify, prevent, or penalize problematic point source discharges. Thus, online pathogen sensors would not technically be part of an ESCP. Online pathogen sensors were then not a focus of this report, which is about integrating real-time collection system monitoring approaches into ESCPs for potable reuse.[1]

2.2 Summary of Related WRF Sensor Efforts

A number of projects have evaluated sensors in raw wastewater or within water reclamation plants. These past and ongoing projects provide an important perspective as the team looks to find the most value in online monitoring systems as they apply to potable water reuse projects. The benefits and challenges of sensor systems tend to appear in multiple efforts, with key references for the summary items below being Liggett et al. 2018, Neemann et al. 2019, and Steinle-Darling et al. 2020. Select findings are presented below.

Liggett et al. 2018 collected surveys from 20 utilities and 20 technology providers, and noted the following important items regarding online monitoring:

- **Reactive Sampling:** Having a sensor system that measures a significant change in water quality and triggers a sampling event initiates documentation of an illicit intermittent discharge that is above permit levels, which may be enforceable.
- **Identification of Water Quality Variation:** A sensor system can assist staff in developing an understanding of baseline water quality and thus identify intermittent pollutant spikes.
- **Direct Monitoring of Known Discharges:** Online monitoring systems allow utilities to monitor industrial dischargers directly and continuously, providing better enforcement and billing.
- **Deterrence Effect:** Knowledge by an organization or industry that they are being monitored improves discharger compliance with local limits.

[1] For further information on online sensors for pathogen monitoring, the team refers the reader to work by the National Water Research Institute (NWRI 2020).

Pepper et al. summarized why utilities would benefit from implementing real time monitoring in support of potable reuse projects. The reasons are these:

- **Cost:** Online systems can better determine where utility-sponsored sampling should occur within the collection system.
- **Regulatory Support:** Online monitoring within the collection system provides confidence to local regulators concerned about industrial and illegal discharges.
- **Event Detection/Response:** The ability to detect a potentially toxic water quality and proactively divert flow to downstream purification systems would reduce water quality concerns. An example of this would be a statistically significant increase or decrease of a particular parameter from baseline data, indicative of an abnormal water quality change.
- **Ease of Maintenance and Data Management:** Online systems would, should they be reliable and of minimal maintenance, reduce the overall collection system monitoring costs and data management challenges.

From various other experiences and surveys, the project team compiled the following additional benefits of real time monitoring, whether in the raw wastewater or within the treatment plant:

- Monitoring Specific Parameters Relevant to Potable Reuse:
  - Example parameters include: TOC, turbidity, pH, and water temperature, organics, dissolved oxygen, chlorine, conductivity, flow, and water level.
  - Pathogens, once the technology is available and validated (NWRI 2020).
- Stringent Control on Discharges: Real-time monitoring induces a deterrent effect, which is expected to limit discharge at the source. Limited discharge benefits a potable reuse program and benefits effluent discharge to receiving water bodies. For example, many potable reuse projects utilize reverse osmosis, which concentrates pollutants. Industrial pollutants that pass through biological treatment, of which there are many, end up concentrated in reverse osmosis concentrate.
- Industrial Discharge Source Tracking: Smart systems can help utilities understand the spatiotemporal variations of contaminants that enter the source water for potable reuse and use that information to reduce or eliminate some industrial discharges should they be problematic.

Past deployments of sensor systems have seen a share of challenges, which has been well documented in Steinle-Darling et al. 2020. Highlights from Steinle-Darling et al. 2020 are listed below and explored in more detail as part of Section 2.4.

Implementing collection system monitoring for potable reuse applications was tested in close partnership with three utilities: Ventura Water in California, El Paso Water in Texas, and Clean Water Services in Oregon. A general assessment was made of the value and readiness of the sensor-based approach to real-time monitoring and decision-making for potable reuse. Lessons learned in these three experiences follow:

- Desirable monitoring locations within the sewershed (e.g., manholes, force mains) can be challenging in practice.

- Limited number of sensor technologies are suitable for monitoring wastewater at the source as raw wastewater presents fouling, corrosion, and ragging challenges.
- Low sensor accuracy leading to low confidence in performance (e.g., pH drifts).
- Data logging can be problematic at times, with some periods not documenting any information (as seen in Steinle-Darling et al. 2020) where wireless connections and low batteries both resulted in lost data).
- Power supply, including battery issues, in remote locations can be challenging.
- Clogged sample lines were frequently seen.
- Corrosion and degradation of components were witnessed during the short trial time period (approximately 3 months), including damage to antennas, broken pH sensors, worn-out strainers, and damaged Electrical Conductivity sensors.
- False positives and subsequent alarms were noted.
- The systems were monitored and controlled based upon wireless communication, and thus present data security and shareability issues.
- Evolution in communication from mobile and LTE networks to faster networks such as 5G for better data transfer capability.
- The maintenance needed on sensor systems within the collection system was substantial, with ragging and grease requiring field maintenance of the sensors several times a week.

Looking to the future, SENG7R16 (Kadiyala and Macintosh 2018) provides several opportunities to enhance real time monitoring, including:

- Implementation of the following would benefit data transfer and management:
  - 5G (for faster data communication).
  - AutoML (i.e., automation of the entire pipeline from raw dataset to deployment of machine learning).
  - Apache Sparke (a unified analytics engine for big data processing).
- Deployment of artificial intelligence for security against Trojans (disguised malware).
- Implementation of quantum computing technology to process Big Data.

Table 2-1. Available Sensor Systems (Non-Exhaustive).

Available Providers	Sensors Provided	Parameters Measured	Level of Maturity	Website
		Physical and Inorganic Parameters		
Aqua COLOR Sensors	Aqua COLOR Algae Sensor	Temperature, Oxygen	R&D	https://www.aquacolorsensors.nl/
Systea S.p.A	Easychem Online	Comprehensive physico-chemical parameters, COD, color, phosphate, nitrite, nitrate, TN, hardness.	R&D	http://www.systea.it/index.php?lang=en
Eco Detection	HTG-1535	Comprehensive Water Quality Monitoring, Flow and Heavy metals monitoring (in development).	R&D	https://www.ecodetection.com/
SouthWest Sensor	DropletSens	Nitrate, Phosphate and ammonia (in development)	R&D	https://southwestsensor.co.uk/
TelLab	Aquamonitrix	Nitrite, Nitrate	Pilot.	https://tellab.ie/product-development/
TriOS	eCHEM	Comprehensive physico-chemical parameters, humic acids, HS, Br, HCO_3^-, Cl, NH_2Cl	Pilot	https://www.trios.de/en/sensors.html
Hach	Individual Sensors	FCl, TCl, Turb., pH, ORP, EC, DO, TOC, SS, F, NH_4^+, NO_3^-, PO_4^{2-}, UVA_{254}, others	Commercial	https://www.hach.com/
Analytical Technologies, Inc. (Badger)	Individual Sensors	Chlorine, pH, conductivity, others	Commercial	https://www.analyticaltechnology.com/
s::can (Badger)	Individual sensors	Total and Free Chlorine, pH, conductivity, ammonia, nitrate, nitrite, others	Commercial	https://www.s-can.at/
Partech	Watertech, Waterwatch series.	ORP, pH, turb., SS, color, EC, UVA_{254}.	Commercial	https://www.partech.co.uk/
Rosemount	Individual Sensors	Pressure, Flow, Level, Temperature,	Commercial	https://www.emerson.com/en-us
Capital Controls	Individual Sensors	pH, ORP, conductivity, Cl, ClO_2.	Commercial	https://denora.com/
Chemtrac	PC Series	Particle Counter	Commercial	https://www.chemtrac.com
Electro-Chemical Devices	Model S80 Intelligent Sensors	pH, ORP, pION, DO, Conductivity, Resistivity	Commercial	https://ecdi.com/portfolio-entries/s80-intelligent-sensors/
Yosemite Technologies	Y532-A digital probe	pH	Commercial	http://en.yosemitech.com/aspcms/product/2020-6-15/154.html
Detection Services	kando	Comprehensive Physico – Chemical Parameters	Commercial	https://www.detectionservices.com.au/technologies/kando/
		Organic Parameters		
Multisensor Systems Ltd.	MS2000 – THM	Total Trihalomethane	Pilot	https://www.multisensorsystems.com/ms-company/
Real Tech	UV254 Online Analyzer	TOC, DOC, Color	Commercial	https://realtechwater.com/
Hach	UVAS Plus	UVA_{254}, TOC	Commercial	https://www.hach.com
Sievers	Individual Sensors	TOC	Commercial	https://www.suezwatertechnologies.com/products/sievers-analyzers-and-instruments
		Biological Parameters		
iBioscan	7000RMS	Real-time Microbial detection, turbidity.	R&D	www.ibioscan.com
Biological Monitoring, Inc.	BioSensor	Toxicity (Biomonitoring)	Commercial	www.biomon.com
Yokogawa Electric Corporation	RAPID	RNA Pathogen Monitoring	R&D	www.yokogawa.com
		Integrated Systems (Multiple Parameter Categories)		
D2K Information Systems	QualitEye	Comprehensive water quality parameters, THMs, Organics, Total Chlorine.	Pilot	https://www.d2kinformation.com/qualiteye/
Ecosen Solutions	Water Lab	Comprehensive Water Quality Parameters, Biological pigments and Particles.	Pilot	https://ecosensolutions.com/products/
Libelium	Waspmote Smart Water platforms	Extensive Water Quality Monitoring	Pilot	https://www.libelium.com/
Proteus	Proteus P	Comprehensive Water Quality Parameters, pH, Temperature.	Commercial	https://www.proteus-instruments.com/
Blue I Technologies	SMART ONE	Comprehensive Water Quality parameters	Commercial (China Only)	https://blueitechnologies.com/product/smart-one-2/
ChemScan	ChemScan Sensor family	Comprehensive water quality parameters, Blue Green Algae, Fluorescein Water Tracer.	Commercial	http://www.chemscan.com/
Krohne	OPTISENS series	pH, ORP, COD, turb., EC, FCl, ClO_2, O_3, DO, SS, Sludge Blanket, Biofilm	Commercial	https://krohne.com/en/products/flow-measurement/
s::can (Badger)	Spectro::lyser V3	COD, BOD, Turb., SS, color, Total Hydrocarbons, UV Fingerprint, Nitrite, Nitrate	Commercial	https://www.s-can.at/
s::can (Badger)	i::scan	BOD, COD, Color, UV254, Turbidity	Commercial	https://www.s-can.at/
Tethys Instruments	UV and EL series.	H_2S, Color, pH, ORP, NO_3, NH_4, turb., PO4, UV254, DOC, EC, Chl A, hydrocarbons, DO, Phenol	Commercial	http://www.tethys-instruments.com/
YSI (Xylem)	Individual sensors	NH4, PhC, PhE, Cl, Chl A, DO, FCl, NO_3, hydrocarbons, ORP, PAR, pH, EC, Rhodamine, turb	Commercial	https://www.ysi.com/
ZAPS Technologies, Inc.	ZAPS LiquID	Comprehensive Water Quality Parameters	Commercial	https://www.environmental-expert.com/companies/zaps-technologies-inc-47664/
Turner Designs	Multiple Types of Fluorometers	Chlorophyll, fluorescent dye tracing, blue-green algae (phycocyanin and phycoerythrin), crude oil, refined fuels, tryptophan, CDOM, optical brighteners, turbidity and pCO2	Commercial	https://www.turnerdesigns.com/fluorometers-and-sensors

2.3 Coordination with WRF 4797

As part of the WRF Project 4797 (Thompson *in process*), specialized online water quality monitoring will be conducted at HRSD. This demonstration project testing is scheduled for a six-month period beginning in July 2022. So, results were not fully available at the time of submission of this report. Nevertheless, the sensors and system problems investigated are described in the table below (Table 2-2).

Table 2-2. Sensor Systems Tested at HRSD as Part of WRF 4797.

Sensor Type	System Problem	Desired Outcome
s::can Spectro::lyzer	The WWTP that provides secondary treatment for the Sustainable Water Initiative For Tomorrow (SWIFT) Research Center is experiencing periodic spikes in organic loading that impacts the final water quality for groundwater injection.	The s::can spectro::lyzer will be used to conduct pollution load tracing upstream of the WTP. Using the 256 wavelengths from the spectral array, it may be possible to identify the chemical classification which would aid in the investigation of the illegal discharger. The team has been effectively transferring the s::can spectral data as a 256-point array every 2 minutes into either a PostgreSQL or Mongo BD in the clients on-premises or cloud servers for over a decade. The spectral array data is then processed into a 3-D image for a 24-hour period Additionally, the s::can con::cube provides real-time spectral alarms using ana::tool. s::can supports clients in identifying specific compound from the raw data. If a specific compound is a recurring problem, the con::cube can be set-up to look for that specific compound and report it every 2 minutes along with the other parameters.
s::can conductivity and bromide	One portion of the wastewater collection system is adjacent to the Chesapeake Bay and is suspected to be impacted by sea water intrusion associated with high tides.	The s::can sensors will be used to validate the source of the high salinity water so HRSD can evaluate the magnitude of the problem and corrective procedures. While conductivity sensors have been used to find sources such as seawater infiltration for many years, the ability to track and communicate the changes real-time has been the value added. Also, one of the goals is to be able correlate the increased salinity and bromide with tidal surges and other potential discharges, such as boat bilge dumping. Understanding the root cause for the water quality issues are required to develop a practical solution.

2.4 Experience with Sensors from Project Partners

The following reports have been used to detail experiences of Project Partners with Sensors.

1. Steinle-Darling et al. 2020
2. s::can experiences.

2.4.1 WRF–17-30/ DRPT- 4908

WRF–17-30/ DRPT- 4908 (Steinle-Darling et al. 2020) was detailed research on implementation of Water Quality sensors for monitoring Potable Reuse system. This project evaluated deployment of pilot sensor networks with three utilities: Ventura Water in California, El Paso Water in Texas, and Clean Water Services in Oregon. The experiences with each utility are discussed in this section.

Pilot demonstration partnership with utilities:
Both s::can and Kando sensors were chosen candidates for the pilot demonstration of real time monitoring of water for Potable Reuse with the three partner utilities. Although s::can sensors were suitable for the purpose on paper, the staff from Ventura or El Paso were unable to choose optimal sites to install the sensors. Therefore, Kando sensors were chosen for this evaluation.

Sensors Used:
The sensors equipped to the Kando units trialed in Steinle-Darling et al. 2020 included ORP, temperature, pH and conductivity. The sensor used for the Steinle-Darling et al. demonstration are shown in the figure below (Figure 2-1). The monitoring station consisted of two sensor probes, a data logger with antenna for data transmission via the cellular network, and an automated sampler. Data is recorded in the loggers and transmitted to the cloud for storage and analysis. Kando provides units which are intended to be networked at multiple strategic monitoring points throughout the wastewater collection system. Using proprietary algorithms, the relative change of the sensors at a single location relative to others across the collection system can be used to infer the location and severity of pollution events, which are normalized to a "pollutant index."

Figure 2-1. Kando Monitoring Equipment.
(with automated sample pump, tubing, reservoir, two sensor probes, a sensor holder, and a data logger)
Source: Steinle-Darling et al. 2020

Summary of Field Deployments:

- **Ventura Water:** Three online sensor stations were installed during the week of October 8, 2018, with automated samplers installed in the first week of January 2019.
- **El Paso Water:** Four sensors were installed on February 18, 2019. This is illustrated in the figure below (Figure 2-2).
- **Clean Water Solutions:** Five sensor stations were installed, and data transmitted from March to October 2019.

Figure 2-2. Placement of Sensors And Sample Collection at El Paso Site.
Source: Steinle-Darling et al. 2020

Sensor Evaluation Approach:

- After a few weeks of initial monitoring, baselines were established for the monitored parameters.
- Two thresholds were set for each measured parameter:

    - Automated collection of samples.
    - System alert for extreme “Pollution events”.
- These deviations from baseline values, if detected, can be correlated with illegal discharges upstream of the collection system. Once triggered the sampling volume drawn is approximately three liters. This allowed the utilities to track upstream pollution events.

The goal of these demonstrations was to evaluate sensor performance based on a set of criteria. The summary of results has been discussed below:

- Sensor Response:
    - Kando sensors have not been consistent in detecting trends of parameters such as EC, pH, oxidation reduction potential (ORP), and temperature.
- Sensor Accuracy:
    - Kando sensors were not necessarily accurate compared to readings from calibrated field instruments.
- Identification of Pollutant Spikes:
    - Sensors were successful at detecting anomalies when kept free of foulants [ragging and FOG].
- Direct Monitoring of Known Discharges:
    - The sensors were able to document several compliant and challenging dischargers at all three sites.
- Deterrent effect:
    - In case of the CWS demonstration, dischargers were notified of the monitoring and thereafter they adhered to the discharge limits. Hence, a deterrence effect was observed.
- Data Management and Software Usability:
    - Kando offers an intuitive map – based dashboard. This was a very useful tool to visualize the collected data.
- Ease of Maintenance:
    - Sensor and hardware failures were common and sensor maintenance was a challenge in all three sites.
- Physical Limitations:
    - Flow was intermittent in several locations. Selecting suitable locations for installation was also a challenge.

Overall summary of the experiences:

- Benefits:
    - The study was able to confirm the ability of a commercially available monitoring platforms to provide 24/7 continued monitoring.
    - Kando’s sensors were able to collect triggered samples but at the time of Steinle-Darling et al. were not necessarily available for full commercial implementation of the system. Early installations did suffer from challenges due to hardware failures and challenges with sensor fouling. Since the pilot described in Steinle-Darling et al., Kando has made improvements to the user interface and system operability..

- Challenges:
    - Of the systems considered, only Kando sensors were suitable for upstream installation. A wider choice of technologies was not available to the partner utilities and the research team.
    - Once installed, the following maintenance issues began to show up in the sensor networks:
        - FOG and moisture buildup.
        - Sample line clogging.
        - Battery issues.
    - Some hardware related challenges (see Figures 2-3 and 2-4) that affected sensor performance include:
        - False positive alarms.
        - No data collections by EC sensors at times.

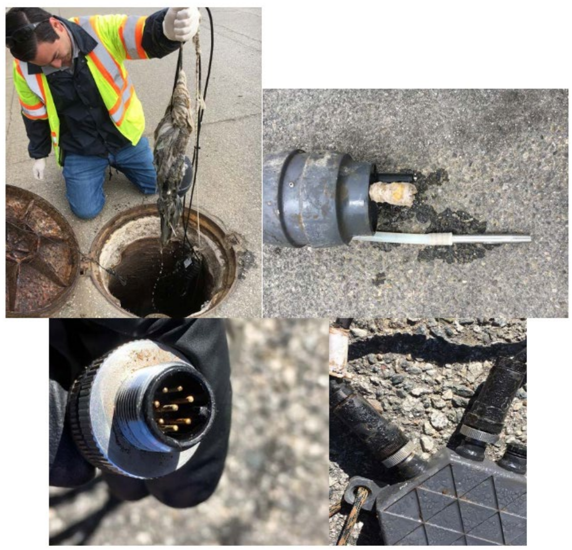


Figure 2-3. Illustration of Hardware Challenges Experienced by Ventura Staff.
(ragging (top photos) and corrosion due to moisture entering cable connections (bottom photos))
Source: Steinle-Darling et al. 2020

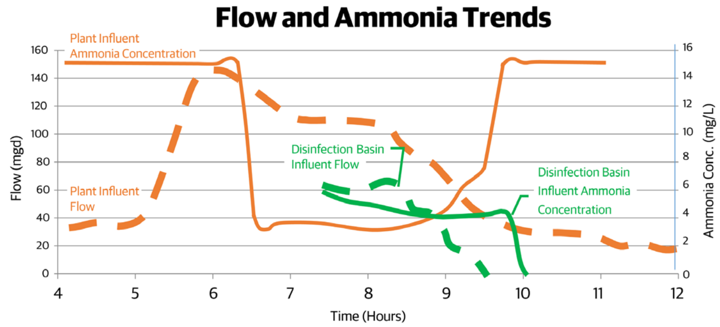
Figure 2-4. Hardware Challenges Encountered by El Paso Water.
(including corrosion on antenna connector (top left), broken pH sensor (top right), a worn strainer (bottom left), and a worn EC sensor (bottom right))
Source: Steinle-Darling et al. 2020

A summary of sensor experiences from Steinle-Darling et al. is in Chapter 5. A more detailed review of sensor experiences from Houston, St. Louis, Omaha, Irvine Ranch Water District, and the City of Phoenix is provided in Table 2-3 and following discussion below.

Table 2-3. Summary of Select Utility Sensor Programs.

Agency	Sensor Types	Application	Maintenance Problems
Houston	**Collection System**		
	Kando System being installed	Influent tracking	No history
St. Louis	No online water quality monitoring in the collection system		
	WWTPs (Collecting data from 4 facilities)		
Irvine Ranch Water District	No online water quality monitoring in the collection		
	The Michelson Water Reclamation Plant uses sensors for tracking influent water quality and optimizing aeration for secondary treatment.		
	pH	Influent Tracking	Daily fouling until floating sensor boat installed
	Conductivity	Influent Tracking	Daily fouling until floating sensor boat installed
	DO	Secondary Treatment	No unusual maintenance experienced
Omaha	No online water quality monitoring in the collection system		
	Missouri River WWTP (New facility)		
	The Missouri River WWTP has large diurnal fluctuations in the influent ammonia that impacts the effectiveness of chlorination. To gauge effectiveness of the chlorination practice, the operations uses influent ammonia levels to control the chlorine feed to prevent the system from transitioning into breakpoint. The use of ammonia sensors for tracking ammonia levels in the influent provides the information for adjusting the free chlorine feed to reduce the ammonia levels in the discharge for compliance.		

Agency	Sensor Types	Application	Maintenance Problems
	Without the influent ammonia level variation, the operators could easily overfeed or underfeed free chlorine, either of which could lead to non-compliance in the effluent.		



Agency	Sensor Types	Application	Maintenance Problems
	Free Chlorine (Capital Controls)	Effluent Disinfection	No unusual maintenance experienced
	Total Chlorine (Capital Controls)	Effluent Disinfection	No unusual maintenance experienced
	pH	Understanding breakpoint curve location	No unusual maintenance experienced
	Bisulfite (ATI)	Dechlorination	No unusual maintenance experienced
	ORP (Hach)	Back-up for Chlorine Analyzer	No unusual maintenance experienced
	Ammonia	Influent ammonia levels are used to determine when the plant produces free chlorine or chloramines	No unusual maintenance experienced

2.4.2 s::can Installation Experiences at City of Phoenix

Summary of Field Deployment:

The City of Phoenix installed s::can spectro::lyser and ammo::lyser (Figure 2-5) in their aeration basins and at the WWTP influent to measure the following parameters: nitrogen dioxide (NO_2), nitrate (NO_3), NH_4, COD (Figure 2-6), soluble COD, and total suspended solids (TSS). The objective of the installations was to monitor and detect illegal COD load specific to industrial discharge. The s::can spectro::lyzer on the WWTP influent was configured to produce alarms based on abnormal values in the ultraviolet (UV)-Visual spectra data (Figure 2-7).

Figure 2-5. s::can ammo::lyser Installed at the Aeration Basin Site (City of Phoenix).

Sensor Evaluation Approach:

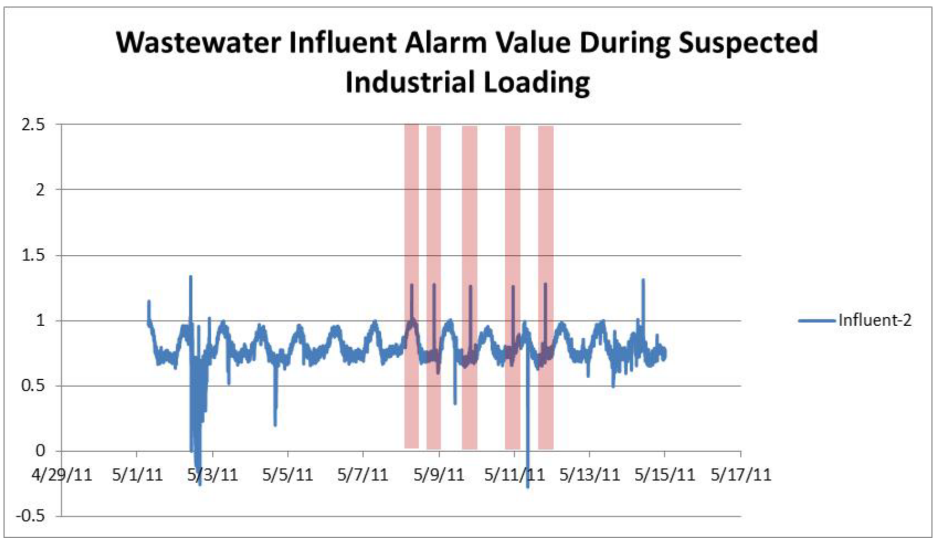


Figure 2-6. The Red Highlighted Parts of the Graph Denote the Spikes (Spectral Alarms) in COD Concentration at the WWTP Influent.

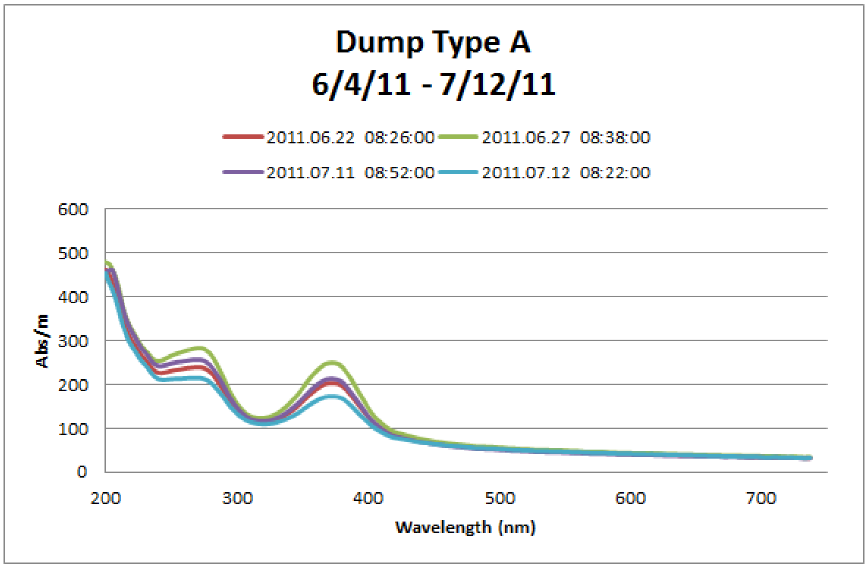


Figure 2-7. Multiple Individual Spectra Detected Which Implies Multiple Industrial Discharges into the System.
Dump Type A refers to a spill the utility had observed, later identified as a heavy metal.

- The COD concentrations are measured at the wavelength range of 250-400 nanometers (nm) of the UV-Visual spectra. S::can reviewed the spectral data collected at the WWTP influent after reports of a possible industrial discharge. Apart from the expected normal diurnal fluctuations, abnormal peaks were identified for short durations over a two-month period.
- A closer investigation revealed that abnormal UV-Visual spectral peaks were visible in the raw spectral data during which spectral alarms were also raised by the sensors. The alarms also occurred during the same time of the day on multiple days over a two-month period.
- An added feature of spectro::lyzer is the Spectral Fingerprint alarm which is helpful in investigating individual spectra. Multiple abnormal spectrums were identified suggesting multiple industrial dischargers.

Overall Summary of Experiences:

- s::can spectro::lysers have been able to detect and differentiate between different industrial discharges over the operation period of 2 years.
- Ammo::lyser installed in the Aeration Basin has been able to detect changes in NO_2/NO_3 following detected industrial discharges.
- This information has been useful to the City in tracking down and issuing fines to illegal dischargers. The City of Phoenix now has plans to install an extended network of sensors in their collection system.

CHAPTER 3

New Bench and Full-Scale Field Research

The goal of this chapter was to provide a focused evaluation of sensor technologies by deploying them at various locations in four wastewater treatment facilities and their sewersheds, and analyzing the data. For these efforts, sensor systems were provided from multiple technology suppliers (Table 3-1). These efforts looked to define and understand accuracy, precision, drift, and fouling aspects of a select group of sensors. Another goal of the field research was to understand if and how probe containment devices can be employed to protect the sensors for issues related to FOG and ragging when the probes are installed in the collection system.

Generally speaking, the results presented below indicate:

- pH and EC probes, under the right circumstances, can be reliably used within the collection system and at the influent to a WWTP to measure water quality in real time.
- Spectrometer based probes that measure multiple parameters (e.g., COD, biological oxygen demand [BOD], UV254) also demonstrated direct value in measuring pollutant variation and spike events, and if placed in a screened or primary effluent can be reasonably accurate and reliable.
- ORP probes did not function well.
- The probe location and the probe operations and maintenance (O&M) program, including staff consistency, matters.
    - Cleaning of spectrometry-based systems may be needed weekly, several times per week, or daily, depending upon the location of the sensor and site-specific water quality. Some sensors with self-cleaning systems might still reliably detect large water quality events with manual cleaning less than weekly.
    - Changing of O&M staff impacts results, as cleaning approaches will differ. Having a detailed and repeatable standard operating procedure for sensor maintenance and calibration is essential.
- Pilot testing of any sensor system should be performed prior to purchase and installation. That pilot testing should define the accuracy and long-term O&M of any sensor system for the specific tested location and water quality.
- In total, a combination of low maintenance probes within the collection system (e.g., pH, EC) coupled with a spectrometer-based probe in screened or primary effluent can effectively monitor a broad range of raw or partially treated wastewater qualities and provide an early warning system to water quality for a potable reuse system.

Table 3-1. Overview of Sensors Tested at Three Utilities for Bench and Full-Scale Field Research in Chapter 3. ✔ indicates generally successful trial. X indicates tested but did not meet the utility's criteria for continued use for real-time monitoring. *Successful when installed within CWS's rag guard sensor holder.

Brand	Sensor	CWS	Morro Bay	LACSD
Yosemitech	Y532-A	✔*		
ECD	ORP Pt Cap peek, two-tang probe	X		
ECD	Extended Life pH Electrode RADEL body	✔*		
s::can	Spectro::lyser	✔	✔	
s::can	condu::lyser	✔	✔	
s::can	pH::lyser	✔	✔	
Real Tech	Titanium Ba-X Series SA2010 multi-wave sensor			X
Sentry	Sentry-AD			X

3.1 Clean Water Services Field Research

3.1.1 Introduction

This chapter included two research tasks:

1. Conduct an evaluation of continuous sensors in a continuous-flow flume environment.
2. Develop a sensor containment device to minimize ragging and test the device in the collection system environments to determine its effectiveness.

This section summarizes the methods and findings of these experiments. A more complete report on these experiments can be found in Appendix A.

3.1.2 Methods

A Plexiglass flume was constructed near the headworks of the Forest Grove WWRF to perform controlled experiments on the sensors (Figure 3-1). The influent flow was diverted to the flume immediately after passing through the bar screens, grit removal, and a wet well. The flume was 12 inches by 12 inches by 72 inches and received 150 gallons per minute (gpm) of flow. Six different sensors were installed in the flume including three pH probes (ECD, Yosemitech, and s::can), one ORP probe (ECD), one conductivity probe (s::can), and one spectrometer (s::can) that detected COD, BOD, TSS, nitrate, and ultraviolet absorbance (UVA) at 254 nm wavelength (UVA_{254}). Sensors were maintained, calibrated, and installed using manufacturer recommendations.

Figure 3-1. Flume in Operation During Experiments.

Ten experiments were run using the flume (Table 3-2). Three different velocities and depths were used to determine the effects these have on sensor performance with two replicates each (Experiments 1-6). Three fouling tests were also conducted where the probes were exposed to increased FOG and/or rags to observe the effects of fouling on their performance and their ability to recover from a fouling event (Experiments 7a-7c). Finally, a long-term experiment was conducted with an increased duration to test the probe performance over a longer period. Each experiment began with cleaning the probes and calibrating (if necessary) followed by setting up the conditions (velocity, depth, fouling) for that experiment. An initial "spike test" was then performed where pH, conductivity, BOD, COD, and ORP were increased in the flow using a spiking solution of sodium hydroxide, table salt, humic acid, and hypochlorite. The spike test typically lasted 20 minutes. The sensors were then operated for a certain duration (typically 7 days) while handheld readings and samples for laboratory analysis were collected every other day for comparison with the sensor readings. Finally, a final spike test was performed similar to the one at the beginning of the experiment. Some experiments also included an additional spike test in the middle of the duration.

Table 3-2. Summary of Experiments Conducted Using the Flume and Sensors.

Experiment Number	Velocity (ft/s)	Depth (inches)	Spikes	Duration (days)	FOG/rag introduction
1	1.3	3	2 (start/end)	7	Normal
2	0.8	5	2 (start/end)	7	Normal
3	0.6	7	2 (start/end)	7	Normal
4	1.3	3	2 (start/end)	7	Normal
5	0.8	5	2 (start/end)	7	Normal
6	0.6	7	2 (start/end)	7	Normal
7a	0.6	7	3 (start/mid/end)	2	Grease dipped
7b	0.6	7	3 (start/mid/end)	2	Rag wrapped
7c	0.6	7	3 (start/mid/end)	2	Increased FOG
8	0.6	7	3 (start/7-day/14-day)	24	Normal

CWS developed a sensor holder called the 'rag guard' designed to decrease fouling of the sensors in the sanitary collection system (Figure 3-2). The rag guard is a section of 4-inch diameter polyvinyl chloride (PVC) pipe with a 3-foot radius 90 degree bend with slats cut into the final 1-foot of the pipe that get progressively wider at the submerged end of the tube (Figure 3-2). The rag guard is suspended by stainless steel cables so that the slats are parallel with the flow, and the sensors are installed to be inside the slatted area. For this experiment, two identical ECD pH probes were installed in a manhole in the collection system. One of the sensors was placed inside the rag guard, and the other was placed in a typical sensor holder as designed by the manufacturer's representative in previous deployments. The performance of the two sensors was compared over the combined duration of the flume experiments.

Figure 3-2. Rag Guard Sensor Holder Showing the Entire Device.

3.1.3 Results and Discussion

Experiments 1 through 6 showed that changes in velocity and depth did not affect sensor performance. This was surprising as general experience has shown that sensors placed in manholes with a higher velocity tend to build up less FOG. However, the range of velocities available in the flume while maintaining a minimum 3-inch depth to keep the sensors wet was small, with a maximum of 1.3 feet per second (ft/s). Higher velocities than measured are certainly present in the collection system, and it is possible that such higher velocities may reduce fouling.

The largest factor affecting sensor performance was the sensor type itself. More specifically, those sensors that were more resistant to fouling performed better than those more susceptible to fouling. While all sensors detected the spikes well and matched the lab, handheld, or SCADA values well when they were clean, they differed greatly in their ability to resist fouling and their response to being fouled during the course of each experiment. The s::can spectro::lyser had similar results for each of the parameters it measured (COD, BOD, TSS, nitrate, and UVA_{254}), so only the results for COD are summarized here. The sensor detected the diurnal pattern of COD consistently, detected real spikes that occurred, had a consistent baseline without drifting, and consistently detected all spiking events during each experiment. Figure 3-3 shows the COD measured by the spectro::lyser at the end of experiment 6 which had the lowest velocities. After 7 days, it detected the spike event very accurately. Even in experiments 7a-7c, the tests with intentionally increased foulants, the spectro::lyser recovered from the deliberate fouling within minutes and accurately detected the spiking events. Overall, the spectro::lyser showed that it can be used for detecting patterns in all of the optical parameters very well and was very resistant to fouling. The only exception was during experiments 1 and 8, when an unknown event occurred in the influent that affected the lens. The coating on the lens caused the baseline to drift upwards over time. In experiment 1, cleaning the lens fully corrected it, but cleaning the lens in experiment 8 did not correct the problem. It is unknown what caused this, but other utilities have had similar experiences with iron addition. This was not found to be the case in these experiments.

While the spectro::lyser captured the patterns and temporal changes very well, laboratory COD tests on the grab samples often differed from the value reported by the spectro::lyser by up to several hundred milligrams per liter (mg/L). The sensor was calibrated to ten samples, but it is likely that additional calibration could have corrected the issue. TSS measured by the spectro::lyser agreed very well with the laboratory measurements, for example.

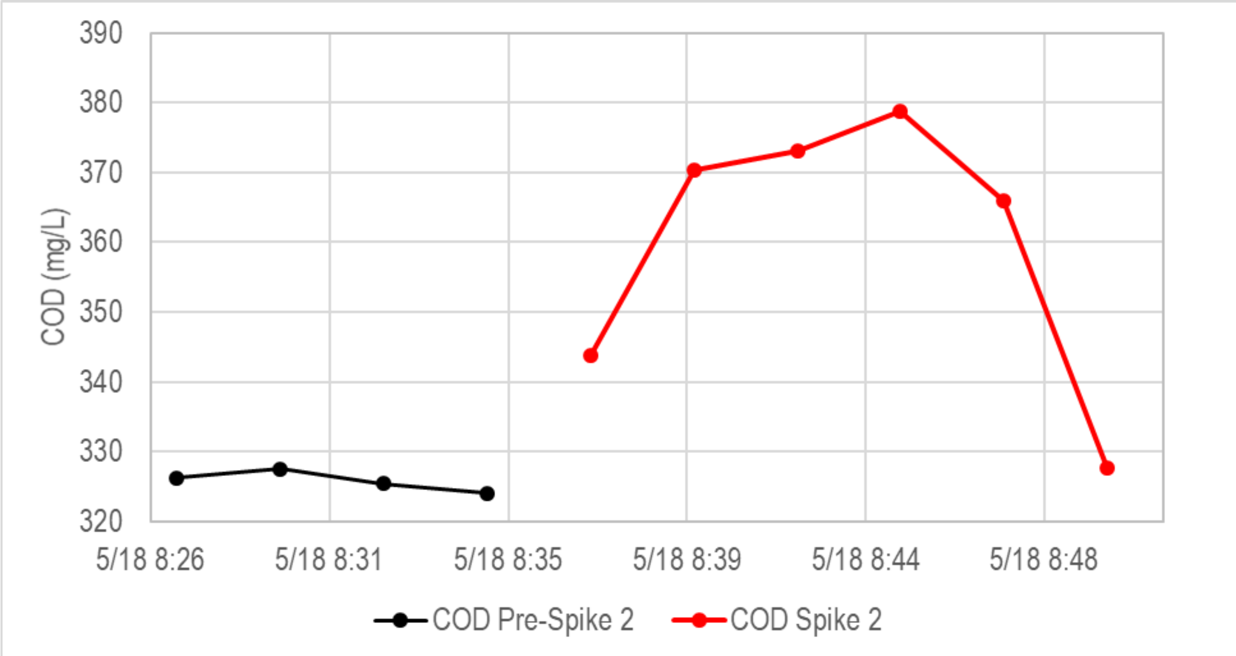


Figure 3-3. COD Measured by the s::can Spectro::Lyser Before and During A Spike Test In Experiment 6 After 7 Days In Influent Without Cleaning.

The s::can condu::lyser consistently detected discharge events, matched handheld values, had a stable baseline and detected all spiking events in all experiments. It was not affected by the deliberate fouling in experiments 7 or even the events in experiments 1 and 8 that caused the spectro::lyser to drift. It is well suited for detecting conductivity accurately and precisely in wastewater with minimal maintenance.

The three pH sensors in the flume performed differently in these experiments. While all three detected the initial spikes and matched the temporal pattern measured by the influent pH probe when they were clean, the s::can pH::lyser was much more consistent in its ability to detect spikes at the end of experiments and match the SCADA values over time than the other two sensors (the influent pH sensor where the SCADA came from was cleaned and maintained daily by treatment plant staff). Figure 3-4 shows the pH values reported during experiment 2 compared to the SCADA pH sensor in the influent which was cleaned daily. The observations during this experiment were typical of the other experiments, as well. The s::can pH sensor matched the SCADA pH accurately and precisely, detected the real pH events that occurred on April 6th through 8th, and detected all spike tests. The ECD sensor detected the initial spike, but fouled and drifted after 1-3 days depending on the experiment and typically did not detect the final spike event. It also had much higher variability than the other sensors. The Yosemitech sensor had periods where it matched well and others where it didn’t, but it was less predictable than the ECD sensor in that these periods were not consistent (i.e., it didn’t always match well for the first few days, then drift away for the rest of the experiment). It sometimes detected the final spike, but was not consistent. It also had a consistent offset of 0.5 units from the other sensors even though all were calibrated. The squared Pearson correlation coefficients with the SCADA pH were 0.84, 0.04, and 0.03 for the s::can, Yosemitech, and ECD pH sensors, respectively (all had p values $<<0.05$).

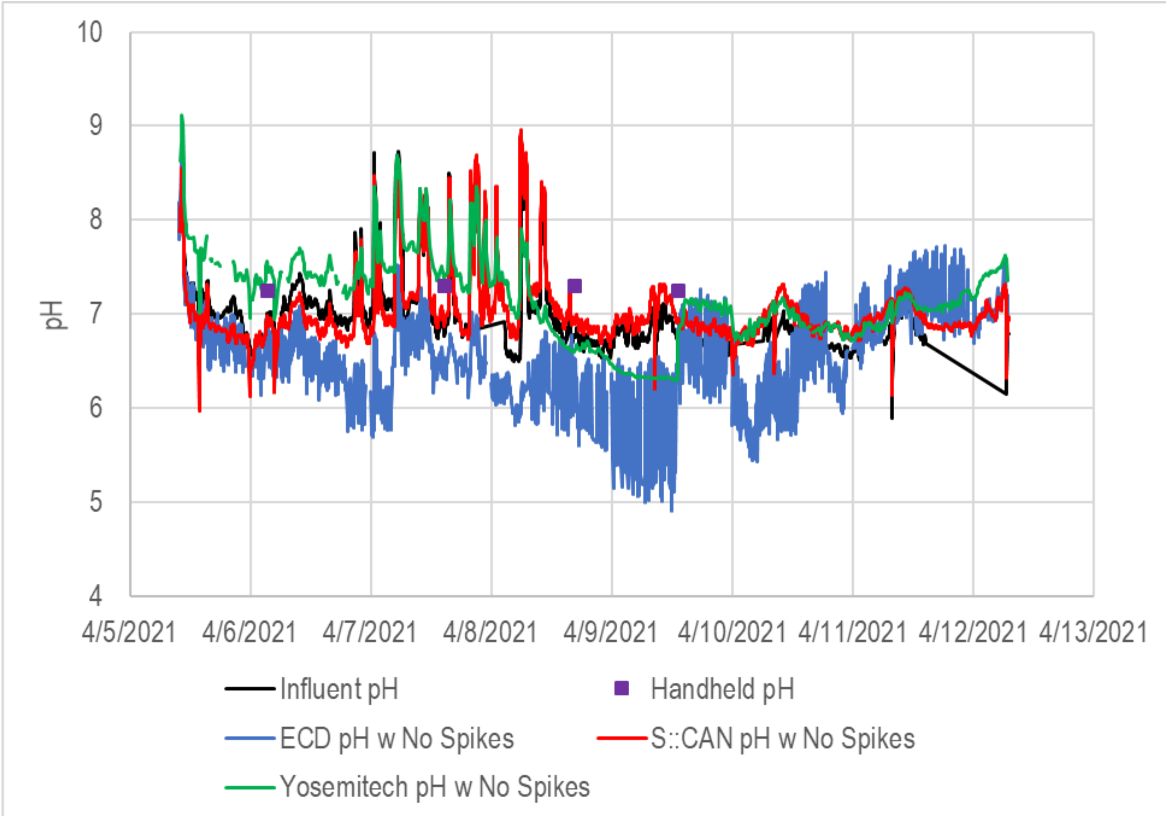


Figure 3-4. Timeseries of pH in Experiment 2 as Measured by All Three pH Sensor Types.
The Handheld Multimeter, And The Influent SCADA Ph.

Comparison of the performance of the two ECD pH sensors installed in the manhole demonstrated the effectiveness of the rag guard. Both of these were the same as the ECD pH sensor installed in the flume. Figure 3-5 shows the pH timeseries for the sensor in the rag guard and the sensor in the generic sensor holder from when it was cleaned on June 8th until it was cleaned again on July 2nd. Issues with the datalogger, hardware, and installation caused data before this period to either not be available or not be reliable. Nevertheless, the figure illustrates the differences caused by the rag guard over nearly a month-long period. In the generic sensor holder, the performance of the ECD pH sensor was similar to that in the flume, where it drifted a few days after cleaning, sometimes periodically recovering, presumably when a collected rag was dislodged. It was generally able to detect the diurnal pattern, but the measured values themselves were well below the real value, as has typically been observed for pH sensors in the CWS collection system in multiple pilot studies. In the rag guard, however, the ECD pH probe had a much more stable baseline and detected the diurnal pattern and several small events very well. There were short periods of drift such as on June 27th, but it recovered much more quickly and returned to the stable baseline. This suggests that the rag guard may enable the maintenance interval for this probe to be approximately one month, which is CWS's target interval for feasibility. As to why or how the rag guard specifically helps, it is speculated that initial contact with strings, hair, and small debris start to foul the sensors, which then leads to a micro-environment where fouling can occur more aggressively. Because there are no surfaces perpendicular to the flow and nothing to catch strings, hair, rags, etc. when the rag guard is in place, it helps to prevent rapid fouling.

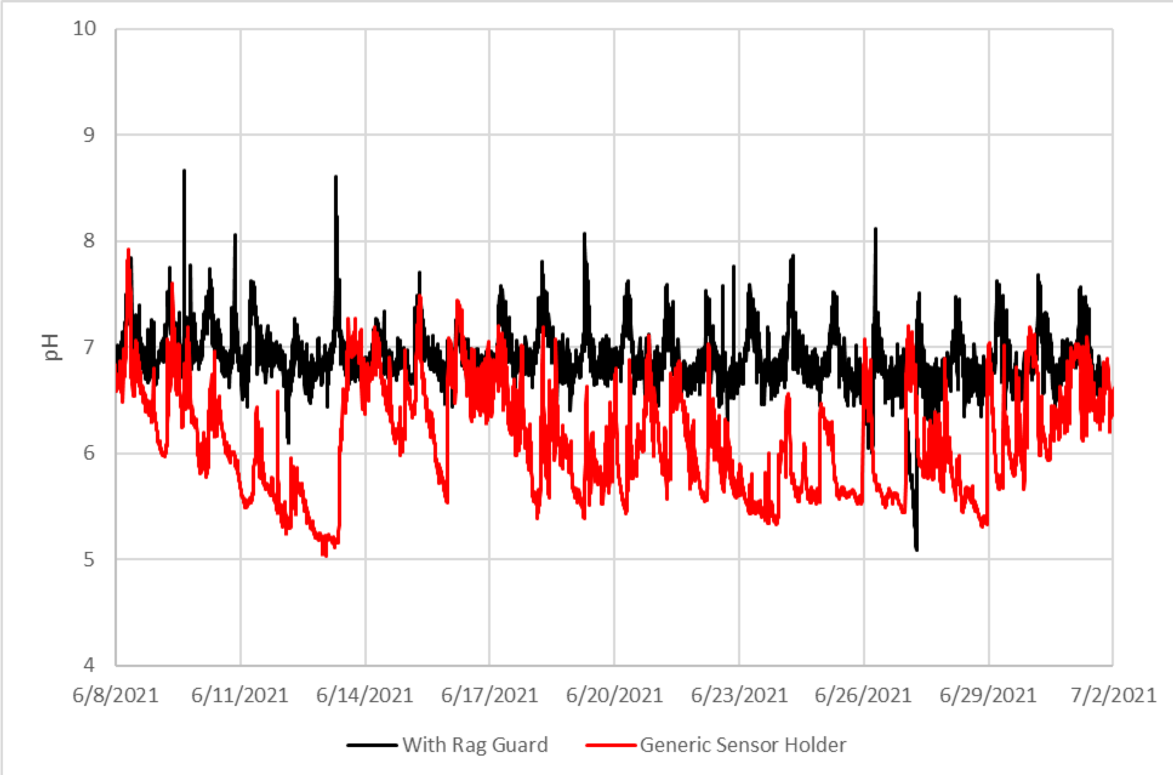


Figure 3-5. Timeseries of pH Measured by Both pH Sensors at a Test Manhole June 8 Through July 2, 2021.

Overall, the field research at CWS demonstrated that water quality sensors can be successfully deployed in sanitary sewers without an excessive maintenance burden. The main factors affecting the success of the deployments were the ability to limit fouling and the type of sensor used. Fouling was shown to be limited by air-blast self-cleaning in the case of the s::can sensors deployed in the flume and by the rag guard deployed at a manhole in the collection system. Resistance to the effects of fouling was also a function of the type of technology used by the sensor (see Appendix A). Velocity and depth were not shown to be major variables in the success of the different deployments, but this may be limited by the range of velocities available to test at this flume and/or the counter effects of changes in depth.

In general, the s::can sensors were successful except when events occurred that affected the lens/sensors. They were consistently able to track measured values and detect spikes after more than a week in the influent, even when impacted by FOG and rags. The combination of the self-cleaning and the technology made them very resistant to typical fouling, and caused them to outperform the other sensors. However, they are expensive, require more power, and cannot currently be deployed in manholes (though this may change in the future). These could be an excellent choice for sensors at the influent or at established monitoring stations in the collection system equipped with power and a utility box. While the spectro::lyser detected the temporal patterns in the influent consistently, it sometimes varied greatly from measured values despite calibration. This may be solvable with better calibration as shown in the case of TSS and at other cities, but it may also mean that the sensors can be used more for detecting patterns rather than replacing laboratory samples of COD and other parameters.

The ECD pH probes were not successful for more than a few days at a time except when deployed inside the rag guard. With the rag guard, the ECD pH probes performed well for a month with no maintenance. These can currently be deployed in the collection system inside manholes.

The Yosemitech probes showed promise that these inexpensive probes could perform similarly to the ECD probes in the collection system. However, they will likely also require a rag guard to avoid excessive maintenance, and their dataloggers cannot currently be deployed inside manholes (though that may change in the future).

The ORP probes did not perform well in any of the applications regardless of the methods used to prevent fouling and will not be used by CWS in their continuous monitoring program.

3.2 Morro Bay Field Research

3.2.1 Introduction

The Morro Bay WWTP is a relatively small facility, with flow ranging from around 0.2 million gallons per day (mgd) to 1.2 mgd. Morro Bay is implementing a potable reuse program with the capability to recharge the groundwater with 1 mgd of purified water. The pilot testing for Morro Bay was focused upon a less controlled experiment, where the sensor systems were deployed in the screened primary influent, calibrated, inspected daily, and left to run without spiking or challenge events.

3.2.2 Equipment and Installation

The same probe system from s::can used for CWS was also deployed to Morro Bay. A summary of the equipment and installation of the s::can equipment is as follows, and is described in more detail in the subsequent sections:

- Probes installed:
  - s::can spectro::lyser was supplied by s::can for measuring COD, BOD, TSS, UV254, bisulfide (a surrogate for hydrogen sulfide [H_2S]) and NO_3. The sensor was connected to a con::cube for data logging, telemetry, and control.
  - s::can pH::lyser was supplied by s::can for measuring pH. The sensor was connected to a con::cube for data logging, telemetry, and control.
  - s::can condu::lyser was supplied by s::can for measuring conductivity. The sensor was connected to a con::cube for data logging, telemetry, and control.
  - Various hardware for automatic cleaning.
- Human Machine Interface:
  - V3 con::cube with integrated event detection and data validation.
  - NEMA 4X, outdoor installation (Figure 3-6).
- WIFI & Remote Monitoring:
  - Terminal is equipped with external modem for remote connection.
  - The system is monitored remotely in real time and data can be downloaded for evaluation (Figure 3-7).
- Installation:
  - Various hardware for in situ installation and handrail mounting.

- Equipped with a sun shield.

- Sensors were installed at the Morro Bay WWTP before primary treatment but after the bar screens and aerated grit to minimize fouling. The probe system was installed outdoors.
- Morro Bay doses approximately 1-2 mg/L ferrous chloride at their headworks (targeting approximately 1.65 mg/L). Iron has caused lens fouling issues on s::can probes at other sites. So, the high frequency of (daily) maintenance required at this location might not be representative of the maintenance requirements at other sites. Generally, daily maintenance could be considered overly burdensome in practice, but was deemed acceptable for this short duration trial.

Figure 3-6. s::can Probe System at Morro Bay.

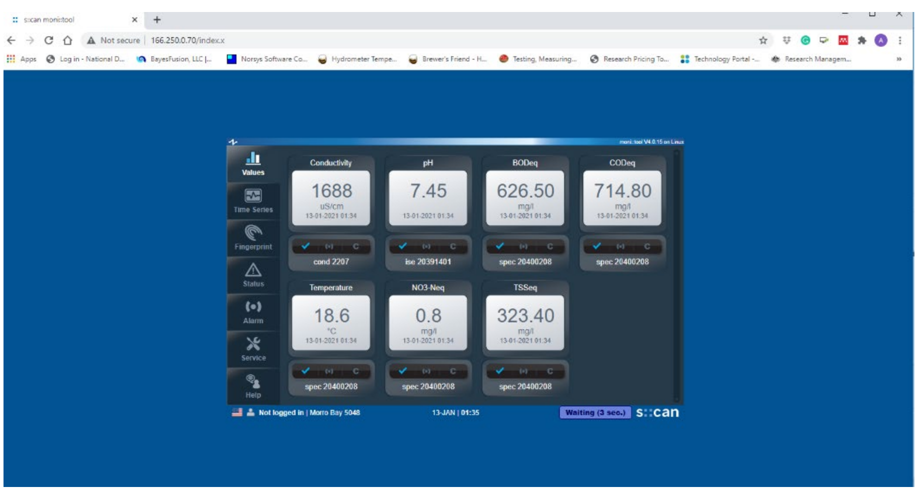


Figure 3-7. Morro Bay s::can Data Presentation on Web Browser.

The s::can system at Morro Bay underwent startup and calibration on site, with direct supervision by s::can personnel, following these procedures:

- Day 1 is for equipment installation and configuration.

- Day 2 and 3 are for testing and calibration.
  - Up to four grab samples are/were taken to verify probe accuracy for COD, BOD, TSS, NO_3, and Bisulfide. These samples were shipped to an off-site laboratory.

3.2.3 Methods

The Morro Bay installation of the s::can probes in screened primary effluent was in the influent to primary treatment, and was intended to evaluate probe accuracy, precision, and fouling in a real-world environment. The trial also enabled a comparison to the CWS laboratory flume experiments with much of the same equipment (Section 3.1).

Morro Bay experiences diurnal changes in water quality and flow. The water velocity fluctuation in this experimental approach was intentionally uncontrolled to provide a comparison with the controlled velocity flume at CWS.

Daily over 3+ months, the probes were removed from service, examined, and manually cleaned. Recordings pre- and post-cleaning were noted as well as other visual observations in the daily checklist below (Figure 3-8).

It should be noted that the s::can equipment had an automated air blast sensor cleaning system. However, unless specifically noted, the sensors were manually cleaned on a daily basis (with measurements taken, cleaning performance, and measurements retaken).

Water Research Foundation Project 5048
Morro Bay Daily Checklist

Name	
Date	
Time	

Visual Observations of Fouling	

Sensor Probe Readings			
	Cleaned Daily	**Before Cleaning**	**After Cleaning**
COD	Yes, clean daily with 3% HCl		
BOD			
TSS			
Nitrate			
H2S			
Hs- (bisulfide)			
UV254			
pH	no, clean as needed based upon data trend		
EC			
ORP			

Benchtop/Handheld Readings (minimum twice monthly)

UV254	
pH	
ORP	
EC	

Other Notes	

Figure 3-8. Morro Bay Daily Field Checklist.

Over the 3+ month period of operation, both the daily checklists and weekly online data downloads were collected and shared with the project team for evaluation and potential modification of efforts at Morro Bay. The con::cube sent data from the s::can sensors to their proprietary cloud, allowing for project team access.

Twice monthly grab sampling and calibration for COD, BOD, TSS, NO_3, UV254, pH, EC, and ORP, with all other samples were sent off site. The off-site laboratory work was performed by Abalone Coast Analytical using standard methods (Table 3-3). The online s::can data were adjusted based upon the results of these calibration results. Furthermore, the Morro Bay sampled their effluent for TSS and BOD twice weekly.

Table 3-3. Analytical Methods to be Used for Laboratory Analysis for Morro Bay.

Analysis	Method	Laboratory
Chemical Oxygen Demand (COD)	Standard Methods 5220 D	Abalone Coast Analytical
Total Suspended Solids (TSS)	Standard Methods 2540 D	Abalone Coast Analytical
Biochemical Oxygen Demand (BOD)	Standard Methods 5210 B	Abalone Coast Analytical
UV254	Standard Methods 5910B	Abalone Coast Analytical
pH	Standard Methods 4500H + B	Abalone Coast Analytical
Nitrate	EPA 300.0	Abalone Coast Analytical
Hydrogen Sulfide	Standard Methods 4500-SF	Abalone Coast Analytical

3.2.4 Probe Results

The s::can equipment at Morro Bay ran nearly continuously from mid-January 2021 and was shut down May 10, 2021.

3.2.4.1 Probe Accuracy

Grab sample data is plotted in the figures below alongside hand recorded values from the online probes over the test period (Figures 3-9 through 3-14). The first full calibration of the probes was completed on February 6, 2021. Values were hand recorded immediately after performing a manual cleaning of all but the pH and EC probes. The pH and EC probes were cleaned only one time, on February 16, 2021. Lab samples taken on and after February 6, 2021 were used to calibrate the probes, which can be seen in some of the data sets as the online meters level out and read closer to lab measured values. Notable observations include:

- Online EC values matched well with grab samples, noting that there is substantial EC variation inherent to the raw wastewater.
- Online pH values over time trended low compared to grab samples. The pH probe was only calibrated at startup of the pilot, and monthly calibration may have corrected for the sensor drift.
- COD, BOD, and TSS online values had initially wide variations in value and were well off of grab sample values (i.e., by greater than a factor of two). Calibration was completed which dramatically reduced the variability and brought the values reasonably close to grab sample values.
- UV transmittance (UVT) readings (calculated based upon UVA readings) were continuously higher than lab samples, with only one exception. The team believes that this was due to the timing of the calibration sample being below peak values seen online. More frequent calibration is needed.

- Generally, calibration of probes monthly or more frequently would be recommended for primary influent based on this dataset.
- Bisulfide and nitrate results were not informative and so are not shown below for brevity.

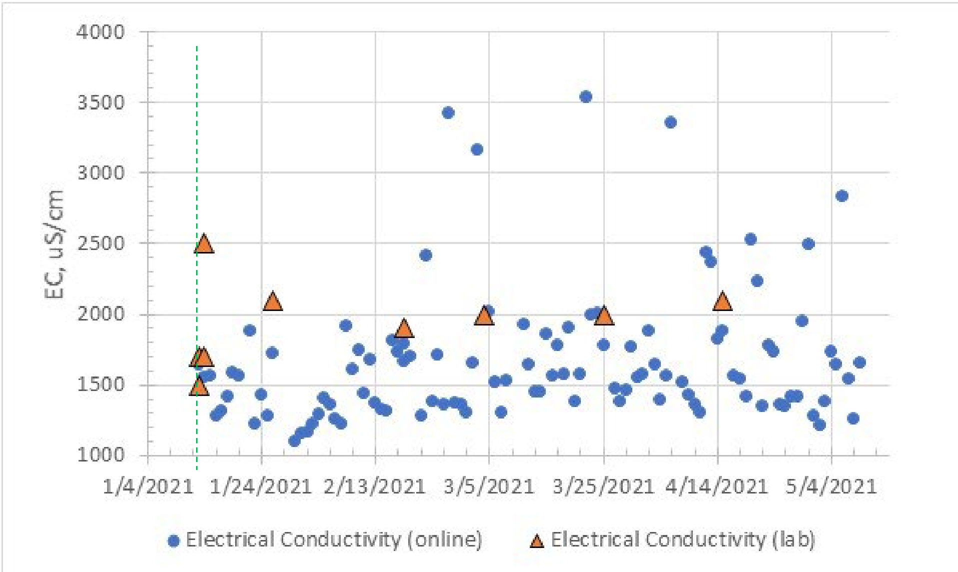


Figure 3-9. Electrical Conductivity Lab and Online Data Comparison.
Dashed green bar represents calibration.

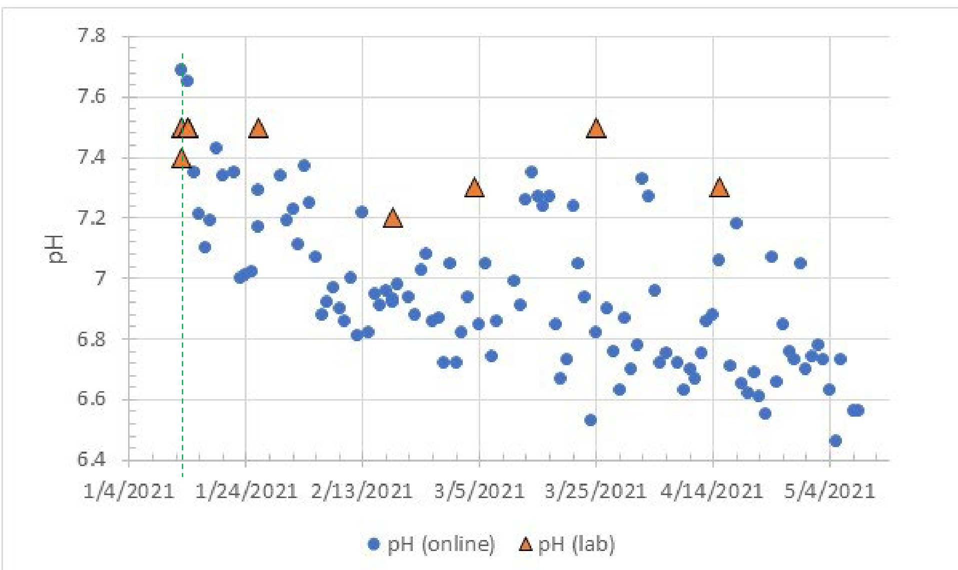


Figure 3-10. pH Lab and Online Data Comparison.
Dashed green bar represents calibration.

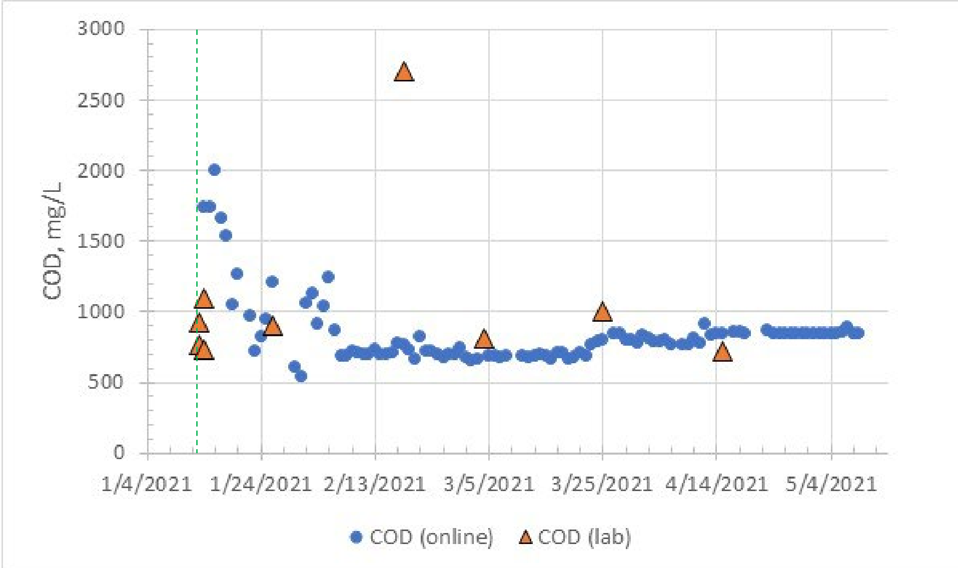


Figure 3-11. COD Lab and Online (After Manual Cleaning) Data Comparison.
Dashed green bar represents calibration.

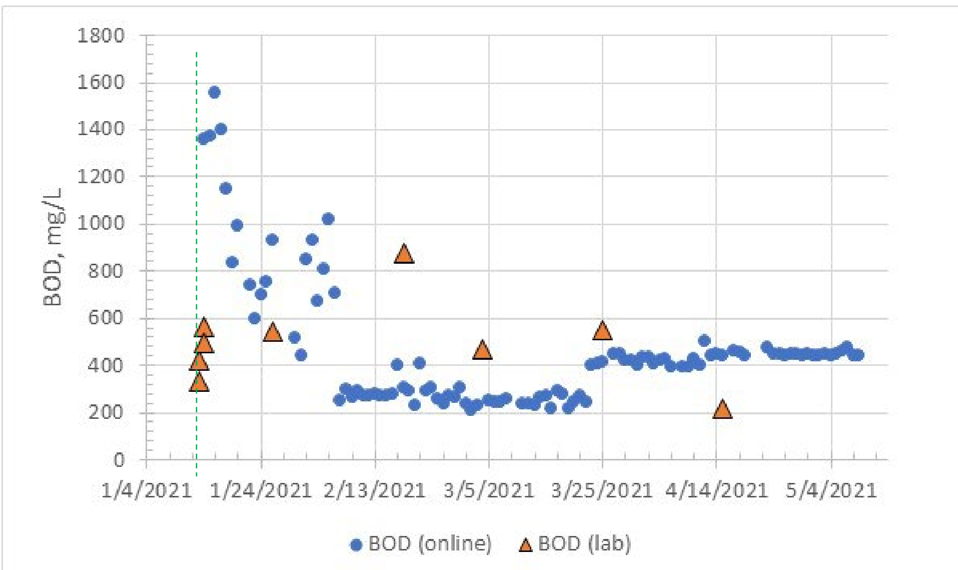


Figure 3-12. BOD Lab and Online (After Manual Cleaning) Data Comparison.
Dashed green bar represents calibration.

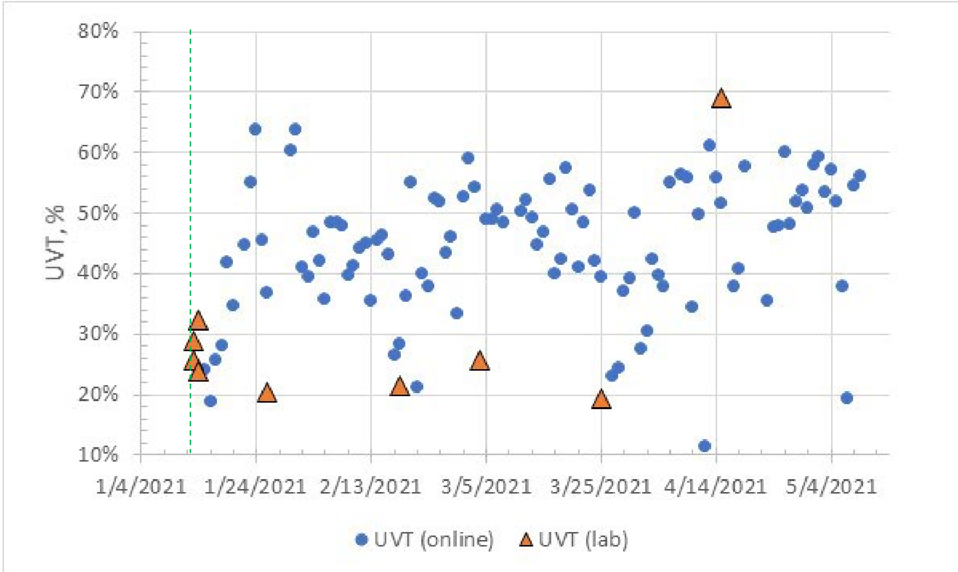


Figure 3-13. UVT Lab and Online (After Manual Cleaning) Data Comparison.
Dashed green bar represents calibration.

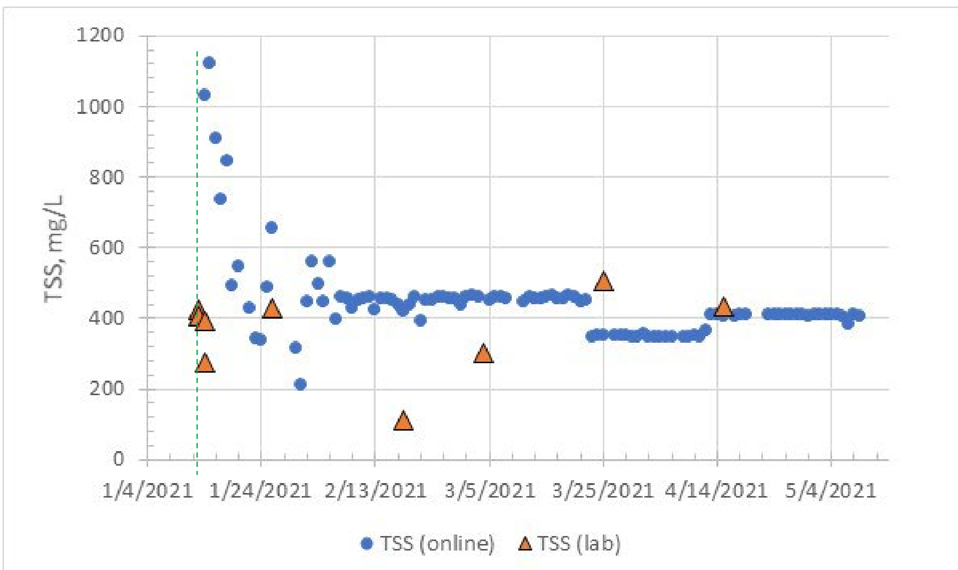


Figure 3-14. TSS Lab and Online (After Manual Cleaning) Data Comparison.
Dashed green bar represents calibration.

3.2.4.2 Measurements and Trends

A week-by-week comparison of data sets was generated, looking to define wastewater quality patterns for the monitored parameters. Six consecutive weeks of data after initial operation and calibration were used with all data presented below starting at 12:00 a.m. Monday morning. The spectro::lyser was cleaned daily, typically between 8 a.m. and 2 p.m., unless noted otherwise. The pH::lyser and condu::lyser were only cleaned once over the test period. More detail on fouling is presented below.

These data and related investigations indicated:

- Electrical Conductivity:
    - Monday through Friday saw multiple spikes of EC of more than double the baseline EC of approximately 1400 microSiemens per centimeter (µS/cm) (Figure 3-15).
    - Saturday and Sunday saw one EC spike each day in the mid-morning, though these spikes were relatively low.
    - Results, week by week, were comparable, clearly indicating an industrial and community pattern. Variations from this pattern, which have not yet been seen, would indicate an abnormal or new discharge to the system.
    - The weekday EC spikes were the result of brine discharge from a large bottled water company. Morro Bay is currently working with that company to develop both equalization of salt spikes (near term) and direct ocean disposal (future).
- pH:
    - Both weekdays and weekend days saw pH spikes at about 8 am, raising pH from approximately 6.7 to 7.2 or 7.3 (Figure 3-16).
    - Some week to week changes in the "baseline" pH were evident, in the range of 6.6 to 6.9.
    - The pH varied diurnally, with the higher spikes following the diurnal pattern from municipal flows. pH was historically a concern, with a commercial laundry discharging substantial flow with elevated pH in the past. Recent discussions have led to better collaboration and pretreatment by the industry, with the industry adding acid to drop the pH in their discharge, which this data shows was successful.
    - While there was some variability, the profile of pH day by day and week by week indicated a stable and predictable water quality. Potential variations, either patterns or spikes, would be indicative of abnormal or new discharge to the system.
- UVA and UVT:
    - Online UVT averaged 44 percent while lab UVT averaged lower, 30 percent (Figure 3-13).
    - UVA254 (showed variability daily within a band with no discernable daily pattern (Figure 3-17).
    - UVA appeared to be highest on Monday.
    - For the 3rd week of study and the 6th week of the study, both Wednesday through Thursday or Friday, there were substantial, prolonged, and atypical upward trend in UVA, followed by a steep drop back to "normal" levels (Figure 3-18). A similar trend was seen in the online readings for COD, BOD, and TSS. While it may at first appear to be a water quality event, these "events" directly correlated to test periods where the spectro::lyser was intentionally not cleaned. Thus, the witnessed events simply attest to the need for daily cleaning for this application.

- TSS:
  - Except for one point, the online TSS data did not match the lab TSS closely (Figure 3-14).
  - TSS showed some variability daily with less of a consistent pattern compared to other variables.
  - The week to week data sets did not overlap consistently. Further evaluation for reasons for this difference is needed.

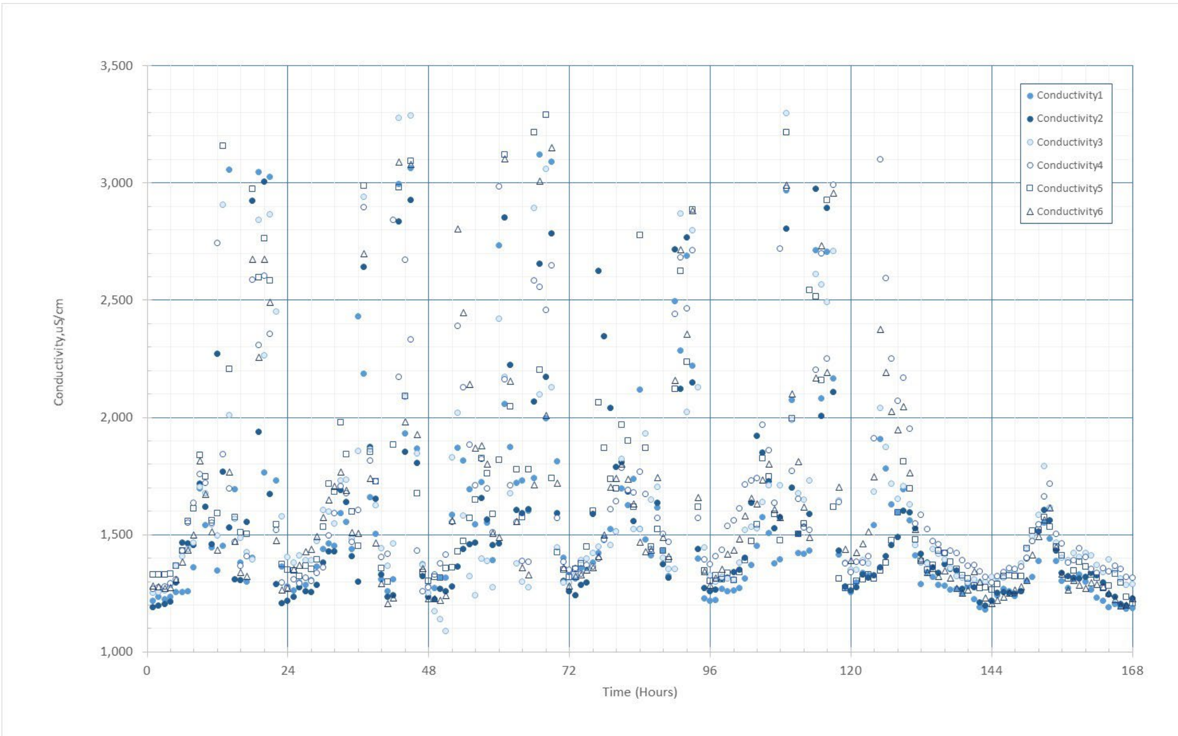


Figure 3-15. Week by Week EC Profiles.
(notes: Conductivity1 is Week 1 Values, Time 0 is 12 am Monday Morning)

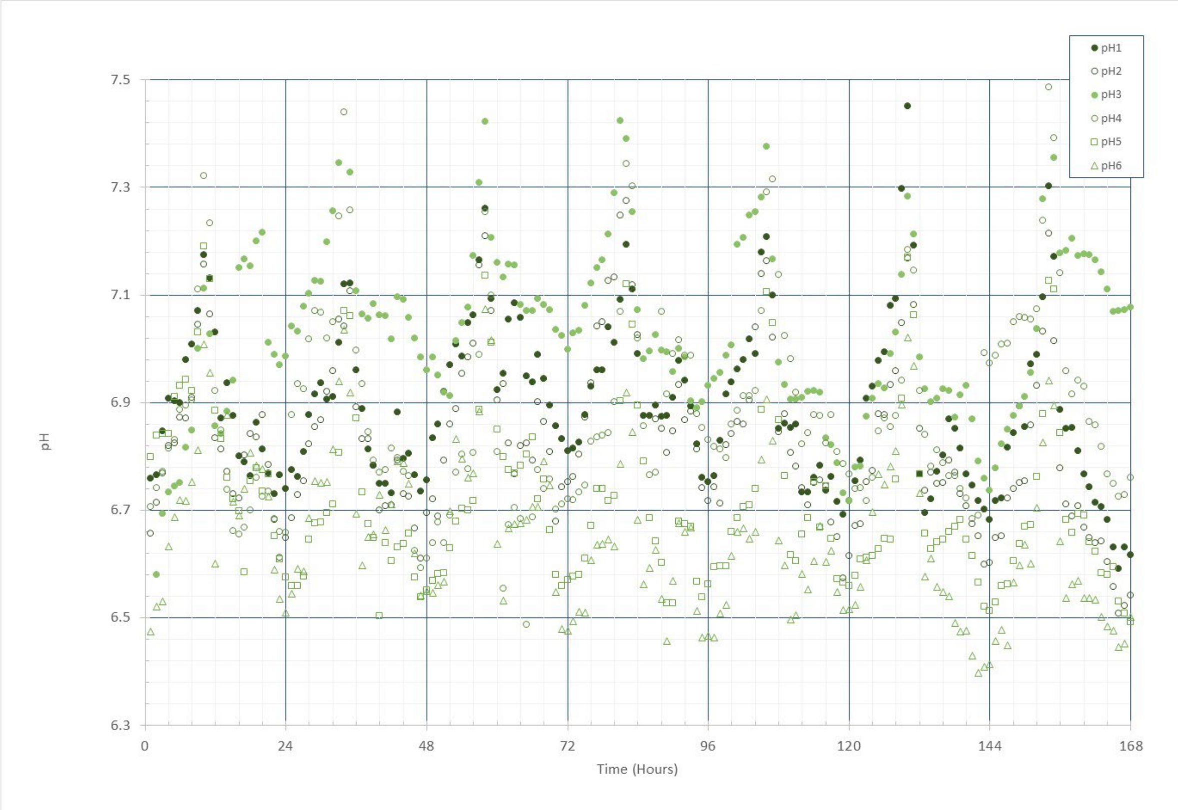


Figure 3-16. Week by Week pH Profiles.
(notes: pH1 is Week 1 Values, Time 0 is 12 am Monday Morning)

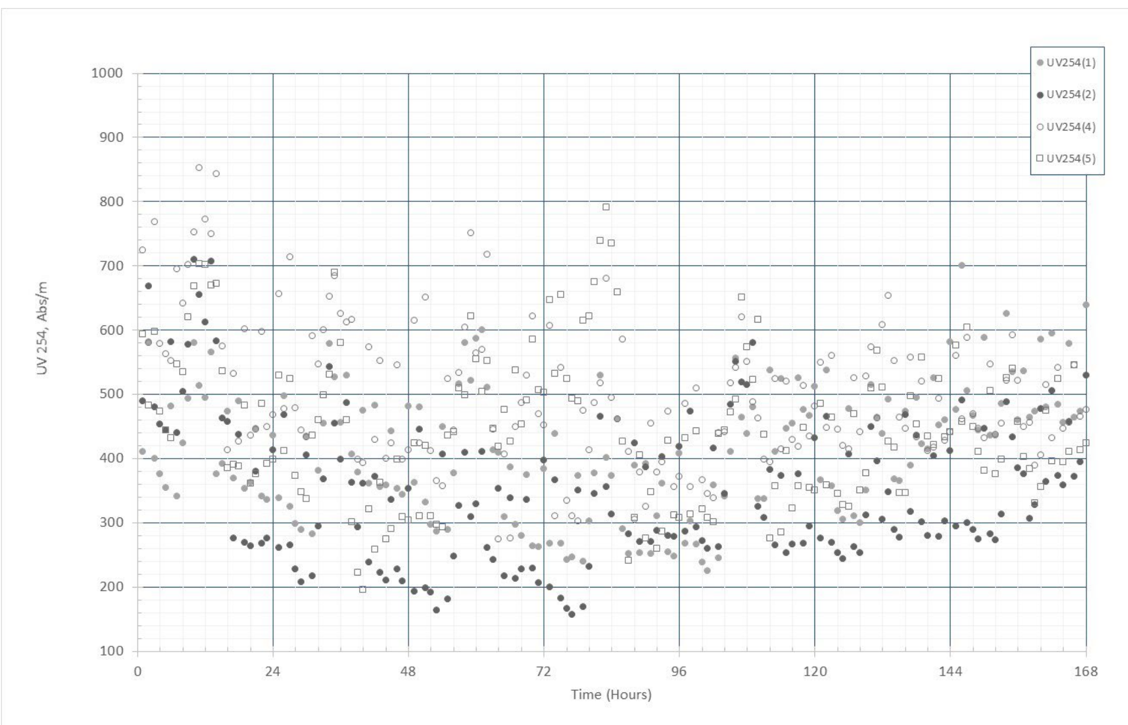


Figure 3-17. Weekly UVA Profiles for Weeks 1, 2, 4, and 5.
Time is hours since 12 am Monday morning. Parenthesis after UV254 indicate week number.

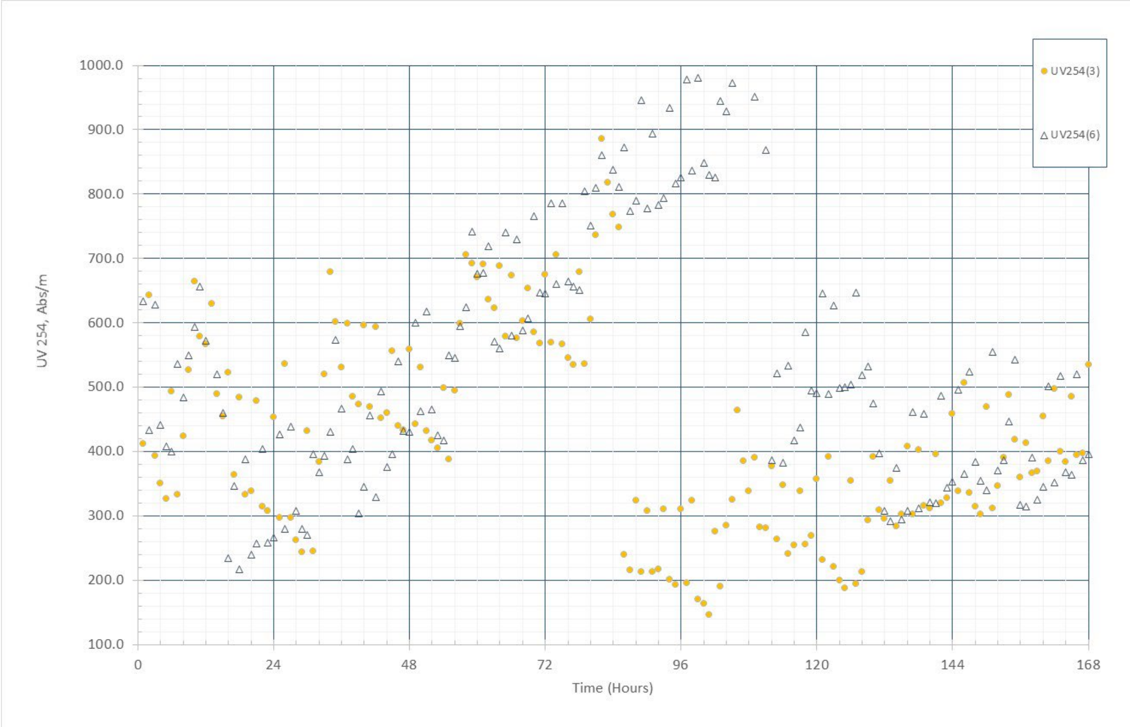


Figure 3-18. Weekly UVA Profiles for Weeks 3 and 6.
Time is hours since 12 am Monday morning. Parenthesis after UV254 indicate week number. Cleaning was performed after 3 to 4 days of operation, as noted by the drop in UV254.

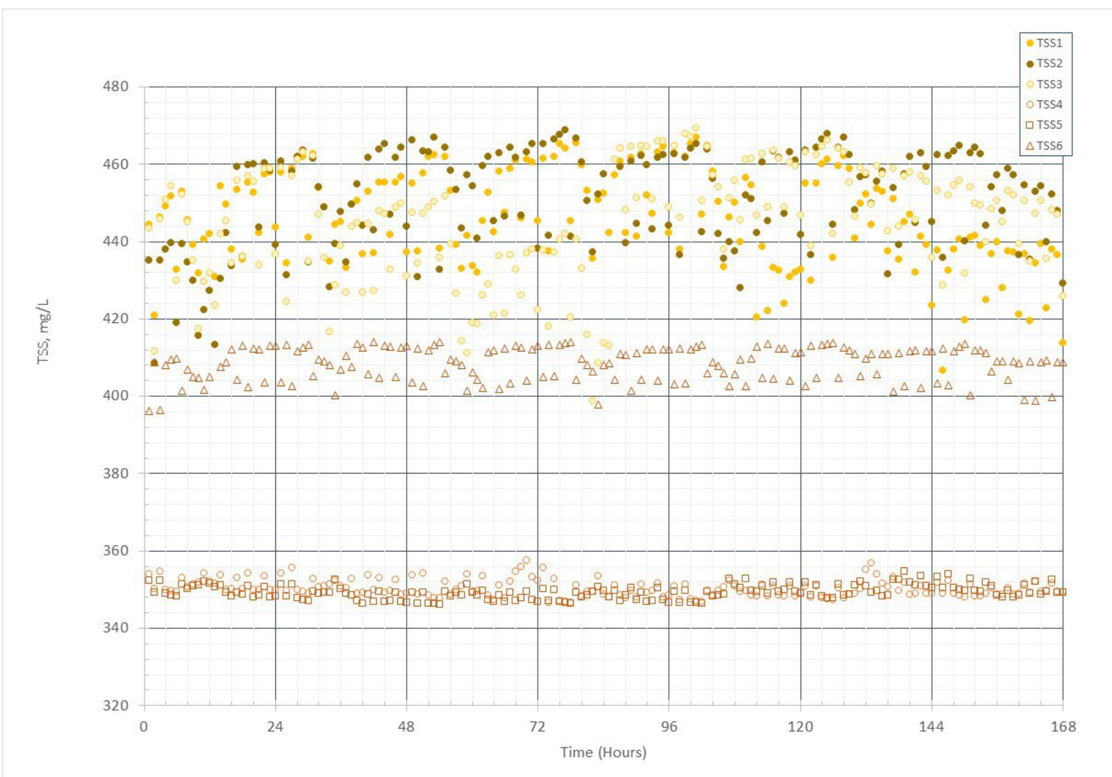


Figure 3-19. Week by Week TSS Profiles.
(notes: TSS1 is Week 1 Values, Time 0 is 12 am Monday Morning)

3.2.4.3 Probe Fouling

As stated previously, Morro Bay doses ferrous chloride at 1 to 2 mg/L (targeting approximately 1.65 mg/L) into the feed of the WWTP to control H_2S and improve primary clarification performance. The expectation from the project team was that this dosing would lead to fouling of the s::can spectro::lyser [COD, BOD, TSS, UV254, bisulfide (a surrogate for H_2S), and NO_3] but not the other s::can probes because of differences in the sensor window. The operations staff cleaned the spectro:lyser sensor daily, but not the other s::can probes (which were only cleaned once). As expected, minimal fouling of EC and pH probes was seen.

Evaluation of fouling results for the spectro:lyser was based upon daily comparisons of recorded values, pre and post cleaning, referred to as a “Fouling Ratio”, which is the pre clean value divided by the post clean value for a given time point. These results include a 3-day period from March 8th to March 11th where the ferrous was turned off (see highlight in Figure 3-20), the probes where cleaned, and the fouling rate was evaluated with the hopes of seeing less fouling. These results also included several periods of time where there was no sensor cleaning. The results indicated:

1. Turning off the ferrous for 3 days did not appear to reduce probe fouling. The daily cleaning of the spectro:lyser was sufficient to minimize fouling impacts, even while ferrous was being dosed.
2. Daily cleaning minimized fouling, as evident by a large cluster of data at the Fouling Ratio of 1, with data spread both above and below this value.
3. Extending the cleaning interval to two or three days did not show dramatically higher Fouling Ratios, but all of these longer test period data showed Fouling Ratios >1, so there was clearly a negative impact on fouling due to longer days between cleaning (Figure 3-21).
4. TSS and COD measurements saw daily fouling ratio of up to approximately 1.2.
5. UVA and BOD measurements saw a daily fouling ratio of up to approximately 2.0.
6. The daily variability shown in the earlier figures for UVA, COD, BOD, and TSS may be more related to fouling than true underlying variability in the water quality or random analytical variation.

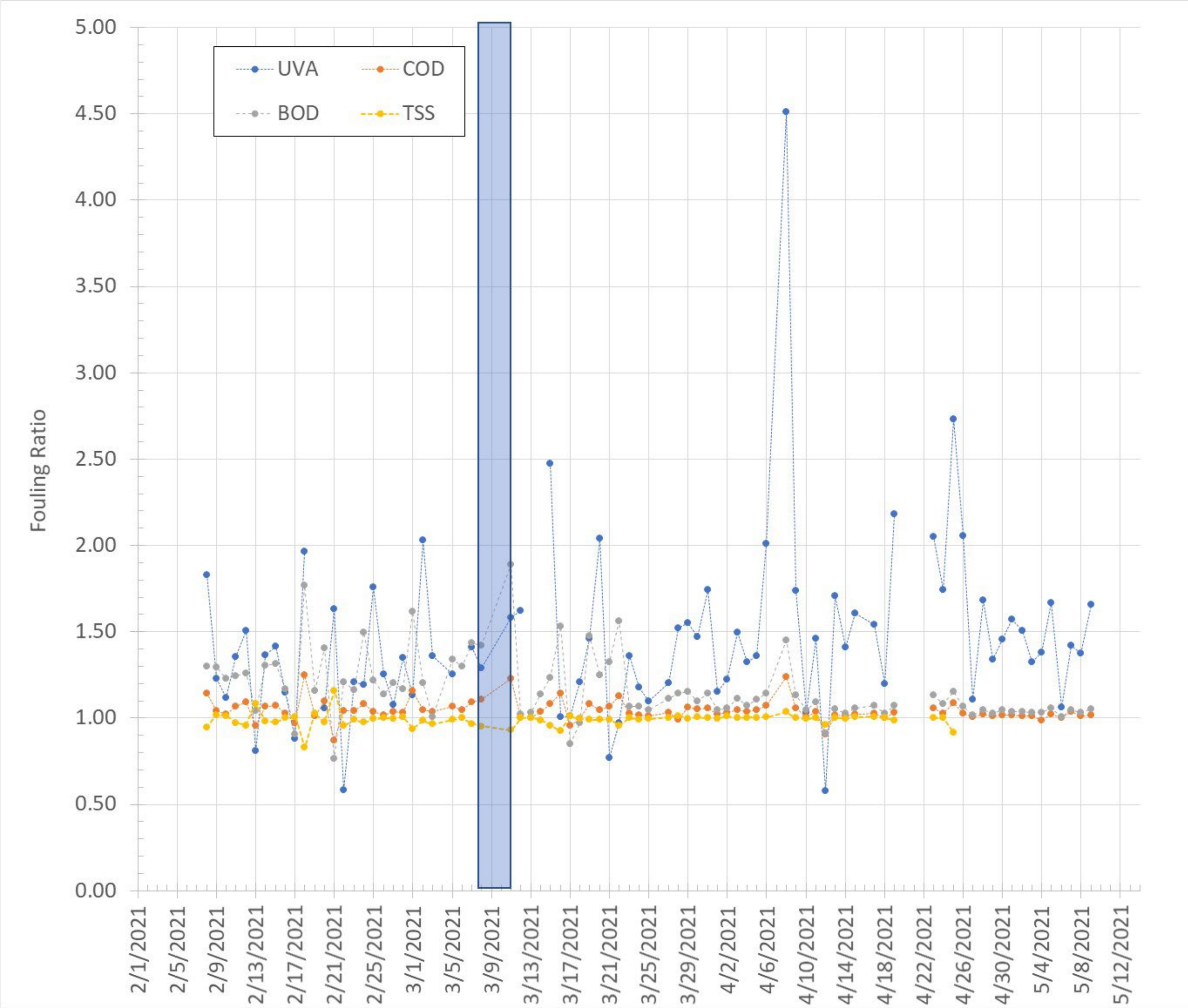


Figure 3-20. Evaluation of spectro::lyser Probe Fouling Ratio.
(3 day period with no ferrous dosing highlighted)

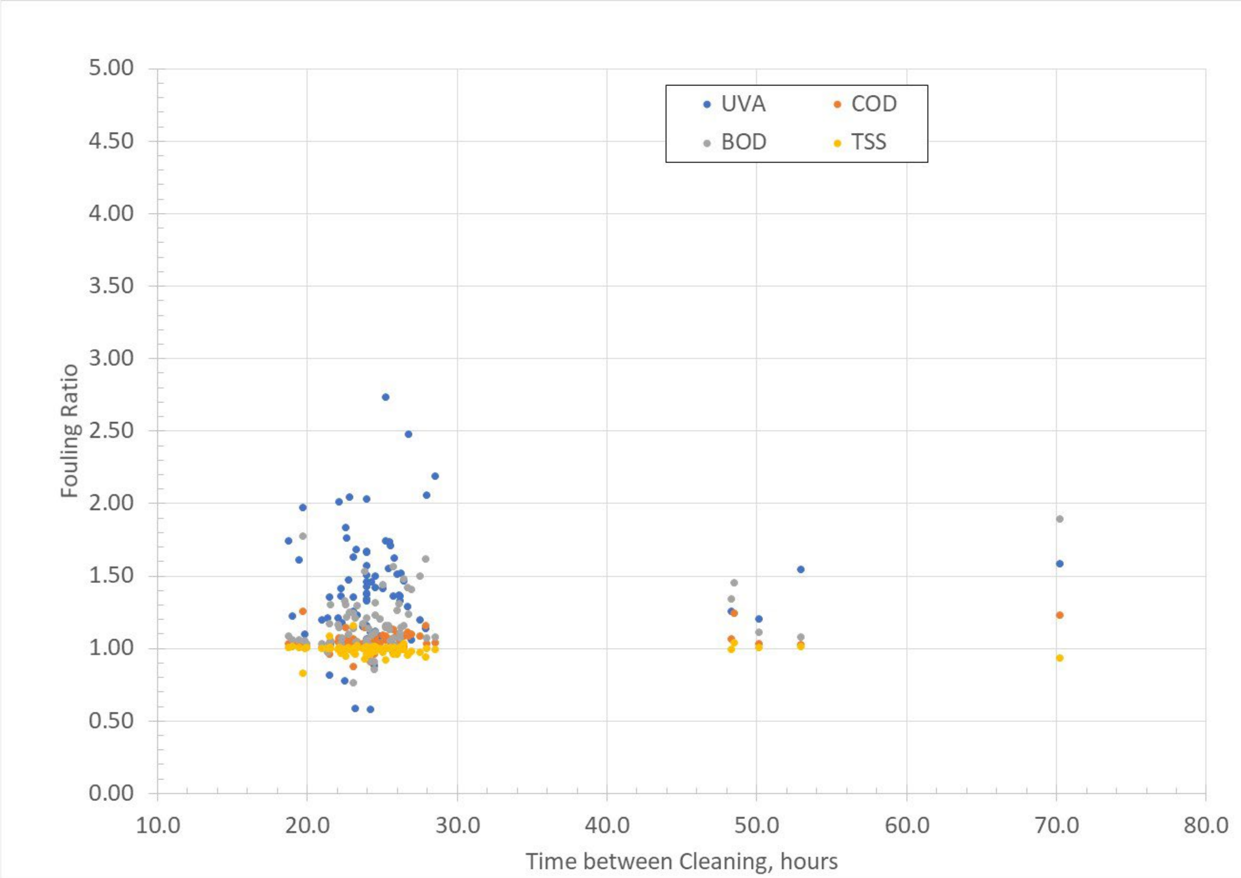


Figure 3-21. Evaluation of spectro::lyser Probe Fouling Ratio Based Upon Cleaning Interval.
(48 hour and 72 hour cleaning intervals had no ferrous chloride dosing)

3.3 Los Angeles County Sanitation Districts Field Research

3.3.1 Introduction

The Sanitation Districts' wastewater pretreatment program is broad, intended to protect all of the Sanitation Districts' water reclamation plants and the downstream permit compliance for both NPDES discharge and for potable water reuse. The program includes:

- Collection of wastes from 850 square miles and 78 cities and unincorporated territory within Los Angeles County.
- Serving 5.7 million people in Los Angeles County.
- Tracking of all wastes to one ocean discharge facility and 10 water reclamation plants, respectively: Joint Water Pollution Control Plant, La Cañada, Lancaster, Long Beach, Los Coyotes, Palmdale, Pomona, San Jose Creek (East and West), Saugus, Valencia, and Whittier Narrows.
- Approximately 400 categorical industrial users (CIUs).
- Approximately 1,000 significant industrial users (SIUs).
- Approximately 1,500 other industrial dischargers.

Because of the extent of the Sanitation Districts wastewater treatment program, their research department actively pursues innovation, including online monitoring of water quality.

In parallel to this research, the Sanitation Districts pilot tested two Sentry online water quality analyzers. In direct support of this research, the Sanitation Districts pilot tested two Real Tech spectrometers. All pilot testing was done within specific wastewater treatment plants (e.g., primary effluent). Details and results from all tested sensor systems are provided below.

3.3.2 Equipment and Installations

Two probe systems composed of a single probe and communication unit, provided by Real Tech, were installed and started in April 2021 by LACSD. Each probe system consisted of a Titanium Ba-X Series SA2010 multi-wave sensor (which utilizes select wavelengths to monitor water quality), a Controller (communication unit), an air clean system, and a backboard and mounting stand (Figure 3-22). Real Tech systems offer remote monitoring for COD, BOD, and TSS. One Real Tech probe system was installed at the Joint Water Pollution Control Plant (JWPCP, Figure 3-23) and one at San Jose Creek East Water Reclamation Plant (SJCEWRP, Figure 3-24). Both were initially installed in a primary effluent channel. The probe system at JWPCP was moved to secondary effluent on September 2, 2021 to compare performance of the probe in primary and secondary effluent. The probe system at SJCEWRP was moved to a different location in the primary effluent channel on September 9, 2021 due to reactor shutdowns.

Figure 3-22. Real Tech Sensor (right) and Communication Unit (left).

Figure 3-23. Real Tech in Primary Effluent (left) and Secondary Effluent (right) at JWPCP.

Figure 3-24. Real Tech in Primary Effluent at SJCEWRP.

At initial location adjacent to Unit 3 (left) and secondary location west of the primary sedimentation tanks (right).

One Sentry system composed of two probes and one communication unit for remote monitoring were also provided to LACSD (Figure 3-25). Sentry probes, which work by measuring signals from biological activity of exoelectrogenic bacteria on the surface of a sensor, were installed in SJCEWRP in February 2021 in the primary effluent channel and in an anoxic zone of a secondary treatment reactor (Figure 3-26). On September 17, 2021, the Sentry system and one probe were relocated to the raw influent of Los Coyotes Water Reclamation Plant (LCWRP), which is known to have the largest industrial waste loading as a percentage of total flow upstream and had more opportunities to observe toxic shocks (Figure 3-27). On November 4, 2021, the Sentry system and one probe was relocated into primary effluent at Lancaster Water Reclamation Plant (LANWRP), a remote facility in a separate collection system from SJCEWRP and LCWRP (Figure 3-28).

Figure 3-25. Sentry Probe (left) and Communication Unit (right).

Figure 3-26. Sentry Probe in Primary Effluent (left), Communication Unit (middle), and Sentry Probe in an Anoxic Zone (right) of SJCEWRP.

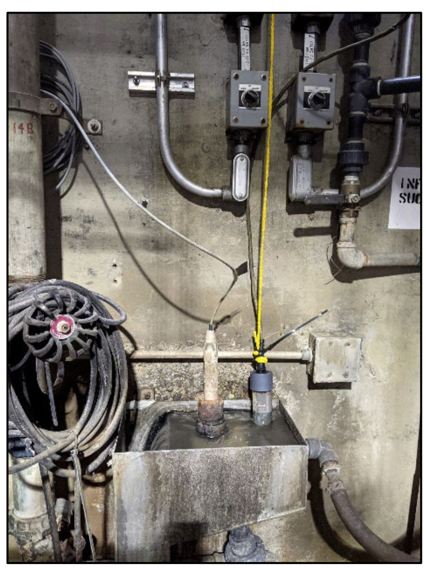


Figure 3-27. Sentry Probe in LCWRP Raw Influent.
(attached to the yellow cord)

Figure 3-28. Sentry Probe in LANWRP Primary Effluent.

3.3.3 Probe Test Methods

3.3.3.1 Test Method for Real Tech

The Real Tech sensors required initial calibration, and each instrument was monitored to ensure it continued to read accurately over the measurement period. For the first two weeks after installation, starting May 10, 2021, three samples a week were collected to establish the calibration of the sensors. Analyses for these samples included soluble chemical oxygen demand (sCOD), soluble biochemical oxygen demand (sBOD), soluble total organic carbon (sTOC), and TSS.

Following the first two weeks after installation, one sample a week was collected to check the accuracy of each analyzer for 24-weeks. sCOD analyses for each sample were performed for a 24-week duration.

In agreement with Real Tech's recommendations, the following cleaning/maintenance was done on the sensors:

- Automated Cleaning:
  - Air blast was initially done every 10 minutes, but reduced to 5 minutes as the evaluation progressed.
- Manual Cleaning:
  - The sensor was initially cleaned weekly. However, weekly cleaning was insufficient and manual cleaning was performed daily for a short period at JWPCP. Though conducted for the sake of this study, in practice, many utilities would consider a daily cleaning frequency overly burdensome. After observing relatively stable signals following a period of daily manual cleaning, the cleaning interval was changed to three times per week from August 18, 2021 onward.

- To clean the sensor, it was removed from the channel, rinsed off with wash water, the optical lens was wiped a few times with a cloth soaked with the mild acid/CLR (calcium, lime, & rust remover) in a "flossing" type action, and re-rinsed with wash water before replacement to the channel.
- The maintenance was done on the same days of the week to try to keep cleaning intervals as consistent as possible. Initial weekly manual cleanings for the sensor at SJCEWRP were on Tuesday mornings. Weekly manual cleanings for the sensor at JWPCP were on Wednesday afternoons. Three times a week manual cleanings for both sensors were on Monday, Wednesday, and Friday mornings.

Calibrations were performed on June 16 and September 15, 2021. Calibrations could only be performed by Real Tech through their proprietary model and not by LACSD staff. Real Tech testing was completed in November 2021.

3.3.3.2 Test Method for Sentry

As the Sentry was designed to detect signals of toxicity, the results from the sensor were monitored and compared to the Real Tech sensor, monitored flows, and routinely monitored constituents to see if there was a correlation with COD, BOD, or TSS and changes in microbial activity. Ideally, this sensor would have provided an early warning to nitrification inhibitory events.

No sampling or calibration was done for this sensor at it measures microbial activity and there is no standard lab method for comparison. The Sentry system and probes were rented for one year so testing may continue through early 2022.

3.3.3.3 Challenge Testing

Challenge testing was conducted for the Sentry and Real Tech sensors at SJCEWRP through batch tests. The purpose of challenge testing was to determine if the sensors detect known concentrations of low molecular weight organics that may pass through advanced treatment and increases in soluble COD through the addition of sodium acetate. The compounds selected were mostly common industrial chemicals that cannot fully be removed by reverse osmosis. In some cases, the concentrations selected were thought to be inhibitory for nitrification. In other cases, no inhibitory concentration was found in literature. Notably, neither sensor was necessarily designed with the goal of detecting these specific chemicals.

For testing the low molecular weight organics, Sentry and Real Tech sensors were placed in a 70-gallon tank with a mixer (Figure 3-29). The tank was filled with 20 gallons of primary effluent and spiked with compounds at low and high concentrations (Table 3-4). Following a short stabilization period (baseline), a single compound was spiked into 20 gallons of primary effluent at a time. After a few minutes of mixing the primary effluent with the spiked compound, the tank was drained and refilled with primary effluent and the next compound was spiked. Duplicate testing of each compound was performed. Low concentrations were tested first in an abundance of caution (in case a specific compound permanently inhibited the Sentry), followed by higher concentrations.

For testing of soluble COD through sodium acetate addition, the tank with the mixer and sensors was filled with 20 gallons of primary effluent and the sensors were given a few minutes to develop a baseline. Following a baseline, sodium acetate was added to the desired increase in concentration, a few minutes of mixing was allowed for sensor acclimation, and additional sodium acetate was added to the next desired increase in concentration. Additions of sodium acetate to achieve 100, 200, 300, 500, and 1000 mg/L were tested.

Figure 3-29. Challenge Testing Setup (left) and Plan View of Sensors in the Tank (right).

Table 3-4. Challenge Tested Compounds and Concentrations.

Compound	Low	High
	Concentration (mg/L)	
Acetone	20	100
Chloroform	5	10
Methylene Chloride	10	50
Toluene	10	100
1,2-Dichloropropane	5	100
Carbon Disulfide	10	20
Formaldehyde	100	200

3.3.4 Probe Results

3.3.4.1 Real Tech Performance

The Real Tech at SJCEWRP exhibited a clear diurnal pattern following installation (Figure 3-30). Prior to the initial calibration, the laboratory grab sample results did not appear to match the sensor results. Following calibration on June 16th, the fit with the grab samples improved. There was a considerable amount of noise in the signal –the source of this was unknown. The channel where the instrument was located had adequate velocity and should have been well mixed.

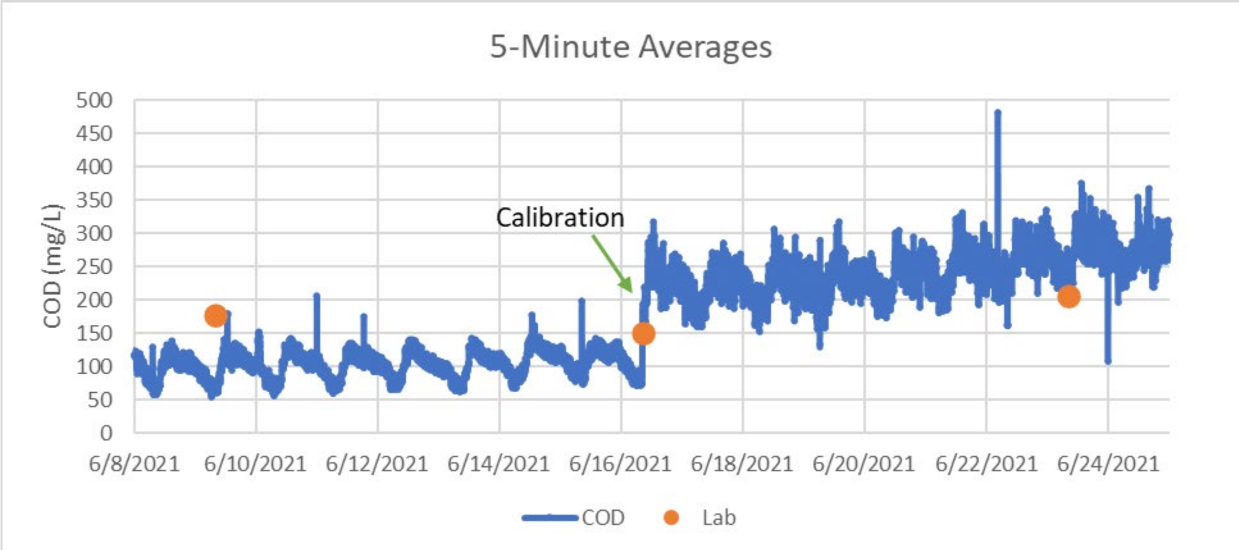


Figure 3-30. SJCEWRP Real Tech and Grab Sample COD Results from June 8 to June 25, 2021.

When re-calibration was performed on September 15th to improve accuracy, the diurnal pattern did not seem to be as pronounced (Figure 3-31).

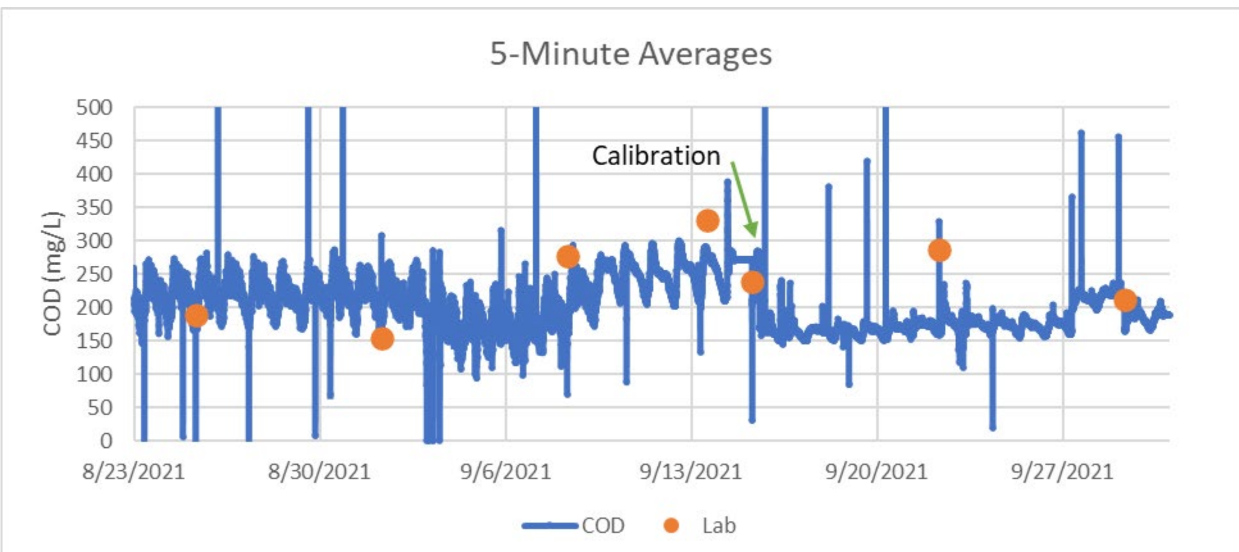


Figure 3-31. SJCEWRP Real Tech and Grab Sample COD Results from August 8 to October 1, 2021.

After the SJCEWRP Real Tech was moved from one location in the primary effluent channel to another, the sensor signal was smoothed out (Figure 3-32). It appears that the source of the noise in the COD signal was not from the instrument but from a hydraulic condition in the channel.

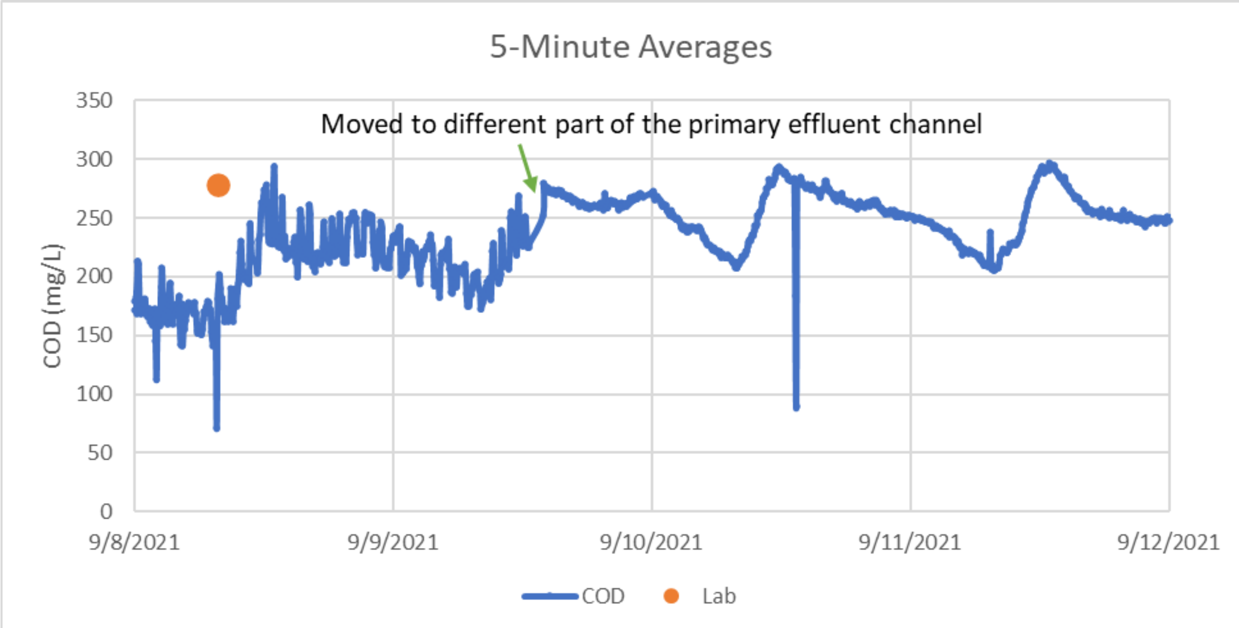


Figure 3-32. SJCEWRP Real Tech and Grab Sample COD Results from September 8 to September 12, 2021.

The Real Tech at JWPCP initially exhibited a clear diurnal pattern following installation in primary effluent (Figure 3-33). However, following the June 16th calibration, the diurnal pattern disappeared. In addition, the signal pattern appeared to indicate some accumulated fouling.

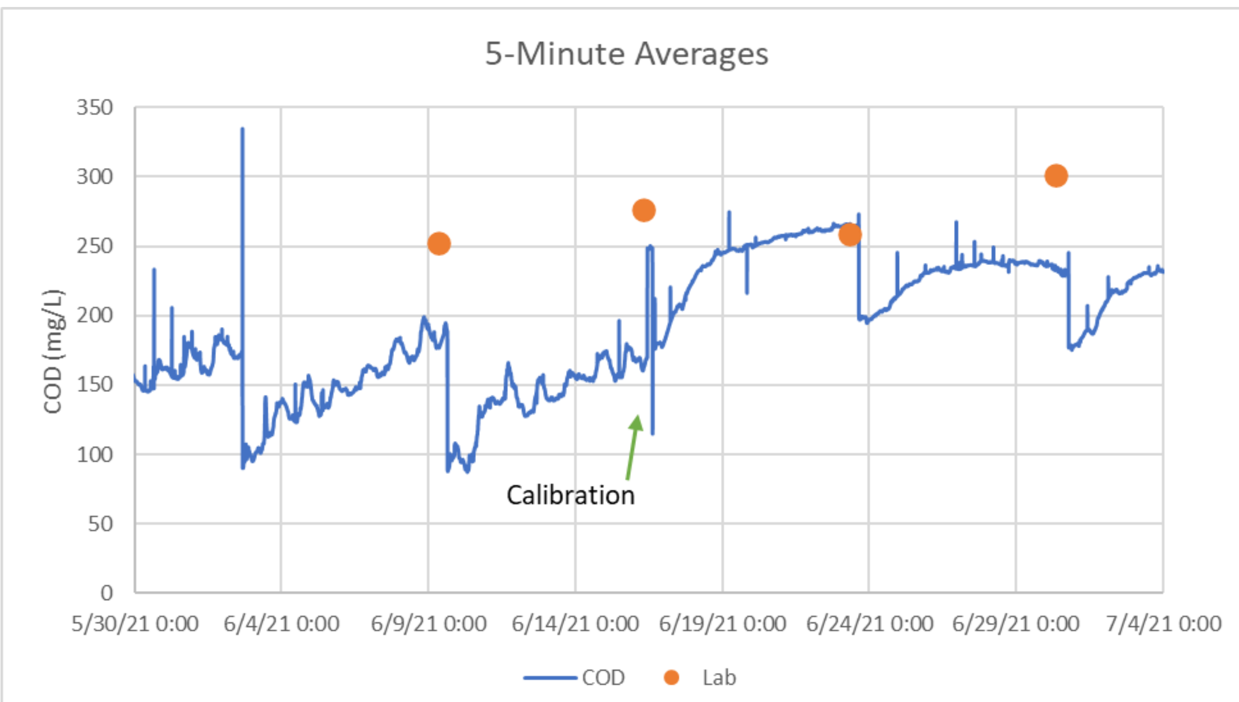


Figure 3-33. JWPCP Real Tech and Grab Sample COD Results from May 30 to July 4, 2021.

The Real Tech did not appear to trend well with laboratory grab samples in JWPCP primary effluent (Figure 3-34).

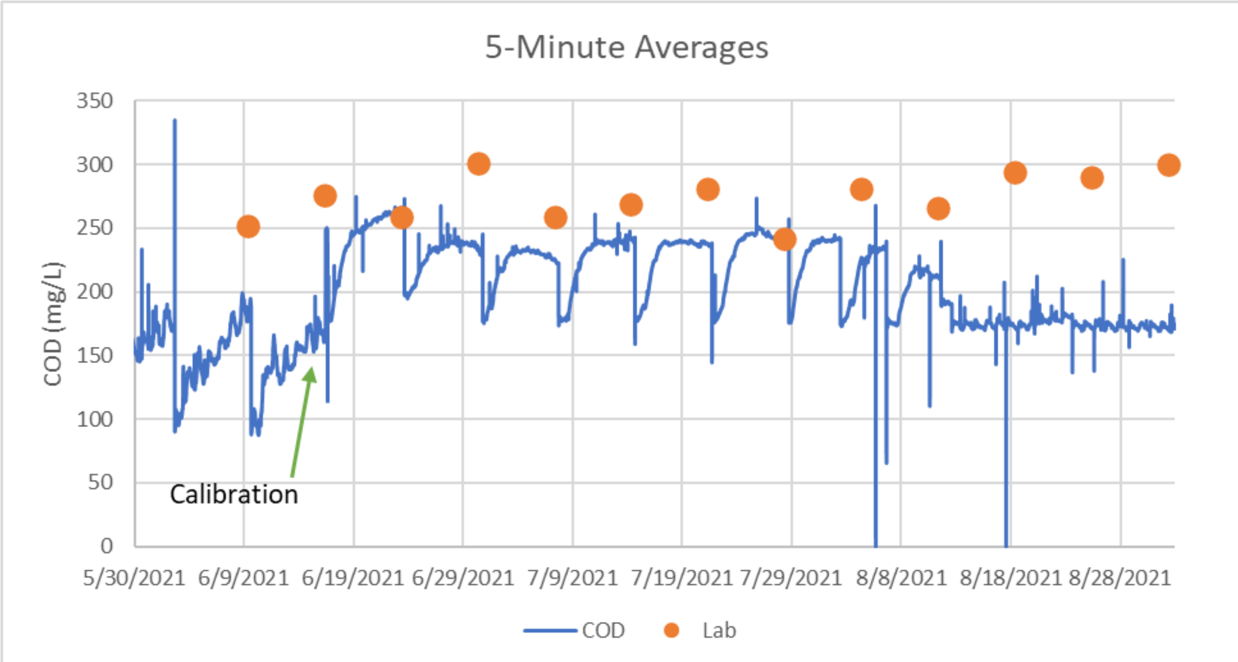


Figure 3-34. JWPCP Real Tech and Grab Sample COD Results from May 30 to September 2, 2021.

It was thought that perhaps the Real Tech frequently cleaning to prevent calibration drift in JWPCP primary effluent was due to the chemical interferences from iron and caustic dosing upstream. However, when the sensor was moved from primary effluent to secondary effluent, the fit with the grab sample data did not appear to improve, even after calibration on September 15th, which appeared to increase the noise in the sensor results (Figure 3-35). Perhaps the matrix effects of JWPCP wastewater were too great. The secondary effluent COD was also on the low end of the specified range of the instrument. Lower range models of this instrument are also available from Real Tech but were not tested for this study.

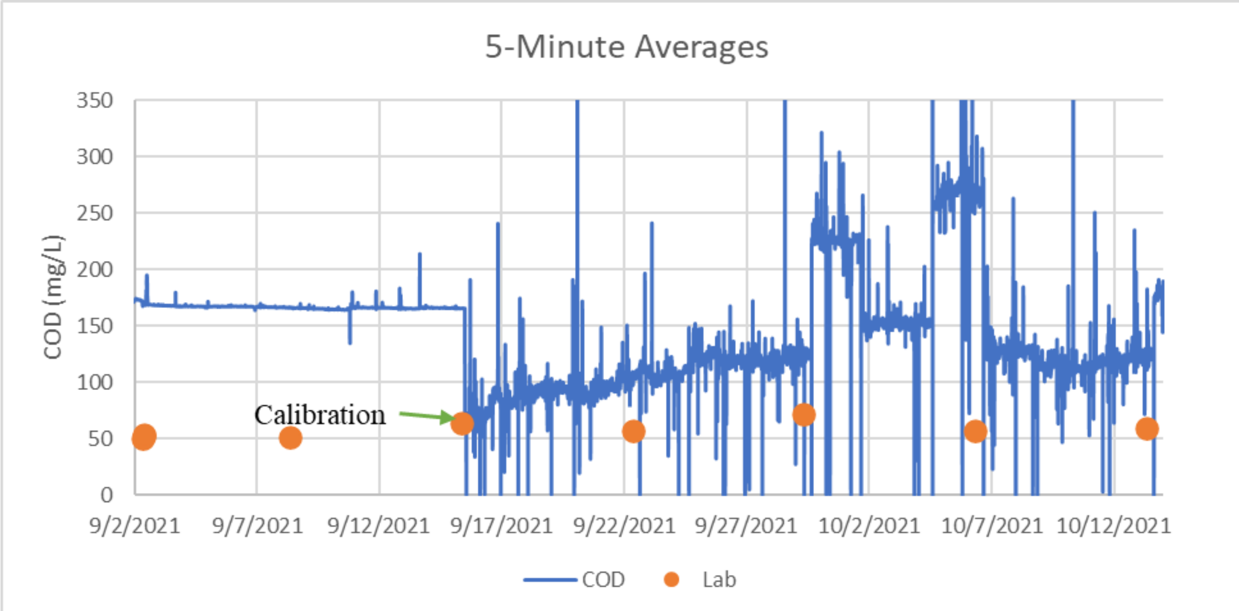


Figure 3-35. JWPCP Real Tech and Grab Sample COD Results from September 2 to October 14, 2021.

3.3.4.2 Real Tech Maintenance

For the SJCEWRP Real Tech, drift was observed starting mid-June (Figure 3-36), when new staff was assigned to clean the sensor. This highlights one of the risks of multiple individuals responsible for sensor maintenance. Starting on June 19th there was a clear pattern that appeared to represent sensor fouling. Deep cleaning (CLR soaking) performed on July 23rd assisted with recovery of some of the signal, but not to the level prior to drift. The increased cleaning frequency, from weekly to daily, in early August helped with maintaining a steady signal. Cleaning with 7.25 percent hydrochloric acid on August 6th helped the signal to recover to values prior to the signal drift. For instances of significant drift, cleaning with hydrochloric acid may be the best remedy for signal recovery. This indicates that after fouling is allowed to occur, more aggressive cleaning measures may be needed to restore the instrument.

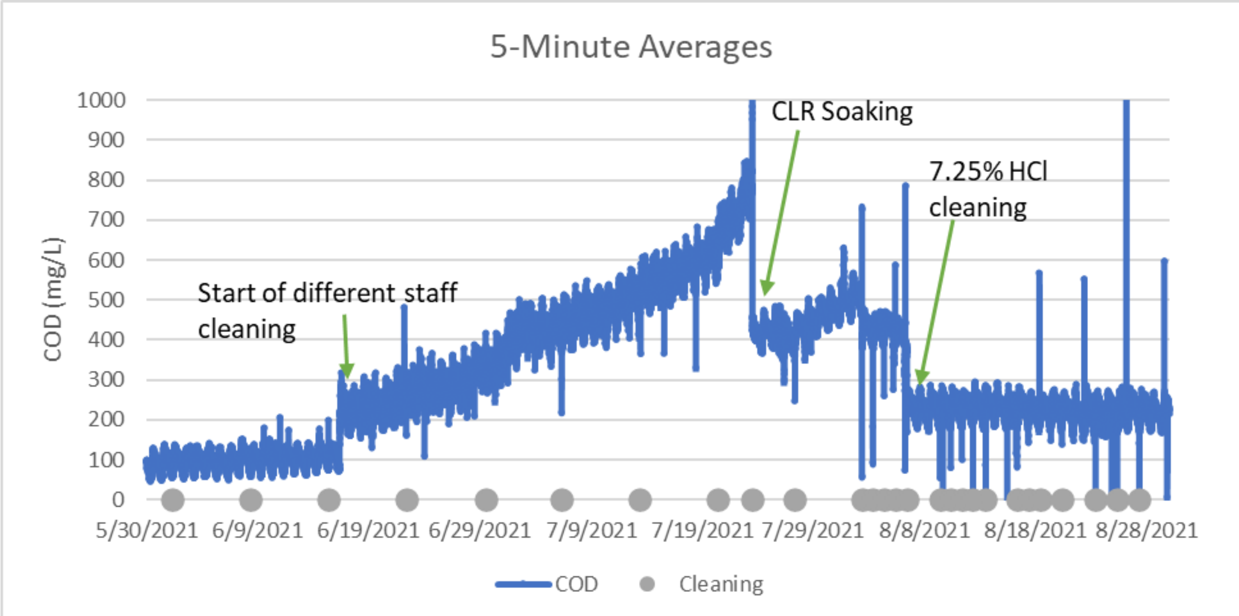


Figure 3-36. SJCEWRP Real Tech Manual Cleaning.

Weekly cleaning was not sufficient for the Real Tech in JWPCP primary effluent. At each manual cleaning event, the signal would drop and slowly increase for a few days (Figure 3-37). The sensor appeared to consistently foul rapidly after cleaning, which was suspected to be due to the ferrous chloride dosed upstream of JWPCP.

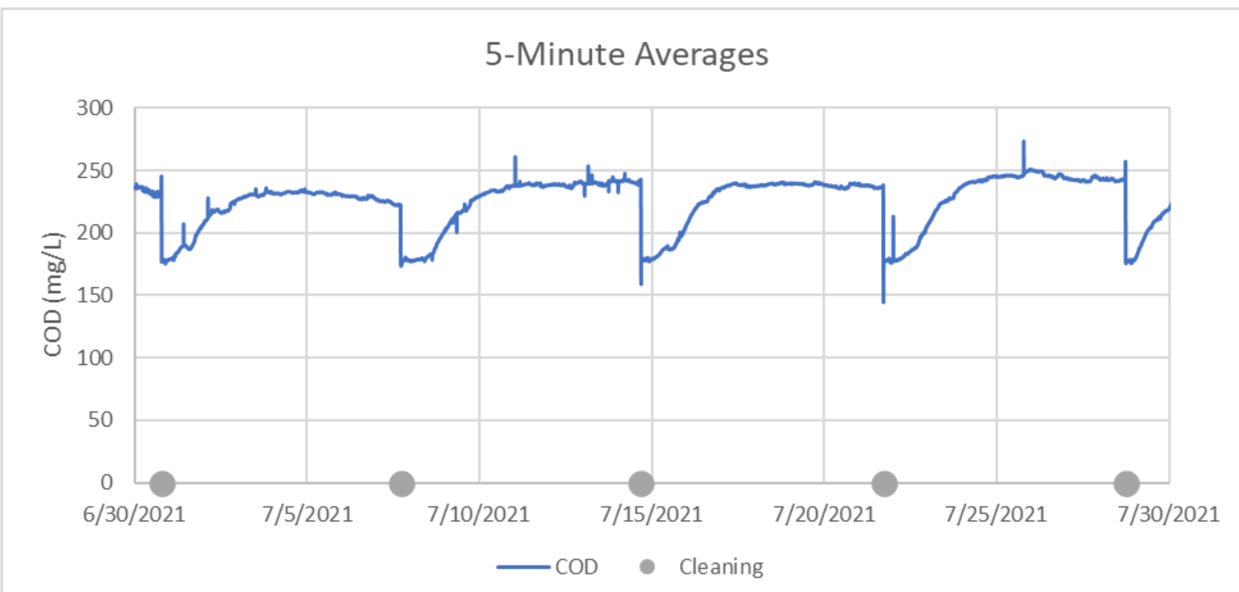


Figure 3-37. JWPCP Real Tech Manual Cleaning from June 30 to July 30, 2021.

When cleaning was changed to daily, the signal appeared to be maintained relatively stable (Figure 3-38). At cleaning three times a week, the signal also seemed to be maintained. This suggests that a minimum manual cleaning interval of three times a week was necessary to maintain stable measurements.

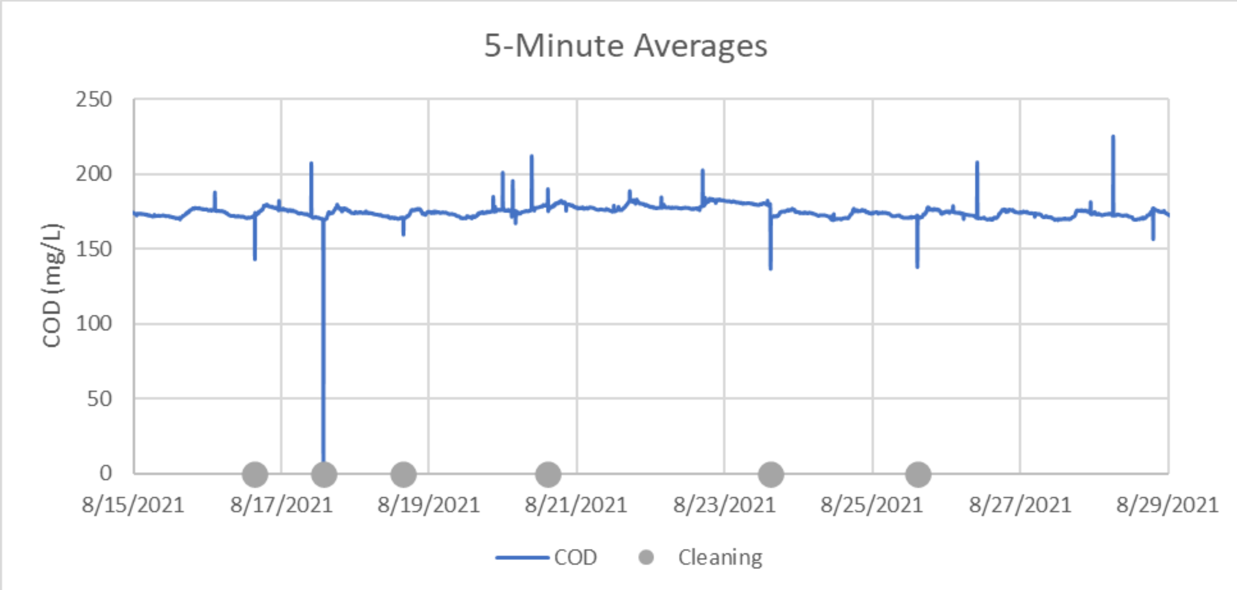


Figure 3-38. JWPCP Real Tech Manual Cleaning from August 15 to August 30, 2021.

3.3.4.3 Sentry Results

No correlation was found between the Real Tech parameters and the Sentry toxicity signal (microbial electrical signal [MES]). No nitrification inhibitory events occurred during the trial at SJCEWRP to assess the Sentry probes' sensitivity in those scenarios directly.

When overlaid with flow patterns, significant signal fluctuations in primary effluent due to plant shutdowns were notable but did not translate to significant fluctuations in the mixed liquor (ML) (Figure 3-39). This suggests that while minor plant shutdowns may affect the signal in primary effluent, there is little effect observed in ML. It is likely that when flow stops in the primary channel, this leads to a localized depletion of substrate in the immediate vicinity of the sensor. This suggests good sensitivity to substrate variations.

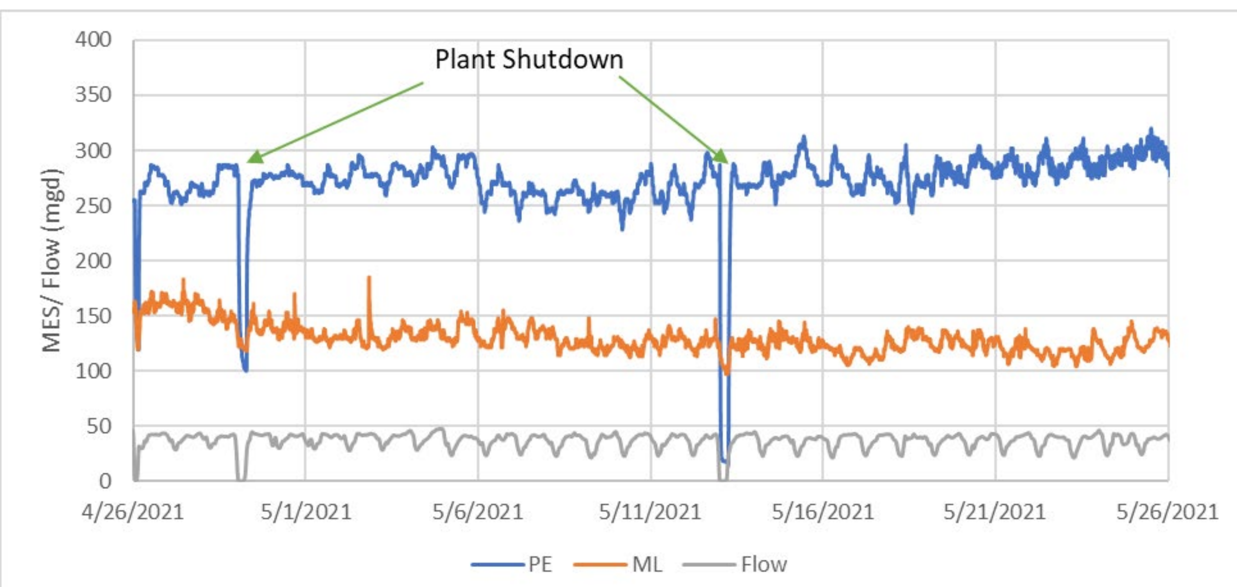


Figure 3-39. SJCEWRP Sentry Primary Effluent and ML MES from April 26, 2021 to May 26, 2021.

At LCWRP, the Sentry detected a weekly pattern in the raw influent (Figure 3-40). On Friday afternoons, between 14:00 and 17:00 weekly shock dosing with sodium hydroxide is added directly to the sewers for sulfur-oxidizing bacteria control. This disrupted the MES, dropping it significantly. Following the drop, the MES increased over the course of a week, decreasing again on Friday afternoon. The recovery time required for the sensor from the high pH wastewater, and thus the time the sensor has lost sensitivity, suggests that high pH shocks may be problematic for this type of biological sensor.

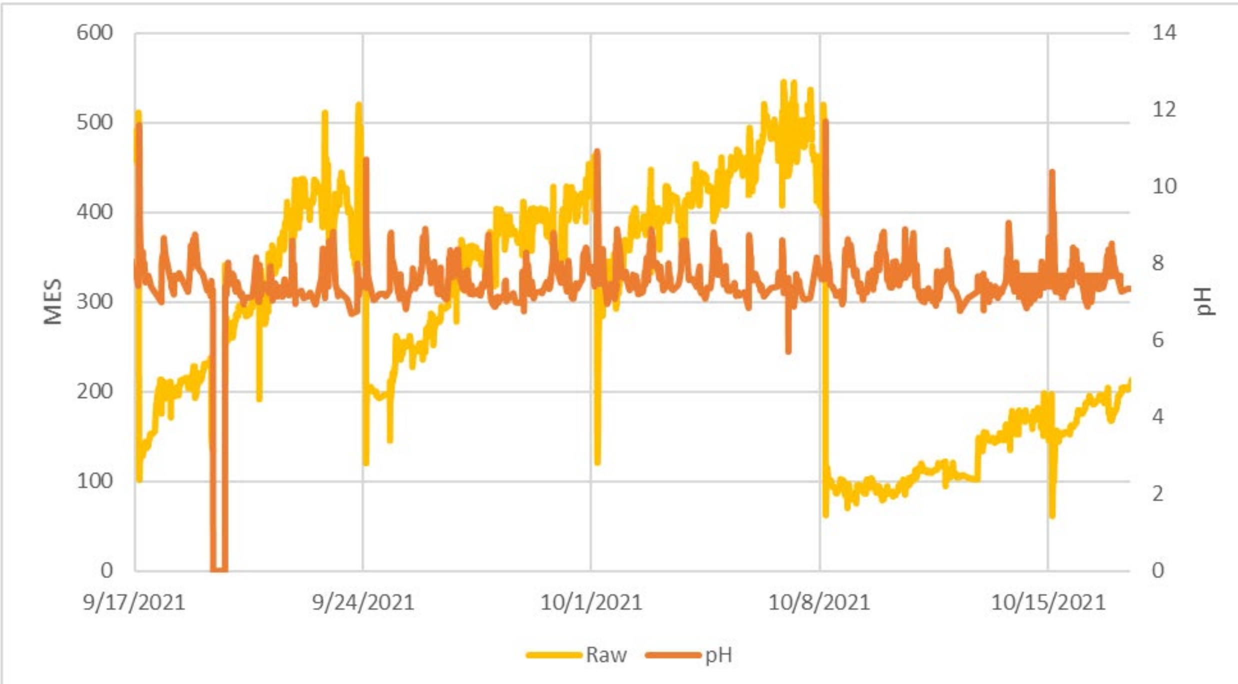


Figure 3-40. LCWRP Raw MES and pH from September 17 to October 18, 2021.

Results of the Sentry at Lancaster WRP were compared with routine 24-hour composite COD laboratory data and ambient air temperature (Figure 3-41). Although ambient temperature is not an ideal comparison (water temperature would be much better), it was only available temperature parameter for this facility.

The COD data and the Sentry data appear to have a similar trend, which suggests that Sentry can be used to try to monitor influent substrate loading. However, it appears that the Sentry is affected by temperature, and the decline in signal may be due to a decline in temperature, which suggests that substantial analysis to isolate temperature effects is required to use the Sentry results.

Figure 3-41. Lancaster WRP Influent COD, MES, and pH from November 4, 2021 to January 19, 2022.

3.3.4.4 Challenge testing Results

While some of the compounds tested were thought to be at inhibitory levels for nitrification, challenge testing of the low molecular weight organic compounds on the Sentry resulted in no significant changes in all compounds except for 200 mg/L formaldehyde. Figure 3-42 shows a typical result from most compounds (left) and the result from high concentration formaldehyde (right). Biological life on the surface of the sensor was more robust and did not demonstrate a signal change at the concentrations tested. It is possible that the Sentry was not designed to specifically detect some of the compounds and concentrations tested, which is important to know prior to procurement or full-scale implementation.

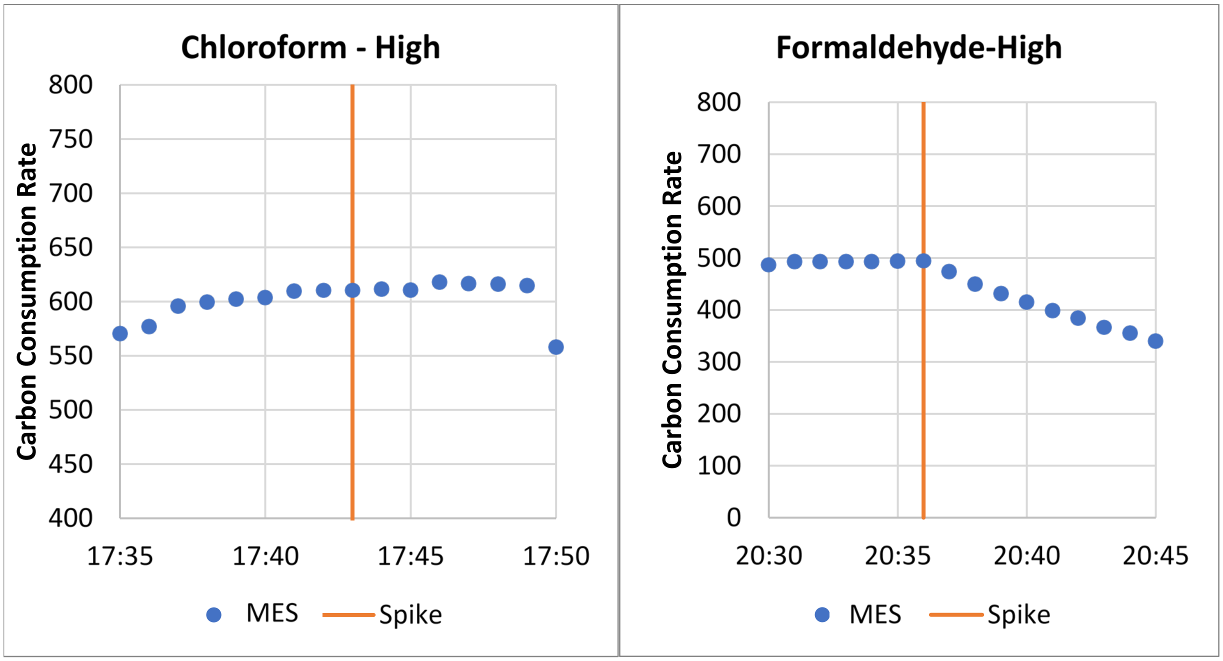


Figure 3-42. Sentry Lightweight Organic Compound Challenge Testing.

The Real Tech sensor exhibited no significant changes in COD results for all compounds tested. The results of the COD for the high concentration formaldehyde test is shown in Figure 3-43. The reason for the lack of detection is likely that the tested chemicals do not significantly absorb light at the wavelengths measured by this Real Tech sensor.

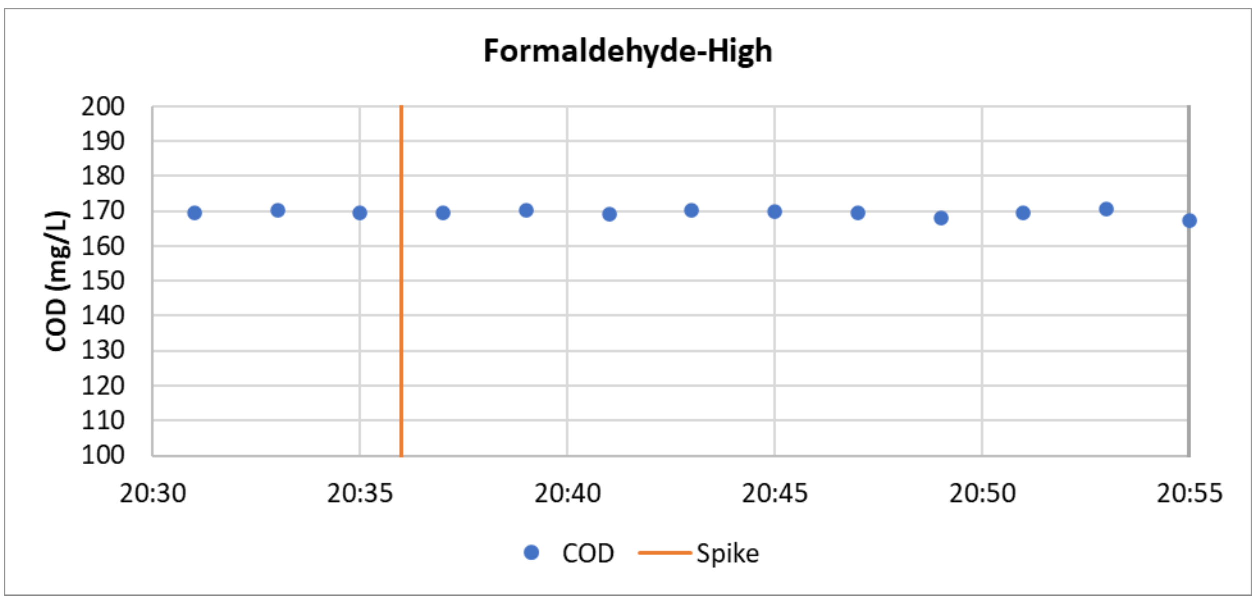


Figure 3-43. Real Tech Lightweight Organic Compound Challenge Testing: Formaldehyde.

In the acetate addition experiment, no significant changes were observed in the Real Tech COD, suggesting that the sensor was not responding to soluble COD as acetate (Figure 3-44).

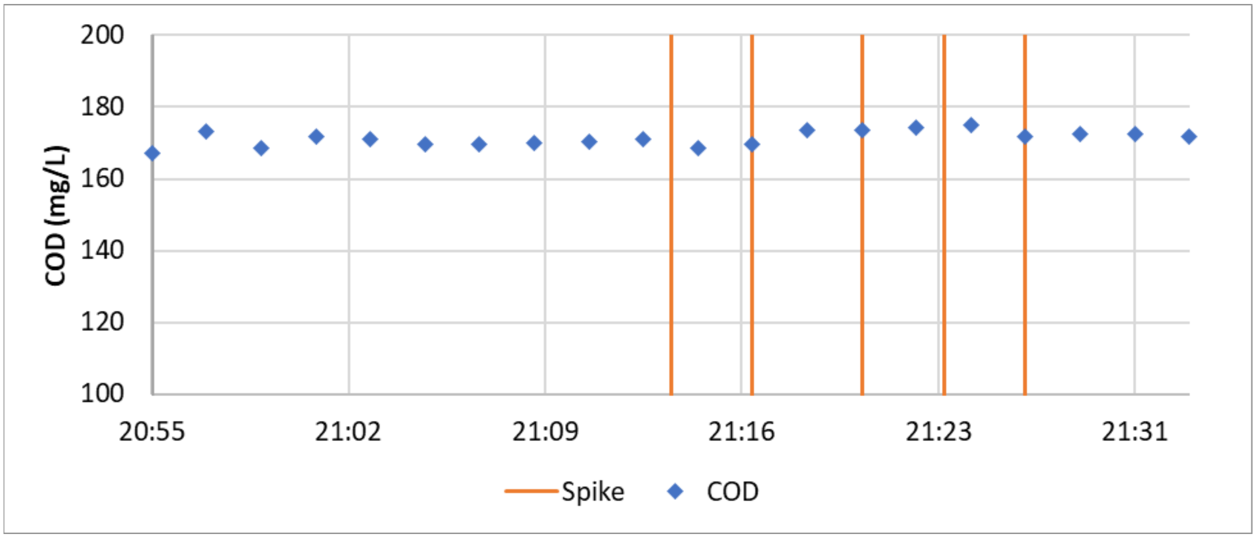


Figure 3-44. Real Tech Acetate Addition Challenge Testing.

CHAPTER 4

Review of Enhanced Source Control Programs

Only a handful of utilities have robust ESCPs for potable reuse right now, though a number of projects are developing such programs. Further, the industry has relatively little experience—which has yielded mixed results—with sensors installed in wastewater collection systems and within WWTPs. This WRF report compiled existing industry practices and information including:

- Examples of typical ESCPs that are currently planned or implemented by utilities today [Chapter 4 (this chapter)].
- Utility experiences deploying real-time monitoring within collection systems in Steinle-Darling et al. 2020 (Chapter 5).
- Utility experiences with online sensors within wastewater treatment plants, along with an analysis of how existing online WWTP and advanced water treatment facility (AWTF) monitoring data and machine learning could be used for enhanced source control to detect industrial discharges (Chapter 6, Appendices B and C).
- New case studies of utility experiences deploying real-time monitoring within collection systems (Chapter 7).

Below are examples of two ESCPs developed in support of existing and future potable reuse programs. The reviews below highlight the key aspects of an ESCP that are important to all potable reuse projects and examines how an effective online monitoring program could benefit the ESCP (Table 4-1). A more detailed review of ESCPs for potable reuse is provided in Nading et al. 2022, which will be completed in early 2022.

Table 4-1. Aspects of ESCP and Benefit of Effective Online Monitoring.

ESCP Concept	Details	Online Monitoring Benefits Include…
Regulatory Authority	The Sewer Use Ordinance (SUO) provides the authority of a utility to develop and enforce an industrial pretreatment program, including requirements to protect potable water reuse.	No applicable
Industrial Dischargers	The type and abundance of different industrial dischargers will define the level of effort and cost of a robust ESCP.	Online monitoring systems can provide greater confidence in industrial discharger compliance.
Enforcement Response Plan	The Enforcement Response Plan (ERP) outlines the procedures followed by pretreatment program staff and management to identify, document, and respond to pretreatment violations	Online monitoring systems can be used to track abnormal discharges up into the collection system and to the point of origin.
Monitoring Program	A robust ESCP relies upon a combination of industry led and utility led sampling efforts, and a tailored sampling campaign that adjusts with time to minimize laboratory analytical costs while closely monitoring water quality that can be of concern	Online monitoring systems can reduce the amount of utility led grab sampling and composite sampling through development of a database of "standard" wastewater quality downstream of industrial dischargers.

Table 4-1. Aspects of ESCP and Benefit of Effective Online Monitoring. (Continued)

ESCP Concept	Details	Online Monitoring Benefits Include…
Outreach Efforts	An effective outreach plan includes: • Communication between government departments (e.g., planning department and wastewater department) • Engagement of businesses • Development and sharing of Best Management Practices (BMPs) • Rewarding and acknowledging model industry partners • Notifying and enforcing non-compliant industries	Online monitoring results can be used to more rapidly contact industrial dischargers to alert them of water quality changes.
Interagency Agreements	Potable water reuse projects often cross jurisdictional boundaries or require collaboration of water and wastewater utilities. The development of clear roles, responsibilities, and financial commitments from participating parties is central to long term project success	Online monitoring of wastewater from partner utilities provides for better cost recovery based upon flows and loads.

4.1 Ventura

4.1.1 Potable Reuse Project and ESCP Overview

The City of Ventura (population: 110,000) is located along the California Coast. The City is in the process of planning an IPR (groundwater injection) project with the potential to add on a future DPR component.

The City of Ventura owns and operates a publicly owned treatment works (POTW) that includes 375 miles of sewer mains, 14 lift stations, and the Ventura Water Reclamation Facility (VWRF). The VWRF currently treats an annual average influent flow of approximately 7.4 mgd of combined domestic, commercial, and industrial wastewater. Six SIUs—including four CIUs—one groundwater remediation discharger, and two external jurisdictions contribute flows to the POTW.

Ventura is in the process of planning for a 4.8 mgd potable reuse project consisting of a new advanced water purification facility (AWPF) and injection wells for IPR. The AWPF will be designed with the potential to expand production capacity to 6.0 mgd for DPR via treated water augmentation.

In preparation for its planned potable reuse project, Ventura augmented its existing USEPA-approved industrial pretreatment program. Ventura's ESCP includes routine monitoring of IUs, within the sewer system, and within the WWTP and AWPF. Ventura also intends to increase its public outreach program, including both more frequent communication with IUs, and a public campaign that will include potable reuse-specific information on flushing domestic products. Ventura is also in the process of updating its Local Limits (as of 2021), potentially including additional pollutants of concern that are relevant specifically to a potable water reuse project. Ventura also has a plan in place to trace contaminants through the collection system if necessary.

4.1.2 Industrial Users

Ventura has six SIUs, including four CIUs.

Ventura's CIUs, SIUs, groundwater discharger, and contributing jurisdictions are summarized in Table 4-2 along with their regulatory classifications, wastewater type, pretreatment, and potential contaminants. Ventura also has two major hospitals that contribute to its collection system that are not formally permitted under its pretreatment program, but the collection system downstream of their discharge points is monitored under the ESCP.

Table 4-2. Ventura's Industrial Users.

Industrial User ID	Regulatory Classification of Industrial User	Wastewater Type	Pretreatment	Potential Contaminants from Subject Discharger
1	SIU & CIU	Metal Finishing (Categorical standard 40 CFR 4330)	Clarification	Cu, Zn, pH
2	SIU & CIU	Metal Finishing (Categorical standard 40 CFR 4330)	Metals Precipitation, Filter Press, and pH Adjustment	Ca, Mg, Na, Fe, Ni, Cr, Zn, pH, TDS
3	SIU & CIU	Polishing Operations (Categorical Standard 40 CFR 469 Electrical and Electronic Components) (CFR 2023a)	Trench floor drain with clarifier pit	HF, HCl, H_2O_2, NH_4OH, KOH, HNO_3, pH
4	SIU & CIU	Metal Finishing (Categorical standard 40 CFR 4330)	Primary: Metal precipitation with MgOH pH Control, Settling, Filter Press. Sulfite Salt to Reduce Cr(VI). Auxiliary: pH Adjustment with NaOH Controller and Settling.	Ni, Cu, Cr, Zn
5	SIU	Fruit Washing	Screens and Solids Filters	TDS, Fixed Dissolved Solids (FDS), pH, Fungicide, Chlorides, Solids
6	SIU	Resin Regeneration and Service Tank Rinse Water; RO Reject	Clarification	TDS, Chlorides
7	Contributing Jurisdiction	Domestic	Air Injection, Line Flushing	Sulfides, Settleable Solids
8	Contributing Jurisdiction	Municipal	None	Municipal sewage only
9	Groundwater Discharger	Petroleum Hydrocarbon Impacted Groundwater	GAC Media Filter	TPHg, BTEX, MTBE, TBA

4.1.3 Enforcement Response Plan

Enforcement procedures for industrial dischargers are in place to ensure that out-of-compliance industries bring themselves into compliance, or their service terminated. For minor incidents of non-compliance, Verbal Warnings or written notices of non-compliance and correction (NONCs) may be issued. If an SIU has clearly violated its permit, a notice of violation (NOV) is sent to the SIU. The SIU then has 14 days to submit an explanation of violation and a plan for correction. For further exceedances, increasing enforcement action is taken as necessary. Such actions can include public notification, fines, cease-and-desist orders, civil and criminal actions, and termination of service, including emergency severance of the POTW connection.

4.1.4 Monitoring Program

Ventura's ESCP, once fully implemented, includes routine monitoring and monitoring response plan. All ESCP activities are overseen by one program point person. The key components of the program are summarized below.

4.1.4.1 Routine Monitoring

The routine monitoring program is summarized in Table 4-3. Samples of specific constituents are routinely collected (2X to 4X per year) within the collection system during discrete sample events; by contrast, online monitoring of surrogates only occurs within treatment facilities.

Table 4-3. Ventura ESCP: Routine Monitoring Program.

Monitoring Type	Description
Industrial Sampling Program	Industry-specific sampling that occurs at each industry's sample port. Includes both self-monitoring and city sampling.
Sector Sampling Program	Sampling for a number of constituents (including some by not all MCLs) is performed at 10 manholes throughout the service area to capture representative contributions from residential, commercial/industrial, and hospital dischargers within the collection system. Constituents are also sampled in the VWRF plant influent.
Hospital Monitoring	Limited (2X) sampling of 29 pharmaceuticals, hormones and personal care products collected downstream of the hospital discharge points. Select chemicals were detected above the monitoring trigger levels (MTLs) set by SCCWRP (Drewes et al. 2018) but those same chemicals were below the MTL values in the VWRP secondary effluent.
Collection System Drainage Zone Nodes	Sampling of drinking water constituents (e.g. all regulated MCLs and NLs within the State of California) and Local Limits at the four major nodes within the collection system. The results serve as a baseline so that elevated levels can be detected in the case of a contaminant tracking event.
VWRF Influent	Sampling for all drinking water constituents, NPDES constituents, and Local Limits.
VWRF Effluent	Sampling for all drinking water constituents, NPDES constituents, and Local Limits.
AWPF Effluent	Sampling for all drinking water constituents and Local Limits
Treatment Facilities Online Monitoring	Online monitoring of surrogates (e.g. electrical conductivity) at the influent to the VWRF, effluent to the VWRF, and within the AWPF for surrogates.

4.1.4.2 Monitoring Response Plan

The ESCP include provisions for AWPF effluent monitoring data to trigger response actions. A response is triggered if a constituent is detected in purified water at a level higher than 10 percent of its applicable level (maximum contaminant level [MCL], sMCL, AL, NL, or MTL). The first response is to reach out directly to potentially culpable IUs. In parallel with industrial outreach, additional monitoring is triggered for the problematic constituent. A confirmation sample is collected and analyzed; the confirmation and initial samples are averaged. If the average of the two samples is still above 10 percent of the applicable level, samples are collected within the VWRF effluent, VWRF influent, and at the four main nodes within the collection system. If one of the four main nodes appears to have abnormal loading, the constituent will be traced through the collection system using the minor nodes that contribute to the implicated major node.

The constituent is then monitored at a higher frequency in both AWPF effluent and VWRF effluent for at least six months until the average of six consecutive months of sampling is lower than 10 percent of the applicable level.

Figure 4-1 shows how the collection system is divided up into four drainage zones that each drain to a single point in the collection system (node). Figure 4-2 provides an example the smaller sewersheds and their minor nodes that comprise one of the four zones.

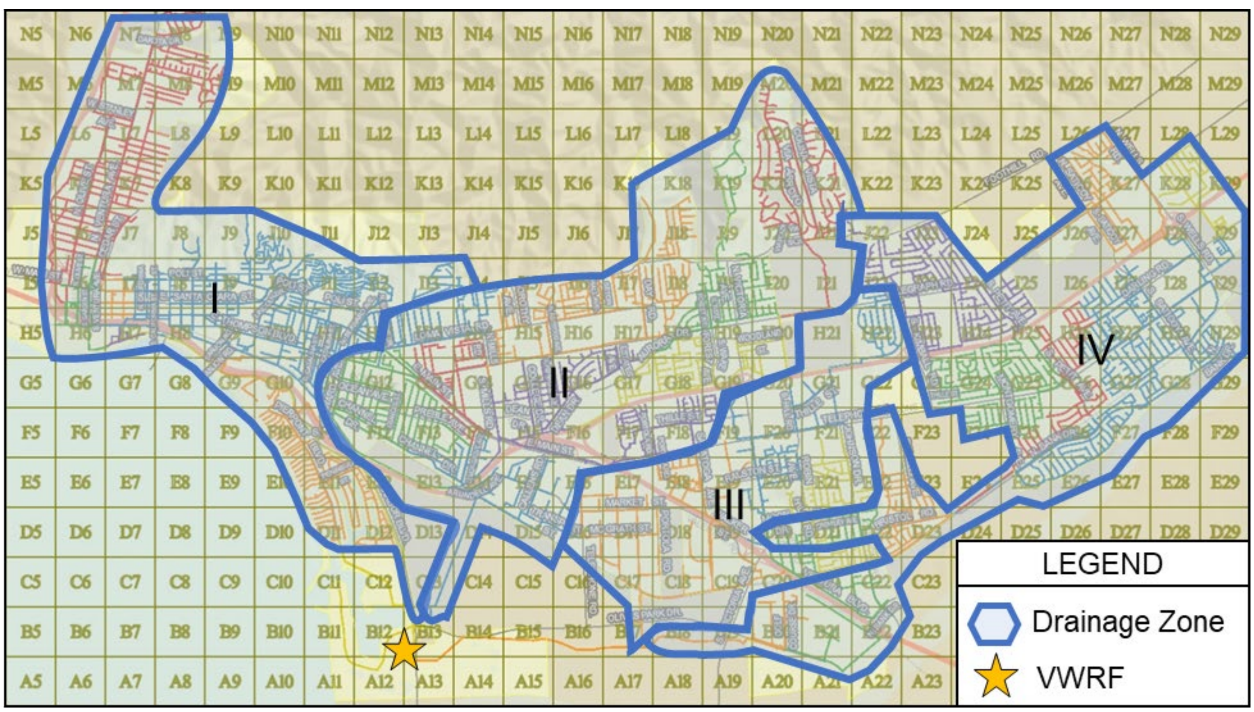
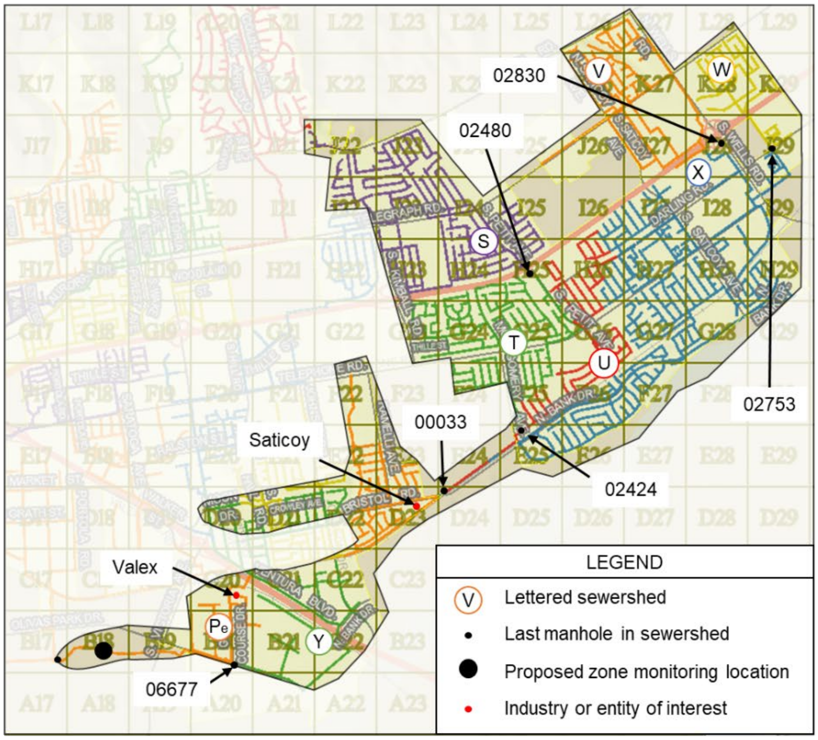


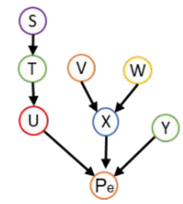
Figure 4-1. Ventura Collection System: Four Drainage System Zones.



Map of Monitoring Zone IV

Additional Sampling Locations within Zone IV

Lettered Sewershed	Last Manhole(s) of each Sewershed (Lucity ID)	Drains to Sewershed
Pe	00245	n/a
S	02480	T
T	02424	U
U	00033	P
V	02830	X
W	02753	X
X	00033	P
Y	06677	P



Zone IV
Flow Diagram

Figure 4-2. Example of Ventura Collection System Drainage Zone and Flow Path.

4.1.5 Outreach Program

Industrial outreach includes regularly scheduled meetings with IUs including reminders of the potable water reuse program, and any updates including monitoring trends, plant upsets, or concerning constituent detections.

The residential outreach program is focused on community education rolled out up front as part of the broader outreach program for the project. Planned educational materials include a website developed to address safe disposal practices of household items such as pharmaceuticals.

4.1.6 Potential Opportunities for Incorporating Online Sensors into Ventura's Source Control Program

The potential benefit to Ventura's ESCP from an effective sensor system is shown in Table 4-4.

Table 4-4. Specific Benefits of Online Monitoring to Ventura ESCP.

ESCP Concept	Specific Benefits of Online Monitoring to Ventura
Industrial Dischargers	With the small number of industrial dischargers in Ventura, a probe network could be installed at select nodes in their collection system which would be in proximity to each of the industrial dischargers. With a small number of industrial dischargers, online sensors could potentially be placed directly into the effluent from each discharger.
Enforcement Response Plan (ERP)	Online sensors can be used in two ways as part of an ERP. First, online sensors can be used to track a pollutant back to a source. As part of Steinle-Darling (2020), Ventura staff used a network of Kando sensors to track a large daily TDS discharge back to the source. Second, online sensors can be used to verify that a discharger has amended/adjusted their operation to be back in compliance.
Monitoring Program	While industry led monitoring and reporting should be maintained, with sufficient online data and correlations with grab sample data, the "baseline" water quality in different sections of the collection system can be determined. Sampling of industrial dischargers could be minimized to events when the "baseline" water quality is not maintained, such as a spike in one or more of the online monitored parameters.
Outreach Efforts	Immediate communication between Ventura staff and industrial dischargers can be made in the event of abnormal collection system wastewater quality.
Interagency Agreements	Not applicable for Ventura, as they own and maintain their collection system.

4.2 Oxnard

4.2.1 Potable Reuse Project and ESCP Overview

The City of Oxnard owns and operates a regional POTW that serves the City, City of Port Hueneme, the Naval Base Ventura County, and several surrounding unincorporated communities. It is comprised of the Oxnard Wastewater Treatment Plant (OWTP) and its associated wastewater collection system and outfall line. The OWTP is a secondary treatment facility with a design flow of 31.7 mgd and an average daily flow of 20 to 22 mgd. Approximately 75 percent of the influent to the OWTP is residential. The remaining 25 percent is derived from commercial and industrial users.

The City's AWPF can divert 8 to 9 mgd of biologically treated secondary effluent for purification using MF, RO, and UV/advanced oxidation process with hydrogen peroxide, resulting in up to 6.25 mgd of advanced treated water. Oxnard has been granted regulatory approval for a groundwater recharge project with the purified water and is in the process of constructing aquifer storage and recovery (ASR) wells that will both inject and extract the advanced treated water into and out of an aquifer for potable reuse.

Oxnard updated its source control program for the potable reuse project including evaluating the local limits and developing a collection system monitoring and response plan.

4.2.2 Industrial Users

Thirty-five SIUs discharge to the OWTP collection system, including 11 CIUs (categories include: aluminum forming; metal molding and casting; steam electric power generating; metal finishing; pulp, paper, and paperboard). The City also permits 2 non-SIUs with effluent limits and monitoring requirements.

Oxnard previously hosted one of the largest centralized waste treatment (CWT) facilities in California within their service area, Santa Clara Wastewater (SCWW). CWTs treat hazardous and nonhazardous wastes (e.g. industrial tank residuals called “tank bottoms”, oil field operations wastes). They are regulated under 40 CRF 437 (CFR 2023b), and are managed by POTWs through their industrial pretreatment programs. The major issue surrounding the acceptance by POTWs of the discharge from CWT facilities—especially Subcategory D facilities like SCWW that accept multiple wastestreams—is their potential impact on water reuse programs. Oxnard has experienced the discharge of chemical pollutants, such as grossbeta, that are challenging to treat at their AWPF, leading to banning certain CWT facilities from discharging to their collection system.

4.2.3 Enforcement Response

Enforcement procedures for out-of-compliance industrial dischargers are built into the City’s Code of Ordinances. If an SIU violates its permit, a written NOV is sent to the SIU. The SIU then has 10 days to submit an explanation of violation and a plan for correction. For BOD and TSS limit violations, the SIU is surcharged based on a predetermined formula. For other exceedances, increasing enforcement action is taken as necessary. Such actions can include discontinuing sewer or water service, a cease and desist order, issuance of a fine, or termination of permission to discharge to the system.

4.2.4 Monitoring Program

4.2.4.1 Routine Monitoring

Oxnard’s monitoring program provides necessary information for evaluating industry compliance, assessing OWTP loading and operation, and determining illicit discharges. SIUs are monitored via three mechanisms: self-monitoring, monitoring by the City, and surveillance sampling. Self-monitoring is performed by industries as required by their permit. In addition to industry self-monitoring, the City conducts facility sampling twice per year. The sampling location is outlined in each SIU’s permit. To facilitate detection of illegal discharges of prohibited materials into the collection system, surveillance monitoring is also conducted. Such monitoring is performed if the City suspects illegal dumping or if there are complaints.

The City also regularly monitors the advanced treated water and secondary effluent for local limits, MCLs, sMCLs, NLs, and CECs. Monitoring of the raw influent wastewater and within the collection system is not conducted regularly, but is done to establish baseline trends and if there is an issue with the purified water.

4.2.4.2 Monitoring Response Plan

Similar to Ventura, Oxnard prescribes increased monitoring in response to elevated levels of routinely monitored constituents or odd online monitoring data.

Routinely monitored purified water quality data triggers actions for enhanced source control. A response is triggered if a constituent is detected in purified water at a level higher than 10 percent of its applicable regulatory level (e.g. MCL). The first response is direct outreach with potentially culpable industries. In parallel with industrial outreach, increased monitoring is triggered for the problematic constituent. A confirmation sample is collected, and the initial and confirmation samples are averaged. If the average of the two samples exceeds 10 percent of the applicable limit, and outreach efforts have not yet identified a culpable industry, samples are collected in the purified effluent, secondary wastewater effluent, raw wastewater, and within the collection system.

The collection system is broken up into six monitoring zones as shown in Figure 4-3, with major trunk lines and industries identified.

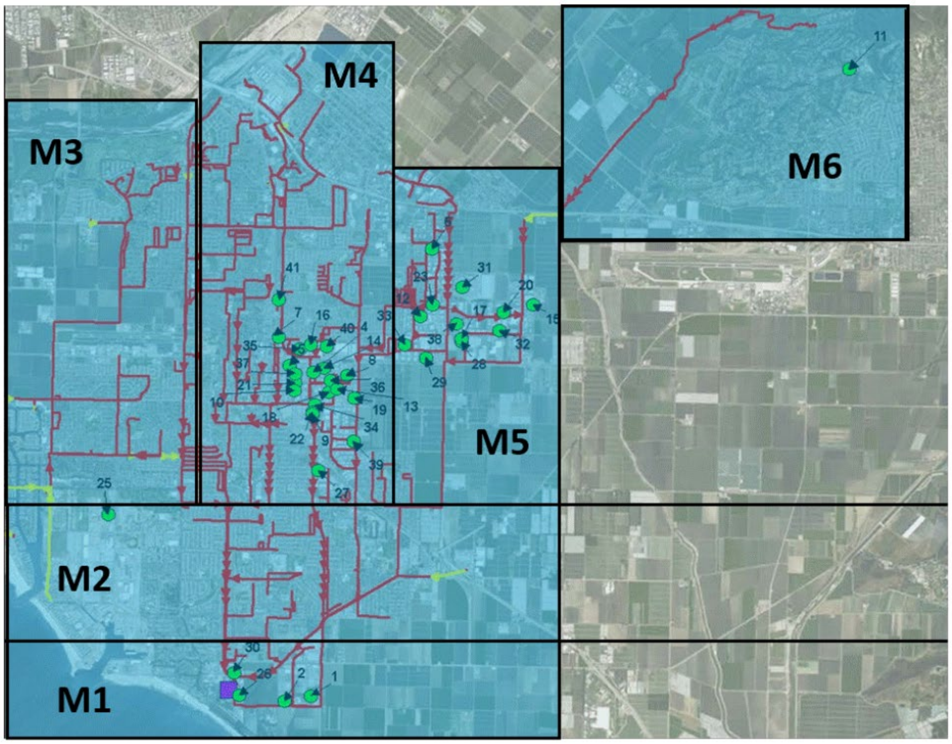


Figure 4-3. Oxnard's Collection System Monitoring Zones and Industries.

4.2.5 Outreach Program

Oxnard conducts outreach to its IUs. Industrial outreach includes regular meetings where the industries are reminded of the potable reuse project and the impacts of violating their permit requirements, as well as updates to the program about any noted slug discharges or monitoring concerns. To encourage engagement, annual awards – Enhanced Source Control Responsible Partner Awards – are given to companies who have not had a discharge violation that year.

4.2.6 Potential Opportunities for Incorporating Online Sensors into Oxnard's Source Control Program

The City of Oxnard has historically seen some challenging industrial discharges worth documenting below.

- CWT:
  - In 2014 the OWTP was witnessing high gross-beta values in the effluent that was eventually tracked to a CWT facility. That facility was receiving, diluting, and discharging

a broad range of liquid wastes. The City of Oxnard issued a cease and desist to this particular CWT, ending its operation within the City.
    - An online monitoring program would not have detected the gross-beta, but considering that the gross-beta wastes were co-mingled with other wastes, abnormal discharge patterns could have potentially been detected by a broad spectrum online monitoring system.
- Large Industrial Discharger:
    - In 2020, a large industrial discharger was improperly discharging waste to the Oxnard collection system, not incorporating the necessary pretreatment ahead of discharge leading to BOD and pH violations in their discharge.
    - Online monitoring, if in place, could have effectively detected such improper discharge.

The potential benefit to Oxnard's ESCP from an effective sensor system is shown in Table 4-5.

Table 4-5. Specific Benefits of Online Monitoring to Oxnard ESCP.

ESCP Concept	Specific Benefits of Online Monitoring to Oxnard
Industrial Dischargers	Like Ventura, an online sensor network could be installed at select nodes in the Oxnard collection system allowing for better monitoring of industrial dischargers. However, Oxnard has a much larger collection system with many more industrial dischargers than Ventura, and thus would likely not implement online monitoring at each discharger.
Enforcement Response Plan	Online sensors could provide both a better early warning system and be incorporated into a more rapid response plan. With the right network of probes, the aforementioned non-compliant discharges could be detected in real time.
Monitoring Program	Similar to Ventura, installation of a sensor network would reduce utility led analytical sampling needs and provide for a clear understanding of "baseline" and abnormal water quality within the collection system.
Outreach Efforts	Immediate communication between Oxnard staff and industrial dischargers can be made in the event of abnormal collection system wastewater quality.
Interagency Agreements	Not applicable for Oxnard, as they own and maintain their collection system.

CHAPTER 5

Review of Real Time Collection System Monitoring Results

Steinle-Darling et al. 2020—*Demonstrating Real-Time Collection System Monitoring for Potable Reuse*—deployed a network of probes into the collection systems of three agencies and demonstrated the ability to track and trend water quality including pollutant spikes and to positively impact industrial discharger water quality. The project also saw substantial challenges related to probe system accuracy, precision, and reliability as well as fouling and ragging complications. Many of the participants in Steinle-Darling et al. 2020 carried over into this WRF 5048 project, looking to improve upon the results of the first project, such as minimizing ragging and fouling and better understanding online analyzer accuracy and the ability to confidently detect changes in wastewater quality.

5.1 Overview of WRF 4908 Test Locations

The three to five Kando sensor stations were deployed at each utility monitored for EC, pH, ORP, and temperature. Initial grab samples were collected at the location of the station and online data was analyzed to establish "baseline" conditions. Following the initial monitoring periods, online sensor data that deviated from baseline triggered the collection of an automatic grab sample to be analyzed by a laboratory for confirmation. These triggered samples can ideally capture non-compliant industrial discharges or illegal industrial discharges. The three participating utilities – Ventura Water, El Paso Water, and CWS – each focused on different phenomena within their sewersheds.

Ventura's pilot successfully tracked the source of EC spikes well up into the collection system. Ventura's three sensor stations were deployed for a total of eight months, and were moved to different locations several times over the duration of the trial. During the first phase, Ventura traced EC spikes upstream in the collection system and confirmed the spikes to be attributable to residential water softener regeneration "slugs". During the next phase, a sensor station was installed downstream of the largest CIU in Ventura's collection system; the data from the sensor station did not indicate deviation significantly above baseline during the deployment period, providing confidence in prior sampling data that indicated consistent permit compliance by this CIU.

El Paso Water installed four sensor stations downstream of a copper refinery and several discharge locations used by an oil refinery. The deployment helped confirm the significant impact that the oil refinery was having to pollution spikes of hydrocarbons in a downstream manhole, as well as the compliance of the copper refinery.

CWS installed five consecutive sensor stations along a sewer pipeline route that receives discharge from a number of industries. A high pH signal from a sensor station directly downstream of a major industrial user triggered a grab sample collection that was analyzed and found to be low in organics and solids but high in metals and ammonia. The pH deviation was suspected to be a result of the major industry's (permitted) cleaning practices, and the elevated levels of pH, metals, and ammonia spurred further query by CWS to determine if the local limits or permitted levels for the industry should be adjusted to provide additional plant protection.

CWS also investigated suspected sensor failure/fouling. CWS intentionally spiked the collection system with brine on two occasions and noted that the signals from one of the downstream sensors did not register a change, even while CWS' handheld EC meter indicated persistent high EC. CWS concluded fouling mitigation for the sensors to be critical to the usefulness of sensor deployment.

5.2 Summary of Key Challenges from WRF 4908

While the results of Steinle-Darling et al. 2020 demonstrated the value of real-time monitoring, there were also challenges, some of which were a direct tradeoff (accessibility versus security, for example) that apply differently depending upon the type (supplier) of sensor systems, as illustrated in the bullets and photos below. Further discussion on recommendations to the challenges identified in Steinle-Darling et al. 2020 can be found in Chapter 8.

- Sensor Locations – Sensors can be placed within manholes or adjacent to manholes.
    - Safety - Manholes are often in the middle of streets. Accessing the manholes disrupts traffic and can present a safety risk to staff. Minimizing the frequency of access for maintenance is crucial to minimize impacts to the public and risks to the staff. Further, manhole entry is a confined space entry, which requires time and effort to safely access equipment.
        - Fouling of sensors required weekly or more frequent site visits to sensor locations, deemed too frequent by some (but not all) project team members.
    - Power - Most sewer access points, other than pump stations, do not have access to power, requiring battery power for sensor systems. Battery life and reliability thus become important factors for remote installation. Some sensor systems, including automated cleaning, cannot be sustained on battery power.
        - The sensor system from Kando was able to run entirely on battery power, including sample stations. Battery life was challenging during the demonstration, requiring frequent staff visits to replace (and recharge) batteries.
    - Security – The sensor systems and sampling stations are costly, and above ground installations are often not secure. The entire sensor systems and sampling stations are located within the manhole, which reduces security risk.
    - Depth – If the water quality sensors require a sample brought to the surface via suction, there is a depth limitation.
- Intermittent Flow - Some parts of the collection system have intermittent flow that can be dominated by periodic industrial discharges.
    - Sensors not consistently immersed in liquid appear to lose accuracy.
- Fouling and Maintenance – Considering the time and impact of accessing sensor systems, fouling of sensors (e.g., FOG, rags, hair, debris, metals, precipitates, and other solids that can buildup on sensors and affect their performance) presents challenges.
    - For Steinle-Darling et al. 2020, fouling and ragging of sensors required multiple staff visits per week at some locations to clean and de-rag system (Figure 5-2).

    - In turbulent flow regimes, sensors can be subject to breakage.
- Sensor Accuracy and Precision – The value of sensor systems as part of an ESCP includes monitoring for upset events, monitoring for compliance, and monitoring for revenue collection.
    - Pertaining to the first item, monitoring for upset events, accuracy and precision are not critical. The goal is to witness an "event" which is a water quality change significantly over the "baseline" value. For Steinle-Darling et al. 2020, the project team did see the direct benefit of the systems to monitor for large changes in water quality, such as the detection of salt spikes in the system.
    - For the latter two methods, accuracy and precision of sensor systems is needed. For Steinle-Darling et al. 2020, the project team did not gain confidence in either the accuracy or precision of the online readings, with repeated laboratory samples or other online meters showing the inaccuracy of the installed sensor systems.
- Communication – remote locations often have poor signal availability for telemetry.
    - Hardware Maintenance – In addition to the battery longevity, damage and corrosion to system components (including sensors) presents challenges (Figure 5-1).
- For Steinle-Darling et al. 2020, humidity impacts and corrosion were seen on system components. Over the trial period, hardware connectors had to be replaced due to corrosion.
- Data Security – All data for the Steinle-Darling et al. 2020 trials was handled by a third party. How data is protected is an important consideration.

Figure 5-1. Corrosion and Damage to the Probe System.

Figure 5-2. Ragging of the Sensor Probes and Connective Wiring.
Source (right photo): Steinle-Darling et al. 2020
Photo on left: Courtesy of Carrollo Engineers Inc.

5.3 Summary of Challenges from Independent Analysis of Kando Systems at LACSD

LACSD, as part of their own independent evaluation, installed and operated the Kando sensor systems at 2 locations for 6 months.

The Kando system online monitoring system demonstrated potential for:

- Determining the dynamic variation of wastewater quality.
- Characterizing wastewater quality.
- Capturing daily and longer-term wastewater quality patterns.
- Determining "out of ordinary" episodic events.
- Determining sources, their loading contribution, and discharge patterns. Sources can include industrial waste dischargers, residential patterns, and sewer interconnections.
- Providing advanced notice to the treatment plant depending on:
  - Whether the parameter of interest can be measured online.
  - The time it takes the "pollution" front to reach the plant.
- Collecting a sample automatically.

Challenges associated with the Kando system included:

- Sensor reliability: Sensors needed frequent maintenance, which can be costly and labor intensive as the number of units deployed increases. An algorithm correcting for drifting appears promising and could reduce the maintenance frequency. The benefit of the system to LACSD must outweigh the maintenance cost. EC, pH and temperature (T) were useful parameters but required: (a) proper sensor maintenance, and (b) proper sewer conditions. ORP was not reliable. Over the test period, there were no extensive periods when all four sensors, EC, pH, ORP, T, functioned properly at the same time.

- Sewer conditions: Sewer depth and velocity may not allow for monitoring at some manholes.
- Manhole location: The monitoring location must have a good signal for data transmission, and also allow for proper installation (e.g. datalogger and antenna) and maintenance (not suitable for high traffic areas).

CHAPTER 6

Data Analysis for Online Monitoring Systems within WWTPs or AWTFs

6.1 Introduction

Upsets in WWTPs caused by transient industrial discharges can lead to exceedances of discharge permits. These upsets may have human health relevance at WWTPs that are water sources for advanced treatment facilities for potable reuse. Hence, proposed regulations for direct potable reuse in California would require: “on-line monitoring instrumentation at critical locations that measure surrogate(s) that may indicate a chemical peak” (SWRCB 2021). However, the best strategy for analyzing the on-line data for accurate, proactive, real-time alerts has not yet been determined. Said another way, once the team gains confidence in the ability to collect meaningful data in real time from the sewer system or from within the WWTP (which is a focal point of this project), how can the team best apply this data to detect adverse water quality events?

One application for the data from instruments in the sewershed, WWTP, or AWTF would be for alert or alarm systems. An alarm would indicate a high degree of confidence that an event is occurring that could pose a risk to the public health, requiring the shutdown or diversion of water from the AWTF. Due to the high consequences of a false positive, alarms should arguably only be based on bench-scale data or reliable, redundant online instruments sampling the same location in parallel. In contrast, an alert would indicate a reasonable probability (e.g., greater than 50 percent) that an event may be occurring that requires attention or corrective action (e.g., increased ozone dose), but not a treatment shutdown. An alert would be more sensitive (i.e., triggered by smaller changes) compared to an alarm. Thus, an alert could trigger prior to an alarm during the early onset of an event, allowing time for corrective action and potentially preventing alarm-level changes to the treated water quality (Figure 6-1). To improve upon the status quo (i.e., data visualization monitored 24/7 by human operators), an alert system would need to detect an event before or equally as soon as it would become visually apparent to a human operator. An alert system with this capability would: (1) allow corrective action to be conducted more promptly or with greater confidence and justification, and (2) serve as a redundant measure to human monitoring.

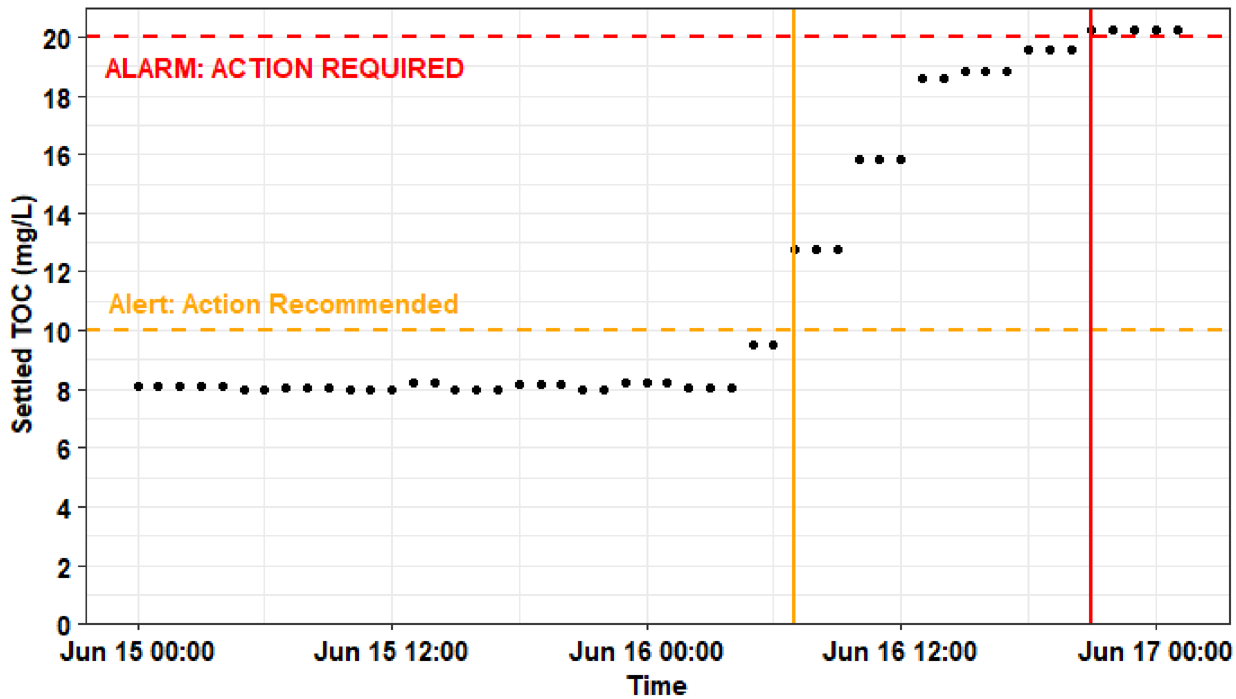


Figure 6-1. Example of an Alert and Alarm Based on a TOC Threshold.
Settled TOC data is from the HRSD proof-of-concept below.

Machine learning could be applied for alert systems in drinking water, wastewater, or reuse. Machine learning is the study of algorithms that improve automatically through experience with data. Specifically, supervised machine learning (SML) creates mathematical models to predict outputs based on a set of labelled input data. In the context of wastewater and drinking water treatment facilities, input variables could include water quality variables, such as pH, and operational information, such as ozone dose. SML requires a training dataset to construct models and a testing dataset to evaluate and compare their accuracy. The training and testing sets must both have known outputs or labels for the models to be constructed and so their predictive accuracy can be compared in a meaningful way. Labels in the water context could be categories, such as "Normal" or an "Industrial Discharge Event", or numerical, such as percentage of influent coming from industrial wastewater. Once the SML models have had their accuracy confirmed on the testing set with known labels, they can then be applied in the field on new data with unknown outputs. This training and testing procedure avoids overfitting, which is when an increasingly complicated model more closely matches the data upon which it was trained, but makes less accurate predictions with new data.

SML models could be more accurate for detecting and categorizing upsets than simpler alternatives, such as fixed thresholds on single variables (e.g., pH above 8 indicating an industrial discharge event). SML models can recognize high outliers on a single variable—while other variables remain near average—as likely instrument malfunctions or maintenance rather than true water quality events. Contrastingly, if all variables differ from the average only slightly but in directions associated with a particular type of upset, SML models could detect low levels or early onsets not yet apparent to human operators or fixed threshold-based alarms. Additionally, unlike a fixed threshold on a single variable or calculated metric, many SML models can categorize data into three or more categories. This could be beneficial, for example, for distinguishing among industrial discharges from different sources. Furthermore, SML models are nonlinear and more flexible than thresholds. That is, thresholds essentially categorizing anything within a rectangular space as Normal, and anything outside that rectangular space as an Event (Figure 6-2A). In contrast, SML models, such as k-nearest neighbors or support vector machines with radial basis kernels, can draw boundaries as any variety of complex, curving shape as dictated by the data (Figure 6-2B).

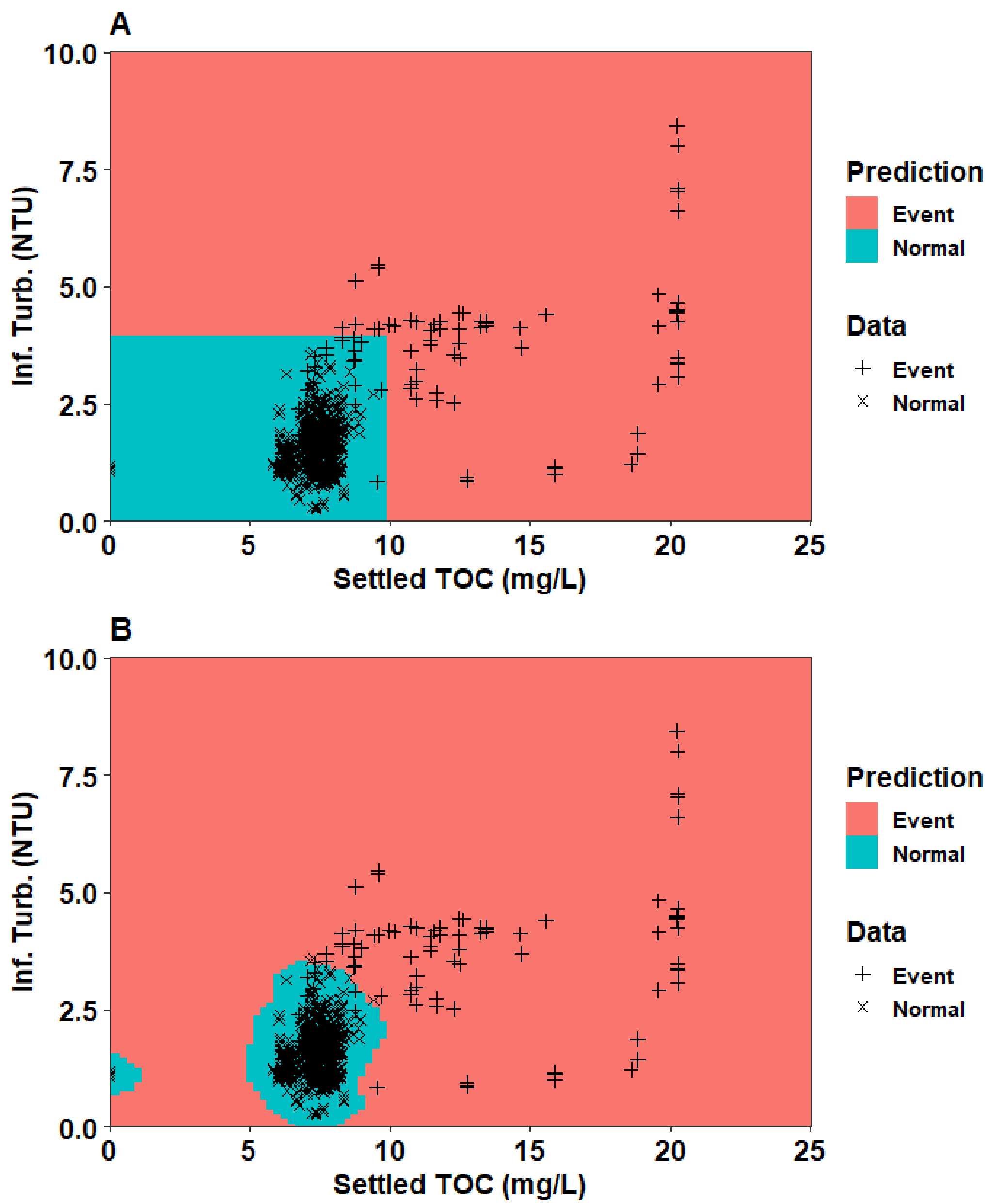


Figure 6-2. Water Qualities that would be Predicted as Event or Normal According to (A) Fixed Thresholds or (B) A Support Vector Machine with Radial Basis Kernel.

X's represent Normal data and +'s represent Event data. Blue areas represented Normal predictions and Red areas represent Event predictions. Influent turbidity and settled TOC data are from the HRSD proof-of-concept below.

In this chapter, SML was applied to historical online sensor datasets from two utilities as a proof-of-concept for SML-based alert systems. There were important differences between these utility datasets (Table 6-1). The online sensor data from CWS was from a grit screened primary effluent. The online sensor data from Hampton Roads Sanitation District (HRSD) was from multiple locations in the treatment system, but mostly from secondary effluent or advanced purified water after ozonation or biofiltration. Thus, the water being monitored at HRSD was cleaner and more equalized, and subject to both less instrument fouling and fewer non-event water quality fluctuations.

Table 6-1. Differences Between HRSD and CWS Datasets Used for SML Alert Systems.
Normal data in this table refers to datapoints that were not part of an industrial event.

	HRSD Case Study	CWS Case Study
Sensor Locations	Mostly secondary effluent or advanced purified water	Grit screened influent
Number of Sensors or Variables	30	10
Industrial Discharge Origin	Real/Unknown	Simulated/Known
Industrial Discharge Types	1	2
Classification Type	Binary (Normal v Event)	Multiclass (Normal v Spike 1 v Spike 2)
Data Frequency	Hourly	10-minute
Total Sample Size Used	878	5275
Percent “Normal” Data in the Testing Set	76.7%	99.5%

Instruments located in the sewershed or influent like at CWS would require more maintenance but could provide more advanced warning in terms of the time between first detection and when the detected water plug would exit the treatment facility. Since HRSD had a “cleaner” dataset, one might expect SML models to achieve greater accuracy, but simpler methods or human judgement might also perform relatively well. For example, if there is less random noise in a variable, the early onset of a true, new, upward trend would become visually clear to a human observer sooner. Furthermore, in HRSD’s dataset, all variables had data for all times. In CWS’s dataset, certain sensors had missing data for much of the time. Sensors with missing data could be omitted, sacrificing number of variables but increasing the usable number of observations. This tradeoff needed to be investigated for CWS, but not HRSD.

The noisiness of the data can also affect which machine learning algorithms perform best (Atla et al. 2011). At CWS, the industrial discharge events were simulated by spiking known contaminants into a flume, so the true beginning and end of the events were fully known. At HRSD, the industrial discharge events were real and from an unknown source, so their beginning and end were labeled based on human judgement after visual interpretation of the data. Furthermore, the HRSD dataset had more variables while the CWS had more sample size of observations. Thus, HRSD would be considered a relatively “wide” dataset while CWS would be considered a “long” dataset. The length vs wideness of the dataset can affect, for example, which SML algorithms can be trained quickly (Lindgren et al. 1993; Rännar et al. 1994). Furthermore, the HRSD data was labelled and predicted among two classes (Normal vs Event) while the CWS data was labelled and predicted among three classes (Normal vs Spike 1 vs Spike 2, where Spike 1 was a blend of bleach, NaCl, and NaOH, and Spike 2 was humic acid). Lastly, the CWS had a higher percentage of data labelled “Normal,” making it inherently easier for models to achieve higher accuracy (overall percent of correct predictions). For example, a model that predicts every day is not Christmas would be over 99 percent accurate, but not useful. So, higher test set accuracies on the CWS dataset should not be perceived as indicating that specific models or SML methods in general were more successful on that dataset. Considering all the above, it is not surprising if the overall success or best SML methods differ among the datasets.

The SML analysis with HRSD’s and CWS’s datasets are described in detail in Appendix B and Appendix C, respectively. Key conclusions from these analyses that could be broadly applicable are summarized below.

6.2 HRSD Machine Learning Case Study Conclusions

- The model type Boosted Tree (bstTree) had the highest testing set accuracy for this dataset. bstTree would have detected the industrial event in about an hour with zero false positives over the 5-day testing set. So, bstTree would have been selected for future monitoring and alerts among the SML models investigated in this study.
- If a fixed threshold had been set for influent UVT using a data-driven, SML-style training approach, this could have resulted in a testing set accuracy of 98.3 percent. This accuracy would be below bstTree, but only by about 1 percent. This influent UVT trained threshold or the actual threshold for influent turbidity both would have detected the event in about two hours, one hour slower than bstTree. However, the actual threshold on influent turbidity had two false alerts within the five-day testing set unrelated to the industrial event.
- The most beneficial preprocessing method differed among the SML model types. Two models performed best without preprocessing, one with principal component analysis (PCA), and three with raw data and differences from the rolling median. So, one best preprocessing method could not be universally recommended based on this analysis.
- In many cases, some variables could be omitted to decrease training time without loss in accuracy. However, like preprocessing, the optimal set of variables depended on the model.
- Certain SML model types (e.g. ORFsvm, rfRules) had testing set accuracies that depended on random chance. So, the accuracy of these models would be more uncertain in full-scale applications, even with appropriate validation and testing procedures.

Looking to the future, the team would make the following recommendations:

- As next steps to engineer an accurate, practical, SML-based alert system at HRSD, the team would recommend repeating the above analyses, but with greater sample size, including multiple instances of the industrial events in the testing set. This would provide greater confidence about the relative performance of the models, particularly whether the highest-performing model would be best for detecting all events of this type, not just the individual event in this testing set. After that, a small number of high-performing SML models could be piloted in real-time, until an additional event occurs. The time until first detection of the SML models could then be compared in the field against human monitoring and other alert approaches.

- Since Event and Normal datapoints in this dataset were distinguished based on human judgement, the best the models could possibly do would be to match—not exceed—human judgement. On the other hand, a human monitoring the data in real-time might not have concluded that an event was occurring as soon as a human evaluating the whole dataset retrospectively. In future research, machine learning for wastewater or reuse alert systems could be achieved by simulating industrial discharges in a pilot or flume like the one at Clean Water Services (CWS) (see Appendix C). Alternatively, real full-scale industrial events could be labelled objectively if the industrial source is known and keeps records of discharge flow (e.g., the landfill that discharges limited quantities of leachate to the WWTP that feeds SWIFT RC) (Gonzalez et al. 2021; Nading et al. 2022).
- A limitation of SML-based alert systems is that they are designed to detect events of a known, previously documented type. If a new type of industrial discharge were to occur associated with difference responses from the online instrumentation, this may or may not trigger an SML-based alert. Changes in the water quality pattern at the AWTF during industrial discharge events could also occur due to changes in the treatment operation response at the WWTP. So, a strategic solution would be to employ both SML-based alerts, and alerts or alarms based on fixed thresholds. This would combine the sensitivity of SML with the generalizability of thresholds. These additional thresholds could be set based on training set data, health-based goals, or operational considerations. Advanced multivariate statistical methods for fault or outlier detection other than SML also merit further research in the context of wastewater and reuse (Klanderman et al. 2020).

6.3 CWS Case Study Conclusions

- pcaNNet, bagEarth, and C5.0Rules were the most accurate models for CWS's dataset.
- pcaNNet, bagEarth, and C5.0Rules had zero false positives and at least one detection of both spike types in the testing set.
- Considering the relative importance of types and locations of errors, pcaNNet, bagEarth, and C5.0Rules were equally successful.
- However, considering training time and interpretability, the team would recommend C5.0Rules for this application.
- Omitting ECD ORP to increase the effective sample size of the other variables did not decrease bagEarth's accuracy and increased C5.0Rules's accuracy. So, omitting variables with missing data could be a worthwhile tradeoff, especially for relatively non-informative variables like ORP at CWS (see Appendix A).
- Preprocessing with PCA or median-based methods to remove diurnal patterns and drift did not increase the testing set accuracy of pcaNNet, bagEarth, or C5.0Rules. So, preprocessing might not always be needed or beneficial, even with relatively noisy data.

6.4 Overall Conclusions and Considerations

- SML models using online sensor data from influent or treated wastewater could be applied for useful, practical alert systems.
- For both datasets studied, multiple SML models performed adequately in terms of accuracy, sensitivity, false positive rate, and Cohen's Kappa.

- However, there were key differences in the results between the datasets in terms of the most accurate model or the helpfulness of the preprocessing methods tested. This indicates the importance of site- and application-specific model comparisons and the need for more research on data preprocessing techniques for online water quality data.
- Models for the CWS influent dataset had higher overall testing set accuracy. However, this was because there was a higher proportion of data labelled Normal. So, a model for the CWS dataset with zero false positives and 60 percent sensitivity would have higher accuracy than a model for the HRSD dataset with the same false positives and sensitivity. Models of the HRSD treated effluent dataset performed better in other performance metrics (e.g., sensitivity, balanced accuracy, and Cohen's Kappa). So, as expected, it was easier for models to make predictions for the less noisy treated effluent dataset. However, an alert system using data from sensors in influent would provide earlier warnings. In either case, models could detect at least half of the event datapoints with zero false positives over the course of the testing sets.
- While the SML studies in this report were proofs-of-concept only, the next step would be to test the models on live, real-time data. Models could then be continuously retrained with additional data and retested on new data in real-time. While the highest performing model would be applied for the alert system, other high-ranking models could also be continuously retrained and retested. These alternative models might then be implemented if they perform best with the additional sample size.
- For the scenario with two categories (HRSD, Normal vs Event), SML performed only slightly better than a data-driven fixed threshold on a single variable. However, given the noisiness in the CWS dataset and the fact that there were three categories (Normal, Spike 1, and Spike 2), it is difficult to envision a "simple" approach that could achieve similar accuracy, specificity, and low false positives.
- Despite overall satisfactory performance, SML models for the CWS dataset did not detect some of the spikes and did not detect suspected real industrial events which they were not trained to detect. This demonstrates the importance of a hybrid approach with both SML-based alerts and threshold-based alarms.
- Simulated spikes can provide more control and allow more definitive datapoint labelling. In some cases, simulated spikes may provide more true positive sample size or events for models to detect. On the other hand, data from past real events may better prepare models for detecting future real events.
- Developing a ML model requires skill in data science and coding. Implementing one in practice also takes expertise in information technology, SCADA, or controls. For a utility not only to implement but also maintain, refine, or retrain such a model for an alert system could require either a dedicated staff member or a Data-as-a-Service (DaaS) contract.
- Utilities may also need to consider whether to house the model in the cloud or onsite (e.g., an edge server). Modeling in the cloud facilitates model refinement by a remote employee or within a DaaS contract. Onsite modeling provides the greatest cybersecurity.

CHAPTER 7

Potential Implementation of Real-Time Monitoring for ESCPs

Three partner utilities (Morro Bay, CWS, and LACSD) were interviewed to determine how their ESCPs could benefit from real time monitoring, utilizing the results of the demonstration work at each site. These utilities provide a useful comparison ranging from a very small collection system with minimal industry (Morro Bay), to a mid-sized system with some industry (CWS), to a very large multi-jurisdiction collection system with a large number of industries (LACSD).

Below, an overview was provided on the existing ESCP at each utility. Then, looking to the future, the three different agencies were interviewed to best understand their perspective on use of real-time sensors, how an online monitoring system could best benefit them and how would it be implemented. These interviews considered the points below and were used in the development of the ESCP Framework (Chapter 8):

- Goals of online monitoring.
- Preferred locations of online monitoring.
- Operations and maintenance.
- Sampling.
- Permitting and enforcement.
- Security and integration.

7.1 Morro Bay

As part of Nading et al. 2022, *Review of Industrial Contaminants Associated with Water Quality or Adverse Performance Impacts for Potable Reuse Treatment*, a detailed case study of Morro Bay's developing ESCP was completed. Excerpts from that case study are summarized below.

7.1.1 Existing ESCP Overview

Morro Bay California is implementing an approximately 1 mgd potable reuse project, called Our Water (City of Morro Bay 2020). The project includes the construction of a new WWTP and advanced treatment facility. Morro Bay is a small town with few industrial users (IUs), and does not currently have a formal, EPA-approved pre-treatment program. The City does currently implement a small pretreatment program, including having a list of prohibited substances from discharge into the sewer, policies and procedures to enable the City to address the potential need for pretreatment of conventional pollutants, and several specific discharge restrictions.

The City has developed, but not yet implemented, an ESCP. Morro Bay's ESCP includes an industrial waste survey (IWS), Local Limits, a Sewer Use Ordinance, Raw Wastewater Sampling, Narrative Limits, IU Discharge Sampling, Source Control Outreach Program, Enforcement Response Plan, Source Mapping, a Funding and Resources Report, and a Collection System and Treatment Plant Monitoring Program Manager.

As part of the development of the ESCP, a thorough evaluation of all industrial dischargers was completed (this is the point of the IWS). This IWS concluded that there were essentially two industrial dischargers of concern, as follows:

- Bottled Water Company - Salt is clearly an important challenge for Morro Bay, as high salt loads must be removed by the City's future new RO system. In particular, the bottled water company currently discharges variable and sometimes high salt spikes into the sewer, which will challenge the production capacity of the future new RO system. The bottled water company has been collaborative on future solutions, with a likely result that discharges will be either equilibrated into the discharge to the sewer or removed entirely from the sewer.
- Industrial Laundry - The industrial laundry discharges more than 25,000 gallons per day, which is greater than 5 percent of the City's wastewater flow. The laundry can provide both high and low pH wastewater, ranging from a pH of 6 to 10. In some cases, they overdose acid. As part of the ESCP, the laundry will need to install a dedicated and reliable pretreatment system, to which the laundry has indicated a willingness to follow the ESCP requirements.

The argument could be made that Morro Bay's industrial source control can be small and cost effective, focusing upon only these two dischargers. However, current California regulations and a precautionary principle lead to a comprehensive grab sampling monitoring program as part of the ESCP.

Morro Bay prepared a Funding and Resources Report, which specifically details what staff is needed to implement the ESCP, what sampling will be done, who will do it, and how frequently it will be done. In total, the effort is extensive, estimating approximately $150,000 in annual personnel costs, $60,000+ in new monitoring equipment, and annual laboratory costs of greater than $5,000. The City notes that today, before the ESCP and potable reuse program is implemented, staff is already the City's highest expense. The inclusion of online monitoring, if proven accurate and reliable, will reduce staffing and sampling needs in the ESCP and provide a real-time early warning system ahead of advanced treatment for potable reuse.

7.1.2 Future Use of Online Monitoring Systems to Detect Pollution Events

Key points from the exit interview with Morro Bay were:

- Maintenance:
  - Daily maintenance of a single probe system within the WWTP is manageable.
- Location of Probes:
  - Inside the WWTP – For a small utility such as Morro Bay, having the probes out in the collection system may not provide much greater value than being within the WWTP, where O&M is the most readily done.
  - At Major Industries – There may be value in placing targeted probe systems at the discharge of some industries. However, for a small community where there are only a few SIUs, the value is limited.
- Probe Accuracy:
  - Probe accuracy was within acceptable range.

- Actionable Value:
  - From Morro Bay's perspective, the immediate actionable value of the probe system has yet to be demonstrated. Potential value includes:
    - Water quality alarm and response. If it can be demonstrated that one or more online values correlate to a risk to water quality, then a probe system could have immediate value to provide alarm and diversion.
    - Automated sampling. If the sensor system was tied to a reactive sampling system, significant variations in water quality could result in sampling and thus determination of the constituents within subject water.
    - Regulatory. The costs of an ESCP are substantial for small utilities such as Morro Bay. If the installation of a real time monitoring system allowed for a reduction in ESCP sampling (e.g., within the collection system, at IUs, etc), then a direct value could be assigned to the online system.

7.2 Clean Water Services

7.2.1 Existing ESCP Overview

CWS's four water reclamation facilities serve a population of over 600,000 people and 75 permitted IUs, some of which are very large. These IUs include many high-tech and semiconductor manufacturers, food processors, a landfill, metal finishers, wineries, and many more which contribute approximately 13.3 mgd of the average dry weather total flow of 57 mgd to the facilities. CWS has a non-potable reuse program that currently applies 1 mgd to agricultural fields, golf courses, and other irrigation needs. This is planned to expand to 5 mgd in the next 4 years. In addition, CWS has a small pilot-scale intermittent potable reuse of approximately 7,200 gallons per day delivered to a brewery. Biosolids from the water reclamation facilities are land-applied to agricultural fields, and struvite is recovered and sold as commercial and residential fertilizer.

CWS implements a robust EPA-approved source control program to protect the water quality of the facility discharges, reuse water, and biosolids, as well as the personnel and infrastructure at the facilities and the collection system. Local limits were developed and are regularly updated to expand upon the minimum requirements for EPA's source control program. CWS regularly conducts inspections, collects samples, and manages the permits for each IU, and requires permitting evaluations for new IUs. Some IUs are also required to collect samples regularly or have continuous monitoring and report these to CWS. CWS also has a local cost-recovery program that surcharges IUs with COD greater than 800 mg/L and/or TSS greater than 400 mg/L. Inspections are regularly conducted to assess industrial and commercial users with excess FOG in their discharges, and requirements are in place for management practices to reduce FOG in the discharges.

The source control program is also in charge of responding to pollution events at the treatment plants. These occur somewhat sporadically, and the sources have been difficult to detect. The facilities notify the source control personnel of an event once it is detected, and source control personnel immediately collect samples from the headworks and key manholes in the collection system which are analyzed by the laboratory. However, in most cases, whatever caused the negative effects to the facility is not evident in the collected samples either because the 'slug' of pollutants has already passed or because the compound(s) responsible are not identified. To overcome these deficiencies in the responses to pollution events, in recent years CWS has been developing a continuous sensor network in the collection system to detect sources of the events and provide advanced warning to the facilities. Currently, four telemetered pH sensors are permanently installed at key manholes in the collection system of the Forest Grove facility which has been most frequently impacted by pollution events. However, use of continuous sensors has already been used to track down and mitigate two consistent sources of pollution events that affected the facility. These include frequent very high pH events caused by a beef jerky manufacturer, and frequent high nitrate events caused by a circuit board manufacturer.

Source control personnel regularly work with IUs to address other pollutants of concern. The pollutants of most concern for CWS have been metals, nitrates, azoles, peroxide, high-strength COD, pH, temperature, and several currently unregulated pollutants (e.g., PFAS). As part of the expansion of the non-potable reuse program to 5 mgd by 2025, additional pollutants have become pollutants of concern including fluoride and TDS. CWS is working with a large semiconductor IU with high fluoride in their discharge to ensure their future growth does not threaten the planned expansion of our non-potable reuse program. CWS is requiring the IU to decrease their fluoride loading and working with them to help them do so.

CWS has not yet developed a formal ESCP because our potable reuse program is still only in the pilot-phase. While two of our facilities are advanced treatment facilities and a third includes polishing from a constructed treatment wetland, no advanced treatment to potable standards is currently done outside of the pilot facility. The water quality necessary for our non-potable reuse programs is supplied by our existing water reclamation facilities. However, as our non-potable reuse and nutrient recovery programs continue to grow, online monitoring is becoming an increasingly important element of the source control program. CWS is working on expanding the network to many other locations in the collection system and adding additional sensor types to the locations. To help decrease the costs of installing and maintaining such a network, CWS has been developing and testing different technologies including a customized sensor holder to reduce maintenance frequencies and open source dataloggers and sensors to decrease equipment costs.

7.2.2 Future Use of Online Monitoring Systems to Detect Pollution Events

- Goals of online monitoring:
  - The goals of online monitoring should be reliable data, cost-effectiveness, moderate O&M, and actionable data.
  - Specifically for CWS, goals included customizability, integration into the existing SCADA system, mobility, and feasibility for manhole installations.
- Preferred locations of online monitoring:
  - CWS's preferred location was within the collection system, followed by headworks then primary effluent. The degree of advance notice drove this preference.
- Operations and maintenance:
  - In order, the biggest O&M challenges were ragging/FOG, equipment malfunctions, loss of signal, and vandalism.

- Despite being CWS's preferred location, the collection system resulted in the biggest maintenance challenges.
- Sampling, Permitting, and Enforcement:
  - Integrating traditional sampling and online sensors allows for faster understanding of water quality.
  - Real-time monitoring could be implemented to support Local Limits.
    - Rather than on the effluent of specific known industries, it could be more useful as a network throughout the collection system, leading to targeted sampling within identified problem areas.
    - Real-time monitoring could determine where compliance sampling should be done, leading to site-specific limits being installed.
    - However, the accuracy of online sensors might not be enough for enforcements, so the local limits might then be based on traditional sampling.
    - Real-time monitoring data can be used in discussions with industries to help explain how it was determined that the industry was responsible.
  - For example, CWS was experiencing pH spikes. The s::can at CWS's influent detected nitrate spikes. Composite sampling and benchtop analyses confirmed the nitrate spikes. The responsible industry was traced, and turned out to be discharging nitric acid, causing both the pH and nitrate spikes.
- Integration:
  - Real-time monitoring involves large data transfers. Dashboards may need to be complex to convey all of this information. Ideally, the data would be brought into one system, so the data loggers and other tools need to be uniform, or ideally open source.
- Cybersecurity:
  - CWS's sensor network would have three vectors of concern: Device, Transmission, and Endpoint
  - CWS is considering leveraging their Microsoft managed key for data encryption, then adding a second layer of encryption with Azure Key Vault over their data at Representational State Transfer.
  - This approach would enforce a managed identity with CWS's portal that could quickly create, rotate, disable, and revoke access should an issue arise.
  - To add security within the Cloud, CWS's queue will be wrapped in a virtual network, network security group, and Firewall to mitigate who can access their resources and what resources this portion of their infrastructure can touch.
  - Transmission is beyond control to a degree. The traffic should be encrypted before entering the utility's vendor network, then encrypted again once on the sensor company's network.
- Benefits:
  - CWS has already seen better plant performance, less downtime, better source tracking, and better local limit enforcement.
  - The research conducted as part of this WRF project has already helped CWS determine the value and challenges of different probe systems. CWS will continue to use the constructed test flume (Section 3.1, Appendix A).
  - Operators see the value of some the analyzers, e.g., nitrate.

- Areas for Improvement:
  - Ongoing challenges include standardizing or automating maintenance and overcoming operator skepticism about analyzer accuracy. Long-term performance reliability demonstration is needed.

7.3 Los Angeles County Sanitation Districts

As part of Nading et al. 2022, *Review of Industrial Contaminants Associated with Water Quality or Adverse Performance Impacts for Potable Reuse Treatment*, a detailed case study of LACSD's long-running ESCP was completed. Excerpts from that case study relevant to the use of online monitoring systems are summarized below.

7.3.1 Existing ESCP Overview

LACSD collect waste from 80 square miles within Los Angeles County, serving 5.7 million people. As of 2019, LACSD oversee 378 CIUs, 945 SIUs, and 1552 other industrial dischargers. Average annual wastewater flows of 390 mgd are treated at 11 wastewater treatment/reclamation facilities.

LACSD support a broad range of water reuse projects, including potable water reuse projects – the Montebello Forebay Groundwater Recharge Project and the Metropolitan Water District of Southern California Advanced Purification Center Demonstration Project. As such, LACSD have an extensive and successful ESCP.

Historically, LACSD have seen upsets at several of their plants, including Pomona and Whittier Narrows WRPs. The aggressive pretreatment program now in place has minimized plant upsets. As one recent example, Sanitation Districts staff were able to detect a water quality problem at the influent to one of their WRPs and track the pollution back to its source and stop the violation within 24 hours.

With such a large and geographically distant sewer collection system, implementation and integration of a sensor network is anticipated to be a large challenge, one that at least early on would add instead of reduce cost and effort. As such, Sanitation Districts staff is examining how online sensor systems could provide an early warning system ahead of biological treatment and potable water reuse, placing sensor systems into the primary effluent.

7.3.2 Future Use of Online Monitoring Systems to Detect Pollution Events

Key points from the interview with LACSD are shown below.

- Location of Probes:
  - Inside the WWTP - Due to challenges with maintenance and installation of probes within the collection system, the first location for further testing should be within the WWTP, either in screened raw wastewater or in primary effluent. Having the probes in this location provides an early warning system to the biological process and would allow for the rapid collection of samples to characterize the water quality.
  - At Major Industries – Placement of sensor systems immediately downstream of major industries provide direct benefit in terms of industrial compliance and potentially enforcement (if sensors are sufficiently accurate). Placement of probes directly

    downstream of industry also provides a deterrence effect, leading to better industrial effluent because the industry is aware that it is being monitored.
  - Within the Collection System – Based upon the current performance of tested sensor systems, LACSD believe that distributed sensor systems in their large collection systems is not viable from an O&M and cost standpoint.
- Probe Accuracy:
  - The ideal sensor system measures chemicals of direct relevance either to water quality, to biological performance, or to industrial compliance. Precision and accuracy of the probes is relevant to all three listed values but is of most importance pertaining to industrial compliance.
  - The secondary value of the probes, which remains significant, is the ability to document changes in water quality, even if the probes are not fully precise or accurate.
- Automated Sampling:
  - Having an automated sampling system, based upon real time data excursions (called pollution events by Kando), provides significant benefit toward capturing and understanding water quality variations. This could only be realized if the sensors remained responsive for adequate periods of time.
- Mobility:
  - A probe system that can be readily collected and moved to different locations allows for rapid deployment to track pollution at the WWTP or in the collection system back to the source.

CHAPTER 8

Real-Time Monitoring Framework

This chapter presents recommendations for how utilities could implement real-time monitoring to support enhanced source control efforts for potable reuse programs summarized within a twelve-step framework. This framework incorporates the results of the preceding chapters and integrates recommendations from other guidance documents (e.g., Steinle-Darling et al. 2020, Liggett et al. 2018, Nading et al. 2022) and the project team's experience.

The objectives of the real-time monitoring framework are to:

- Help utilities understand how to incorporate real-time monitoring into existing pretreatment program or ESCP activities.
- Incorporate best practices from this research and other research projects into their programs.

The proposed framework steps are separated into three phases that represent planning, design, and operations. While it can be beneficial to implement real-time monitoring prior to implementing potable reuse, this framework is intended to be implemented at any stage of a project or to improve ongoing operations. Figure 8-1 presents the three phases and an overview of the framework. The following subsections present each of the three primary phases and the steps recommended during each phase.

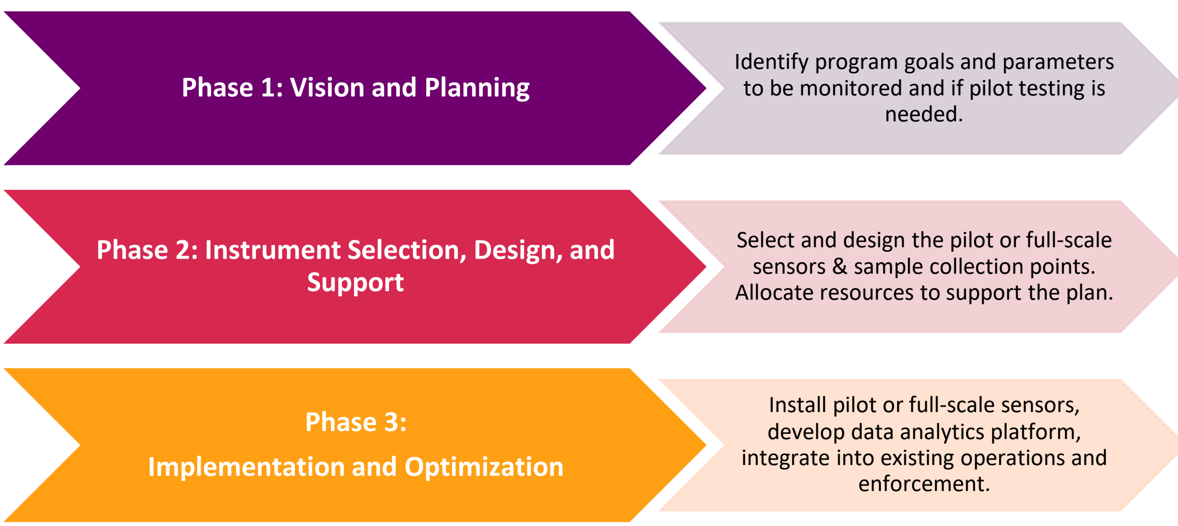


Figure 8-1. Phases of the Industrial Enhanced Source Control Program Framework.

8.1 Phase 1: Vision and Planning

Figure 8-2 presents an overview of Phase 1, which includes four steps and is intended to establish the goals, monitoring parameters, and overall approach for implementing a real-time monitoring program for ESCPs. These are all high-level steps that should occur prior to design and implementation and should align and support the goals of a utility's potable reuse ESCP. The following sub-sections provide a more detailed explanation of each step.

PHASE 1: VISION AND PLANNING

1. Identify clear goals and targets of online monitoring program
2. Select the water quality parameters to be monitored and identify criteria for sensors and monitoring locations
3. Consider a pilot test (highly recommended, especially for larger/more expensive programs)
4. Identify online monitoring locations and approach
   - Selecting sensor locations
   - Static vs. flexible approach

Figure 8-2. Phase 1 Framework Steps.

8.1.1 Step 1 – Identify Clear Goals and Targets of Online Monitoring Program

The successful implementation of a sustainable online monitoring program is more than a technology implementation solution. The program also needs to consider the workforce (end-users), how the program will fit into the organization's culture, and the level of training required for staff (operators, maintenance, water quality, engineering, etc.) to use the new technology to enhance their jobs and improve operations. Once the program is operational, an appropriate amount of governance is needed support the program. Developing a program that integrates the workforce, governance, and technology into a common framework will lead to a highly successful program.

This step is intended to distill the vision of the online monitoring program into realistic and attainable goals. Online monitoring can appear to be a "silver bullet" solution for source control, but past studies and the results of Chapters 3 through 7 have indicated that if the monitoring program does not have clear goals and responsibilities outlined, the sensors (a.k.a, analyzer or probe) will quickly become expensive maintenance items that do not provide value to the utility and will not be supported by operations. The utility should consider forming an inter-departmental team for the development of the vision. This will help to develop cultural buy-in to the program that is critical for long term success.

> The real-time monitoring program requires clear goals and direct action associated with the sensors. If this is lacking, the program will consume resources without providing value.

Prior to purchasing sensors, it is very important to determine the reasons why the data is needed and how the data will be actionable or informative. Potential value in providing real-time monitoring to support ESCPs include:

- Implementing critical control points (CCPs) at the WWTP or AWTP so that immediate action can be taken if the measured values are outside of the allowable range.
- Advance notice of influent WWTP slugs that will allow the operational staff to take early action.
- Determining potential sources within the collection system of WWTP influent contaminant slugs.
- Confirmation that treated water regulatory requirements are being achieved.
- Enforcement of limits or development of local limits, either at strategic locations in the collection systems or at the discharge of industrial dischargers.
- Data collection, research, and creation of actionable information for future projects.

The stakeholders of the ESCP should discuss the goals of each online sensor and ensure that it will provide value to the program and is worth the resources needed to make its installation successful. It is okay to have some sensors that are research focused and do not have assigned direct actions, but the utility is advised to limit the number of non-actionable sensors. If the utility already has some installed sensors and is considering adding more, it is still recommended to review the goals of the new sensors with the rest of the ESCP stakeholders. The installation of additional online sensors can in some cases provide diminishing returns (e.g., when new sensors increase the O&M costs without providing significant value). This research project has documented that the time and resources needed to maintain these sensors can be substantial in some cases, so it is vital that clear goals are defined for each individual sensor.

This is the first step of the framework as it is the most important. The O&M requirements for these sensors can be extensive; Chapter 3 of this report documented that many sensors require weekly (or even more frequent) cleaning to be reliable. The utility is cautioned to go into this decision with a full understanding of the costs and resources needed and impact on the operations and maintenance staff to be successful.

8.1.2 Step 2 – Select the Water Quality Criteria to be Monitored

Once the goals and data uses have been identified, the next step is to select the water quality criteria that can achieve those goals and uses within the ESCP. This step may seem obvious and in many cases it will be. If the utility has decided that conductivity must be monitored for a CCP, an online conductivity sensor should be installed. But there will be cases where direct measurement of the targeted contaminant is not possible. In these cases, the utility should consider if a surrogate measurement can be informative. Chapter 2 provides a review of available online sensors. Two examples of where this may be relevant are as follows:

- The utility has identified that infrequent WWTP influent slugs of metals results in WWTP effluent violations. While direct measurements of the metals are not possible, the utility has identified via grab sampling that the slugs correspond with high conductivity events and a conductivity sensor can be reliably installed as a surrogate.
- Influent TOC slugs into an AWTP exceed the CCP target and require the AWTP to divert and go offline. Grab sampling during these events has identified that influent total phosphorous

(TP) slugs precede the high TOC events and cause poor WWTP performance. An online TP sensor can help notify and prepare for subsequent events.

Each water quality parameter that is measured should be directly linked to one of the goals of the real-time monitoring program (Step 1). If there is not a reliable sensor that can achieve the stated goal (Step 5), it is not recommended to install a different sensor; in this case the utility should consider different ways to gather information for that specific goal (e.g., grab sampling, composite sampler, etc.).

8.1.3 Step 3 – Plan A Pilot Test for Select Sensors

In this step, the utility is evaluating the value of a pilot test. Given the challenges associated with implementing online monitoring in wastewater, pilot testing is strongly recommended to gain confidence in the ability of the specific types of sensors to meet the goals and requirements identified in Steps 1 and 2. Pilot testing should always be performed for sensors that utility staff are unfamiliar with, where multiple sensors for a single parameter are being evaluated, and for large programs where the utility intends to purchase and install many sensors. Note that it may be possible to request a trial, rental, product demonstration, or other options from the vendor prior to purchasing some instruments.

The costs of a pilot test are easier to justify for larger, more expensive programs; however, the case study interviews from Chapter 4 of this report recommended that pilot testing for smaller programs is also beneficial in reducing the risk that the real-time monitoring program is unsuccessful. The value of pilot testing online sensors include:

- Troubleshooting and improving the installation in a challenging location (e.g., manhole, sewer, force main, etc.) prior to purchasing several sensors for installation in similar locations.
- Testing a specific sensor prior to purchasing multiple of the same sensors.
- Verifying that power, communication, and integration into the utility's system is achievable.
- Identifying the resources needed for sensor maintenance, data collection, data verification, and data analytics to estimate the total resources needed for the full real-time monitoring program.
- Identifying if a certain site or sensor is infeasible and needs to be reevaluated.
- Verifying that the sensor can actually detect the events that it is intended to detect.

If the utility selects to pilot test specific sensors, phase 2 of this framework is still relevant. The utility should follow Framework steps 5 through 12 (Phases 2 and 3) for the pilot sensor and then return to Framework Step 4 (Phase 1) when implementing the remaining sensors and proceed through the remainder of the framework. Establishing the value of the pilot sensors is essential to appropriately selecting the rest of the sensors so it is important that dedicated resources are committed to the success of the pilot.

8.1.4 Step 4 – Identify Online Monitoring Locations and Approach

The intent of this step is two-fold: (1) to identify the locations within the collection and/or treatment systems to implement monitoring, and (2) to identify whether the instruments will be installed in a fixed location (static approach) or will be mobile for use in multiple locations (flexible approach). These two decisions are important because they will impact the design and cost of the RTM program.

8.1.4.1 Selecting Sensor Locations

This evaluation is often driven by the outcomes of Framework steps 1, and 2. Now that the goals have been set and the key water quality parameters have been identified, the exact locations for monitoring can be determined to achieve those goals for both pilot and full-scale implementation.

This requires a solid understanding of the key dischargers, the collection system, sub-sewersheds, and the flow paths to the treatment facility during wet and dry weather. Ideally, the fewest number of locations can be determined to achieve the utility's goals and then additional locations can be selected for redundancy, to target problem areas, or wherever additional data is desired. Table 8-1 provides a list of advantages and disadvantages of different locations for real-time monitoring.

Table 8-1. Advantages and Disadvantages of Different Sensor Locations.

Location	Advantages	Disadvantages
WWTP Secondary Effluent	• Steadier values than WWTP influent make it easier to see long-term trends • Less fouling potential from FOG and rags compared to WWTP influent • Typical sampling location for future AWTP influent which may serve multiple purposes for the potable reuse program • Identifies if a water quality parameter is a challenge for AWTP • Typically can find accessible installation locations • Onsite locations are easier to maintain than remote locations	• Secondary effluent, like all locations in this table, still contains substances that can cause sensor fouling • Does not provide an indication of the source of challenging contaminants • No advance notice of potential threat to WWTP processes or ability to divert influent to avoid them. • Difficult to divert contaminated flow away from AWTP • Diurnal patterns and discharge events visible at influent are often muted or not visible in effluent
WWTP Influent	• Provides more time to react to influent events compared to secondary effluent • In combination with secondary effluent, provides data on the removal of a parameter or contaminant within the WWTP • Onsite locations are easier to maintain than remote locations • Easier to divert contaminated flow into a holding basin or away from WWTP • Easer to observe diurnal patterns and detect discharge events	• Can be more challenging to find installation locations than secondary effluent • Often significant fouling due to more rags and grease which can make maintenance challenging • Greater variability in incoming water quality can make long-term trends harder to see • Does not provide an indication of the source of challenging contaminants

Table 8-1. Advantages and Disadvantages of Different Sensor Locations. (Continued)

Location	Advantages	Disadvantages
Strategic Locations in Collection System	• Provides an indication of where the sources are coming from and helps with source tracking • Provides an early indication of an upset event allowing plant operations to take proactive action • Can be located directly downstream of a specific, challenging industry to confirm compliance	• May be more suitable for a flexible installation during source tracking; a permanent installation may not provide ongoing value • More FOG and rags can lead to more rapid fouling and greater maintenance burden • Maintenance is more of a burden because of the remote location away from the WWTP • Requires remote power and communication • Can have access issues (e.g. traffic, high water, safety, etc) • Can be more prone to theft/vandalism
Industry Discharge	• Can potentially be used for regulatory enforcement at specific dischargers or guide sampling effort for regulatory enforcement • Cleaning and calibration can be written into the permit and performed by the industry • Remote power and communication may not be needed (though it is still recommended for the utility to have direct access to the data) • Onsite locations are easier to maintain than remote locations	• Only provides data from one discharger and would be expensive to install at all dischargers • Industry could still be out of compliance for a different parameter than what is installed • Industry can find a way to by-pass sensor

8.1.4.2 Static vs. Flexible Approach

Once the locations have been selected to achieve the program goals, it is easier to compare the complexity and cost of static installations at these selected locations vs. mobile installations that can be used at multiple locations.

Information from the preceding chapters indicate that there are pros and cons to each approach. In general, fixed locations are recommended when there is a long term need to characterize the water quality. This is often relevant for sensors installed at the WWTP or at industrial discharges. As shown in Table 8-1, these locations provide value for long-term installations as they are easier to maintain than remote, offsite locations. Key manholes may also be useful static installation locations if the utility wants earlier detection of an event and has identified a suitable location.

Mobile sampling systems are recommended when they are used for temporary source tracking. This typically occurs at locations within the collection system. Once the utility has identified that there is an upset or a water quality event at the WWTP or at key manholes, operations staff can use mobile sensors connected to a data logger (or telemetered) and a battery pack. These can be moved sequentially upstream towards higher concentrations until a point source is identified or narrowed down. This does not require remote power or communication and is calibrated before the event. Fixed installations are not able to track events to individual point sources during temporary source tracking events as it would be cost prohibitive to have online sensors everywhere that is needed to fully characterize the system. There may be instances where fixed installations are the best option for monitoring within the collection system, but the utility should make sure that long term data is really needed for those sensors.

> **Best Practice**: in general, fixed installations are recommended at the WWTP, at key points in the collection system, and at industrial dischargers; mobile sampling is recommended within the collection system during source tracking events.

8.2 Phase 2: Instrument Selection, Design, and Support

PHASE 2: INSTRUMENT SELECTION, DESIGN, AND SUPPORT

5. Select and procure the sensors and instrumentation to implement goals
6. Develop station design and ancillary equipment
    - Physical/mechanical design
    - Integration, communication, and security
7. Develop O&M plan for sensors
    - Cleaning, calibration, and verification protocol
    - Troubleshooting guidelines and resources
8. Consider cost and resources needed for successful system operation
    - Identify maintenance resources and responsibilities
    - Identify how online monitoring will impact existing sampling program
    - Identify data management and reporting responsibilities

Figure 8-3. Phase 2 Framework Steps.

Figure 8-3 presents an overview of Phase 2, which includes steps 5 through 8 of the framework and is intended to include selection and procurement of sensors, station or installation design, sensor integration, development of an operations and maintenance plan, response plan, and additional resource planning. Phase 2 activities will typically come after Phase 1 activities have been completed; however, there is often some iterative decision making that occurs once you get into the details of the station selection and design). So, it may be necessary to revisit the goals and parameters initially selected in Phase 1. The following sub-sections provide a more detailed explanation of each step.

8.2.1 Step 5 – Select and Procure the Sensors and Instrumentation

Now it's time to select and procure the specific sensors best suited to achieve the program goals. This is a good time to review Chapter 2, particularly the master sensor list (Table 2-1). It may also be helpful to review the new sensors listed in Chapter 2 and the experience of other utilities using real-time monitoring for ESCP in Chapters 5 and 7. There may be other sensors available that are not listed in this report. If that is the case, ask for references and call the operations teams that use the sensor to make sure it will be a good fit for your application. Asking around for recommendations prior to purchasing is always a good idea. However, references may not be sufficient in all cases as matrix-specific interferences may make certain sensors non-suitable for your application.

Best Practice: It's easy to focus on the "shiny new technology" and end up spending time and money on stranded assets. It is important to stay focused on the goals established in Phase 1 before procuring the sensors. This will improve the likelihood of getting value out of the investment.

When selecting the sensor, make sure to pay attention to the design details and coordinate with mechanical and instrument technicians so that the correct options are selected. This is particularly important for installations in remote locations. Before purchasing, make sure you know how it will be powered, how it will communicate (e.g., telemetry and integration into SCADA or data management platform), and how the sensor will receive and discharge flow (if it is not a submerged sensor).

The installation location will help identify what sensor is selected. Can reagents be avoided for remote installations? Are the sensor and any reagents suitable for the environment? Are there any physical restrictions to installing the sensor at the identified site? Additional criteria to evaluate are the detection limits, accuracy, and precision. Will the recorded data meet the goals of the monitoring? Lastly, it is always recommended to include commissioning and startup support from the supplier to help get started up correctly as the suppliers often have best practices to share.

8.2.2 Step 6 – Develop Station Design and Ancillary Equipment

The most important part of this framework is designing a station for the sensor to collect and transfer reliable data. This is critical whether the utility is installing a pilot or a full-scale sensor. Much of this report, and previous reports, have focused on installation best practices, demonstrating the importance of this step. The design should typically not begin until Steps 1-5 are complete as it is specific to the selected sensor and location. However, it is sometimes helpful to perform steps 5 and 6 in parallel since design details can impact sensor selection.

8.2.2.1 Mechanical Design

It is impossible for this framework to provide best practices for every possible installation since the details related to station design will be site and sensor specific. The reader is referred to Section 3.1 of this report, Steinle-Darling et al. 2020), and Liggett et al. 2018 for compilations of best practices and lessons learned from utilities. This section can provide some high-level recommendations that should be considered for every design. The designer should always consider:

If the sensor will be installed in a manhole, sewer, or other gravity location:

- Where does the sensor need be installed so that it is always submerged?
- Are the equipment installed in a manhole C1D1 certified for explosive environments?
- Is there any part of the sensor or associated equipment that cannot be submerged or exposed to high humidity conditions?
- If the sensor and its appurtenances are installed in a corrosive environment (i.e. manhole), are all connections corrosion-proof?
- Is the sensor installed in a location that is representative of the intended flow?
- How will the sensor be supported and protected? Is the support sufficient for all flow conditions? Is the sensor secured in a way that it cannot be lost into the system?
- How will the sensor and the support be installed?
- How will the sensor avoid ragging and FOG accumulation?
- Does the sensor need to be protected against large items that may flow through the sewer?
- Can the sensor be easily removed for cleaning and calibration and then put back into service without the need for a confined space entry? Is it lightweight enough to be removed by a single person? Will specialty tools be required to remove the sensor?
- Will the sensor be completely hidden from the public or is a cabinet or other security needed?
- Will access be provided so that samples can be collected for sensor calibration and verification?

When a pump or pressurized pipe is used to feed the sensor:

- Will the pump be submerged and is it C1D1 certified for hazardous environments?
- Is there risk of clogging the inlet tubing to the pump or to the sensor?
- When maintenance of the sensor is needed, can it be taken offline easily or isolated from flow? Will the shutoff/isolation valve controls be near the sensor?
- How will the sensor discharge flow be collected? Will it be pumped back into the system?
- Does the system have a sensor and alarm to notify when the flow to the sensor(s) is restricted?
- How will the connection to the pressurized pipe be made without creating a clogging situation and/or impacting the flow?
- Will access be provided so that samples can be collected for sensor calibration and verification?
- Will the sensor be located in an exposed, outdoor environment? If so, what security and weatherproofing is needed?

Lessons learned from WRF 4908 (Chapter 5): While the results demonstrated the value of real-time monitoring, there were also significant challenges, as summarized here:

- Sensor Locations – sensors are often placed within manholes or adjacent to manholes, which can create safety issues. Minimizing the frequency of access is crucial to minimize impacts to the public and staff.
- Fouling – required weekly or more frequent site visits to many sensor locations. (e.g., FOG, rags, hair, debris, metals, precipitates, and other solids that can buildup on sensors and affect their performance).
- Power – most sewer access points do not have access to power, so battery life and reliability are important factors, especially for remote installations. Note: some features, including automated cleaning, cannot typically be sustained on the battery power provided by the manufacturer. Solar power has often been used in these types of deployments.
- Security – Sensors and sampling stations are costly, and above ground installations are often not secure, so locating it within manholes reduces physical security risk.
- Depth – depth limitations impact sensors that require a sample brought to the surface via suction unless a submerged pump is deployed.
- Intermittent Flow - Some parts of the collection system have intermittent flow which can create O&M and data quality issues. Some of these locations can be dominated by periodic industrial discharges. Also, most sensors not consistently immersed in liquid appear to lose calibration or can be damaged.
- Sensor Accuracy and Precision – accuracy varies so some sensors are sufficient to monitor for upset events, while more accurate sensors may be needed to monitor for compliance.
- Communication – remote locations often have poor signal availability for telemetry and there are hardware challenges related to battery life, humidity, and corrosion of the communications portion of the equipment.

8.2.2.2 Integration, Communication, and Security

Two key elements to a successful sensor installation are how it will be powered and how it will communicate data. There are many decisions that will likely be made in coordination with several groups including operations, water quality, controls and SCADA integration (OT), Engineering, IT, and security. Input from these groups will drive the design decisions for the selected sensors, communications protocols, cyber security protections, and data.

- Powering the sensor is a relatively simple decision. During sensor selection, the design team should review what options are available for the installation location. Mains Power: If the sensor is located where there is already power or wiring can easily be pulled, a wired connection can be considered. It is recommended that the monitoring station have its own circuit breaker if possible. If this is an option, the team then needs to confirm how the wiring is connected to the sensor and whether or not the design location is suitable for a hard-wired power connection.
- Remote Power: Sensors installed in the collection system, particularly in manholes, usually do not have mains power available. Battery or solar power is typically used for remote locations. Battery power is the lowest capital cost solution and is easier and safer to install. In situations where the use of batteries is not practical due to power requirements, solar

becomes the preferred option. However, many locations will not be suitable for solar panel installation (e.g. some roads, sidewalks, dense vegetation, snow, etc).

Identifying how to transmit the sensor data to the central system architecture (Cloud or on-premises) can be a more complex challenge. Large monitoring programs will often use a hybrid communication architecture as the best options for the individual sites. The following criteria should be considered for all offsite locations:

- How will the sensor transmit data (Cellular, Fiber, T-1, LoraWan, radio, other proprietary wire mesh systems)?
    - Has a communication architecture been developed for the program?
    - What types of communications are available for each station?
    - Can the sampling stations be integrated into another communications system, such as street lighting or advanced metering infrastructure (AMI)?
- Can a data logger be used to avoid data integration?
    - Is real-time information required for operations?
    - How often will the data need to be manually collected and uploaded into the central system?
    - How will the goals of the monitoring program be impacted if data is only available after retrieving data from the data logger?
- What cybersecurity is needed to meet the utility's standards for remote data?
    - Will the data be transmitted and hosted in the SCADA system, Business network, or in an external cloud?
    - What level of cyber security protection does the cloud provide?
- Is programming or proprietary software needed to interpret, analyze, or store the data?
    - Front-end: JSON.
    - Back-end: Python (Most prominent), R, C#.
- How will data be stored? Will it be stored in the cloud or on-premises?
    - What frequency should data be transmitted and stored?
    - Does the data need to be cleaned and validated?
    - Does the system need to incorporate event detection and alarms?
    - Is a database needed to receive and/or store the data?
        - Example 1: NoSQL like Mongo for on premise.
        - Example 2: Cosmos in Azure for time series.
- Are analytics (machine learning or user optimized) needed to clean and/or analyze the data?
- Can the monitoring platform be programmed to identify sensor problems and send out alerts?
- Is the sensor connected to a data processor locally that can store, display, and download the data?
- How will the industry dischargers be able to access the data (where sensors are installed onsite)?

Developing a monitoring program that provide real-time actionable information is not a trivial exercise. The number and location of stations, type of data communications, Business

Intelligence Architecture, workforce training, and governance need to be evaluated and consistent with the mission and vision of the utility.

8.2.3 Step 7 – Develop O&M Plan for Sensors

Step 7 focuses on developing an O&M plan that sets up the real-time monitoring program for long-term value. While it is acknowledged that O&M plans are important for all equipment in treatment plants and collection systems, it is particularly important for sensors installed to support ESCPs. As discussed throughout this report, there are significant challenges to getting reliable and accurate data from these sensors and a robust O&M program is needed. This step focuses on the different items that should be included in the O&M plan while the next step discusses the resources needed for the plan to be successful. While the recommendations in this section are high-level, Chapter 4 provides specific details and best practices for sensor maintenance in the lab and in the field. The approach for implementing a sustainable O&M plan can be internally implemented, a Data as a Service contract, or a hybrid.

8.2.3.1 Cleaning, Calibration, and Verification Protocol

The requirements for cleaning, calibration, and verification will vary for different sensors, so the sensor manufacturer's documentation and representatives should be consulted as necessary to develop O&M plans. O&M requirements also differ based on the site, the installation configuration, and the automatic cleaning devices installed. While each plan will be different, it is important that a clear protocol and schedule is established for each sensor. Weekly cleaning and monthly calibration can be used for general best practices for most sensors, but pilot testing is the only way to understand the true requirements to produce reliable data.

Prior to installing the sensor, make sure that the chemicals and solutions needed to perform the cleaning and calibration are available and that staff have been fully trained on how to properly clean and calibrate the sensors. Some sensor types (e.g. permeable membrane bulbs on a pH sensor) can be damaged by touching, scrubbing, or even by contact with soap, so an understanding of the proper maintenance and calibration protocols is needed and should be provided by the manufacturer or representative.

Sensor maintenance and calibration should be assigned to a limited number of staff (2 to 3) to ensure consistent technique. The approach often attempted by utilities, often referred to as 'maintenance by committee', where the utility just finds someone available to do the work, often leads to more overall data variability.

Recordkeeping is also very important since these types of sensors can lose calibration or report inaccurate data at unplanned times. It is very important to retain historical data, keep accurate calibrations and service logs, and record past issues that required troubleshooting and the fixes that solved the issues. Note that if an event does occur that could have an impact on public health (even though this is unlikely based on all the permit provisions put in place for potable reuse), records such as these will be likely be reviewed for proper recordkeeping.

It is also frequently useful to have a source of the 'correct' value that the sensor should be reading to evaluate performance. This verification should be done on a regular basis in addition to calibration. For example, taking a handheld pH sensor measurement during a maintenance visit helps to determine if the sensor had drifted away from the 'true' value, and if cleaning or calibration restored it. Other sensor types may require a grab sample for laboratory analysis. Verification measurements help to determine if events seen by sensors are real or an artifact of the sensors and if long-term trends are real or due to sensor drift.

8.2.3.2 Troubleshooting Guidelines and Resources

As you can infer from the preceding chapters, troubleshooting is synonymous with implementing long-term, real-time monitoring within wastewater collection and treatment systems. Therefore, it is helpful to develop simple guidelines so that staff know when and who to contact to perform troubleshooting when needed. It is recommended to develop standard operating procedures (SOPs) that staff can perform if the calibration is poor, if the sensor loses power, if the sensor stops transmitting data, or other common challenges. It is also recommended to maintain contact with the sensor representative for situations that cannot be solved by internal staff.

Sensor suppliers and representatives have a vested interest in the sensors performing well and they can be a valued resource in the success of real-time monitoring systems. Each utility should diagnose its in-house capabilities for sensor maintenance and troubleshooting. If they do not have either the resources available or the technical expertise, it is recommended to maintain an ongoing contract with the supplier. This allows the supplier access to the data and provides a mechanism for them to regularly monitor the performance of the sensor. While the utility may think this is a cost that can be avoided, it has been demonstrated to be very valuable for producing reliable data.

Step 7 is also a good time to consider investing in a mobile monitoring skid or back-up sensors that can provide redundancy if O&M activities, troubleshooting, or replacement of sensors or ancillary equipment is taking longer than expected and accurate data is not being collected as needed. There will definitely be times when the sensor is offline so redundant sensors or a mobile system should be considered if the data is needed for continual operation. Such redundant sensors may also be used for verification.

8.2.4 Step 8 – Consider Cost and Resources Needed for Successful System Operation

Step 8 is the final step of Phase 2 and the final step before the sensors are installed. The goal of this step is to estimate and identify the budget and resources that are needed for the sensors that will be installed and to reconvene the stakeholder team to make sure that the installation and O&M plan will achieve the overarching goals of the monitoring program. Typically, the purchase price of a sensor is much lower than the total cost to install, integrate, and maintain the stations. This provides an opportunity once the utility has a more accurate total cost to confirm that the information provided will be worth the budget and resources required.

The sensor design was completed in Step 6 and should have produced cost estimates for mechanical, electrical, and integration efforts. These costs should be documented to inform decisions when additional sensors are considered in the future.

> **The Hidden Costs of Real-Time Monitoring.** Make sure your cost estimate includes costs for:
>
> - Design installation and integration.
> - Installation: mechanical, electrical, and integration.
> - Ongoing maintenance.
> - Increased sampling for data verification.
> - Ongoing service contracts.
> - Communications.
> - Data analysis, storage, and visualization platform.

Resources for ongoing sensor maintenance is the most important part of this step. Many of the sensors tested in this study required weekly, or more frequent, cleaning. Prior to installing the sensor, there was no way for a utility to know exactly how often it will need to be cleaned or calibrated, unless the sensor was pilot tested. Still, it is important to set reasonable expectations so that resources are planned and available. For any sensor installed in the collection system or the WWTP influent, it is recommended that the utility assume a minimum of weekly cleaning and monthly calibration will be required. The cost for this maintenance program, with contingency for unanticipated challenges, should be estimated before the sensor is installed. This should include time for troubleshooting and repairing the installation, particularly for sensors in manholes or sewers. This cost should also include the annual service contract with the sensor supplier, if the utility decides this is valuable.

In addition to costs associated directly with keeping the sensor running properly, the impact to the existing sampling program should be considered. Although real-time monitoring provides continuous data, it likely will also increase the amount of analytical sampling needed to verify sensor accuracy, especially if the sensor is used for regulatory compliance. Additionally, if the online sensor identifies an upset event, it might trigger additional sampling to capture periodic variations or for operational, enforcement, compliance, or other reasons. While this hopefully provides improved operation, it is important to note the increased burden on resources, including the staff performing the sampling and the analytical cost. These additional requirements highlight why the cultural buy-in of the program is so critical.

Lastly, a perfectly operating and calibrated sensor will produce a tremendous amount of data but likely will have limited value unless someone has been identified to review, interpret, and report the data appropriately. Chapter 4 discusses applied machine learning techniques that can be used to detect upsets and establish a baseline for performance. But even if this is not needed, a resource should be identified that routinely evaluates the data. Decisions on data storage frequency are important so that efficient data storage, access, and management is provided for the team, and it does not become too expensive to store the data. The problem with the manual approach with data cleaning and analysis is that the data velocity and quantity quickly become overwhelming for traditional tools. The result is the utility having a voluminous amount of unprocessed data that quickly loses its value over time and the monitoring program failing to achieve the goals. The approach of automating these processes inside the Business

Intelligent Architecture (BIA) is being used globally to alleviate this problem and maximize the value of the utility investment.

This step creates an "all-in" estimate for capital and operating costs. After the sensors and system have been installed, it is best practice to update these costs, so they reflect the actual costs incurred. This enables to the utility to make informed decisions on expanding the real-time monitoring program. If the total cost of the sensor installation exceeds the available budget, innovative cost strategies can be pursued. For example, using a DaaS approach places the burden on the third-party contractor (often the vendor). DaaS contracts have been developed so the utility only pays when quality data has been delivered.

8.3 Phase 3: Implementation and Optimization

Figure 8-4 presents an overview of Phase 3, which includes Steps 9 through 12 and focus on implementing the real-time monitoring program and continuous improvement. This phase is relevant for implementing a pilot test or an entire program.

PHASE 3: IMPLEMENTATION AND OPTIMIZATION

9. Install sensors and conduct training
10. Integrate sensor data into existing operations and reporting procedures
    a. Data storage frequency
    b. Data display and interpretation: SCADA/DCS, alarms and CCPs
    c. Data evaluation and verification
    d. Communication and reporting
11. Define source tracking strategy and compliance monitoring
    a. Integrate into industrial permits
    b. Use for enforcement of permit violations
12. Continuous improvement
    a. Implement lessons learned and new best practices
    b. Follow advancements in sensor or other technology to improve data accuracy

Figure 8-4. Phase 3 Framework Steps.

8.3.1 Step 9 – Install Sensors and Conduct Training

Now that the planning and design phases have been prepared, it's time to execute the installation as planned. If your utility has determined to conduct a pilot test, this may include installing the pilot equipment to inform previous steps that were not finalized. If your utility has already performed a pilot test or elected to skip a pilot test, then this step is installing all the stations as part of the full-scale program.

Similar steps are recommended for starting up sensors as would be performed for other equipment. Personnel is needed to make sure it is installed correctly with a reliable water source. And personnel is needed to make sure that it is communicating the data effectively. It is recommended that an increased sampling program is established for the first week or month after installation and then can be decreased after reliable results are proven.

Chapters 5 and 7 has more detailed information and case studies of utilities' experiences installing and utilizing real-time monitoring equipment for ESCPs. Note that it is common for field conditions to require tweaks to designs and plans since all parts of a collection and treatment system will be different than others, so be prepared to communicate with the equipment representative, designer, IT, and other support staff as needed once the work is under way.

Once the sensor has been installed, staff must be trained on its maintenance and calibration requirements. Initial trainings are often conducted by the equipment supplier; however, it's likely that additional staff-led trainings will be required to get other key staff familiar with the technology and how it is integrated and monitored within a system. It is also recommended that staff be informed of the importance of the sensors for the overall program. O&M staff buy-in is key to a successful real-time monitoring program. If staff do not understand the value of the data, they will be less likely to perform all the tasks needed to get reliable data.

8.3.2 Step 10 – Integrate Sensor Data into Existing Procedures

Once the stations are operating and producing quality data, the operations staff need to be trained on how to evaluate the information produced from the data. This is a very important step for the staff to understand how the data can be integrated with their intuitive knowledge of the system. The staff can use this process to rapidly assess the situation and respond accordingly.

The sensor data likely will not initially be displayed in the perfect way for interpreting the data and additional steps will be needed to streamline the data analysis process. Some of these tasks may have been completed as part of the sensor installation but many of them cannot be completed until data is being transmitted from the sensor.

- Data storage, access, and organization: data storage frequency may need to be reevaluated and optimized as the full program is implemented to keep it both useful and manageable. Data storage should be integrated into the rest of the facility's data storage or in a cloud environment that offsets the cost of purchasing and maintaining additional servers on-premises.
- Data evaluation and verification: this refers to confirming the quality of the data, which inherently includes sensor calibration, and to establishing a baseline for the data. Once a baseline is developed, upset events or spikes are easier to detect

Chapter 6 Machine Learning Assessment Conclusions:

- ML models using online sensor data from influent or treated wastewater could be applied for useful, practical alert systems.
- Multiple ML models performed adequately in terms of accuracy, sensitivity, false positive rate, etc.; but ongoing research will refine which models are best for which applications.
- Next steps could include demonstrating trained ML models with live, real-time sensor data and training regression ML models to predict quantitative values such as ozone demand, recommended operational settings, etc.

and alarms can be set accordingly. Chapter 4 provides a detailed review of different machine learning techniques that can be implemented to achieve this.

- Advance data analytics, trending, machine learning, and artificial intelligence: transforming the validated data into actionable information is an important step to maximize the value of the investment and enhance the operations of the utility. Some of the elements include:
  - Creating trends to evaluate long-term sensor performance (leveraging the plant historian or other data storage and trending capabilities).
  - Creating trends that link the sensor data with other WWTP data to evaluate the impact of an event on other parameters.
- Data visualization and interpretation: the data is only useful if it is correctly interpreted and presented, so these activities include:
  - Displaying the current sensor value at SCADA, plant DCS screen, or a separate platform. An independent platform allows the information to be shared across the organization and maintain the security of the SCADA system.
  - Establishing alert and alarm values to notify the operator of upset events.
  - Establishing CCPs so that direct action can be taken if needed.
- Communication and reporting: once the data is verified and analyzed, it is important to set up the communication pathways to get the data trends, reports, and key conclusions to the right people in the right timeframe. This may include:
  - Expanded user access to the data analytics platform to allow users throughout the organization to evaluate the monitoring system that relates to their areas of responsibilities.
  - Protocols that trigger immediate communication, data accuracy verification, and/or event confirmation (e.g., exceedance of a threshold or an event that triggers an action, sampling to confirm the results, and notification of operations staff).
  - Updates to keep stakeholders aware of observed trends and the value of the data. Note that it might be useful to implement these updates to a wider audience only after there is confidence in maintaining data accuracy to reduce negative impacts of false alarms.

What to do if your real-time monitoring detects an excursion?

- If possible, verify that the detection is a real excursion and further investigate the cause with a grab sample.
- Identify what treatment decisions need to be made:
  - Is this a critical control point and does water need to be diverted?
  - Should treatment decisions or chemical dosing change to address the contaminant(s)?
- Determine if stakeholders or regulators need to be informed. Do so promptly if needed.
- Initiate a source tracking investigation, such as with a mobile real-time monitoring unit.
- Document the event.

8.3.3 Step 11 – Define Source Tracking Strategy and Compliance Monitoring

Once accurate data is being reported and integrated into utility operations, it is time to identify how this data can potentially be used to support and bolster compliance monitoring, if this was a goal set in Phase 1. There has been significant effort and resources to get to this point and this is where value can be demonstrated for the real-time monitoring program. If steps 1-10 were the "how to", Step 11 is the "so what."

Best practices for source tracking:

- Be prepared for events by having sensors available and resources trained.
- Confirm the data with an immediate grab sample, if possible, before mobilizing resources.
- Identify how to track the contaminant upstream through the collection system, which may require high sample size or composite sampling if it is not a continuous discharge.
- Proceed upstream through the collection system until the source is identified or the event ends.
- Identify if it is a one-time event or a repeat event; if a repeat event, document progress so that you can continue during the next event.
- Document the results so that they can be used for compliance purposes.
- Identify the type or class of contaminant and evaluate upstream dischargers for possible violators.

The goal of most real-time monitoring programs for ESCPs is for early detection and notification of upset events. These may trigger direct action at the WWTP or AWTP (via alarms or CCPs), sampling to verify the data, and source tracking events. It is unlikely that a utility will install a real-time monitoring program so comprehensive that they will immediately know the source of an event. As each event or challenge will be unique (although some may be repetitive), it is important to have a defined strategy for locating the source. Best practices are summarized in the text box to the left for developing a source tracking strategy and program.

If the source tracking results identify a single source that is responsible for the event, the data should be documented and discussed with the source or discharger. This can result in a violation if it was already in the permit or can result in a modification to a permit. The goal is to systematically identify and eliminate the sources of upset events. If the source is an industrial discharger, the utility should consider having the industry install an online sensor and using it for compliance or installing a station immediately downstream of the discharger. If sensors are installed at a permitted industry's facility, the utility should still receive the data and require the industry to perform and document cleaning and calibration. See Chapter 4 for examples on how the cities of Ventura and Oxnard have utilized real-time monitoring in their compliance permitting.

If the identified source is not a permitted discharger, compliance monitoring and enforcement can be more challenging. There are many anecdotes from utilities of infrequent discharges from residents or commercial businesses that were not permitted where the utility identified the source and discussed directly with the person or business and avoided enforcement. This is often enough to educate the offender and stop the discharge, though it does not provide ongoing enforcement. Some examples from utilities of these situations include:

- A resident performing household oil changes and dumping the waste at the end of the month, resulting in high TP and poor nitrification at the WWTP.
- Weekly regeneration of in-house water softeners that resulted in high conductivity at the WWTP.
- Infrequent discharges of hauled waste at an RV park that resulted in upsets at the WWTP.
- A septic hauler dumping into manholes to avoid fees.
- Gas station washing down an oil or gas spill into the sewer.

It should be noted that identifying the actual source of the upset is the best-case scenario and should be considered a major success. Utilities often refer to "chasing ghosts" in the collection system where it can take years to track down the cause of a monthly upset event, and many times there is no identified cause. It is important to quantify the challenges that are caused by the upset and make sure that the resources spent investigating the cause do not exceed the consequence of the upset.

8.3.4 Step 12 – Continuous Improvement

Steps 1-11 have identified the key items needed to establish a successful real-time monitoring program. Now it's time to begin reviewing and learning from both the positives and negatives and using this information to refine and improve the program. This research led to three recommendations for continuous improvement which are to:

1. Implement lessons learned and best practices from O&M, IT, OT, water quality, engineering, and equipment vendors.
2. Follow advancements in sensor technology to improve reliability and data accuracy and to be able to monitor new contaminants.
3. Follow industry trends for specific suppliers or products (e.g., discontinued products, vendors entering or existing the market, new offerings that reduce costs or O&M requirements, etc.).
4. Follow advancements in data cleaning and machine learning to identify upset events.
5. Track the "true cost" of sensor installation, not just the purchase price of the sensor, so that future sensors can be evaluated appropriately.
6. Document success stories in the annual ESCP report to justify the cost of the program.
7. Evaluate on an annual basis if the real-time monitoring program is achieving its goals or if it has become too costly.
8. Confirm that the program has the trained staff and resources available to be successful.

Best practice: Schedule a time each year (or more frequent if needed) to revisit Steps 1-11 and:

- Confirm the program is meeting the stated objectives.
- Document success stories.
- Review the cost of the program against the goals.
- Identify if the appropriate resources are trained and available.

Many of the utilities that contributed to this project noted that for these systems to achieve their benefits, three “pillars” need to be followed. Pillar 1 is referred to as Workforce, which requires both trainings for and a culture of acceptance from the staff that will implement the program. Pillar 2 is referred to as Governance, which requires clear policies and procedures to implement the necessary tasks. Pillar 3 is Technology, which requires selection, design and implementation of a technology that works to achieve the intended goals. This type of organizational commitment is recommended for establishing real-time monitoring programs and for focusing on continuous improvement.

CHAPTER 9

Conclusions

9.1 Summary

Many online water quality sensors are now commercially available, including optical sensing (e.g., absorbance, fluorescence) that correlate with other water quality measures (e.g., TOC, COD) or specific compounds (e.g., nitrate, optical brighteners). Many utilities planning or practicing potable water reuse already have comprehensive ESCPs in place, but real-time sewershed monitoring would further strengthen detection, identification, and mitigation of industrial or illicit discharges. Several utilities have experimented with real-time sewershed monitoring as part of this and prior research projects as well as established monitoring programs. Several have successfully detected or tracked down industrial or illicit discharges. All have found high maintenance frequencies are necessary (i.e., every other day to weekly).

Not all sensors proved sufficiently resilient in the sewer environment, so thorough pilot testing in real wastewater are recommended prior to purchase. Piloting in the intended installation environment (e.g., sewer) may be necessary to test not only the sensor, but also the auxiliary equipment and power supply and data transfer plan. Constructed flumes allow for analyzing the effects of different variables on and comparisons between sensors in a controlled experiment in a relatively accessible location, while still subjecting sensors to real wastewater or even intentional pollutant spikes or fouling.

Placing sensors further down the treatment train (e.g., primary effluent as opposed to raw wastewater) allows greater data reliability, but later warning in the event of a spike. Installations at the head of, or within, a WWTP allows for easier access, better maintenance, and better security, but less early warning compared to installations within the collection system. Mobile systems allow for tracking sources in the collection system. Sensor holders designed to prevent clogging or fouling in the sewer improve the sensor's precision and resilience. Placement of optical sensors downstream of water treatment chemical additions (e.g., ferric coagulants) at WWTPs often increases the rate and fouling and data drift. This should be avoided to the maximum extent possible.

Machine learning can detect events in wastewater influent or treated water with greater accuracy and specificity than conventional threshold-based alerts. However, machine learning requires several, diverse sensors and multiple recorded instances of the type of event(s) to be detected. So, machine learning should be used in tandem with, not instead of, threshold-based alerts and alarms.

Real-time sewershed monitoring is expensive, with unforeseen costs such as time for travel to and from remote sewer locations and data analysis costing more than the sensors themselves. Even the early step of determining if a probe is accurate and its needed cleaning and calibration frequency requires much effort. Without a clear and actionable data management strategy, the utility will struggle to realize any benefit from its investment in real-time sewershed monitoring.

9.2 Recommendations for Integrating Real-Time Monitoring into ESCPs

- Real-time monitoring requires clear goals and direct action(s) that would be associated with the sensors (e.g., alerts and responses that improve water quality and reduce risk). If this is lacking, the program will consume resources without providing value.
- Select the water quality parameters to be monitored and identify criteria for sensors and monitoring locations that directly allow for enforcement action, troubleshooting, and water quality improvement.
- Select online monitoring locations and a monitoring approach (i.e., sensor locations and permanent vs mobile installations) based on clear goals and before procuring sensors.
- Sensors without consumables or moving parts are generally lower maintenance and more resilient in challenging environments like sewers.
- Perform a pilot test, especially for larger or more expensive programs. Sensor vendors often allow free trial periods or rentals.
- The utility should have full control over the sensor calibration for ongoing programmatic sustainability, but with vendor support.
- Develop an operations & maintenance plan and a data quality assurance and governance plan before proceeding with installation and update as needed.
- Consider cost and resources needed for successful implementation, including less obvious costs like travel time to and from the sensors for maintenance, and data storage and analysis.
- Having a specific operator in charge of each sensor or a maintenance apprenticeship program—as well as an SOP—results in more consistent maintenance and thus more consistent data. Maintenance by committee leads to inconsistent sensor upkeep, reduced performance, and compromises data quality.
- Have a mindset and culture of continuous improvement. Implement lessons learned and new best practices. Follow advancement in sensor technology, data infrastructure, and data analysis.
- Data cleansing and validation should be the first step before attempting to process data through machine learning or artificial intelligent algorithms.

9.3 Recommendations for Future Research

Further research and innovation remain necessary to improve wastewater monitoring systems, including reducing maintenance frequency, improving sensor accuracy and precision, expanding measured parameters, and addressing concerns like power sources, data connectivity, physical and cybersecurity for remote sensors in sewersheds. Further, the needs go beyond sensor systems, as the industry must continue to examine how to best make use of the data from real-time sewershed monitoring. Three potential further research areas related to real-time sewershed monitoring are outlined below.

9.3.1 Real-Time Sewershed Monitoring Cost-Benefit Analysis

- **Problem Statement.** This project demonstrated that it is possible for a utility to conduct real-time sewershed monitoring given sufficient resources, creativity, and institutional will. Furthermore, research has shown that real-time sewershed monitoring can detect real water quality events caused by industrial activity. Nevertheless, it was beyond the scope of this project to conduct an in-depth cost-benefit analysis (CBA) comparing the cost of real-time sewershed monitoring against the cost of alternative, equally protective approaches (e.g., surveillance in advanced treated water and a larger engineered storage buffer).
- **Research Question.** What is the total financial cost of real-time sewershed monitoring? How does this compare to influent water quality excursions?
- **Desired Outcome.** Such a CBA would require a long term (minimum 1 year) trial of real-time monitoring to determine what level of monitoring was institutionally sustainable and refine their maintenance frequencies. Through that process, the utility would determine the initial costs of the sensors as well as (1) accessories, appurtenances, and apparatuses; (2) operator labor for maintenance including travel time; consumables and sensor or component replacement; (3) vehicle wear and tear for travel for maintenance; (4) staff time for troubleshooting, data analysis, communication, etc.; (5) subscription fees for any software or data platforms; etc. Furthermore, for the cost information to be representative, the study ought to include water reclamation and reuse systems of various sizes and levels of industry within the sewershed, thus requiring multiple (at least 3) different size collection systems. Studies could include sensor systems within the collection system or within the WWTP. The study would conclude on how the costs and benefits of the real-time monitoring system can reduce the burden of conventional source control programs while providing equal or greater water quality protection. Alternative approaches for comparison could include (1) real-time monitoring after the WWTP, within the advanced treatment, or in the purified water, with corresponding larger engineered storage buffers or (2) additional treatment barriers or greater safety factors within existing treatment barriers such that the advanced treatment could successfully purify even the worst-case real water quality excursion observed over many years of monitoring. This CBA would better inform government agencies on the most effective ways to ensure that reuse systems are protecting the public from any risk from intermittent industrial discharges to the sewershed.

9.3.2 Novel Sensors and their Application to Real-Time Sewershed Monitoring

- **Problem Statement.** As shown in Chapter 2, few real-time sensors are currently commercially available for specific chemicals, especially organics. Nevertheless, it is plausible that more sensors will become available for real-time monitoring of specific chemicals or chemical classes within the next 5-10 years. Examples include bromide (Westerhoff et al. 2022), total PFAS (Law et al. 2021; Wang et al. 2021; Faiz et al. 2020) or *N*-nitrosodimethylamine (NDMA) (Roback et al. 2020; Fujioka et al. 2017; Kodamatani et al. 2018). Initial studies about such sensors often measure spiked contaminants in deionized water or relatively high-quality finished drinking water. Before field implementation for sewershed monitoring, sensors should have their precision, accuracy, and maintenance requirements quantified in raw wastewater.
- **Research Question.** Can a sensor be developed to detect specific chemical(s) not previously detectable with commercially available sensors? Can it do so at concentrations that would be representative of real water quality events in raw wastewater with a reasonable maintenance and calibration frequency?
- **Desired Outcome.** If developed, such sensors could detect types and levels of industrial discharges not previously possible. Such sensors could measure water quality parameters more directly relevant to human health or treatment settings than the bulk surrogates currently monitored. Industrial chemicals prioritized for ESCP monitoring in Nading et al. 2022 could also be prioritized for sensor development (e.g., NDMA, NMOR, PFAS, heavy metals, 1,4-dioxane) (Nading et al. 2022). Some chemicals might never be measurable with useful precision in true real-time. However, measurement a larger time period (e.g., hourly) might be achieved with relatively rapid automated chromatography. Such frequencies could still be useful for detecting water quality events. So, for the purposes of a request for proposals on this topic, hourly measurement frequencies could be considered adequately near to real-time.

9.3.3 Machine Learning with Real-Time Sewershed Monitoring Data

- **Problem Statement.** Real-time sewershed monitoring is only as effective and useful as the approach applied to informing decisions with the resulting data. Even when best practices are followed, real-time sewershed data is inherently noisy compared to water quality data from locations with more treatment and equalization. This data noise presents an inescapable limitation on traditional, univariate methods for alerts and alarms. There is inevitably a tradeoff between detecting events at lower levels vs more false alarms. ML—thanks to being multivariate and nonlinear—can "overlook" outliers on a single variable if other variables are behaving as usual. The adage "garbage in, garbage out" is true to an extent for any modeling approach, but the noisiness of sewershed data may actually play to ML's strengths relative to simpler alternatives.

  This study included two ML for alert systems proof-of-concept case studies. However, much work remains to be done in this area. Greater sample size (number of variables, number of timepoints, and number of events) would improve both the accuracy of ML models and the confidence in their relative accuracy. Many more preprocessing methods such as quantile-based or standard-deviation-based outlier screening and omission could be paired with ML models. Cross-validation methods—including ones specifically designed for timeseries data—could be compared for how well they point to the best tuning parameters settings for the testing sets using the training sets.

- **Research Question.** What are the best ML methods (preprocessing, cross-validation, and model types) for event detection in real-time sewershed monitoring? What is the most cybersecure and institutionally sustainable way to implement these algorithms in the field? Would utilities giving different weights to the adverse impacts of false alarms vs false negatives change the final ML model selection?
- **Desired Outcome.** More research on ML for event detection with real-time sewershed data would lead to greater accuracy and fewer false positives and potentially pave the way for more widespread adoption. Applied research could also explore the best means for field implementation, e.g., cloud vs. on-premises computing. This research could work in tandem with the other future research topics above. E.g., the CBA could include the costs of ML implementation or compare the cost-effectiveness of ML using data from the sewershed or later in the treatment system. Data from novel sensors could improve ML accuracy or enable the detection of new types of events.

APPENDIX A

Bench and Field Experiments Examining the Factors Affecting the Performance of Water Quality Probes in the Sanitary Sewer Collection System

A.1 Introduction

Wastewater utilities are in need of high-quality continuous water quality data from their conveyance systems. These data are required for successful source control program that protects treatment plants and the receiving water from harmful discharges, to facilitate the effectiveness of advanced treatment processes for reuse applications and ensure the safety of land application of biosolids. Very few utilities have deployed long-term continuous water quality sensors in in the sanitary collection system, however, because the sewer system environment is a difficult environment for water quality probes and supporting equipment. Some of the factors contributing to such difficulties include:

- Sewers are confined entry locations that limit access and space for location of equipment.
- Wastewaters contains fats, oils, and greases (FOG), rags, hair, debris, metals, precipitates, and other materials that can buildup on sensors and affect their performance.
- Flows vary diurnally and seasonally, sometimes with very low flows at night and/or very high flows during storm events.
- Manhole headspace is typically humid which limits the types of electronics that can be utilized.
- Manholes typically have poor signal availability for telemetry and lack ready availability to a continuous power source.
- Manholes are often in locations difficult to access safely on a regular basis (e.g. roads), and frequently not secure from vandalism and theft of expensive instruments.

Many utilities lack experience in using continuous monitors. Sensors, dataloggers, and other equipment vary greatly in capability and cost between different manufacturers, and often utilities struggle with decisions related to levels of investment due to a lack of previous applications and experiences.

Water Research Foundation (WRF) Project 5048 (Integrating Real-Time Collection System Monitoring Approaches into Enhanced Source Control Programs for Potable Reuse) was initiated to provide better understanding of how real-time continuous water quality data can be collected successfully and used in source control programs. Task 2a of WRF 5048 included bench and field scale experiments to study the factors affecting the performance of water quality sensors and compare these effects on various brands and types of water quality sensors in the sanitary sewer environment. Task 2a also included the testing of a rag guard developed by Clean Water Services (CWS) in side-by-side comparisons with more generic sensor holders. The study was completed by CWS personnel in Hillsboro, Oregon. This report describes the experiments conducted and summarizes the results, analysis, and major findings.

A.2 Methods

Task 2a of WRF 5048 consisted of two subtasks:

1. Conduct an evaluation of continuous sensors in a continuous-flow flume environment.
2. Develop a sensor containment device to minimize ragging and test the device in the collection system environments to determine its effectiveness.

This section describes the methods and equipment used to perform both of these subtasks.

A.2.1 Subtask 1: Conduct Evaluation of Continuous Sensors

Subtask 1 consisted of construction of a flume at the influent of the Forest Grove Wastewater Recovery Facility (WWRF), installation of sensors of various types and manufacturers in the flume, and experimental testing of the impact of selected variables on the performance of the sensors. Each of these experiments is described in the subsection below.

A.2.1.1 Construction of the Flume

A custom, Plexiglass flume was constructed near the headworks of the Forest Grove WWRF to perform controlled experiments on the sensors. The flume is shown in Figures A-1 through A-4. The influent flow was diverted to the flume immediately after passing through the bar screens, grit removal, and a wet well (Figure A-1). The oil and grease concentrations were measured in both the influent upstream of the grit screen and in the flume and were found to be comparable. While the grit screens did remove large debris and some rags, fibrous solids, hair, and FOG still remained and were present in the flume. These materials attached themselves to the sensors forming rag and grease balls that simulated often observed conditions in the collection system. The flume was contained inside of a tent to provide shelter from the elements; a space heater was used, as needed, to prevent freezing conditions.

Figure A-1. Flume located at the Forest Grove WWRF.
Showing the 2-inch intake line from the influent after the bar screens and the tent protecting the flume from the elements.

The flume was built to provide maximum flexibility for deployment of sensors for this and potential future studies. The overall dimensions were 12 inches high, 36 inches wide, and 72 inches long, with an adjustable working width of 12 inches, 24 inches or 36 inches and water depth between 1-inch and 11 inches (Figure A-2). The flume was constructed on a jack on one end to vary the slope. A 2”-diameter flexible tube was tapped into the influent that provided influent flows up to 150 gallons per minute. The inlet of the flume consisted of three 2-inch diameter ball valves to adjust the flow rate. Flow from the ball valve inlets then was passed through a stilling well to reduce turbulence and to aid in the transition of the pipe flow to a more uniform open channel flow similar to that observed in sanitary sewers. The stilling well was constructed by combining a volumetric bucket to redirect flow from the inlet and dissipate energy, and a sluice weir to maintain steady flow across the channel and reduce turbulence (Figure A-3). The active channel in the flume, where the sensors were located, was approximately 48 inches long. A weir, constructed of one to three 2-inch diameter pipes, was placed sideways across the cross section at the end of the active channel to regulate the flow depth (Figure A-4). A 3-inch diameter drain was placed on the bottom at the end of the flume to direct the flow for return to the treatment works.

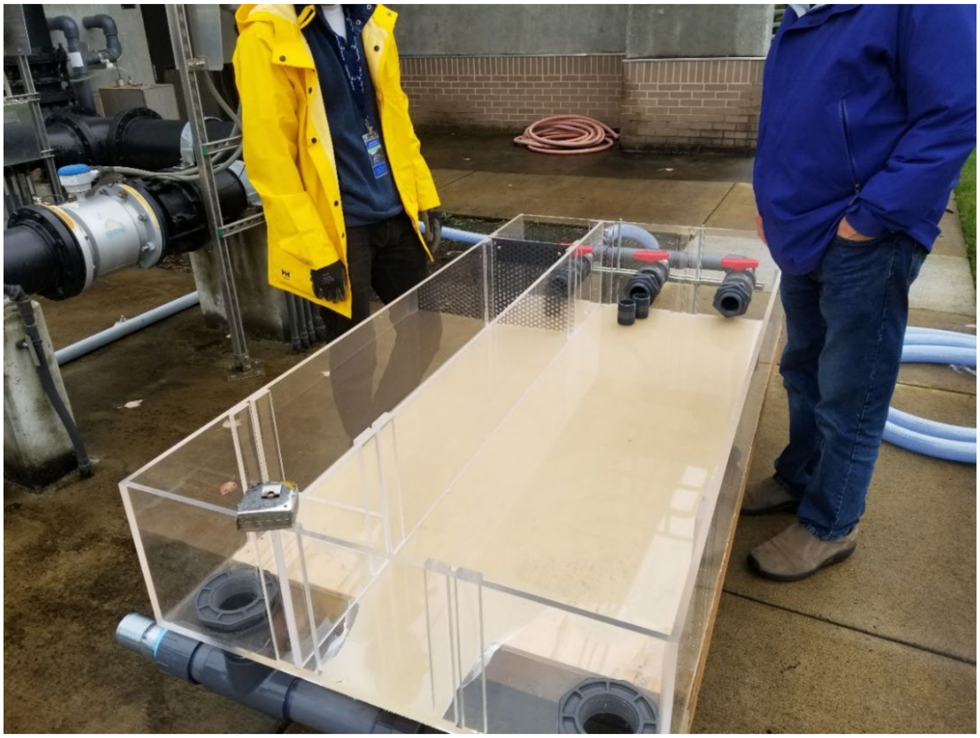


Figure A-2. Experimental Flume in the Initial Construction Phases.

For the experiments, only the 12-inch wide channel was used to simulate the most common pipe size in the Forest Grove collection system and to provide the highest range of depths and velocities. The slope and flow rate were held constant over all experiments. The maximum flow rate of 150 gpm was used, and the slope was increased as much as possible, while still maintaining subcritical flow in the flume. The probes were all mounted on cross bars over the flume and held in the flow at an approximately 45 degree angle pointed downstream to decrease ragging (Figure A-4). All sensors were installed within 1 foot of each other in the channel and were installed so that the tips or sampling location of the sensors were all at a similar depth below the water (typically approximately 2 inches). A utility box next to the flume housed the dataloggers and supporting equipment, and AC power was available from the nearby building. A 5-gallon drum with an adjustable outlet valve was used for dosing during spiking experiments. This was placed on the upstream side and the spiking solution was introduced into the stilling well portion of the flume (Figure A-3).

Figure A-3. Stilling Well on Front End of the Flume with Elevated Humic Acid Spiking Solution Dosing In Progress.

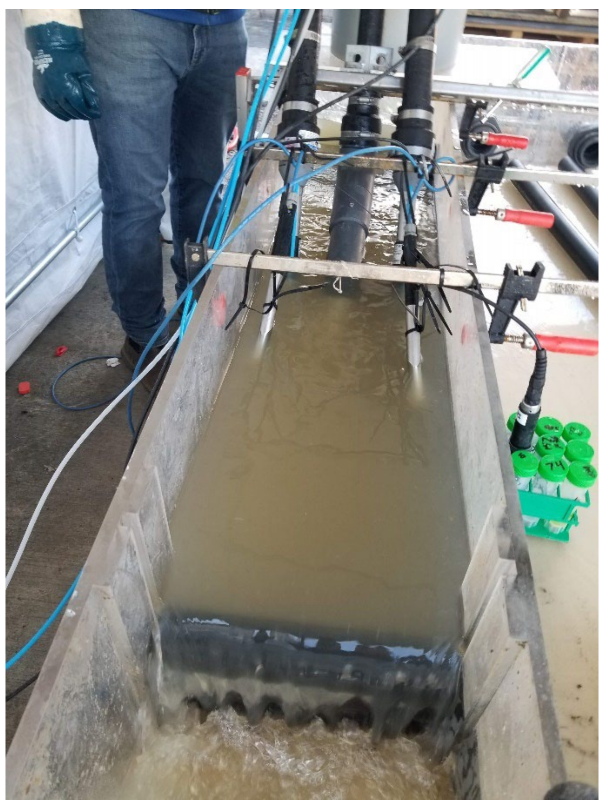


Figure A-4: Flume In Operation During Experiments with Two Pipes as Outlet Weir.

A.2.1.2 Sensors Installed in the Flume

The sensors installed in the flume for all experiments were:

- Electro-chemical Devices (ECD) brand Extended Life pH Electrode RADEL body: CWS has extensive experience with this ECD pH probe in the collection system. The head was recently replaced to restore peak performance. The sensor used a slightly different technology than most pH sensors with a flat glass membrane instead of a bulb and is built to be more resistant to chemical fouling. The sensor records pH and temperature every 10 minutes and is connected to a Telog RU-35 datalogger using battery power. Cellular telemetry was used to automatically send the data to CWS's Enterprise system, and the data were automatically brought into the Aquarius Time Series program for analysis, quality assurance, and comparison with other data.
- ECD brand ORP Pt Cap peek, two-tang probe: CWS had used this ECD oxidation reduction potential (ORP) probe in the collection system for experiments over the past year. The head was recently replaced to restore peak performance. The reading frequency, power source, and telemetry were identical to the ECD pH sensor.
- s::can Spectro::lyser: s::can provided a spectro::lyser instrument with supporting equipment for this study. The spectro::lyser is essentially a spectrophotometer placed in the water that emits light across the visual and ultra-violet (UV) ranges of wavelengths and measures the absorption at different wavelengths. The instrument uses these absorption values across the spectrum of wavelengths to quantify an equivalent of biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), nitrate, UV254 absorption, and temperature. The distance between the light source and the sensor is approximately 2 mm, and an air blast is sent through the slot between the light source and the lens every 2 minutes to reduce fouling and buildup on the lens. A reading is collected every 2 minutes immediately following the air-blast. The data are stored locally on a con::cube that provides the data storage and software for operation of the device and analyzing results. This con::cube is accessible remotely through a web page, so the data were available remotely. In addition, the device could also be controlled remotely. An RU-35 datalogger was also connected to the con::cube to bring the data into the Enterprise system and the Aquarius Time Series program for analysis and comparison with other data.
- s::can pH::lyser: s::can provided a pH::lyser pH sensor for this study. This pH sensor used a different technology than typical pH sensors with a non-porous/non-leaking combined reference electrode. This provides the benefit of being much less affected by fouling from FOG and rags; the sensor is described by s::can as maintenance-free. The pH sensor also collected temperature data and had the same compressed air blasts, reading frequency, power source, and data storage as the spectro::lyser.
- s::can condu::lyser: s::can provided the condu::lyser electroconductivity sensor for this study. This electroconductivity sensor also collected temperature data and had the same compressed air blasts, reading frequency, power source, and data storage as the spectro::lyser and pH::lyser. It was also designed to be maintenance-free and resistant to effects from FOG and rags.

- YosemiTech Y532-A: CWS had been using this pH sensor in surface waters in the watershed for the past year. The head was recently replaced to restore peak performance. This is a very inexpensive sensor used by CWS in its open source, EnviroDIY monitoring of surface waters. This sensor had not been previously tested in a wastewater environment. It uses a traditional sensing method of determining a potential on a porous glass bulb. Readings of pH and temperature are taken every 10 minutes and sent to a Mayfly unit that stores the data. The Mayfly is powered by a battery that is recharged using a solar panel. Cellular telemetry is used to send the data to MonitorMyWatershed, an open source cloud for environmental data, where it is accessible remotely by CWS and read automatically into Aquarius Time Series with the rest of the data from the other sensors.

The sensors were calibrated using methods specific to the sensor as instructed by the manufacturer. The ECD and Yosemitech pH sensors used multi-point calibration using standard pH buffer solutions of 4, 7, and 10. The ORP sensor was likewise calibrated using a standard solution provided by the manufacturer. The s::can sensors were calibrated by collecting laboratory samples of the different analytes (usually 9 grab samples) at different parts of the day over several days. The samples were analyzed for BOD5, COD, TSS, nitrate, and oil and grease using Standard Methods. The results from these samples were used by s::can personnel to calibrate the spectro::lyser remotely. The pH::lyser and the condu::lyser were pre-calibrated and were not calibrated by CWS. However, the readings were frequently compared to a Hach HQ40D portable multimeter that measured pH, temperature, and conductivity. This multimeter was calibrated to standard solutions.

All of the sensors were cleaned between experiments by soaking the bulb, disc, or lens with a Kimwipe soaked in 3 percent HCl solution in DI water for approximately 5 minutes. The lens was also gently wiped using this Kimwipe. The bulb and discs were only wiped as needed very delicately to remove visible buildup not removed by soaking in the 3 percent HCl solution. The flume and the non-sensing parts of the sensors were cleaned thoroughly between each experiment using tap water and a brush to avoid buildup between experiments that could affect flow or sensor readings.

A.2.1.3 Experiments

Ten experiments were conducted to analyze the effect of different variables. It was anticipated that sensor performance and needs for maintenance would be affected by flow velocity. The velocity was controlled in each experiment by maintaining a constant flow rate and width but changing the depth using the outlet weir. Three depths were analyzed (3, 5, and 7 inches), which led to three velocities of approximately 0.6, 0.8, and 1.3 ft/s. These velocities were estimated from the dimensions of the flume and the measured flow rate and verified using a handheld velocity meter. A depth of 3 inches was the smallest possible to keep the sensors consistently submerged.

Each experiment consisted of the following procedures:

1. Clean the sensors and flume and check the calibration.
2. Set all parameters to the desired variable for the experiment (e.g. velocity, FOG, etc.).
3. Conduct an initial spike test as described below.

4. Allow the sensors to operate for a prescribed length of time (usually 1 week) with daily handheld measurements and grab samples every other day to compare to the sensor readings.
5. Conduct a final spike test (with some experiments also having an intermediate spike test).

The conditions of each experiment are summarized in the table below. The timing of the experiments is shown in Figure A-5. Three velocities/depths were tested with two replicates each in Experiments 1-6. For Experiments 7a-7c, additional FOG and/or rags were added to the probes using different methods (see below). For Experiment 8, the intention was to test conditions where the probes were not cleaned or adjusted for 24 days. However, an event occurred at the plant on June 3rd that affected the lens on the s::can as well as the other sensors. They were, therefore, cleaned after 7 and 14 days during this experiment with a new spike test following each cleaning. The spectro::lyser and pH::lyser sensors continued to drift soon after cleaning, so this experiment was abandoned after 24 days without conducting final spikes while the team did troubleshooting to figure out how to restore the probes.

Table A-1. Summary of Experiments Conducted Using the Flume and Sensors.

Experiment Number	Velocity (ft/s)	Depth (inches)	Spikes	Duration (days)	FOG/rag introduction
1	1.3	3	2 (start/end)	7	Normal
2	0.8	5	2 (start/end)	7	Normal
3	0.6	7	2 (start/end)	7	Normal
4	1.3	3	2 (start/end)	7	Normal
5	0.8	5	2 (start/end)	7	Normal
6	0.6	7	2 (start/end)	7	Normal
7a	0.6	7	3 (start/mid/end)	2	Grease dipped
7b	0.6	7	3 (start/mid/end)	2	Rag wrapped
7c	0.6	7	3 (start/mid/end)	2	Increased FOG
8	0.6	7	3 (start/7-day/14-day)	24	Normal

A.2.1.4 FOG introduction methods for Experiments 7a-7c

Three methods were used to introduce additional FOG and rags to the probes for Experiments 7a-7c. In all three experiments, a spike test was performed immediately after introduction of the sensors to the rags/FOG, and after 24 and 48 hours. The specific methods of introducing the rags and/or FOG were:

- Grease dipped: Hardened grease was collected off of the influent stilling basin of the WWRF and warmed in a small bucket to a softer consistency. The probes were then dipped into this grease and then dipped into a bucket of ice water to allow the grease to harden on the probe. While being dipped in the grease, the air blast cleaning process was turned off on the s::can sensors. The probes were then inserted in the flume and the air blasts were allowed to resume.
- Rag wrapped: Baby wipes were soaked in influent water for 10 minutes, then dipped in grease from the influent stilling basins and wrapped around each of the cleaned probes. Rubber bands were used to secure the greased wipes to the probes. During this process, the air blast for the s::can sensors was turned off. The probes were then inserted into the flume and the air blasts were allowed to resume.

- Increased FOG: The FOG content of the influent was increased by dosing 20 gallons of liquified grease from the influent stilling well into the flume over 20 minutes. This simulated a slug of high-FOG water coming from a discharge in the sewershed and allows the FOG to collect on the sensors more organically.

A.2.1.5 Spike Tests

At the beginning and end of each experiment, and in the middle of Experiments 7a-7c, a spike test was done by simulating an event at the plant to test the ability of the sensors to detect rapid changes in the measured parameters when newly cleaned and when fouled over a period of time. Two spikes were done for each spike test. The first consisted of adding five gallons of a spiking solution into the stilling well on the upstream side of the flume over approximately 20 minutes with an approximate dosage rate of 0.25 gpm. The spiking solution had elevated pH, ORP, and conductivity. The mixture varied somewhat between spiking experiments, but generally consisted of 5 gallons of tap water, 26 oz (737 grams) of Morton table salt, 550-600 mL of Clorox bleach (3 percent hypochlorite), and 550-600 mL of sodium hydroxide (25 percent strength v/v). The handheld multimeter was placed in the flume throughout the spiking test to verify that the electroconductivity and pH both increased well beyond the baseline values. Immediately following this first spike test, a second spike test was conducted using 5 gallons of tap water mixed with 550 mL of Aquahume (now named Superhume) fulvic/humic acid concentrate (12 percent strength). This was dosed into the water at a similar rate as the previous spike over 10 minutes (using only half of the 5 gallons). This caused the water to become a darker brown and increased the BOD and COD detected by the spectro::lyser. The humic acid was dissolved, so it did not increase the TSS. No increase in TSS was simulated in the spiking tests. However, TSS was monitored during the BOD/COD spike to confirm that the TSS did not increase during this spiking event as a test of the sensor’s ability to distinguish between dissolved organic matter and suspended solids.

One variable that could not be controlled was the quality of the influent coming into the plant. Throughout the experiments, there were noticeable changes in influent water quality. A diurnal pattern in most parameters was very clear, and there were occasional events where one or more of the parameters would change rapidly in response to some change in the influent, likely a discharge from an industry. The ability of the probes to detect these patterns and events were documented and are discussed in the results section.

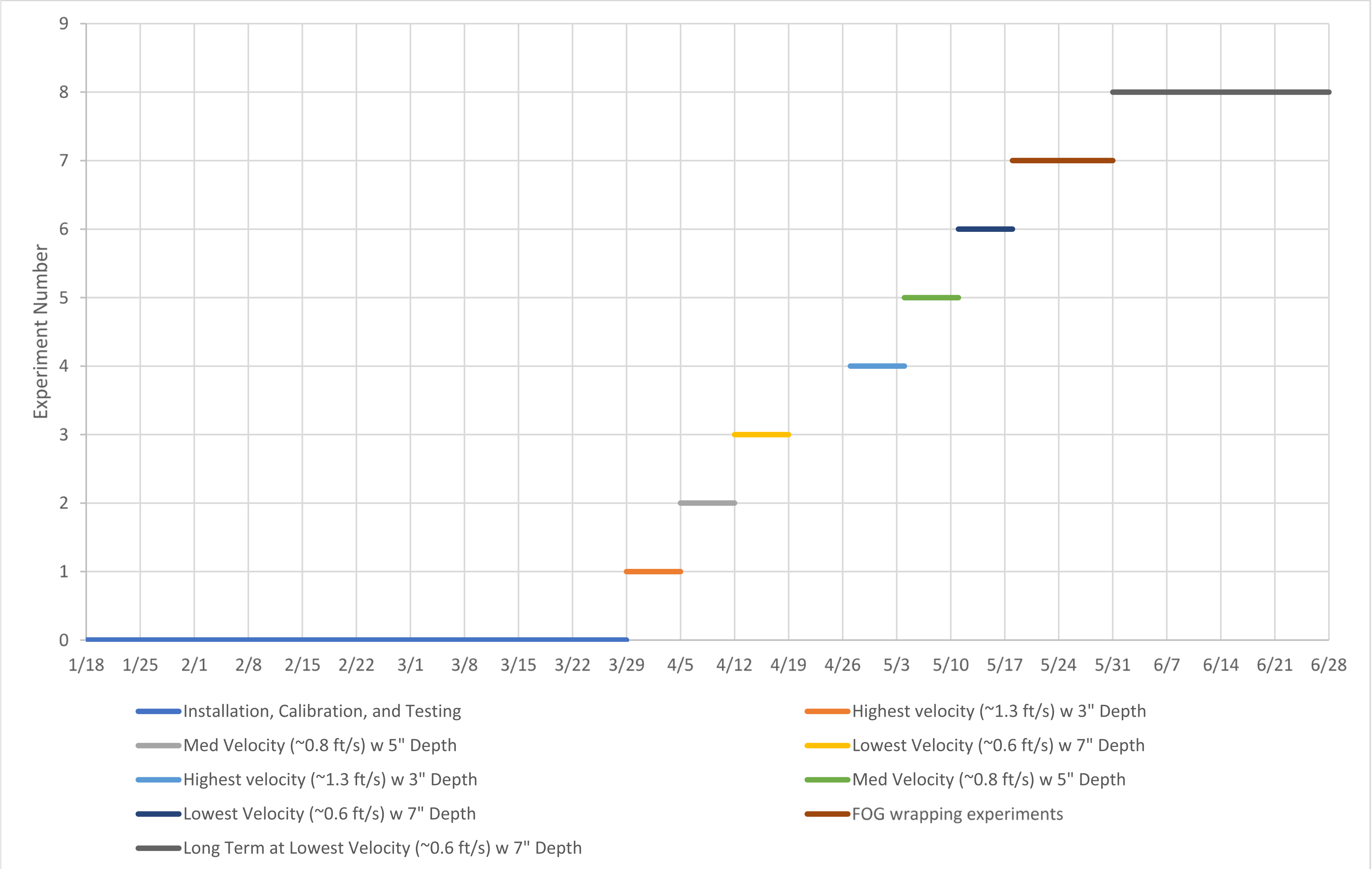


Figure A-5. Schedule of the Experiments Conducted Using the Flume Along with Conditions of each Experiment.
FOG grapping experiments consisted of 3 shorter experiments (7a-7c).

A.2.2 Subtask 2: Develop and Test Sensor Containment Device in Collection System

CWS has developed a sensor holder (called the ‘rag guard’) that had shown promise in decreasing rag buildup on sensors in the collection system. In this subtask, the performance of this rag guard was more rigorously analyzed by installing identical sensors at the same location in the sanitary collection system with one set in the rag guard and the other set in a generic sensor holder like those received from the manufacturer. A manhole in the Forest Grove sewershed was chosen for this testing. This manhole had frequently been shown to have pH spikes based on previous monitoring and had adequate flow so that the sensors remained immersed at night.

The rag guard is shown in Figures A-6, A-7, and A-8. The generic sensor holder is shown in Figure A-8 and Figure A-9. The rag guard is a section of 4-inch diameter PVC pipe with a 3 foot radius, 90 degree bend (Figure A-6). The rag guard device is suspended by stainless steel cables so to hang freely, and the angle and depth is directed by the flow (Figure A-8). The device remains pointed downstream because of the flow but is able to temporarily tilt or twist to avoid collecting debris and rags. The device holds the sensors parallel to the flow and provides no surface perpendicular to the flow for rags to collect, having only a curved, smooth surface facing upstream (Figures A-7 and A-8). The last length of the device has slats that get progressively wider with distance downstream (Figure A-7). This allows the probes to be submerged with water with a similar velocity to the water outside the rag guard but allows any rags that penetrate the slats to be pushed out by the flow.

The generic sensor holder is based on a design frequently used from manufacturers during pilot tests for CWS. It consists of a long PVC pipe that is mounted to the side of the manhole (Figure A-8). This holds the sensors firmly in position, but unable to move in response to debris or rag collection. The sensors are installed at a 45-degree angle pointing downstream (Figures A-8 and A-9). While this angle is designed to help rags slide off, the holder and probes are not perfectly smooth, which provides opportunity for rag collection. The sensors themselves are immersed in the water without any kind of protection (Figure A-9).

Two ECD sensors (pH and ORP, identical to those installed in the flume) were installed in the rag guard and in the generic sensor holder. Both had similar reading frequencies, dataloggers, and telemetry as the ones in the flume. Maintenance was only performed on an “as-needed” basis based on monitoring the data indicating that the sensors were not operating properly. One of the purposes of this experiment was to determine the required frequency of maintenance.

No individual experiments were done, but the sensors were both installed for the duration of the study and their performance with and without the rag guard was analyzed. The sensors were originally installed on March 24th, 2021. However, trouble with the sensors and telemetry caused troubleshooting which meant that only two extended periods of data were representative:

- March 24th-April 19th: Sensors in the generic sensor holder were not functioning, but data was being collected from the sensors in the rag guard. Several real pH events were detected at the WWRF during this period which allowed for testing the ability of the sensor in the rag guard to detect these events.

- June 7th-July 2nd: The pronged generic sensor holder was replaced with a single generic sensor holder to avoid the problems the prong was causing, and both sensors operated normally until the end of the study. This period allowed the comparison of probe performance in the rag guard versus the generic sensor holder.

Each of these periods is slightly less than 1 month, which is less time than was desired for observation, but is sufficient to illustrate the performance of the rag guard. The pronged generic sensor holder shown in Figures 8 and 9 was replaced with a single generic sensor holder on June 7th to avoid the issues the prongs were causing.

Figure A-6. Rag Guard Sensor Holder Showing the Enter Device.

Figure A-7. Rag Guard Sensor Holder Showing the Probes as Installed in the Slats.
Slats get wider with distance downstream.

Figure A-8. Rag Guard Sensor Holder and Generic Sensor Holder Installed in the Same Manhole.
This is showing the 'pronged' sensor holder initially used. The prong was eventually removed during the experiment to provide a gap between the sensor holders.

Figure A-9. Generic Sensor Holder with the Two Probes (pH and ORP).
This shows the 'pronged' sensor holder initially used to hold the pH and ORP probes. The ORP probe was later removed along with the prong to provide more space between the sensor holders.

A.3 Results and Discussion

With continuous data collected over so many parameters and so many experiments, timeseries of all of the data are not shown in this report. Timeseries of all of the data, along with analyses, are available in a spreadsheet by request. A summary of the qualitative and quantitative results of each of the experiments in the flume is shown in Table A-2. Several timeseries of individual parameters for some experiments are shown to illustrate common patterns and observations, and timeseries of pH from both sensors at the test manhole are shown. Visual observation of the timeseries along with the quantitative and qualitative results shown in Table A-2 formed the primary bases for the findings from this study. The results presented in this report are directed mainly towards supporting the reported findings rather than providing a comprehensive record of the results of all of the experiments.

A.3.1 Flume Experiment Results

Table A-2 summarizes the qualitative and quantitative results from each experiment for each of the parameters. ORP is not shown in the table as the ORP sensor performed very poorly and did not provide virtually any reasonable results. The patterns and observations for COD were similar for BOD and UV254, so only the results from the COD data are shown in the table. For each experiment, the table shows several results indicating how the sensor performed. These include:

- Qualitative assessment of how well the sensor detected each of the spiking events. This was based on visual observation of each spiking event for each sensor in each experiment. Some example plots are shown in the discussion below. Notes are provided, as needed, to clarify and justify the qualitative assessment.

- The average difference between the values reported by the laboratory of a grab sample or the handheld multimeter and the values reported by the sensor at the same moment. Data collected during spike test periods were excluded for this calculation.
- Qualitative assessment of the agreement of the sensor with the laboratory or multimeter values. These observations were based on a combination of the calculated difference above and a visual observation of the timeseries for each parameter in each experiment. Justification and additional notes are also provided, where appropriate. Some example timeseries are shown in the discussion below.
- Qualitative assessment of the ability of the sensor to consistently detect the diurnal pattern of the analyte and to detect any real events that occurred during the experiments. This was based on visual observation of the timeseries and comparison with the influent SCADA values.
- The squared Pearson correlation coefficient (R2) of the pH data from each of the three pH sensors in the flume with the influent pH recorded in the SCADA system. The pH sensor in the influent was cleaned daily by treatment plant personnel, but still had periods where it was temporarily fouled. Periods when the quality of the influent SCADA pH data were poor were removed for this calculation. Because the sensors in the flume each collected data at different intervals ranging from 2-minutes to 10-minutes, 10-minute averages were computed for the purposes of computing the correlation coefficient so that a common interval could be used. For all of the correlation coefficients shown in Table A-2, the p-value was computed, and it was always much less than 0.05 due to the large number of observations with 10-minute interval readings over days and weeks.

A.3.1.1 Effects of Velocity and Depth

Experiments 1-3 show the differences in performance observed due to changing the velocity and depth. Experiments 4-6 are equivalent to Experiments 1-3 and provide an additional replicate for observing the effect of changing velocity and depth on sensor performance. As summarized in Table A-2, decreases in velocity did not significantly affect sensor performance, in general. Qualitative and quantitative performance measures were either similar across Experiments 1 through 6 for a sensor or did not show a trend of decreasing performance from 1 to 3 and from 4 to 6. The lone exception to this is the ECD pH sensor, where the ability of the sensor to detect the final spike is very good at the highest velocity experiments and poor at the lowest velocity experiments. Further analysis, however, showed that this was not due to changes in velocity (see section below). The agreement with the handheld pH and the agreement with the influent SCADA pH also do not show a strong trend with velocity and depth for the ECD pH sensor.

The lack of trend in performance with velocity and related depth was unexpected given previous observations in the sanitary collection system where locations with very high velocity tended to have less FOG buildup. One potential explanation for this is the limited range in velocities that could be applied to the flume given the flow rate and flume width. The maximum velocity that could be applied while maintaining at least 3 inches of depth was 1.3 ft/s. Velocities higher than this certainly occur in the collection system at some locations, and it is possible that a wider range of velocities may show more of a relationship with sensor performance. Adjustments to the flume are in discussion for being able to increase the flow rate. A second potential explanation is that the higher velocities in the flume were accompanied by lower depths that provided less cross-sectional area to enable debris to get around or under the sensors. This may have offset the decreased ability for FOG to deposit on the sensors at the higher velocity.

Table A-2: Summary of Results from Each Experiment for Each Parameter and Sensor in the Flume.

All p-values for the correlation coefficients were <<0.05, so are not shown in this table.

* = cleaned sensors after 7 and 14 days due to drift that occurred after an event. ** = sensors cleaned just before this spike, so it does not represent conditions after the experiment period. *** = The handheld pH sensor consistently had values roughly 0.5 standard units higher than the influent SCADA sensor, s::can sensor, and ECD sensor when all were clean and calibrated. For computing the average difference between the handheld sensor and the s::can and ECD sensors, 0.5 was subtracted from the value to create the value for this table. The Yosemitech pH sensor also had values roughly 0.5 higher than the others similar to the handheld, so this adjustment was not applied to that sensor.

	Experiment	**1**	**2**	**3**	**4**	**5**	**6**	**7a**	**7b**	**7c**	**8**
	Velocity (ft/s)	**1.3**	**0.8**	**0.6**	**1.3**	**0.8**	**0.6**	**0.6**	**0.6**	**0.6**	**0.6**
	Depth (inches)	**3**	**5**	**7**	**3**	**5**	**7**	**7**	**7**	**7**	**7**
	Duration (Days)	**7**	**7**	**7**	**7**	**7**	**7**	**2**	**2**	**2**	**24***
	Added Grease	**None**	**None**	**None**	**None**	**None**	**None**	**Grease dipped**	**Grease wipe wrapped**	**Grease dosed**	**None**
s::can pH	Agreement with Influent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent, though a consistent offset of about 0.1.	Excellent, though a consistent offset of about 0.2	Excellent, though a consistent offset of about 0.2	Good, but drifts away from influent after each cleaning
s::can pH	Correl Coeff with Influent	0.94	0.73	0.77	0.76	0.88	0.84	0.74	0.68	0.79	0.52
s::can pH	Agreement with Handheld	Excellent	Excellent	Excellent	Good	Excellent	Good	Only 1 value, but matches well	Only 1 value, but matches well	Excellent	Excellent
s::can pH	Average difference w handheld***	0.01	-0.13	-0.16	-0.29	0.1	-0.29	-0.16	-0.23	-0.13	-0.08
s::can pH	Detected Real Events and Diurnal Patterns	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Very good, but drifts after each cleaning
s::can pH	Detected Initial Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
s::can pH	Detected Second Spike	NA	NA	NA	NA	NA	NA	Excellent	NA	Excellent	Excellent
s::can pH	Detected Final Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
ECD pH	Agreement with Influent	Fair	Poor, except for first day and last day	Poor	Visually looks much better than the coeff,	Excellent for first 3 days, then fair as it drifts upwards	Excellent for first 3 days, then fair as it drifts upwards	Poor, severely affected by the grease	Poor, severely affected by the grease	Fair	Poor
ECD pH	Correl Coeff with Influent	0.27	0.07	0.01	0.11	0.28	0.03	0.08	0.1	0.13	0.02
ECD pH	Agreement with Handheld	Good	Very good	Very good	Excellent	Upward drift after 3 days causes larger difference	Underpredictions early and overpredictions late	Never recovers from grease	Never recovers from grease	Fair	Fair
ECD pH	Average difference w handheld***	0.38	0.2	0.21	0.05	-0.27	-0.05	2	1.34	-0.77	-0.71
ECD pH	Detected Real Events and Diurnal Patterns	Very good when not fouled, but fouled first day and last 2 days	Poor, except for first day and last day	Poor, except for a few days in middle	Excellent except for a few periods of drift, but recovers	Excellent for first 3 days, then fair as it drifts upwards	Excellent for first 3 days, then fair as it drifts upwards	Poor	Poor	Shift partway through and lots of oscillation, but patterns are ok	Pattern is ok, but drifts and offset
ECD pH	Detected Initial Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Poor	Poor	Excellent	Excellent
ECD pH	Detected Second Spike	NA	NA	NA	NA	NA	NA	Poor	NA	Excellent	Excellent
ECD pH	Detected Final Spike	Excellent	Excellent	Poor	Excellent	Poor	Poor	Poor	Poor	Good	Excellent

Table A-2: Summary of Results from Each Experiment for Each Parameter and Sensor in the Flume. (Continued)

All p-values for the correlation coefficients were <<0.05, so are not shown in this table.

* = cleaned sensors after 7 and 14 days due to drift that occurred after an event. ** = sensors cleaned just before this spike, so it does not represent conditions after the experiment period. *** = The handheld pH sensor consistently had values roughly 0.5 standard units higher than the influent SCADA sensor, s::can sensor, and ECD sensor when all were clean and calibrated. For computing the average difference between the handheld sensor and the s::can and ECD sensors, 0.5 was subtracted from the value to create the value for this table. The Yosemitech pH sensor also had values roughly 0.5 higher than the others similar to the handheld, so this adjustment was not applied to that sensor.

	Experiment	**1**	**2**	**3**	**4**	**5**	**6**	**7a**	**7b**	**7c**	**8**
	Velocity (ft/s)	**1.3**	**0.8**	**0.6**	**1.3**	**0.8**	**0.6**	**0.6**	**0.6**	**0.6**	**0.6**
	Depth (inches)	**3**	**5**	**7**	**3**	**5**	**7**	**7**	**7**	**7**	**7**
	Duration (Days)	**7**	**7**	**7**	**7**	**7**	**7**	**2**	**2**	**2**	**24***
	Added Grease	**None**	**None**	**None**	**None**	**None**	**None**	**Grease dipped**	**Grease wipe wrapped**	**Grease dosed**	**None**
Yosemitech pH	Agreement with Influent	Very good for first 4 days, fair after that	Very good for first 3 days and last 2 days, fair in middle	Very good	Visually looks better, but has offset, missing periods, and drift	Matches well, but missing data the vast majority of the time.	Only 1 day where data are present and matching influent	Poor, severely affected by the grease	Poor, severely affected by the grease	Fair	Poor
Yosemitech pH	Correl Coeff with Influent	0.30	0.47	0.61	0.29	0.2	0.04	0.1	0	0.16	0
Yosemitech pH	Agreement with Handheld	First 3 excellent, last one good	First 2 good, last 2 poor	Value is pretty consistent	Excellent	Average is great, but high variability	Average is good, but high variability	Never recovers from grease	Never recovers from grease	Very good	Excellent
Yosemitech pH	Average difference w handheld***	0.05	0.33	-0.3	-0.13	0.06	0.22	1.95	2.1	0.22	0.08
Yosemitech pH	Detected Real Events and Diurnal Patterns	Excellent first 4 days, fair after that	Very good	Excellent	Excellent most of the time except offset by ~ 1 pH unit and missing some periods	Matches pattern well, but 1 unit offset and missing data the vast majority of the time.	Fair. Drifts badly after the first 2 days. Also cuts out a lot	Poor	Poor	Good when it's not fouled, but has several periods of fouling	Fair. Several false positives, and drift within a day or two after each cleaning
Yosemitech pH	Detected Initial Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Not functioning	Poor	Poor	Excellent	Excellent
Yosemitech pH	Detected Second Spike	NA	NA	NA	NA	NA	NA	Poor	NA	Excellent	Excellent
Yosemitech pH	Detected Final Spike	Poor	Excellent	Fair	Poor	Not functioning	Not functioning	Poor	Poor	Good	Excellent
s::can COD (Similar for BOD and UV254)	Agreement with Lab	Good before 3/31, poor after 3/31 (Drift)	Underpredicted both samples	Consistently underpredicts	Consistently underpredicts	2/3 very good	2/3 very good	Only 1 value, but matches well	Only 1 value, but matches well	1/2 very good	Poor due to drift. Better soon after each cleaning.
s::can COD (Similar for BOD and UV254)	Average Difference from Lab Values (mg/L)	-167	283	269	249	66	101	32	36	63	-59
s::can COD (Similar for BOD and UV254)	Detected Real Events and Diurnal Patterns	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
s::can COD (Similar for BOD and UV254)	Detected Initial Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
s::can COD (Similar for BOD and UV254)	Detected Second Spike	NA	NA	NA	NA	NA	NA	Excellent	NA	Excellent	Excellent**
s::can COD (Similar for BOD and UV254)	Detected Final Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent**

Table A-2: Summary of Results from Each Experiment for Each Parameter and Sensor in the Flume. (Continued)

All p-values for the correlation coefficients were <<0.05, so are not shown in this table.

* = cleaned sensors after 7 and 14 days due to drift that occurred after an event. ** = sensors cleaned just before this spike, so it does not represent conditions after the experiment period. *** = The handheld pH sensor consistently had values roughly 0.5 standard units higher than the influent SCADA sensor, s::can sensor, and ECD sensor when all were clean and calibrated. For computing the average difference between the handheld sensor and the s::can and ECD sensors, 0.5 was subtracted from the value to create the value for this table. The Yosemitech pH sensor also had values roughly 0.5 higher than the others similar to the handheld, so this adjustment was not applied to that sensor.

Experiment		**1**	**2**	**3**	**4**	**5**	**6**	**7a**	**7b**	**7c**	**8**
Velocity (ft/s)		**1.3**	**0.8**	**0.6**	**1.3**	**0.8**	**0.6**	**0.6**	**0.6**	**0.6**	**0.6**
Depth (inches)		**3**	**5**	**7**	**3**	**5**	**7**	**7**	**7**	**7**	**7**
Duration (Days)		**7**	**7**	**7**	**7**	**7**	**7**	**2**	**2**	**2**	**24***
Added Grease		**None**	**None**	**None**	**None**	**None**	**None**	**Grease dipped**	**Grease wipe wrapped**	**Grease dosed**	**None**
s::can TSS	Agreement with Lab	Very good, though some baseline drift	Excellent	Excellent	Excellent	Very good. Large difference due to 1 sample during turbulent period.	Excellent	Only 1 value, but matches well	Only 1 value, but matches well	Excellent	Poor due to drift. Better soon after each cleaning.
	Average Difference from Lab Values (mg/L)	-23	13.7	-31	16	-558	50	41	13	53	-566
	Detected Real Events and Diurnal Patterns	Excellent	Excellent	Excellent	Excellent, but noisier than other experiments	Excellent	Excellent	Excellent	Excellent	Excellent	Very good
	Does not respond to dissolved humic acid spikes	Excellent	Excellent	1/2 Excellent, the other is affected	Excellent	Good, variability during spikes inconclusive	Good, variability during spikes inconclusive	2/3 Excellent, the other is affected	Excellent	Excellent	Very good
s::can Conductivity	Agreement with Handheld	Excellent	Very good, difference driven by 1 sample during turbulent period	Excellent	Excellent	Excellent	Excellent	Only 1 value, but matches well	Only 1 value, but matches well	Excellent	Excellent
	Average Difference from Handheld (uS/cm)	-29	-207	-22	-1	-46	-13	3	-10	29	-9
	Detected Real Events and Diurnal Patterns	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Detected Initial Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent
	Detected Second Spike	NA	NA	NA	NA	NA	NA	Excellent	NA	Excellent	Excellent**
	Detected Final Spike	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent**

A.3.1.2 Effects of Sensor Types

The greatest effect on performance of the sensors was the sensor type. While all sensors detected the spikes well and matched the lab, handheld, or SCADA pH values well when they were clean, they differed greatly in their ability to resist fouling and their response to being fouled during the course of each experiment. The performance of each of the sensor types is described in the subsections below.

A.3.1.2.1 s::can spectro::lyser

The spectro::lyser measured BOD, COD, TSS, NO3, UV254, and temperature. The performance was similar for each of these parameters, so only COD and TSS are discussed in this section. Figure A-10 shows the COD timeseries during Experiment 6 (with spiking periods removed), which was fairly typical of the other experiments. The figure shows that the sensor detected the diurnal pattern of COD consistently and detected real events that occurred in all experiments in Experiment 6. This was the case for all of the experiments. While the sensor detected the patterns and spikes closely, the reported values often differed fairly substantially with the measurements of the grab samples despite being calibrated (Figure A-10 and Table A-2). These differences tended to get smaller over the course of the study even though no recalibration was done (Table A-10).

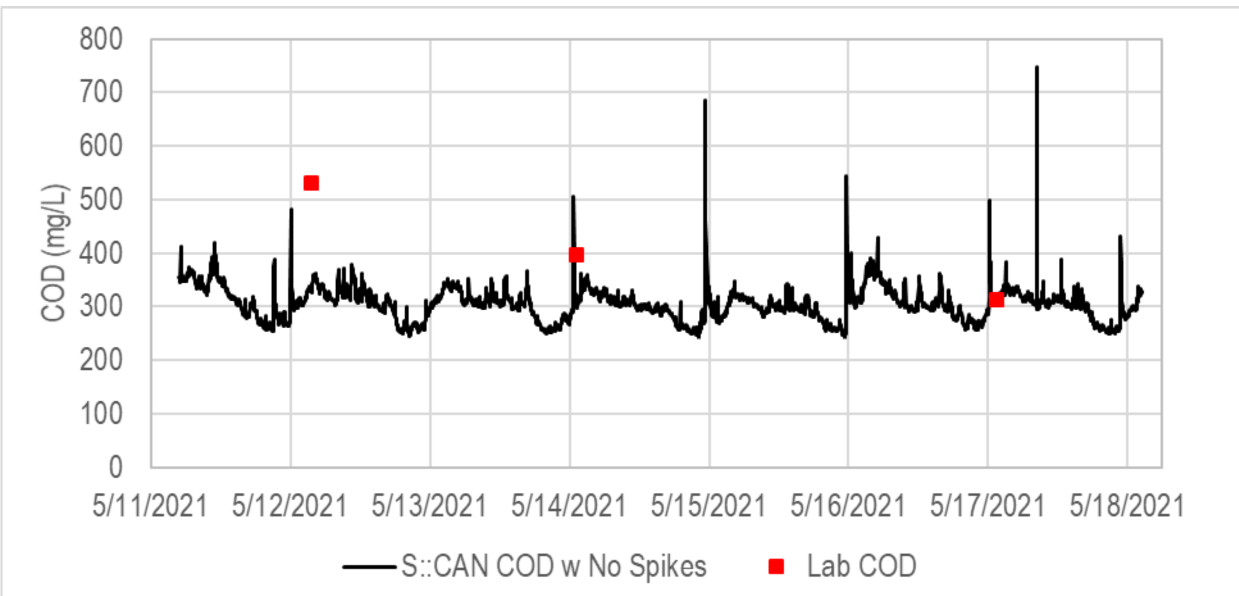


Figure A-10. Timeseries of COD in Experiment 6.
Spiking periods removed as measured by the s::can spectro::lyser and by grab samples.

Figure A-11 shows the COD measured by the spectro::lyser during the final spike test in Experiment 6. The sensor detected the spike closely even after 7 days in the influent with no maintenance. In fact, the sensor always detected the spikes throughout each of the experiments showing a remarkable resistance to fouling. The spike only caused an increase of approximately 50 mg/L of COD over 10 minutes, while real events that occurred where frequently much larger (Figure A-10). This showed that the sensor maintained high sensitivity to rapid changes in influent quality.

The COD baseline reported by the spectro::lyser was generally consistent showing no drift over the course of an experiment while still detecting the diurnal pattern, as shown in Figure A-10. This was typical of all of the experiments except Experiment 1 and Experiment 8 when an event occurred that affected the lens and caused the signal to drift steadily upwards. Figure A-12 shows the COD timeseries form the spectro::lyser in Experiment 1 where the baseline is seen to clearly drift upwards after the event on March 31. This drift was corrected after cleaning between Experiment 1 and Experiment 2 but was persistent even after cleaning during Experiment 8 where cleaning occurred on June 7 and June 14 (Figure A-13). The cause of this change in behavior was not determined before the end of the study where the equipment on loan from s::can was returned. Kimwipes used to clean the lens showed no obvious compound responsible for the fouling in initial tests. Other utilities have reported iron addition as a source of fouling of the lens of the spectro::lyser, but this was not revealed to be the case in this study.

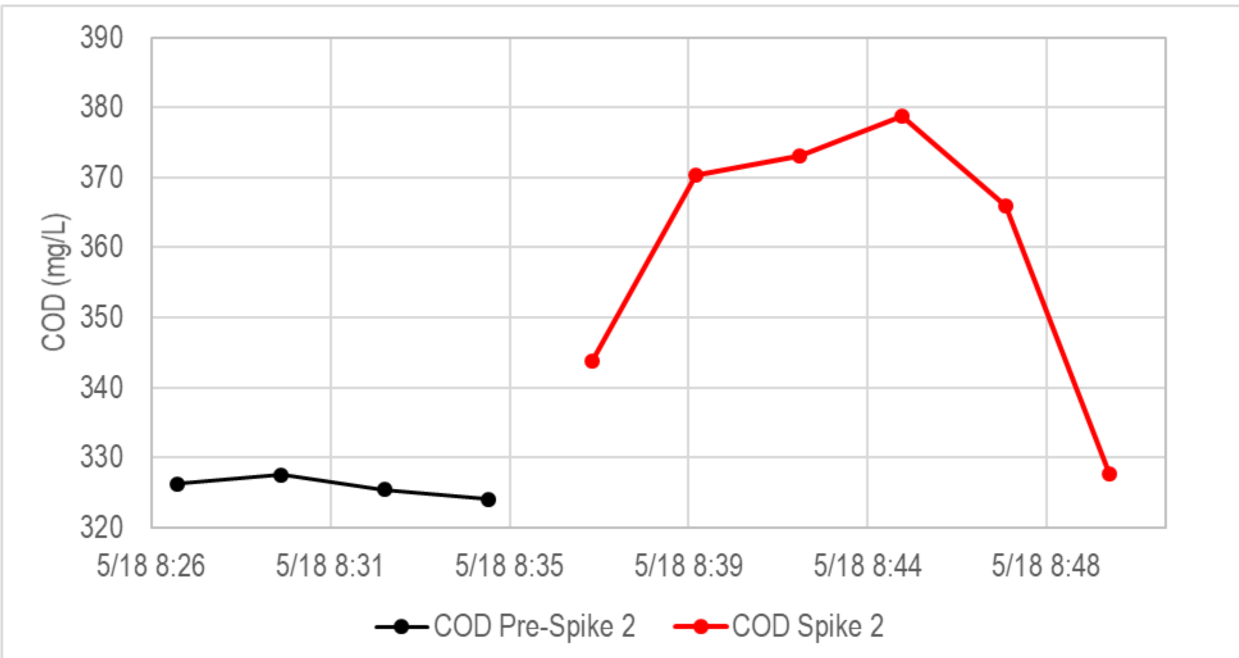


Figure A-11. COD Measured by the s::can spectro::lyser.
Before and during a spike test in Experiment 6 after 7 days in influent without cleaning.

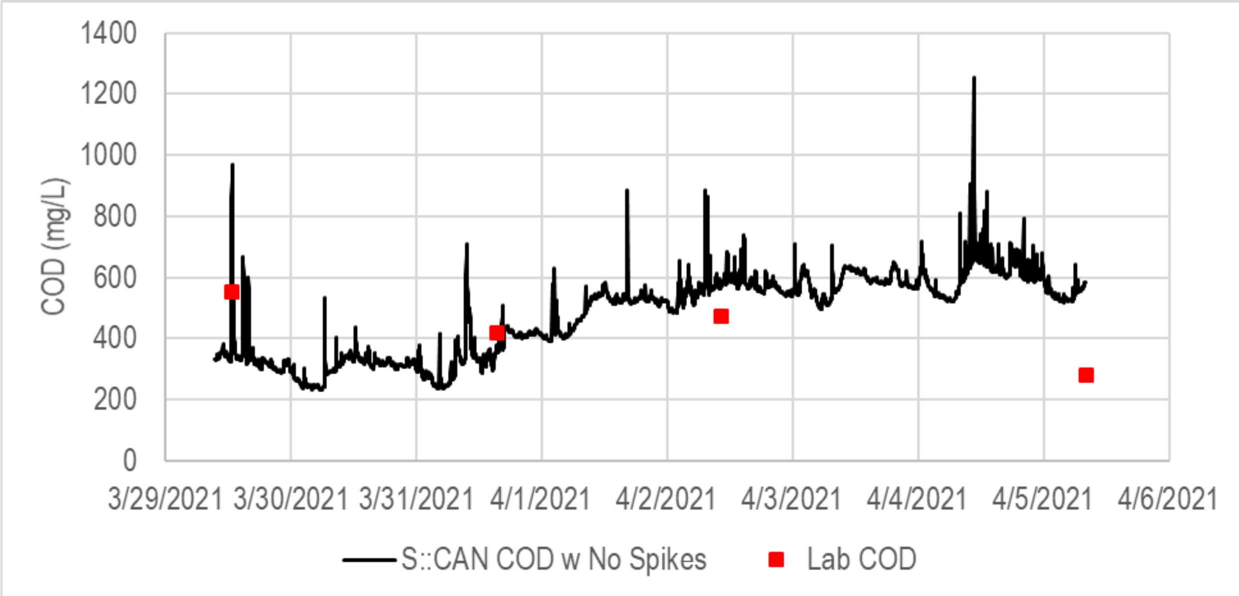


Figure A-12: Timeseries of COD in Experiment 1.
Spiking periods removed as measured by the s::can spectro::lyser and by grab samples.

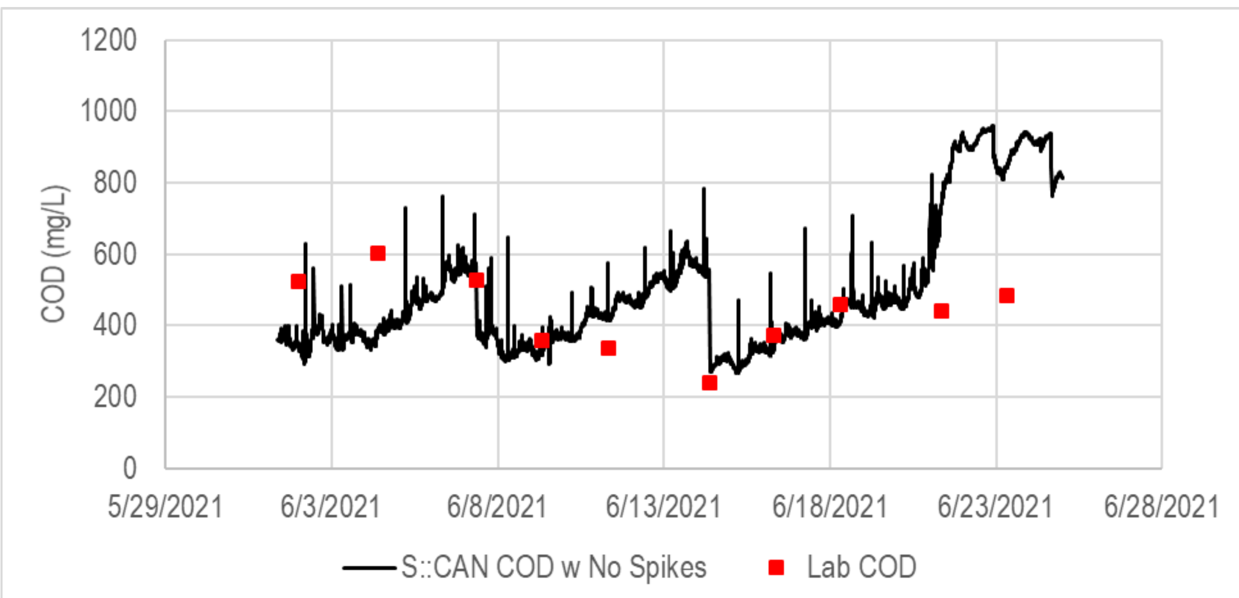


Figure A-13. Timeseries of COD in Experiment 8.
Spiking periods removed as measured by the s::can spectro::lyser and by grab samples. Cleaning occurred on June 7 and June 14.

The spectro::lyser detected TSS values close to the laboratory values and recorded the expected diurnal variations and real events that occurred (Table A-2). Figure A-14 shows the TSS timeseries reported by the spectro::lyser in Experiment 6, which was typical of the other experiments. The diurnal patterns and real events are clearly visible, the baseline is stable, and the values agree much better with the grab samples than did COD and some other optical parameters. In general, the spikes of humic acid did not affect the TSS reading by the spectro::lyser, showing that the instrument could distinguish the two different parameters. However, it was somewhat affected in a few of the spiking events as described in Table A-2. The same events that affected the COD in Experiments 1 and 8 also affected the TSS readings by the spectro::lyser causing similar baseline drift.

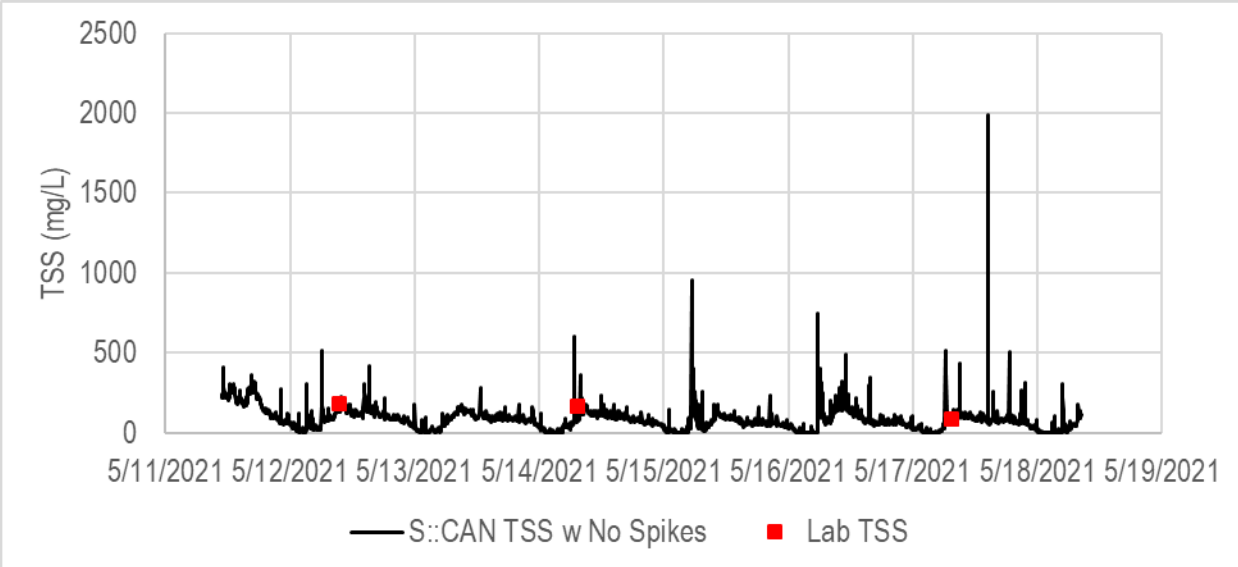


Figure A-14. Timeseries of TSS in Experiment 6.
Spiking periods removed as measured by the s::can spectro::lyser and by grab samples.

The spectro::lyser was able to detect the patterns, spikes, and real events even after at least 7 days without maintenance. This showed its capability of serving as an alarm system for events and analyzing temporal and spatial patterns in water quality at the influent and the collection system. However, for some of the parameters its lack of accuracy in reproducing the absolute values measured by the laboratory may limit its usefulness for applications that require accurate values such as plant operation, loading calculations, and strength calculations for billing purposes. However, improved calibration may help alleviate this. The sensor reported TSS values close to the laboratory samples in this experiment, and several other cities in Arizona and Washington use optical COD sensors to replace BOD and COD samples for billing and other calculations. With improved calibration, it may be capable of providing values that match laboratory values of COD more closely.

The current version of the spectro::lyser requires AC power and the datalogger/computer cannot be installed inside a manhole. It is, therefore, most useful at the influent or at established permanent monitoring stations in the sanitary sewer equipped with power and a utility box. s::can is developing a version that can be run off battery power and be placed inside manholes in the future.

A.3.1.2.2 s::can condu::lyser

The condu::lyser detected conductivity both accurately and precisely on a consistent basis (Table A-2). Similar to the spectro::lyser, the sensor detected patterns and real events, had a stable baseline, and detected all spike events. Figure A-15 shows a timeseries of conductivity measured by the condu::lyser in Experiment 1 that was typical of all other experiments. The reported values consistently agreed with the handheld measurements very well, and the baseline did not drift even in Experiment 1 and 8 when the spectro::lyser drifted after events that occurred at the influent. While the conductivity values measured by the condu::lyser were variable and did not show a strong diurnal pattern like most other parameters, they always agreed with handheld measurements in the flume using the multimeter and where correlated with other parameters such as nitrate measured by the spectro::lyser. Known temporal patterns such as pump station on/off cycles were visible during high conductivity events. Therefore, the variability in conductivity in the influent appeared to be real and not an artifact of the instrument. It is likely that intermittent industrial discharges and pump station cycles cause the large differences that mask the diurnal pattern.

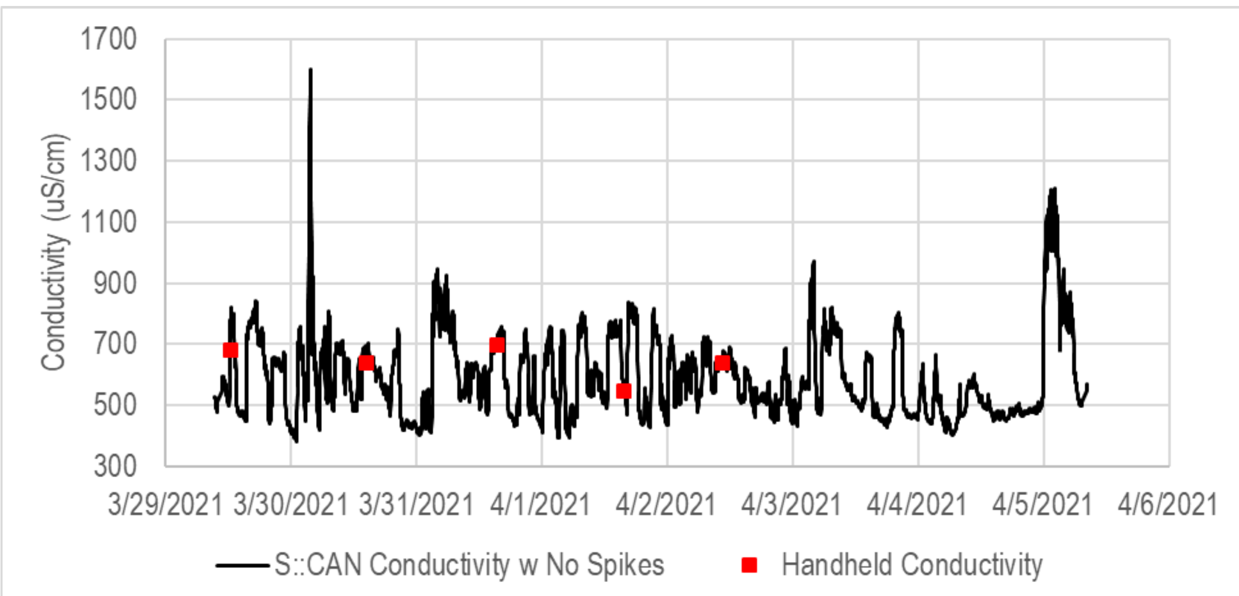


Figure A-15: Timeseries of Conductivity in Experiment 1.

Spiking periods removed as measured by the s::can condu::lyser and by the handheld multimeter.

A.3.1.2.3 pH Sensors

The three pH sensors om the flume differed fairly substantially in their performance in this experiment. While all three detected the initial spikes and matched the temporal pattern measured by the influent pH probe when they were clean, the s::can pH::lyser was much more consistent in its ability to detect spikes at the end of experiments and match the influent values over time than the other two sensors (Table A-2). Figure A-16 shows a timeseries of pH during Experiment 2 from all of the pH sensors. This was fairly typical of the other experiments. At the beginning of the experiment, all of the sensors detected an actual high pH event that occurred at the influent soon after the initial spiking test and matched the pattern of the influent pH sensor from SCADA well for approximately the first 24 hours. The ECD pH sensor showed more variation, and the Yosemitech sensor had a consistent offset from the influent values, but their temporal patterns were similar to the influent pH from SCADA (Figure A-16). The s::can pH::lyser continued to match the influent closely for the rest of the experiment, capturing several other high pH events that occurred. This sensor even captured some likely high pH events that occurred while the SCADA pH sensor was not operational such as the night of April 8th (Figure A-16). It had a stable baseline and reflected the same diurnal pattern measured by the influent pH sensor. The ECD pH sensor began to drift soon after the first day due to fouling, and the variation increased. This was typical of the other experiments, though it sometimes was stable for up to 2-4 days before drifting. Sometimes it corrected itself (likely a rag came off), and it matched the influent values well again temporarily, as occurred in the last 12 hours of Experiment 2 (Figure A-16). The Yosemitech pH sensor consistently had the approximately 0.5 unit offset from the influent, and it also tended to drift away from the influent value at times. However, it usually recovered more quickly and did not drift as significantly as the ECD pH sensor as shown in Experiment 2 (Figure A-16). It did tend to be less predictable with larger differences in behavior observed between experiments, while the ECD was more predictable in that it consistently drifted after the first few days and had more variation than the other sensors. The Yosemitech pH sensor detected most of the real pH events that occurred during this experiment but missed the last few on April 8th as it fouled and drifted downwards. The Yosemitech pH timeseries also showed many gaps including some long gaps in some of the experiments which caused it to miss spiking events (Table A-2). However, this was an artifact of the open source telemetry being employed with this sensor rather than an issue with the sensor itself.

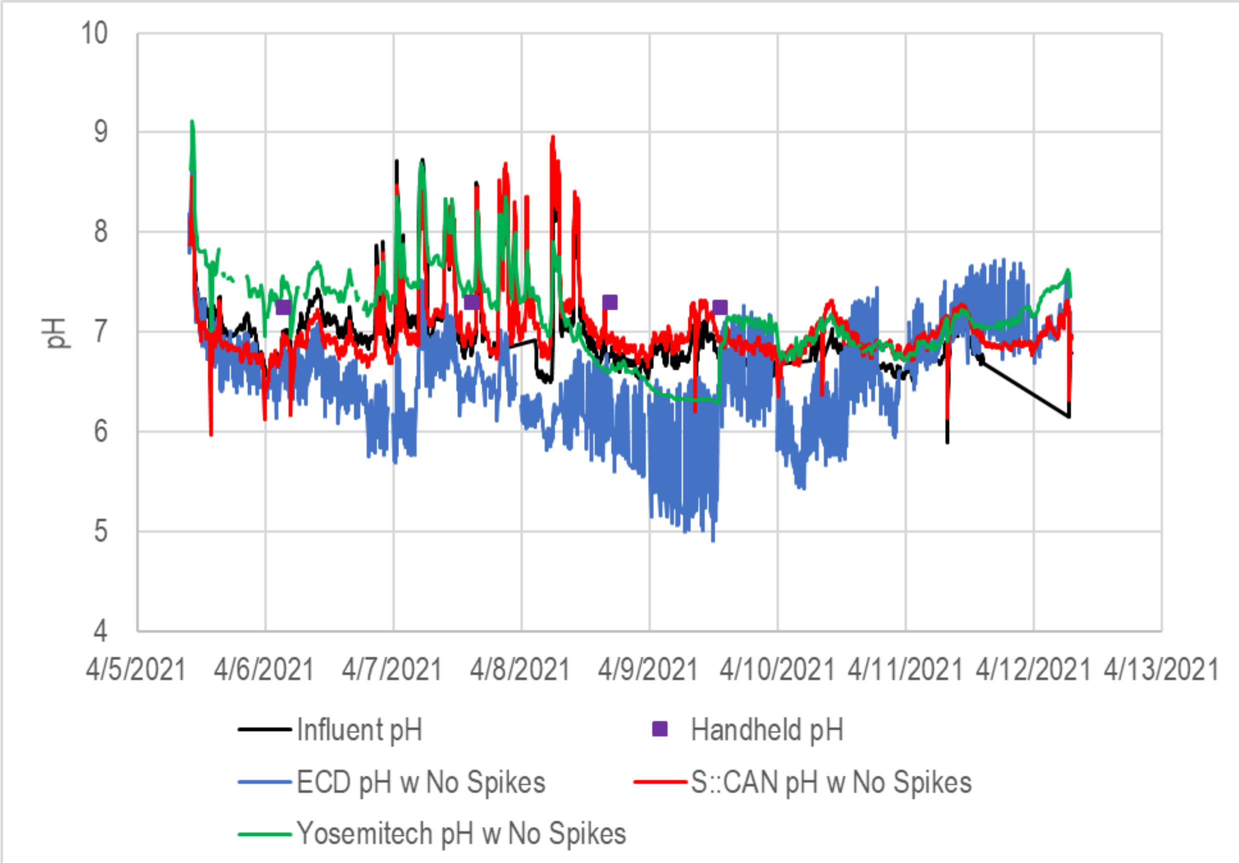


Figure A-16. Timeseries of pH in experiment 2 as measured by all three pH sensor types, the handheld multimeter, and the influent SCADA pH.

The ability of the ECD and Yosemitech sensors to detect the spikes at the end of the experimental period varied between experiments. This was mainly determined by whether or not the sensor was currently fouled when the experiment ended. For example, in Experiment 2, despite being fouled for long portions of the experiment, both the ECD and the Yosemitech sensors performed well for the last 12 hours of the experiment, and both detected the final spike in that experiment much better than in most of the other experiments (Figure A-16). The s::can pH sensor consistently detected all spikes (Table A-2). Figure A-17 shows the 10-minute average of the pH measured by each of the sensors compared to the 10-minute average of the influent SCADA pH in Experiment 6. The s::can pH::lyser matched the influent pH sensor from SCADA tightly, while the other two pH sensors showed much less agreement due to their variation and periods of drift. The differences in performance are apparent in their Pearson correlation coefficients (Table A-2). This pattern was typical of the other experiments.

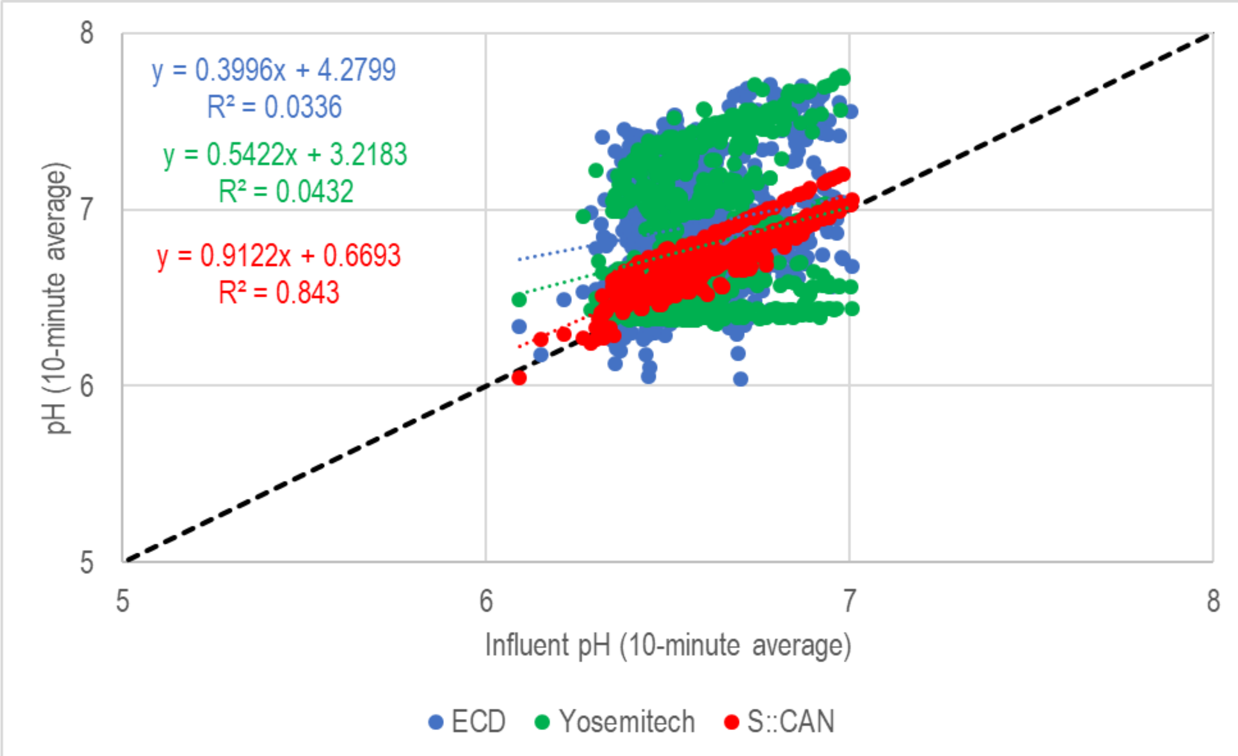


Figure A-17. Agreement Between Measured pH Values for Each Type of Sensor with those Measured by the SCADA pH Sensor in Experiment 6.

Periods of poor data quality by the SCADA sensor were removed for this analysis.

The pH::lyser was less affected by the events that occurred in Experiments 1 and 8 that caused the spectro::lyser to drift (Figure A-18). However, some drift was apparent in response to these events, and that drift continued even after cleaning in Experiment 8, similar to the spectro::lyser but to a much smaller extent (Figure A-18). Therefore, whatever affected the lens on the spectro::lyser likely also had a minor effect on the bulb on the pH::lyser, though it appears to have not affected the condu::lyser. Even when affected, it still matched the influent pattern closely up until the last few days of the experiment and performed better than either of the other two pH sensors. The Yosemitech pH may also have been affected by the event in Experiment 8 as it began having large false positive events not observed at the influent and went to a constant value towards the end of the extended experiment without maintenance (Figure A-18). The ECD sensor mostly behaved similarly in the extended experiment to the shorter-term experiments (Figure A-18).

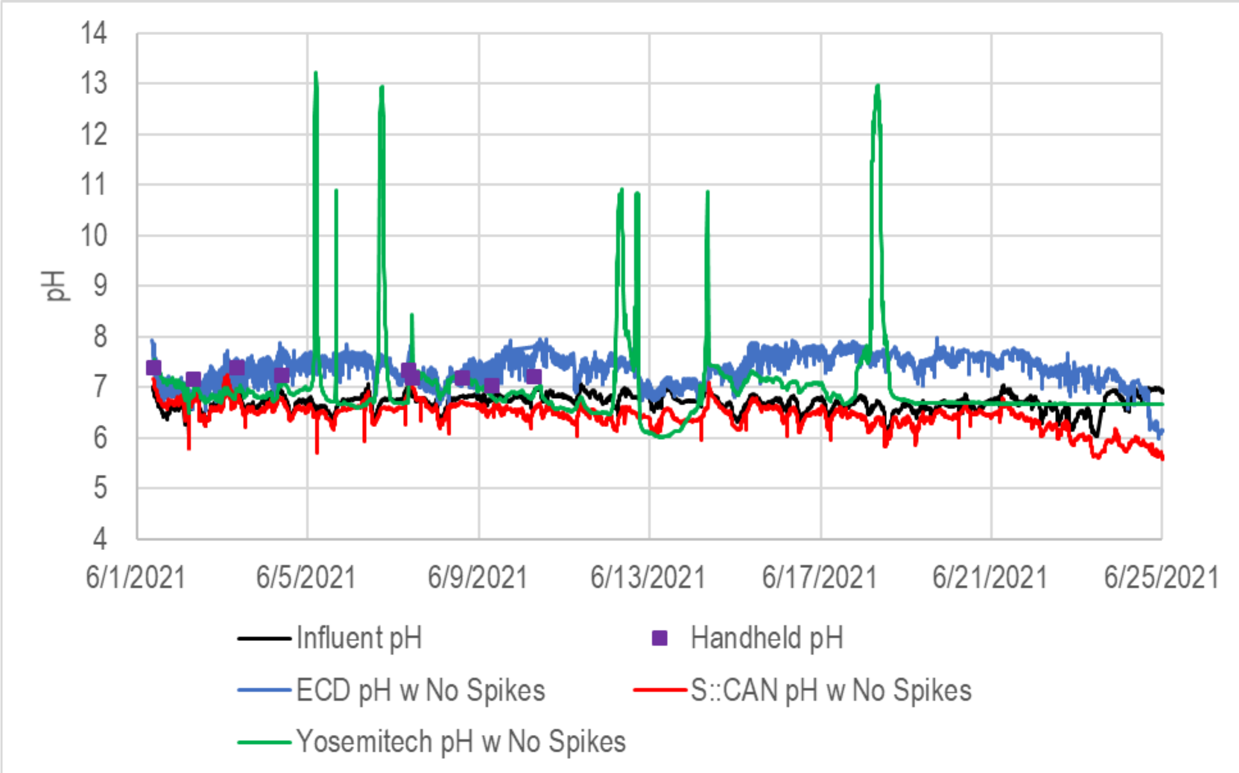


Figure A-18. Timeseries of pH in Experiment 8.

as measured by all three pH sensor types, the handheld multimeter, and the influent SCADA pH.

The s::can pH sensor showed more precision and accuracy than the other two pH sensors in these experiments and showed much more resilience to fouling. This was partly driven by the air blast, as periods during initial testing of the equipment showed that when the air blast was turned off or not functioning the pH measured by the pH::lyser drifted. However, the sensing technology is also different from the other pH sensors making it less affected when the bulb is fouled. It is more expensive than the other sensors, and it requires more power for the air compressor, which keeps it from being suitable for installation in manholes. However, similar to the spectro::lyser, s::can is working on a version that can run off battery power and be installed in manholes.

The ECD pH sensor was affected badly by fouling and had a lot more variation than the other sensors in the flume. However, CWS has used this sensor many times in the sanitary collection system and, as long as it was installed in the rag guard sensor holder, it performed very well for a month or more at a time without maintenance. When not in the rag guard, the performance was similar to what was seen in the flume. This is a good sensor to use as long as the rags and FOG can be controlled.

The Yosemitech sensor is the least expensive of all of the sensors and works with the EnviroDIY system that makes the datalogger and supporting equipment much less expensive, as well. It performed fairly similarly to the ECD pH sensor in the flume with maybe a slight improvement in its resistance to fouling in some experiments. However, the calibration issue needs to correct, and it needs testing in the actual sanitary sewer over a longer period. The current dataloggers also cannot be placed inside a manhole and are solar powered, making them not suitable for many locations in the sanitary sewer. CWS is working on updates that will allow them to run off battery power and be placed inside the manholes. CWS is currently testing the Yosemitech sensor side-by-side with an ECD sensor in the collection system.

A.3.1.2.4 Effects of FOG and Rags

The results of Experiments 7a-7c showed that the s::can sensors were virtually unaffected by the introduction of the FOG and rags, regardless of the method. The other sensors were severely affected by the FOG in Experiments 7a and 7b, but much less so for Experiment 7c. Figure A-19 shows the timeseries of pH by all three types of pH sensors during Experiment 7a (when the sensors were dipped in grease from the stilling well) with the spike events removed. This pattern is typical of that seen at the other sensors in Experiment 7a and 7b, as well as the condu::lyser and spectro::lyser. The s::can pH sensor recovers very quickly with the air blast quickly removing smeared grease from the bulb, and it detected the initial spike and subsequent spikes closely. The other two pH sensors did not recover from the grease within the two days of the experiment period and did not detect any of the spikes. Figure A-20 shows the timeseries of pH by all three pH sensors in Experiment 7c where FOG was dripped into the stilling well at the front of the flume over 20 minutes. The sensors performed fairly similarly to Experiments 1 through 6 in this case, except that the ECD sensor began the experiment ~1 pH unit higher than the influent SCADA pH instead of drifting after the first few days. Therefore, the effect of increase in liquid FOG on the sensors, at least in a typical range, may be limited. However, when rags aid in the FOG collection or the FOG is hardened on the sensor through cold or dry conditions, the effect can be dramatic if the sensors do not have the ability to clean themselves.

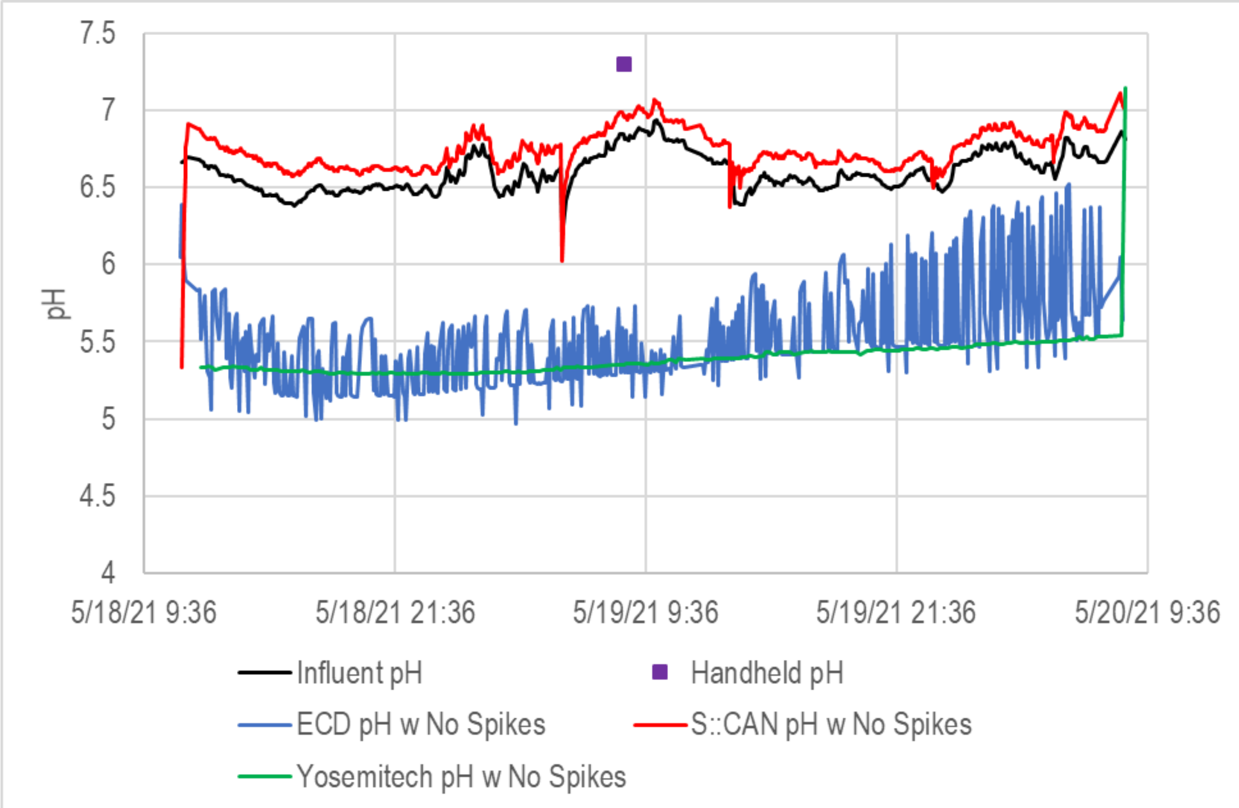


Figure A-19. Timeseries of pH Readings from All Three Types of Sensors During Experiment 7a.
The Sensors Were dipped in grease from the stilling well.

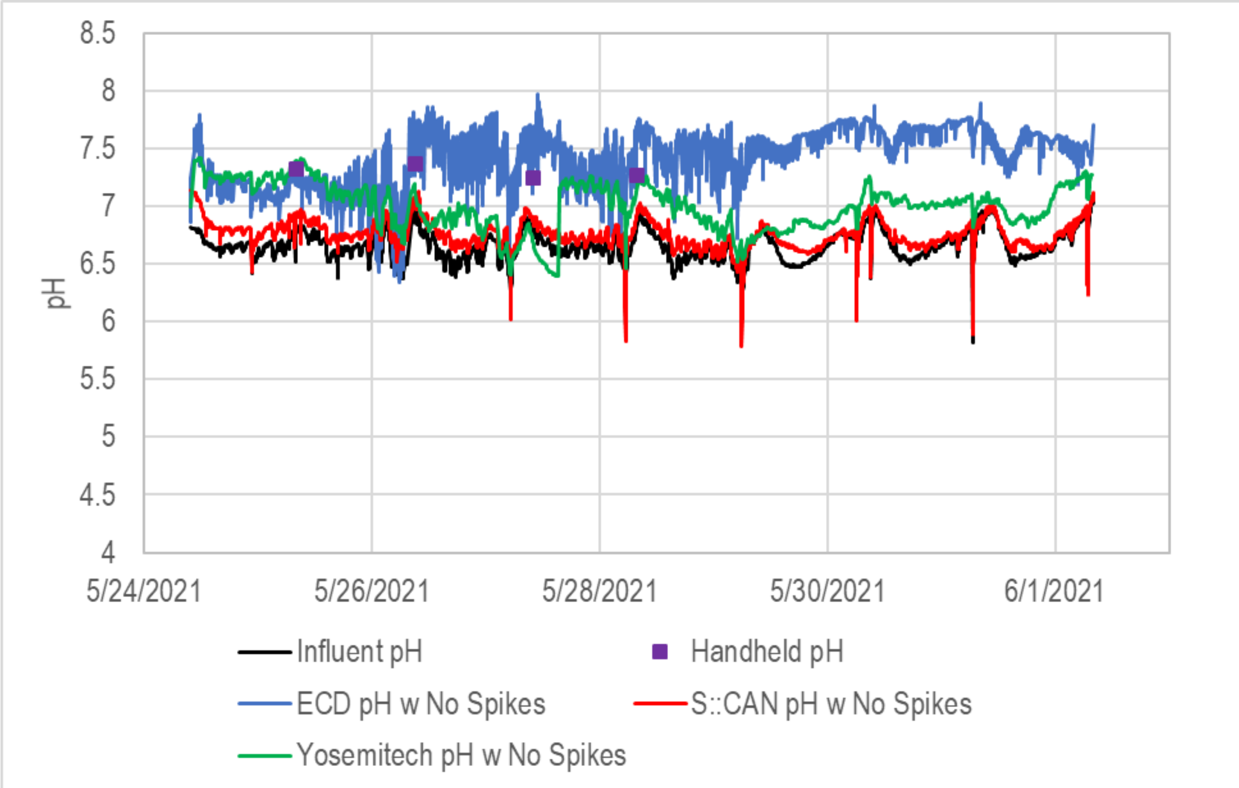


Figure A-20. Timeseries of pH Readings from All Three Types of Sensors During Experiment 7c.
Liquified grease was dripped into the flume over a 20-minute period.

A.3.2 Rag Guard Testing Results

As described in the Methods section, there were two periods of representative results for the rag guard. The period where both sensors (in the rag guard and in the generic sensor holder) were operating normally was June 7th-July 2nd (the end of the study). The pH timeseries for this period from both sensors in shown in Figure A-21. The behavior of the pH probe in the generic sensor holder was typical of many previous pilot studies conducted by CWS with different pH probes in the collection system. Downward drift with much less variability began very soon after installation/cleaning and continued with intermittent periods of returns to an expected pH as rags/FOG dislodged from the sensor temporarily (Figure A-21). The diurnal cycle was typically visible, but the pH range and values were not reliable. In contrast, the probe in the rag guard had a much more stable baseline for the entire period (almost a month) and detected the diurnal patterns consistently (Figure A-21). There were occasional periods where it drifted downwards such as on June 27th, but it recovered much more quickly and returned to the stable baseline. Between March 24th to April 19th, the probe in the generic sensor holder was not functioning, but the probe in the rag guard was functional with no maintenance for almost a month (Figure A-22). During this period, the baseline was stable and many real pH events occurred that were confirmed by the pH sensor at the influent (Figure A-22). Therefore, it appears that the rag guard may make it feasible for a pH probe to last at least a month in the collection system without maintenance while providing a stable baseline and sufficient sensitivity to detect short-duration discharge events. This is a marked improvement over the generic sensor holder where the pH sensor was fouled so frequently tcohat extended periods of reliable data were not obtained. Adjustments are needed to ensure more consistency between locations and further refine the sensor holder. CWS is currently working on these refinements.

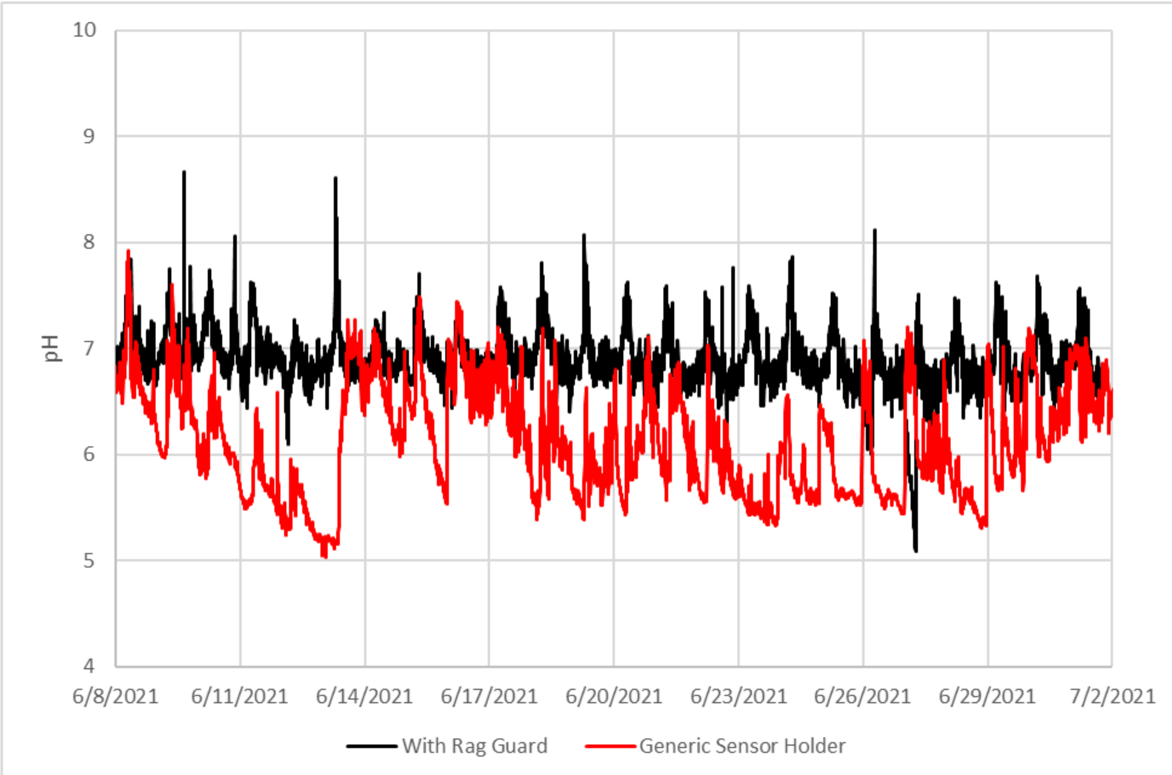


Figure A-21. Timeseries of pH Measured by Both pH Sensors at Hwy 47 Test Manhole June 8-July 2, 2021.

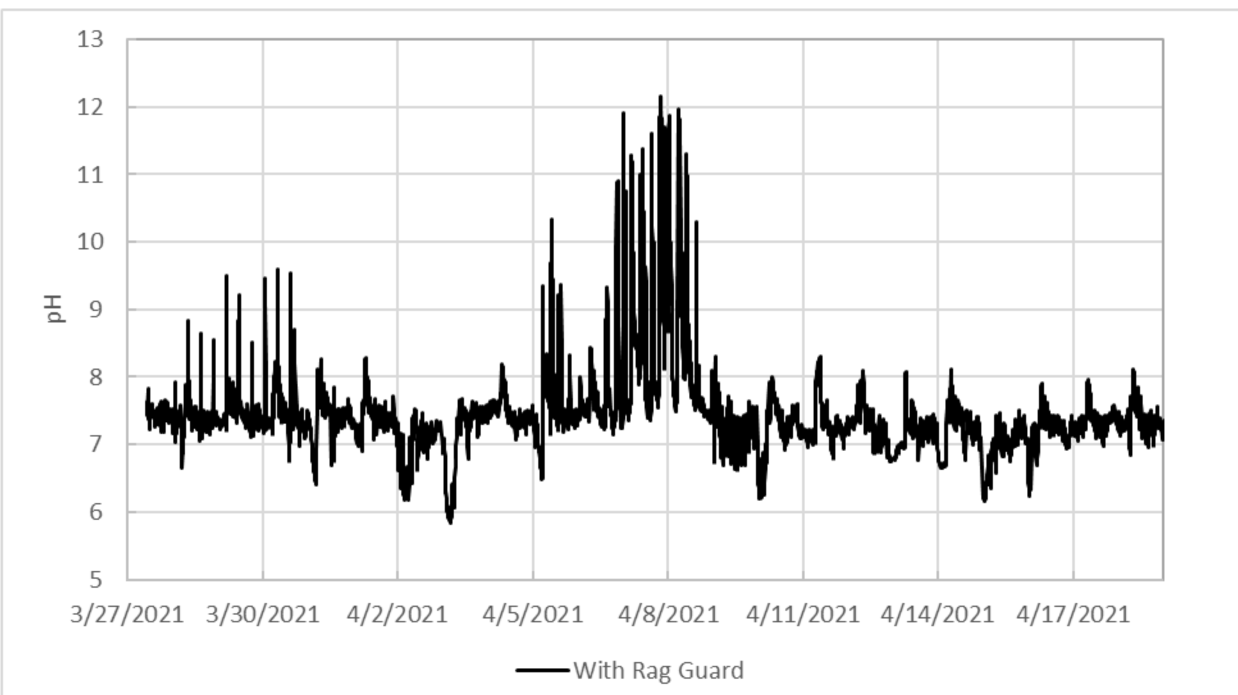


Figure A-22. Timeseries of pH Measured by the pH Sensor in the Rag Guard at Hwy 47 Test Manhole March 27 through April 19th.

A.4 Conclusions

This study demonstrated that water quality sensors can be successfully deployed in sanitary sewers without an excessive maintenance burden. The main factors affecting the success of the deployments were the ability to limit fouling and the type of sensor used. Fouling was shown to be limited by air-blast self-cleaning in the case of the s::can sensors deployed in the flume and by the rag guard deployed at a manhole in the collection system. Resistance to the effects of fouling was also a function of the type of technology used by the sensor. Velocity and depth were not shown to be major variables in the success of the different deployments, but this may be limited by the range of velocities available to test at this flume and/or the counter effects of changes in depth.

In general, the s::can sensors were successful except when events occurred that affected the lens/sensors. They were consistently able to track measured values and detect spikes after at least a week in the influent, even when impacted by FOG and rags. The combination of the self-cleaning and the technology made them very resistant to typical fouling and caused them to outperform the other sensors. However, they are expensive, require more power, and cannot currently be deployed in manholes (though this may change in the future). These could be an excellent choice for sensors at the influent or at established monitoring stations in the collection system equipped with power and a utility box. While the spectro::lyser detected the temporal patterns in the influent consistently, it sometimes varied greatly from measured values despite calibration. This may be solvable with better calibration as shown in the case of TSS and at other cities, but it may also mean that the sensors can be used more for detecting patterns rather than replacing laboratory samples of COD and other parameters.

The ECD pH probes were not successful for more than a few days at a time except when deployed inside the rag guard. With the rag guard, the ECD pH probes performed well for a month with no maintenance. These can currently be deployed in the collection system inside manholes. The Yosemitech probes showed promise that these inexpensive probes could perform similarly to the ECD probes in the collection system. However, they will likely also require a rag guard to avoid excessive maintenance, and their dataloggers cannot currently be deployed inside manholes (though that may change in the future).

The ORP probes did not perform well in any of the applications regardless of the methods used to prevent fouling and will not be used by CWS in their continuous monitoring programs.

A.5 Next Steps

CWS is looking into potentially purchasing and installing an s::can system at the influent of the Forest Grove plant (where the flume is located). This study provided sufficient evidence to make a good business case for installing these at the influent. CWS is also making refinements to the rag guard using 3D printing and some new design elements to improve upon the existing model. CWS is also testing a Yosemitech pH probe in the collection system side-by-side with an ECD pH probe to determine if they can perform similarly in the rag guard. CWS has also deployed four permanent pH monitoring stations in the Forest Grove collection system based on the success of this study and previous pilots using ECD pH probes that are installed in the rag guards. Yosemitech pH probes, as well as other probe types, have been installed permanently in rivers and creeks throughout the watershed by CWS, and CWS will continue testing additional sensors with the EnviroDIY system in the flume and collection system. CWS is also looking into adjustments to the flume in order to allow a larger range of testing velocities.

APPENDIX B

Hampton Roads Sanitation District Machine Learning Case Study

B.1 Introduction

The sensor analysis detailed in this appendix focused upon data collected within a secondary WWTP and in a downstream advanced water treatment facility (AWTF). Online instruments within WWTPs or AWTFs generally undergo less fouling, clogging etc. due to upstream treatment (e.g., screening or primary settling). Thus, their maintenance is expected to be less expensive in terms of staff hours, and they should produce a cleaner dataset in terms of fewer gaps or errors. Slugs of industrial discharge may be relatively diluted or equalized by the time they reach these instruments. This dilution and equalization could be beneficial since there would be lower baseline variance, or a limitation since there would be a smaller true water quality change resulting from an equivalent industrial discharge.

Hampton Roads Sanitation District (HRSD) has begun the Sustainable Water Initiative for Tomorrow (SWIFT), which will purify effluent from many of HRSD's WWTPs to recharge the Potomac Aquifer. The SWIFT Research Center (SWIFT RC) is a 3.8 million L/day demonstration-scale AWTF, which treats secondary effluent from a WWTP with a 5-stage Bardenpho Process for biological nutrient removal. The SWIFT RC treatment train includes coagulation, flocculation, settling, ozonation, biofiltration, GAC, and UV disinfection. SWIFT has a final treated TOC goal of 4 mg/L (Gonzalez et al. 2021).

HRSD has a robust industrial pretreatment program. For example, HRSD has identified sources of bromide, PFAS, acrylamide, and 1,4-dioxane discharged to its WWTPs (Nading et al. 2022). Permits, flow limitations, or innovative industrial pretreatment have been implemented to reduce the concentrations of these chemicals. However, approximately monthly spikes in online monitoring surrogates including secondary effluent TOC have been observed at the SWIFT RC and caused pauses in production.

Elevated TOC has a cascading effect on both downstream treatment and water quality (e.g., higher TOC could cause a larger ozone demand, and higher TOC on the inlet to the AWTF may result in higher TOC in the finished water which may present disinfection byproduct challenges). The chemical(s) and industrial source causing these events has not yet been identified. Since no TOC instruments were located at the WWTP influent, it is unknown whether these industrial discharge events were pass-through (i.e., organic substance(s) not fully removed by the WWTP) or interference (i.e., organic or inorganic substance(s) that inhibited the WWTP's removal of overall TOC). Rapid detection of future such events would be beneficial for (1) for corrective action such as increased ozone dose or GAC contact time and (2) collecting water samples to assist with identifying the chemical signature of these events. Chemical analysis collected in the midst of these events could then provide clues about the responsible industry.

B.2 Methods

In this study, 35 SML models were compared to detect suspected industrial discharge events at the SWIFT RC. Models were trained and tested on real, full-scale, hourly data from 30 variables with a total sample size of 878 (about 37 days). Since the industrial source was unknown, datapoints were labeled "Normal" or "Event" based on retrospective expert human judgement. SML was conducted in R using the caret package. Caret is a package in the R programming language that enables around two hundred different SML model types to be applied using similar code structure (Kuhn 2008). Preprocessing methods were also compared to enhance model accuracy. SML performance was benchmarked against fixed thresholds on each of the 30 variables.

B.2.1 Online Instrumentation

Models were trained on 30 variables that included readings from online instruments or gauges (Table B-1).

Table B-1. HRSD Variables and Instrument Locations.

Location	Variable	Manufacturer	Instrument	Units
Raw Wastewater Influent	Conductivity	Hach	3725E2T	mS/cm
Secondary Wastewater Effluent	Flow	Rosemount	8750W	gpm
	Total Nitrogen	Shimadzu	TOC-4200 FA E ROHS	mg/L
	Total Inorganic Nitrogen	WTW	TresCon	mg/L
	Total Organic Carbon	Shimadzu	TOC-4200 FA E ROHS	mg/L
	Nitrite	WTW	TresCon	mg/L
	Nitrogen Oxides	WTW	TresCon	mg/L
	Nitrate	WTW	TresCon	mg/L
	Ammonia	WTW	TresCon	mg/L
	Conductivity	Hach	D3727E2T	mS/cm
	UV Transmittance	Hach	UVAS	%
	Turbidity	Hach	TU5300	NTU
	pH	Foxboro	871A	
	Temperature	Foxboro	871A	°C
Settled Water (Post-Floc/Sed)	UV Transmittance	Hach	UVAS	%
	Monochloramine	Hach	5500	mg/L
	Ammonium	Hach	5500	mg/L
	Total chlorine	Hach	CL-17	mg/L
	Redox potential	Foxboro	871A	
	Total Organic Carbon	Shimadzu	TOC-4200 FA E ROHS	mg/L
	Total Nitrogen	Shimadzu	TOC-4200 FA E ROHS	mg/L
	Free Ammonia	Hach	5500	mg/L
Ozonation System	Ozone Dose	Wedeco	LC400Plus	lbs/day
	Ozone Sidestream Flow	NA	NA	gpm
	Ozone Residual Setpoint	NA	NA	mg/L
	Ozone Residual	Hach	Orbisphere 410	mg/L
Biofiltration Influent	UV Transmittance	Hach	UVAS	%

Location	Variable	Manufacturer	Instrument	Units
	Total Chlorine	Hach	CL-17	mg/L
	Redox potential	Foxboro	871A	
	pH	Foxboro	871A	

B.2.2 Data Collection

Data was exported hourly on dates May 20th, 2019 through June 4th, 2019; June 15th, 2019 through June 21st, 2019; May 25th, 2020 through June 2nd, 2020; and October 17th, 2020 through October 21st, 2020 (Table B-2). Suspected abnormal industrial discharges occurred during these data periods on June 16th, 2019; June 1st, 2020; and October 18th, 2020. It was assumed these industrial discharges were from the same source or related enough to classify within the same SML output category. Missing data was assumed zero for ozone residual, ozone dose, and ozone output, since missing data for these variables was associated with shutdown of the ozonation system. For other variables (i.e., independent variables), missing data was assumed equal to the most recent previously measured value.

Table B-2. Data Time Periods.

Time Period	Start Date	End Date	Industrial Discharge	Dataset
#1	5/20/2019	6/4/2019	None	Training
#2	6/15/2019	6/21/2019	6/16/2019	Training
#3	5/25/2020	6/2/2020	6/1/2020	Training
#4	10/17/2020	10/21/2020	10/18/2020	Testing

B.2.3 Supervised Machine Learning

Machine learning is the study of algorithms that learn from data. Supervised machine learning (SML) creates a model or function that predicts outputs from inputs based on example input-output pairs. These example input-output pairs are called the training set. The example outputs in the training set are called labels. SML is in contrast to unsupervised machine learning, which finds patterns in unlabeled data; examples include principal component analysis (PCA) or clustering algorithms. There are over 200 types of SML models; some examples include neural networks, nearest neighbors, or random forest (Kuhn 2019). SML models fall into two categories based on their application: classification or regression. Classification SML models predict categorical outputs (e.g., Good vs Bad or Normal vs Event). Regression SML models predict quantitative outputs (e.g., 17 percent industrial wastewater).

In addition to a training set to train the model, SML also requires a fully separate testing set (also known as the verification set) to assess its predictive capabilities (Figure B-1). Many SML are highly complex and flexible—analogous to a linear or polynomial model with many terms—and thus capable of an extreme degree of overfitting if appropriate training and testing protocols are not followed. Overfitting is when a model fits a particular data set too closely, interpreting noise or random errors as if they represent true underlying patterns in the phenomenon being studied. Overfit models have extremely high accuracy for their training set but lower accuracy on new data compared to similar but less overfit models. Overfitting is a risk when dealing with timeseries data such as online water quality data. For example, TOC above 10 mg/L might be a highly reliable indicator of an industrial discharge within a given time window, but over time, the average TOC may decline due to treatment changes, or the average measured TOC or its standard deviation may increase over time due to fouling. Thus, applying models to a new, fully separate testing set from a later time ensures that models are in fact capable of making good predictions using new data from new times, and not just overfit to data from a certain time period.

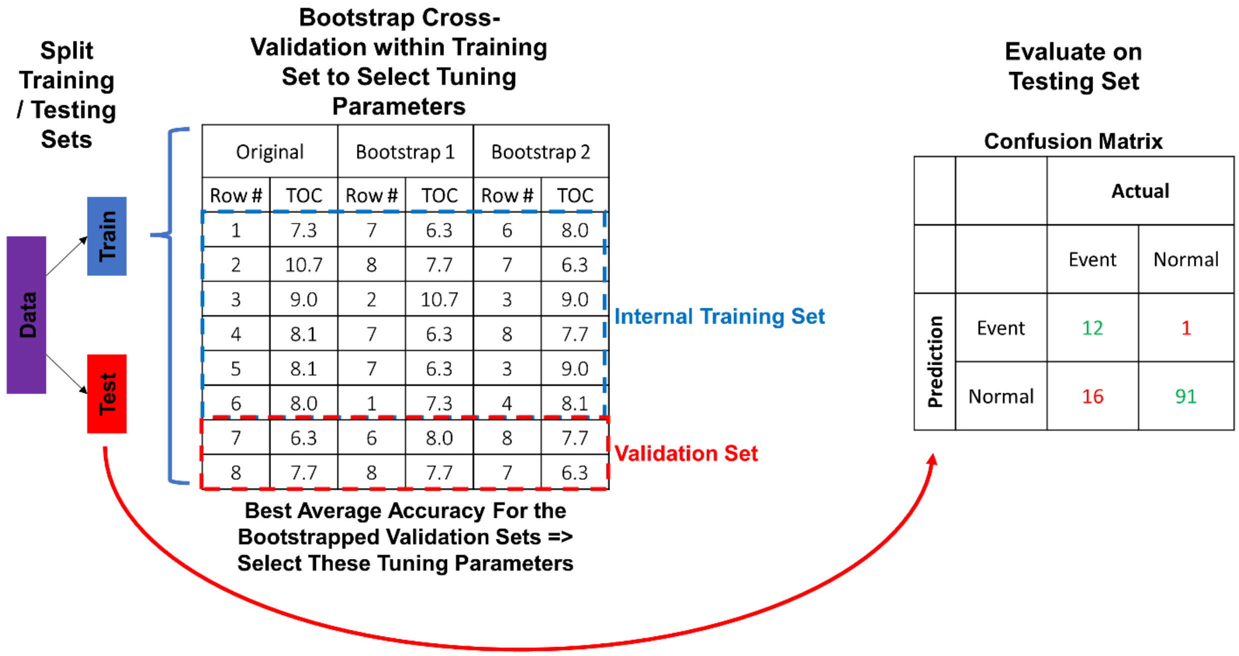


Figure B-1. Data Splitting, Cross-Validation, and Testing Workflow.

The accuracy of certain types of SML models depends in part on random chance. For example, random forest models randomly select a subset of the observations and/or variables and then construct a decision tree based on this random subset. This process is repeated, typically hundreds of times, and an average or consensus is taken of the outputs of the decision trees. If the same random forest model is trained on the same data, but different subsets of data are selected for each tree, different testing set accuracies could result. In programming environments, randomness is simulated using random number generator algorithms such as Mersenne-Twister (Matsumoto and Nishimura 1998). In the programming language R, numerical seeds can be provided to the random number generator. The same seed can be provided to enable reproducible results, or different seeds can be provided to simulate random replication. In this study, the seed was set to 1 unless otherwise noted to ensure reproducibility. For models selected for in-depth evaluation, seeds were set to integers from 1 to 30 or 1 to 100 to check whether the model accuracy was subject to random chance.

Most SML models have parameters that can be adjusted within the model that impact the learning process rather than being determined via the training. These parameters are called tuning parameters or hyperparameters. For example, k-nearest neighbors models assign new data to a class based on the most common label of the most similar datapoints. k is the number of similar datapoints considered in the analysis and is an example of a tuning parameter. Tuning parameters are selected in a step in the machine learning process called cross-validation

(Figure B-1). In cross-validation, the training set is repeatedly split into smaller training and testing sets (sometimes called validation sets in this context). Models with different tuning parameter settings are trained on each internal training set and tested on each validation set. The tuning parameters that result in the best average performance on the validation sets are then selected, and applied when making predictions on the final, fully separate testing test.

SML was conducted in R version 3.6.3 using the caret package (Kuhn 2008). The caret package contains a set of programming functions that streamline the process of generating SML models. It allows a library of over 200 types of SML models to be trained and tested using similar coding grammar. Observations occurring during suspected abnormal industrial discharges were labeled Event. Other observations were labeled Normal. Data from May 20th, 2019 through June 2nd, 2020 (i.e., the first three of the four time periods, see Table B-2) were used as a training set and contained two abnormal industrial discharge events: June 1st, 2019 and June 16th, 2019 (total sample size n_{total}=758, event sample size n_{event}=66). Data from October 17th, 2020 through October 21st, 2020 was used as a testing set and contained one industrial discharge event: October 18th, 2020 (n_{total}=120, n_{event}=28). Thus, the data was split approximately 86 percent training set, 14 percent testing set.

Thirty-five models (Table B-3) were selected for screening based on their accuracy performing a similar classification task—detecting de facto reuse in surface water—using similar online water quality instrumentation (Thompson and Dickenson 2021). Models were screened on raw data (i.e., no preprocessing) using default tuning parameters in the caret package.

Training set accuracy, testing set accuracy, event sensitivity, and total false alerts (i.e., false positives or Normal observations incorrectly predicted as Event), and p-value relative to the no information rate (NIR) were recorded for each model. Accuracy in the context of classification models means the overall percent of the dataset for which the model predicted the correct label. Sensitivity is how often the models were correct when the true answer was an Event. The NIR is the accuracy that could be achieved by always assuming the most common label, which in this case was Normal. The NIR was 76.7 percent. The p-value that the testing set accuracy was above the NIR was calculated using the binomial confidence interval method (Kuhn 2008; Clopper and Pearson 1934).

The training set accuracy was internally cross-validated with 25 bootstraps (Kuhn 2008). That is, 25 random samples were selected from the training set with the same total sample size as the original training set. These random samples were "with replacement," i.e., it was possible for datapoints to occur twice, or not at all. Random samples like these are called "bootstraps." The bootstraps were then split 75:25 into training and validation sets, and the models were trained and validated 25 times using each bootstrap. The average accuracy on the validation sets was then calculated and is referred to simply as "training set accuracy" below. This bootstrapped training set accuracy was used for selecting tuning parameters before final evaluation with the fully separate testing set (Figure B-1).

Table B-3. List of SML Models Screened.

AdaBoost Classification Trees	DeepBoost	Linear Distance Weighted Discriminant	Oblique Random Forest with Ridge Regression	Sparse Distance Weighted Discrimination
Bayesian Generalized Linear Model	Distance Weighted Discrimination with Radial Basis Function Kernel	Linear Support Vector Machines with Class Weights	Oblique Random Forest with Support Vector Machines	Sparse Linear Discriminant Analysis
Boosted Classification Trees	Generalized Additive Model using LOESS	Localized Linear Discriminant Analysis	Penalized Logistic Regression	Stabilized Nearest Neighbor Classifier
Boosted Linear Model	Generalized Additive Model using Splines	Mixture Discriminant Analysis	Quadratic Discriminant Analysis	Support Vector Machines with Linear Kernel
Boosted Smoothing Spline	Generalized Linear Model with Stepwise Feature Selection	Neural Network with Feature Extraction	Random Forest Rule-Based Model	Support Vector Machines with Radial Basis Function Kernel
Boosted Tree	L2 Regularized Linear Support Vector Machines with Class Weights	Oblique Random Forest with Logistic Regression	Rotation Forest	Tree-Based Ensembles
Cost-Sensitive C5.0	Least Squares Support Vector Machine with Radial Basis Function Kernel	Oblique Random Forest with Partial Least Squares Regression	Single C5.0 Ruleset	Weighted k-Nearest Neighbors

Testing set accuracy was used as the primary metric of success in this study. Nonetheless, models from the screening phase were selected for further evaluation and tuning based on ranking in the top two for any of the following criteria: training set accuracy, testing set accuracy, or testing set event sensitivity. This was done because it was hypothesized that (1) models that were overfit (relatively high training set accuracy compared to testing set accuracy) might perform better on the testing set after tuning parameter optimization; and (2) models with high testing set sensitivity but many false positives might perform better after preprocessing to reduce noise.

The models selected for the in-depth evaluation phase were first trained and tested with one hundred distinct seeds (1 to 100) to check whether their high performance was inherent to the model or due in part to random chance (Figure B-2). Next, preprocessing techniques were tested to enhance model accuracy. Then, least important variables were iteratively omitted to investigate whether training time could be improved without loss in accuracy. Finally, models were cross-validated across a greater range of tuning parameter values.

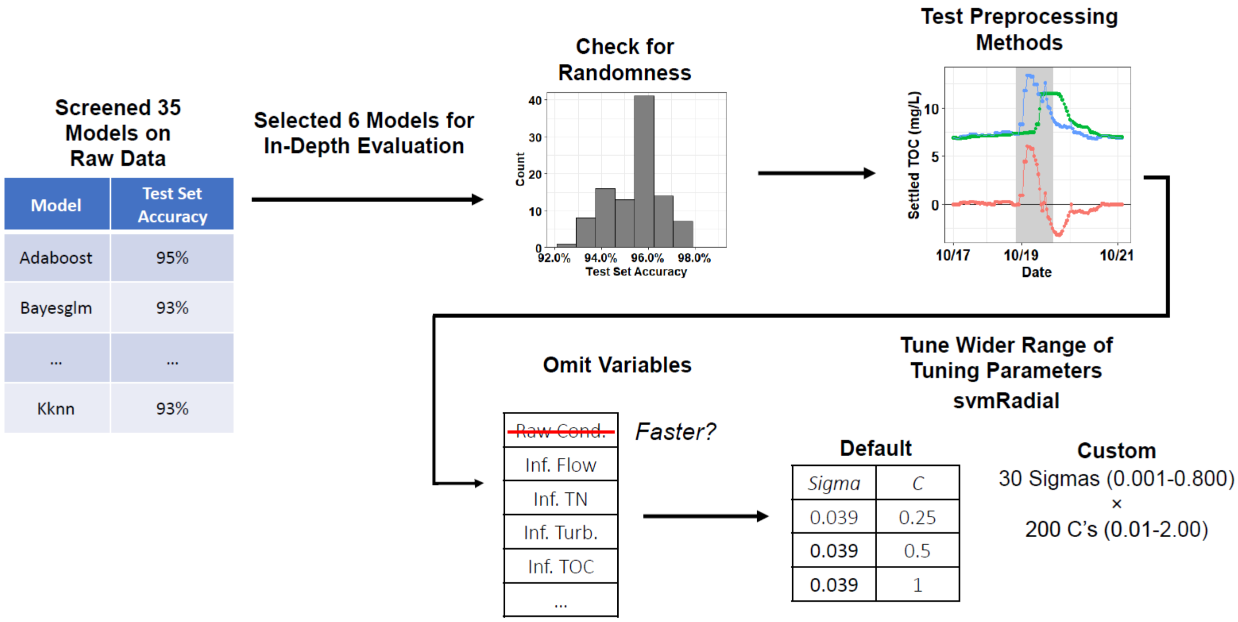


Figure B-2. Workflow of this SML Proof-of-Concept Study: Model Screening, Model Section, Checking for Randomness, Preprocessing, Variable Omission, Tuning Parameter Optimization.

Least important variables were identified using the variable importance (varImp) function from the caret package (Kuhn 2008). The varImp function calculates importance differently depending on the type of the model. However, in all cases, variables are ranked and normalized on a scale of 0 to 100 based on their importance relative to the most important variable. Where the varImp function was not applicable, variables were omitted one at a time. If there was one variable whose omission resulted in equal or greater testing set accuracy, this variable was omitted. If there were multiple variables whose singular omission resulted in equal testing set accuracy, training set accuracy was used as a tiebreaker. If there were multiple variables whose singular omission resulted in equal testing and training set accuracies, one of these variables was selected at random for omission in the next iteration. This process was repeated until no variables could be omitted without a loss in accuracy.

B.2.4 Preprocessing

It was hypothesized that certain preprocessing methods could enhance model accuracy by reducing noise in the data or counteracting the effects of instrument drift. The three preprocessing methods assessed in this case study were: rolling median, difference from rolling median, and principal component analysis (PCA). The rolling median of the past three observations of each variable was calculated to reduce noise in the data and omit non-consecutive outliers. The difference between each observation and the median of the past day (i.e., 24 hourly observations) was calculated to account for the non-stationary nature of real wastewater data and optimize the data for detecting sudden changes (Figure B-3). Differences from the rolling median were provided to the models as variables both instead of and in addition to the raw data. PCA was conducted to promote diversity among the variables, considering that each principal component is perpendicular (non-correlated) with the others. PCA has previously been applied as a preprocessing technique for SML (Rodriguez et al. 2006).

The PCA model was constructed based on the training set and then the scores for each principal component were then also calculated on the testing set.

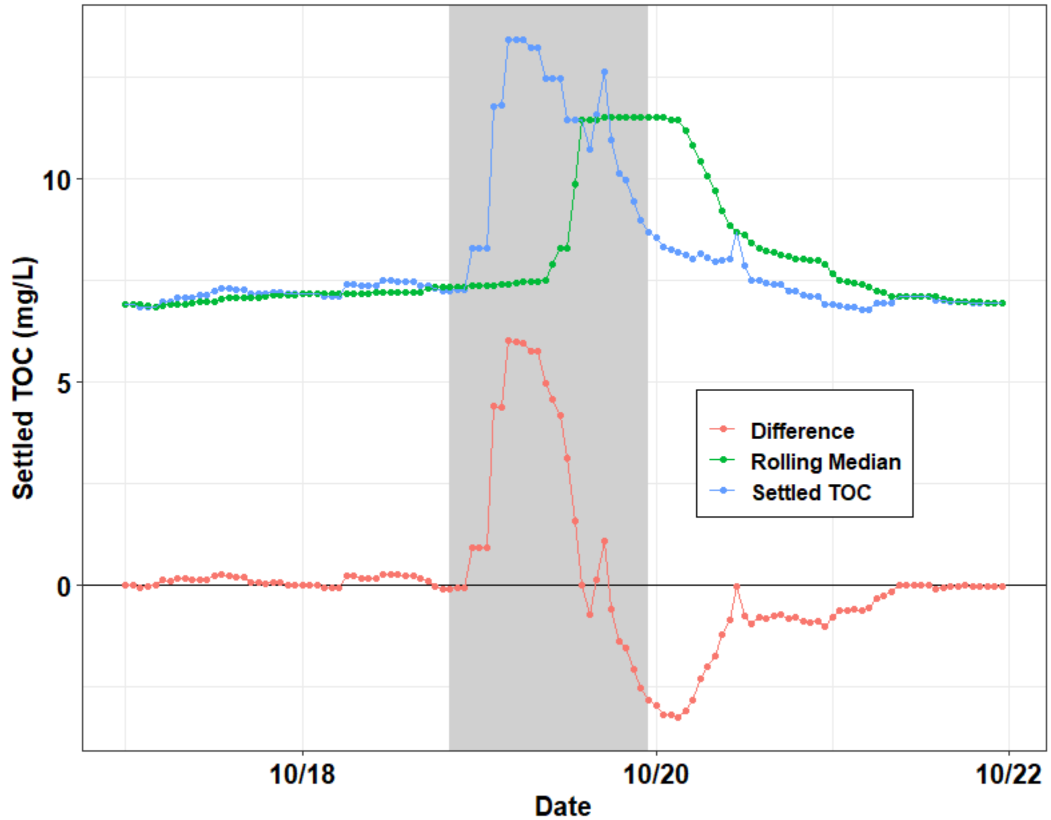


Figure B-3. Example of the Preprocessing Technique of Substracting the 24-hr Rolling Median.
The example data is settled TOC from the testing set.

Raw wastewater influent and the secondary wastewater effluent had hydraulic residence times upstream of the post-floc/sed settled water of 18 and 2 hours, respectively. Thus, any changes or spikes from industrial discharges would be expected to begin at these sensor locations sooner, out of sync with the downstream sensors. Lagging the upstream sensors to align with the sensors in the settled water would provide many synchronized variables, while still providing a degree of advanced warning compared to the final purified water. So, lagging the raw wastewater influent and secondary wastewater effluent based on hydraulic residence time to match the settled water was explored as another preprocessing method.

B.3. Results

B.3.1 Water Quality Data

Descriptive statistics for Normal and Event data are shown in Table B-4 and Table B-5, respectively. These tables include both train and testing set data. Timeseries for all variables are shown in Figure B-4, illustrating the four segments of time from which data were analyzed. Within the figure, graphics #1, #2, and #3 are the training set and graphic #4 is the testing set.

Table B-4. Summary Descriptive Statistics of Normal Data (n=784).

Variable	Mean	SD	Min	Median	Max
Raw Cond. (µS/cm)	1142	178	-500	1113	2278
Inf. Flow (gpm)	857	18	779	854	943
Inf. TN (mg/L)	2.05	0.82	0.54	1.84	4.05
Inf. TIN (mg/L)	1.94	0.99	-0.05	2.19	3.91
Inf. TOC (mg/L)	8.85	2.50	-0.06	9.34	12.04
Inf. NO2 (mg/L)	0.042	0.048	-0.001	0.012	0.190
Inf. NOx (mg/L)	1.99	0.82	0.49	1.78	3.94
Inf. NO3- (mg/L)	1.95	0.82	0.49	1.73	3.94
Inf. NH3 (mg/L)	0.060	0.016	0.025	0.056	0.150
Inf. Cond. (µS/cm)	1030	95	921	1004	1343
Inf. UVT (%)	58	4	46	58	64
Inf. Turb. (NTU)	1.55	0.53	0.28	1.50	3.60
Inf. pH	7.32	0.11	6.93	7.35	7.51
Inf. Temp. (°C)	24.4	1.3	21.8	24.6	26.9
Settled UVT (%)	70.2	2.7	56.1	70.6	74.8
Settled NH2Cl (mg/L)	2.57	0.77	0.00	2.41	4.32
Settled NH4+ (mg/L)	0.727	0.204	0.030	0.694	1.308
Settled Total Cl2 (mg/L)	2.68	0.76	0.00	2.51	4.43
Settled ORP (mV)	278	104	0	311	415
Settled TOC (mg/L)	7.30	0.85	-0.03	7.44	9.41
Settled TN (mg/L)	1.93	0.78	-0.04	2.05	3.68
Settled Free Ammonia (mg/L)	0.219	0.132	0	0.213	0.859
Ozone Sidestream Flow (gpm)	372	69	0	407	446
Ozone Dose (lb/d)	55.9	8.2	0.0	55.8	86.9
Ozone Resid. Setpoint (mg/L)	0.546	0.069	0.350	0.552	0.769
Ozone Resid. (mg/L)	0.407	0.190	0.000	0.409	1.025
Biof. Inf. UVT (%)	81.1	2.1	74.9	81.1	85.1
Biof. Inf. Total Cl2 (mg/L)	0.0539	0.1800	0.0027	0.0453	4.9496
Biof. Inf. ORP (mV)	299	98	0	306	480
Biof. Inf. pH	7.12	0.10	6.78	7.15	7.32

Table B-5. Summary Descriptive Statistics of Event Data (n=94).

Variable	Mean	SD	Min	Median	Max
Raw Cond. (µS/cm)	1080	426	-500	1176	1396
Inf. Flow (gpm)	855	23	777	853	923
Inf. TN (mg/L)	1.23	0.53	0.40	1.19	2.35
Inf. TIN (mg/L)	1.53	0.91	-0.05	1.92	2.78
Inf. TOC (mg/L)	11.3	7.3	-0.1	12.0	20.3
Inf. NO2 (mg/L)	0.00232	0.00250	-0.00077	0.00139	0.00827
Inf. NOx (mg/L)	1.16	0.52	0.35	1.14	2.18
Inf. NO3- (mg/L)	1.16	0.52	0.35	1.14	2.17
Inf. NH3 (mg/L)	0.0631	0.0196	0.0402	0.0552	0.1605
Inf. Cond. (µS/cm)	1080	87	972	1049	1235
Inf. UVT (%)	42.9	8.5	22.1	43.9	60.7
Inf. Turb. (NTU)	3.51	1.48	0.83	3.66	8.44
Inf. pH	7.02	0.12	6.83	7.01	7.38
Inf. Temp. (°C)	24.5	1.4	22.6	24.3	26.9
Settled UVT (%)	59.8	6.1	51.1	58.7	72.7
Settled NH2Cl (mg/L)	2.18	1.35	0.00	2.39	3.93
Settled NH4+ (mg/L)	0.73	0.39	0.00	0.88	1.28
Settled Total Cl2 (mg/L)	2.33	1.40	0.00	2.62	4.16
Settled ORP (mV)	181	157	0	256	409
Settled TOC (mg/L)	12.8	4.5	6.7	11.7	20.3
Settled TN (mg/L)	1.96	0.38	1.26	2.00	2.75
Settled Free Ammonia (mg/L)	0.303	0.162	0.000	0.343	0.592
Ozone Sidestream Flow (gpm)	259	137	0	281	418
Ozone Dose (lb/d)	63	35	0	71	127
Ozone Resid. Setpoint (mg/L)	0.530	0.078	0.391	0.548	0.669
Ozone Resid. (mg/L)	0.434	0.308	0.000	0.445	1.057
Biof. Inf. UVT (%)	73.6	3.5	62.4	72.9	80.5
Biof. Inf. Total Cl2 (mg/L)	0.0452	0.0588	0.0027	0.0449	0.5914
Biof. Inf. ORP (mV)	268	133	0	323	414
Biof. Inf. pH	6.89	0.13	6.68	6.89	7.15

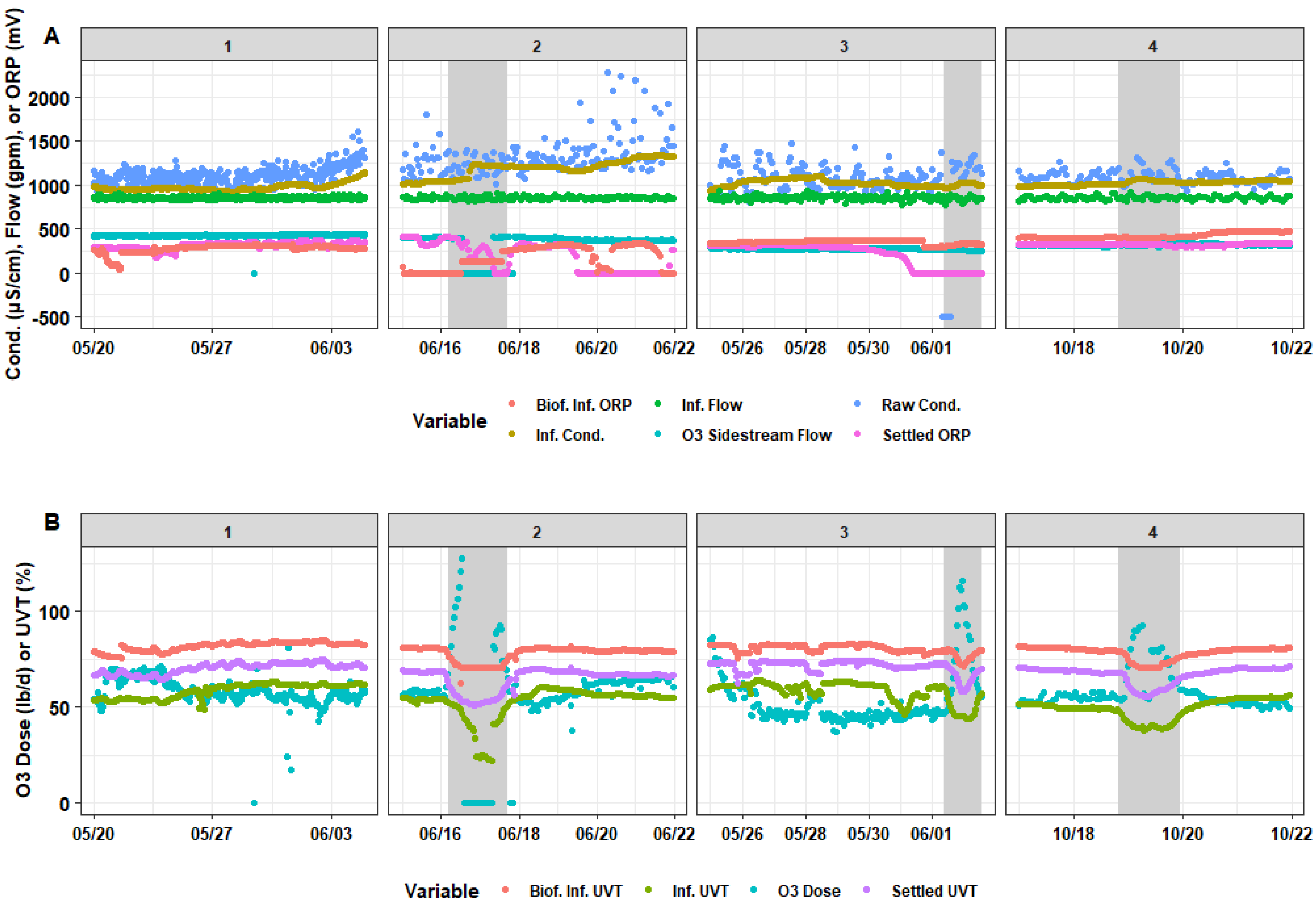


Figure B-4. Timeseries of 30 Water Quality Variables.

Grouped A-F based on similar maximum values. Horizontally, panels are separated into the four time-segments analyzed. #1 and #2 are in 2019. #3 and #4 are in 2020. #1, #2, and #3 were used for the training set and #4 was the testing set. Gray shaded areas represent abnormal industrial slug events.

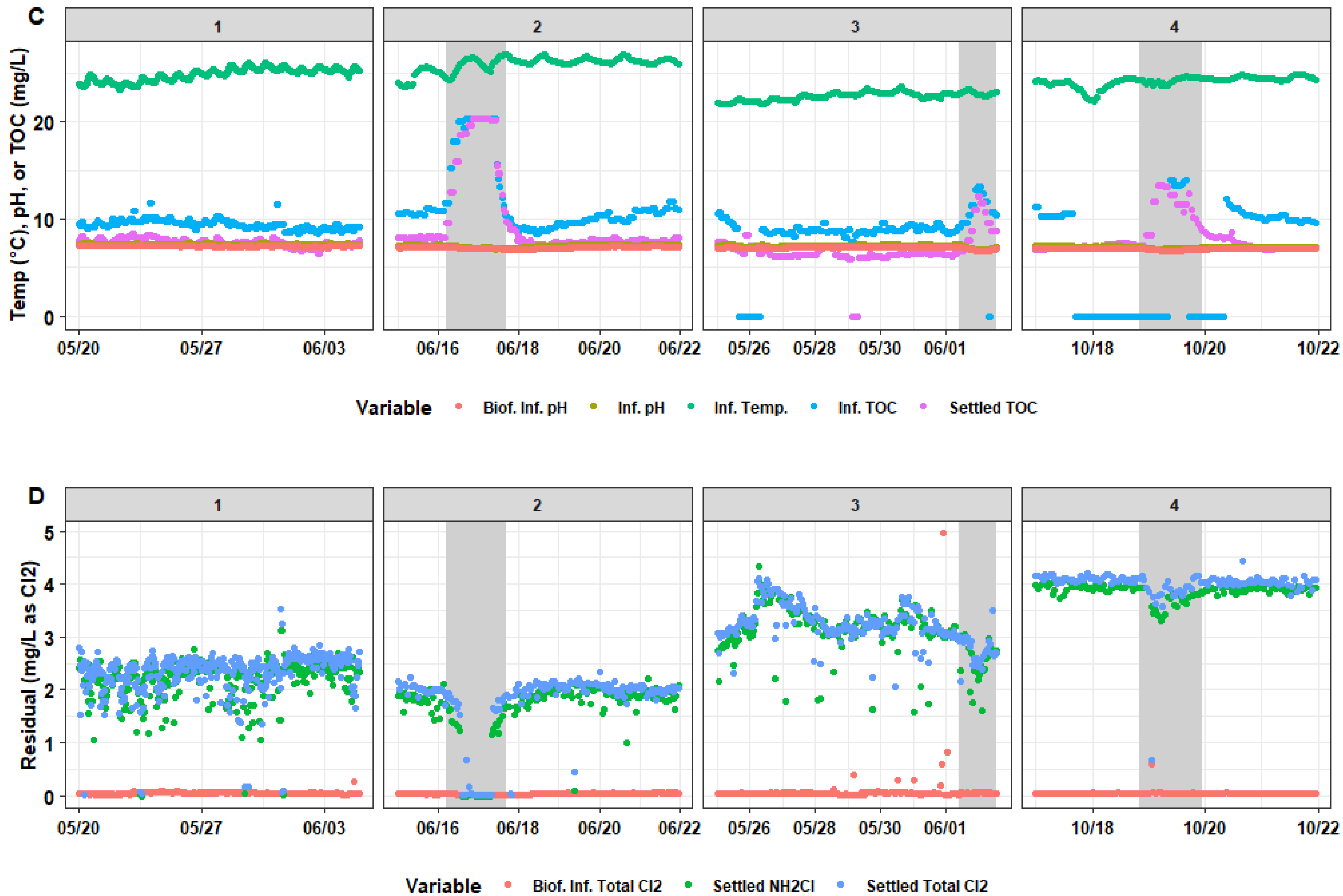
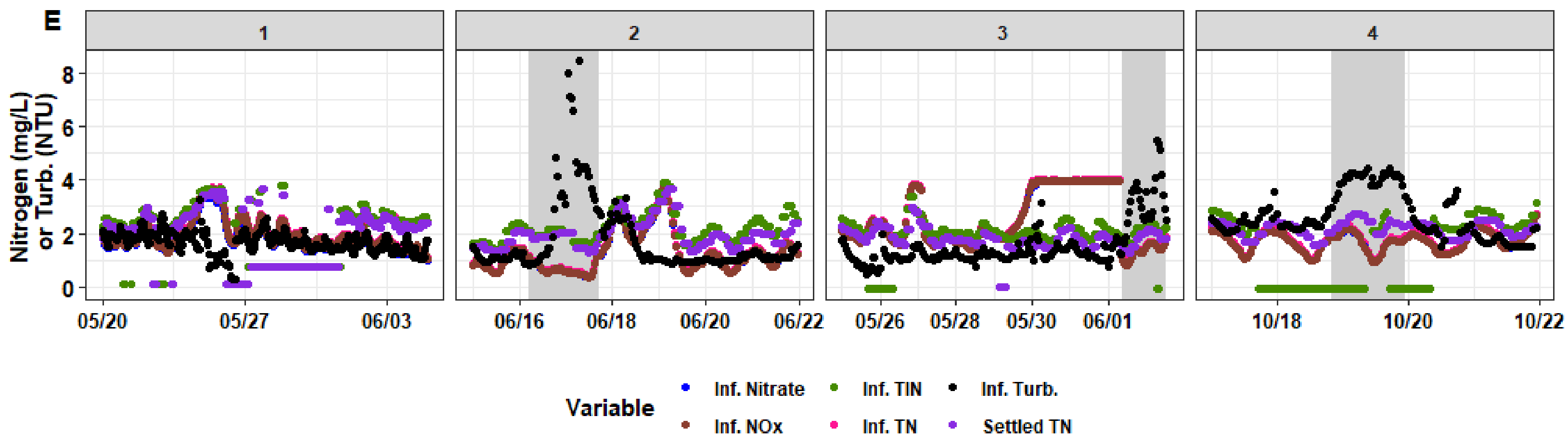


Figure B-4. Timeseries of 30 Water Quality Variables.

Grouped A-F based on similar maximum values. Horizontally, panels are separated into the four time-segments analyzed. #1 and #2 are in 2019. #3 and #4 are in 2020. #1, #2, and #3 were used for the training set and #4 was the testing set. Gray shaded areas represent abnormal industrial slug events.



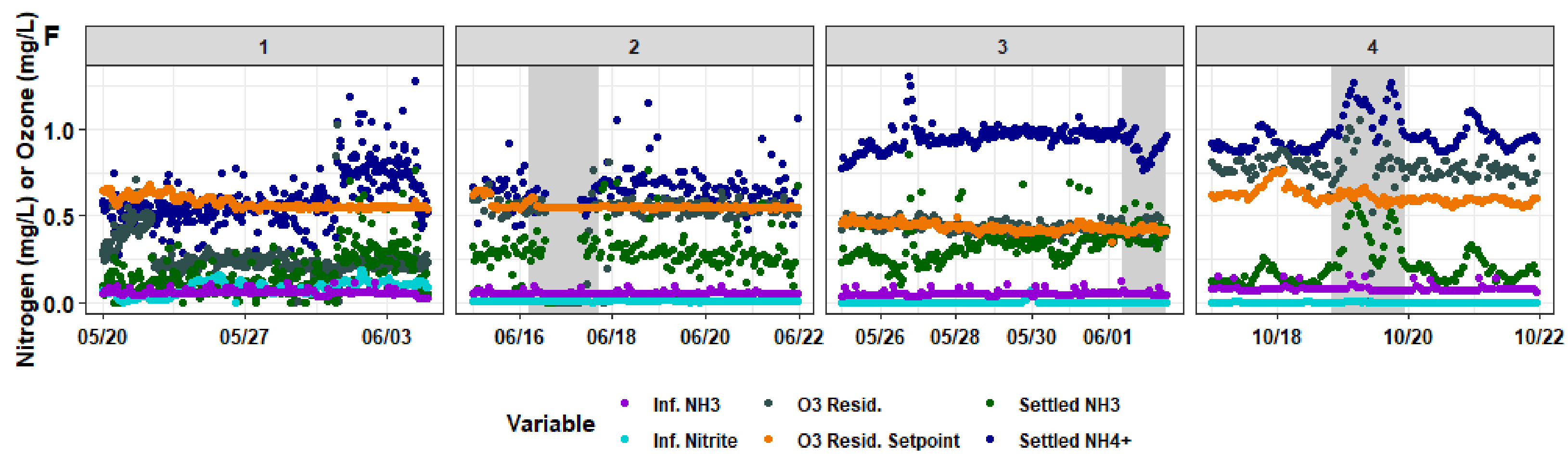


Figure B-4. Timeseries of 30 Water Quality Variables.

Grouped A-F based on similar maximum values. Horizontally, panels are separated into the four time-segments analyzed. #1 and #2 are in 2019. #3 and #4 are in 2020. #1, #2, and #3 were used for the training set and #4 was the testing set. Gray shaded areas represent abnormal industrial slug events.

B.3.2 Machine Learning Results

Based on the screening results, six models were selected for further evaluation: Cost-Sensitive C5.0 (C5.0Cost), Oblique Random Forest with discriminative nodes based on linear support vector machines (ORFsvm), Penalized Logistic Regression (plr), Support Vector Machines with Radial Basis Function Kernel (svmRadial), Random Forest Rule-Based Model (rfRules), and Boosted Tree (bstTree) (Table B-6).

Lagging the raw waster influent and secondary effluent variables resulted in less testing set accuracy for all six of these models. This could be because lagging reduced the training set sample size by n=54 or about 7%—eighteen sample points now had missing data at the start of each of the three non-consecutive periods. The lower testing set accuracy after lagging may also have been in part due to the increased percent accuracy loss per error with the smaller testing set (n=102 instead of n=120).

Table B-6. Model Screening Results.

All models were trained with raw data for all 30 variables using their default tuning parameter options in the caret package. Selected for in-depth evaluation and optimization.

Model	Abb.	Training Set	Testing Set			
		Acc.	Acc.	p-value (Acc. > NIR)	Event Sensitivity	False Positives
Boosted Classification Trees	ada	99%	93%	1.13×10-6	71%	0
AdaBoost Classification Trees	adaboost	99%	95%	4.95×10-8	82%	1
Bayesian Generalized Linear Model	bayesglm	99%	93%	4.46×10-6	68%	0
Boosted Linear Model	BstLm	93%	77%	0.55	0%	0
Boosted Smoothing Spline	bstSm	98%	94%	2.54×10-7	75%	0
Boosted Treea	bstTree	99%	95%	4.95×10-8	89%	3
Cost-Sensitive C5.0a	C5.0Cost	99%	97%	1.12×10-9	86%	0
Single C5.0 Ruleset	C5.0Rules	98%	93%	1.13×10-6	79%	2
DeepBoost	deepboost	99%	93%	1.13×10-6	71%	0
Linear Distance Weighted Discriminant	dwdLinear	99%	93%	1.13×10-6	71%	0
Distance Weighted Discrimination with Radial Basis Function Kernel	dwdRadial	92%	77%	0.55	0%	0
Generalized Additive Model using LOESS	gamLoess	99%	83%	0.049	29%	0
Generalized Additive Model using Splines	gamSpline	99%	82%	0.12	21%	0
Generalized Linear Model with Stepwise Feature Selection	glmStepAIC	99%	83%	0.049	29%	0
Weighted k-Nearest Neighbors	kknn	99%	93%	4.46×10-6	68%	0
Localized Linear Discriminant Analysis	loclda	99%	77%	0.55	0%	0
Least Squares Support Vector Machine with Radial Basis Function Kernel	lssvmRadial	99%	88%	0.00094	50%	0
Mixture Discriminant Analysis	mda	99%	92%	1.57×10-5	64%	0
Tree-Based Ensembles	nodeHarvest	99%	93%	1.13×10-6	71%	0
Oblique Random Forest with Logistic Regression	ORFlog	99%	89%	0.00038	54%	0
Oblique Random Forest with Partial Least Squares Regression	ORFpls	99%	95%	4.95×10-8	79%	0
Oblique Random Forest with Ridge Regression	ORFridge	99%	93%	1.13×10-6	71%	0
Oblique Random Forest with Support Vector Machines	ORFsvm	99%	96%	8.20×10-9	82%	0

Table B-6. Model Screening Results. (Continued)
All models were trained with raw data for all 30 variables using their default tuning parameter options in the caret package. Selected for in-depth evaluation and optimization.

Model	Abb.	Training Set	Testing Set			
		Acc.	Acc.	p-value (Acc. > NIR)	Event Sensitivity	False Positives
Neural Network with Feature Extraction	pcaNNet	99%	92%	1.57×10-5	64%	0
Penalized Logistic Regressiona	plr	100%	88%	0.00094	50%	0
Quadratic Discriminant Analysis	qda	98%	77%	0.55	0%	0
Random Forest Rule-Based Modela	rfRules	98%	54%	1	100%	55
Rotation Forest	rotationForest	99%	95%	4.95×10-8	79%	0
Sparse Distance Weighted Discrimination	sdwd	95%	77%	0.55	0%	0
Stabilized Nearest Neighbor Classifier	snn	97%	77%	0.55	0%	0
Sparse Linear Discriminant Analysis	sparseLDA	92%	77%	0.55	0%	0
Support Vector Machines with Linear Kernel	svmLinear	99%	92%	1.57×10-5	64%	0
Linear Support Vector Machines with Class Weights	svmLinearWeights	99%	90%	0.00015	57%	0
L2 Regularized Linear Support Vector Machines with Class Weights	svmLinearWeights2	98%	83%	0.077	25%	0
Support Vector Machines with Radial Basis Function Kernela	svmRadial	99%	83%	0.077	25%	0

B.3.2.1 Cost-Sensitive C5.0

C5.0Cost is a decision tree algorithm with adaptive boosting and efficient pruning algorithms for relatively fast calculation (Nolan 2002; Peng et al. 2020). C5.0Cost had the highest testing set accuracy in the screening, 96.7 percent (Table B-7). The testing set accuracy of this model did not depend on the seed. Preprocessing by PCA, taking the rolling median of each variable, or the difference relative to the rolling median of each variable did not increase testing set accuracy. Biofilter influent pH was identified as the least important variable but omitting it reduced testing set accuracy. C5.0Cost has four tuning parameters: (1) whether the model is based on associative rules or decision trees, (2) the number of boosting iterations (i.e., trials), (3) the cost of errors, (4) and whether an internal variable section process called winnowing is used. A rules-based model without winnowing with 20 trials (boosting iterations) and cost=1 (weight of one assigned to errors) was selected based on the bootstrapped training set accuracy. Trials over 20 or cost greater than 1 would have led to overfitting, with similar training set accuracy but lower testing set accuracy (Figure B-6). C5.0 had zero false positives and four false negatives, which were consecutive at the beginning of the industrial discharge event (Figure B-7). Thus, there would have been a 4-hour delay between the first hourly datapoint considered to be part of the event and the automated alert (i.e., first true positive).

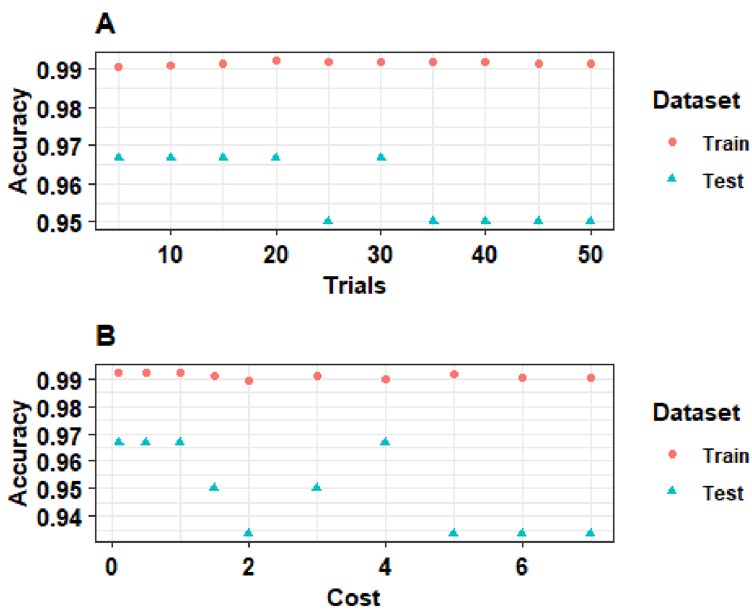


Figure B-6. Train and Testing Set Accuracy of C5.0Cost with Rules-Based Model without Winnowing Across Range of (A) trials and (B) cost.

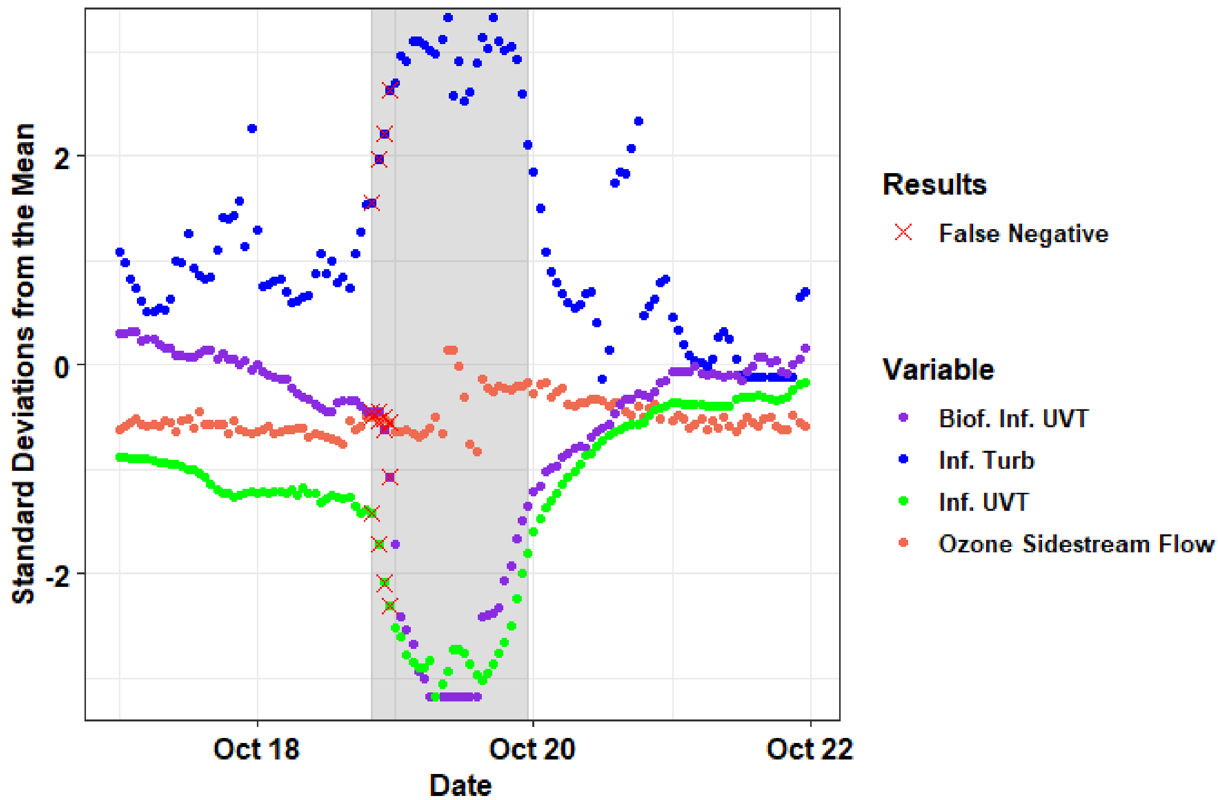


Figure B-7. Testing set results of C5.0Cost.

Using all raw data with default tuning parameters (Cost=1, trials=20, model=rules, winnow=false). The shaded gray area represents the Event. Red X's indicate false negatives. The four most important variables (biofilter influent UVT, influent turbidity, influent UVT, and ozone sidestream flow) are shown and scaled by training set standard deviation and mean.

Table B-7. Summary and Performance Metrics of Six Optimized Models with Their Most Beneficial Preprocessing Techniques, Optimal Tuning Parameters, and Final Variable Selection.

Testing set false negatives and false positives are out of a sample size of n=120 or 5 days of hourly data.

				Training Set	Testing Set					
Model	**Preprocessing**	**Variables**	**Tuning Parameters**	**Accuracy**	**Accuracy**	**Balanced Accuracy**	**Cohen's Kappa**	**Event Sensitivity**	**False Positives**	**Time until 1st Detection (hr)**
C5.0Cost	None	All	Winnow=FALSE, model=rules, cost=1, trials=20	99.2%	96.7%	92.9%	0.902	86%	0	4
ORFsvm	Raw and Differences from the Rolling Median	All	Mtry=31	99.3%	96.7%	92.9%	0.902	86%	0	3
plr	PCA	Principal Components 1 through 22	CP=BIC, lambda=0.001	99.3%	90%	78.6%	0.672	57%	0	5

Table B-7. Summary and Performance Metrics of Six Optimized Models with Their Most Beneficial Preprocessing Techniques, Optimal Tuning Parameters, and Final Variable Selection. (Continued)

Testing set false negatives and false positives are out of a sample size of n=120 or 5 days of hourly data.

Model	Preprocessing	Variables	Tuning Parameters	Training Set Accuracy	Testing Set Accuracy	Testing Set Balanced Accuracy	Testing Set Cohen's Kappa	Testing Set Event Sensitivity	Testing Set False Positives	Testing Set Time until 1st Detection (hr)
svmRadial	Raw and Differences from the Rolling Median	Influent UVT difference, influent pH difference, influent temperature difference, settled UVT difference, settled ORP difference, ozone dose difference, biofilter influent UVT difference, biofilter influent pH difference, raw conductivity, influent NH3, influent conductivity, influent UVT, influent turbidity, influent pH, ozone residual setpoint, and biofilter influent ORP	C=1, sigma=0.015	99.5%	98.3%	96.4%	0.952	93%	0	2
rfRules	Raw and Differences from the Rolling Median	Influent nitrite, settled TOC, influent nitrate difference, raw conductivity, ozone sidestream flow, and influent ammonia difference	mtry=6, maxdepth=4	99.3%	94.2%	88.7%	0.826	79%	1	6

Table B-7. Summary and Performance Metrics of Six Optimized Models with Their Most Beneficial Preprocessing Techniques, Optimal Tuning Parameters, and Final Variable Selection. (Continued)

Testing set false negatives and false positives are out of a sample size of n=120 or 5 days of hourly data.

				Training Set	Testing Set					
Model	Preprocessing	Variables	Tuning Parameters	Accuracy	Accuracy	Balanced Accuracy	Cohen's Kappa	Event Sensitivity	False Positives	Time until 1st Detection (hr)
bstTree	None	Raw conductivity, influent nitrate, influent conductivity, influent UVT, settled NH4+, settled ORP, settled TOC, biofilter influent total Cl2, biofilter influent pH	maxdepth=3, nu=0.1, mstop=150	99.3%	99.2%	98.2%	0.976	96%	0	1

B.3.2.2 Oblique Random Forest with Support Vector Machines

Oblique random forest is a decision tree ensemble in which multivariate trees learn optimal split directions at internal nodes using linear discriminative models (Menze et al. 2011). ORFsvm is a type of oblique random forest in which node splitting rules are based on support vector machines (Poona et al.2016). ORFsvm had the second highest testing set accuracy in the screening, 95.8 percent (Table B-7). Retraining the model with 100 distinct seeds revealed that the testing set accuracy of this model was stochastic (Figure B-8A). Nevertheless, the mean testing set accuracy was 95.5 percent with 0.1 percent standard error, so this model would indeed rank second in testing set accuracy on average. Training the model on both the raw data and the differences from the rolling median increased the ORFsvm median testing set accuracy to 96.7 percent, tying C5.0Cost as the most accurate model (Figure B-8B). The four errors in ORFsvm with this preprocessing were all false negatives, three of which were at the start of the event, and one at the end of the event (Figure B-9). Thus, in practice, this model would have detected the event three hours after the first hourly datapoint considered part of the event. This would have exceeded the performance of C5.0Cost. ORFsvm had one tuning parameter, mtry, which is the number of randomly selected variables for each decision tree within the ensemble. However, varying mtry from 1 to 60 had no impact on train or testing set accuracy when using both raw data and differences from the rolling median.

ORFsvm had a relatively slow training calculation time, about 6 minutes per tuning parameter setting and seed iteration with 60 variables (all raw data and differences from rolling median). The varImp() function was not applicable for ORFsvm, and so could not be used to omit variables. Considering ORFsvm accuracy was stochastic, a sample size of at least 30 seed iterations would be required to determine if small changes in accuracy were the result of variable omission or random chance. Thus, a one-at-a-time variable omission procedure would have taken at least one week of computation time, and potentially months or over a year depending on the number of variables omitted. So, ORFsvm was not evaluated for variable omission. While not necessarily precluding the usage of this model, this slow training time could be a practical limitation, especially if the utility chooses to expand the training set sample size over time.

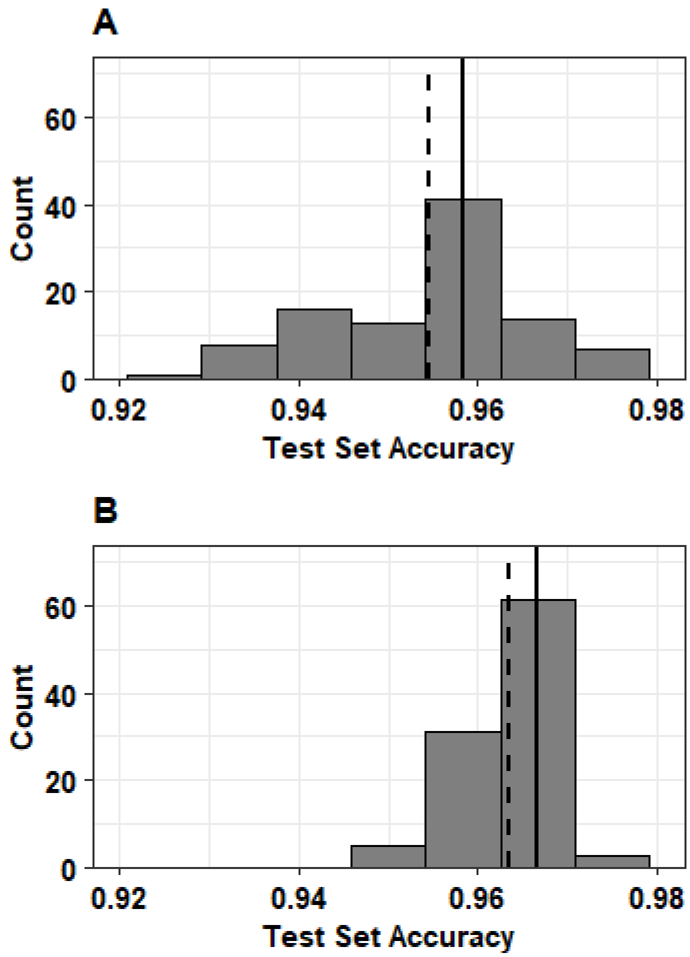


Figure B-8. Distribution of Testing Set Accuracy of ORFsvm.

All variables and with default tuning parameters over 100 distinct seeds, with (A) raw data and (B) both raw data and differences from the rolling median. The solid vertical black lines represent the median and the dashed vertical black lines represent the mean.

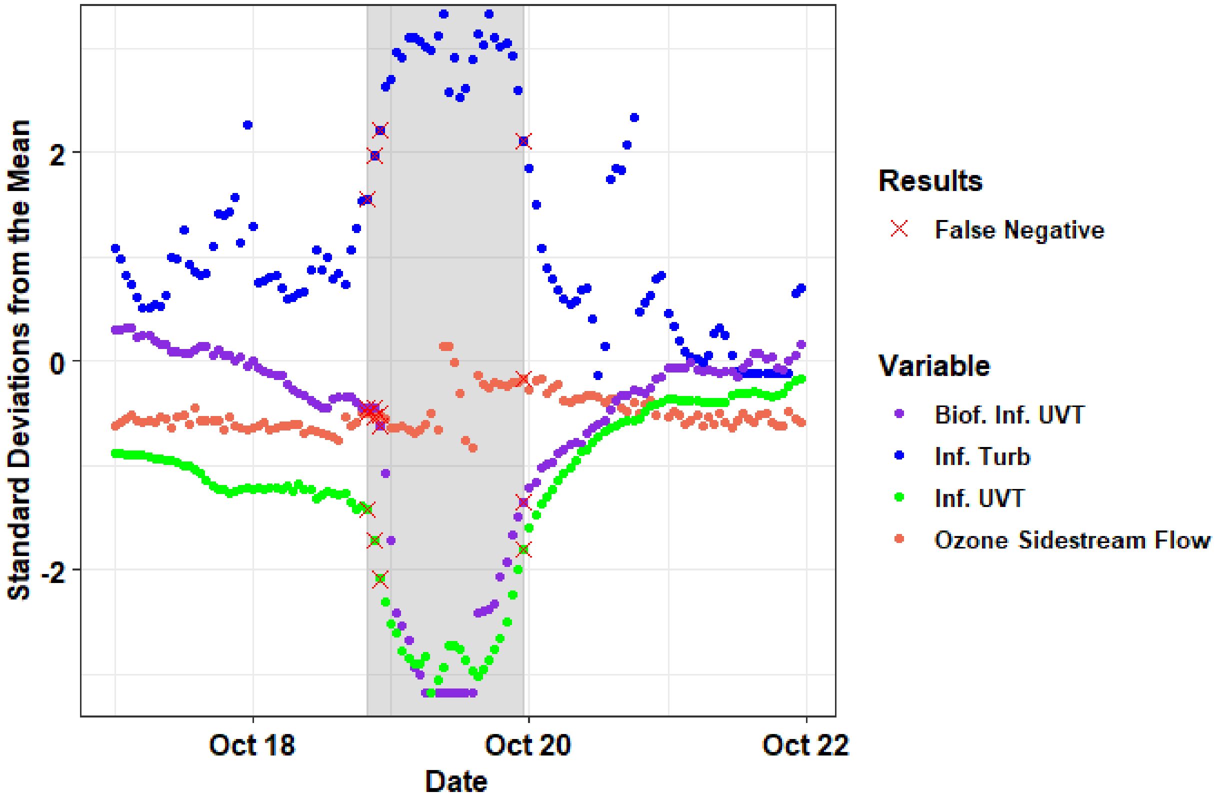


Figure B-9. Testing Set Results of ORFsvm Using All Variables, Both Raw Data and Differences from the 24-hr Rolling Median, and Default Tuning Parameters (mtry=31).

The shaded gray area represents the Event. Red X's indicate false negatives. Biofilter influent UVT, influent turbidity, influent UVT, and ozone sidestream flow are shown and scaled by training set standard deviation and mean.

B.3.2.3 Penalized Logistic Regression

plr is logistic regression with L_2-regularization (Park and Hastie 2008). plr had the highest training set accuracy in the screening, 99.8 percent (see table above). However, its test accuracy was a less impressive 88.3 percent, indicating that under the conditions of the screening, this model was relatively overfit (i.e., mistaking random noise in the training set for true patterns, and thus resulting in a model fit that is more accurate for the training set but less accurate for the testing set). The testing set accuracy of this model did not depend on seed. PCA was the most beneficial preprocessing technique for this model, improving testing set accuracy from 88.3 percent to 90 percent. PCA also decreased the training time per tuning parameter setting from 196 s to 1.7 s. Omitting the 23rd through 30th principal components further decreased the training time to 1.5 s with no loss in testing set accuracy. Training and testing set accuracy were unaffected if the "complexity parameter" (CP) tuning parameter were set to Bayesian information criterion (BIC) or Akaike information criterion (AIC). Testing set accuracy was not affected over L2 penalties ranging from 10-5 to 1. Despite the improvements with preprocessing, the 90 percent testing set accuracy for plr would not be satisfactory compared to other models evaluated.

B.3.2.4 Support Vector Machines with Radial Basis Function Kernel

Support vector machines construct optimal separations in multi-dimensional space using the points that are closest to the boundaries (Schölkopf et al. 1997). svmRadial constructs non-linear hyperplanes based on distances from centers (Schölkopf et al. 1997). svmRadial had the second highest training set accuracy in the screening, 99.5 percent (see table above). However, its test accuracy was a less impressive 82.5 precent, indicating that under the conditions of the screening, this model was relatively overfit. The accuracy of this model did not depend on the seed. The most beneficial processing technique was using both the raw data and the differences from the rolling median, improving the testing set accuracy to 86.7 percent. Omitting 40 variables improved the testing set accuracy to 98.3 percent. The remaining variables after these omissions were differences from the rolling median for influent TIN, influent nitrate, influent UVT, influent pH, influent temperature, settled UVT, settled ORP, ozone dose, biofilter influent UVT, biofilter influent pH; as well as raw conductivity, influent NOx, influent NH3, influent conductivity, influent UVT, influent turbidity, influent pH, influent temperature, ozone residual setpoint, and biofilter influent ORP. Omitting four more variables (influent TIN difference, influent temperature, influent NOx, and influent nitrate difference) resulted in no loss of accuracy and improved the training computation time from 1.8 to 0.97 seconds. With this preprocessing and set of variables, svmRadial had zero false positives and only two false negatives, which were consecutive at the beginning of the event (Figure B-10). Thus, this model outperformed C5.0Cost or ORFsvm.

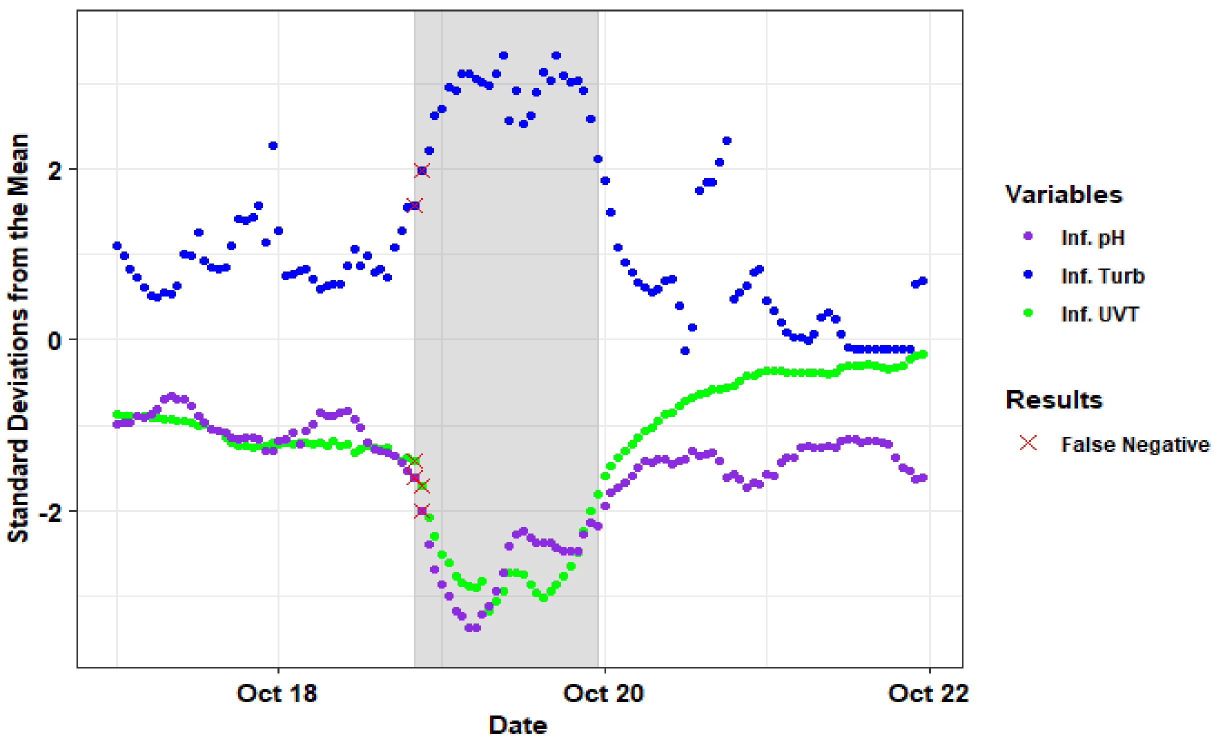


Figure B-10. Testing set results of svmRadial.

(C=1, sigma=0.15, raw and differences from rolling median, 44 unimportant variables omitted). The shaded gray area represents the Event. Red X's indicate false negatives. The three variables whose omission would have resulted in greatest loss in testing set accuracy (influent pH, influent turbidity, and influent UVT), are shown.

Tuning parameters for svmRadial were C, the cost of errors, and sigma, the decay rate as points become more distant from the centers. For all raw data and differences from the rolling median, the optimal settings among the default options were C=1 and sigma=0.015. So, these settings were kept when iteratively omitting variables. Broader ranges of tuning parameters were then tested using the sixteen selected variables. Holding C to 1, highest testing set accuracy was reached with sigma around 0.15, while highest training set accuracy occurred at a slightly higher sigma of 0.23 (Figure B-11A). Holding sigma to 0.15, the highest training set accuracy occurred with C around 1.5, but highest testing set accuracy occurred with C around 1 (Figure B-11B). Thus, the default tuning parameter settings were effectively optimal for predictive accuracy in this dataset.

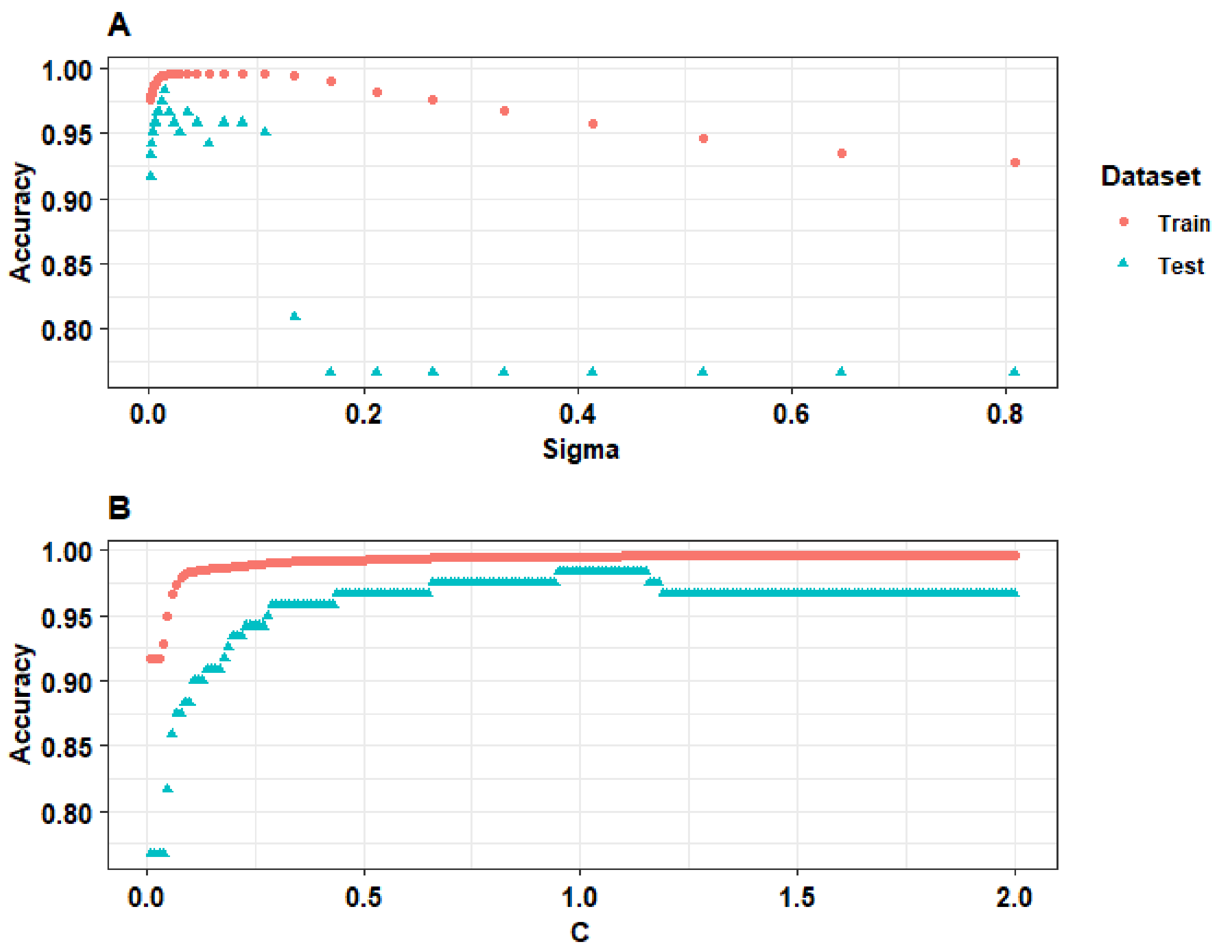


Figure B-11. Training Set (red) and Testing Set (blue).
Accuracy with (A) Sigma Ranging from 0.001 to 0.81 with C=1 and (B) C ranging from 0.01 to 2 with sigma = 0.015.

B.3.2.5 Random Forest Rule-Based Model

rfRules is an ensemble classifier based on associative rules (Deng et al. 2014). In the screening, rfRules had the highest event sensitivity at 100 percent (see table above). However, it had a testing set accuracy of 54.2 percent with 55 false positives, which is clearly unacceptable over the five-day timeframe of the testing set. Replicating with 30 distinct seeds, the testing set accuracy of this model was stochastic, ranging from 54.2 percent to 98.3 percent with a median of 82.5 percent, This indicated the median testing set accuracy was likely better than it appeared in the screening, but more variable compared to ORFsvm. Also, the distribution of testing set accuracies with different seeds was not normally distributed. Based on a paired Wilcoxon test and the same thirty distinct seeds, PCA, rolling median, and differences from the rolling median did not result in a significant increase in testing set accuracy (p-value > 0.05). However, including both raw data and the differences from the rolling median did increase median testing set accuracy (p-value = 0.0085), to 93.75 percent. With that preprocessing, according to the varImp function, all variables had an importance score of 0 except influent nitrite, settled TOC, influent nitrate difference, raw conductivity, ozone sidestream flow, and influent ammonia difference. With just these six variables, the training time did not meaningfully decrease but the testing set accuracy was significantly higher (p-value = 0.0046), 94.2 percent in all 30 seed iterations. rfRules had two tuning parameters: mtry, the number of variables randomly selected for each tree; and maxdepth, the maximum depth of each tree. With the six variables listed above, mtry was varied from 1 to 6 and maxdepth was varied from 1 to 5. Training set accuracy generally increased with higher maxdepth and mtry (Figure B-12A). The maximum testing set accuracy was 94.2 percent, and this occurred with a maxdepth of at least 3 and mtry of at least 5 (Figure B-12B). This maximum testing set accuracy corresponded to six hours until the first detection, which would not be competitive with the models described above.

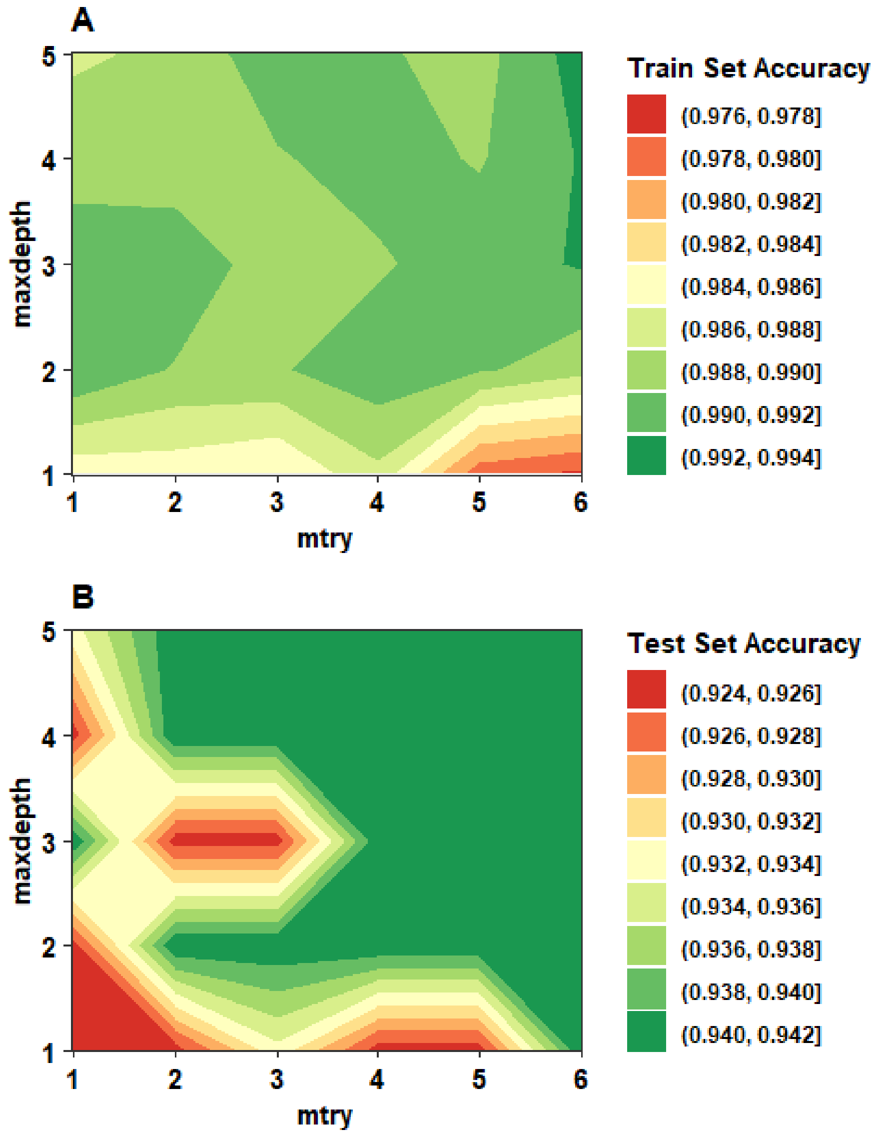


Figure B-12. Contour Plot of (A) Train and (B) Testing Set Accuracy of rfRules with Six Variables.
(influent nitrite, settled TOC, influent nitrate difference, raw conductivity, ozone sidestream flow, and influent ammonia difference); maxdepth from 1 to 5, and mtry from 1 to 6.

B.3.2.6 Boosted Tree

bstTree is a type of decision tree ensemble in which each subsequent tree is adjusted to optimize performance using a truncated loss function for robustness against outliers (Wang 2018). In the screening, bstTree had second highest event sensitivity at 89.3 percent (Table B-7) and a testing set accuracy of 95 percent. However, it had 3 false positives, which could be considered unacceptable over the five-day timeframe of the testing set. bstTree testing set accuracy did not depend on seed. None of the investigated preprocessing techniques improved bstTree testing set accuracy. Omitting influent TOC and ozone residual setpoint increased the testing set accuracy to 99.2 percent. Further omitting variables until only thirteen remained (raw conductivity, influent nitrate, influent conductivity, influent UVT, settled NH4+, settled ORP, settled TOC, biofilter influent total Cl2, biofilter influent pH) resulted in no loss in accuracy and decreased the training time from 32 to 21 s. bstTree had three tuning parameters: maxdepth, the maximum depth of the decision trees; mstop, the number of boosting iterations; and nu, the step size. Maxdepth=3, mstop=150, and nu=0.1 were selected from among the default options based on training set accuracy for the model trainings described above. Ranging maxdepth 1 to 4, nu from 0.1 to 1, and mstop from 50 to 500 revealed that highest testing set accuracy was achieved with maxdepth=3 and either nu=0.1 with mstop=150 or nu=1 with any value for mstop. Thus, the default tuning parameters were among the most accurate for bstTree. The testing set accuracy of 99.2 percent with bstTree was the highest in this study and corresponded to one false negative and zero false positives. The sole false negative occurred on the first datapoint of the event (Figure B-13), so this model would have detected the event after about 1 hour.

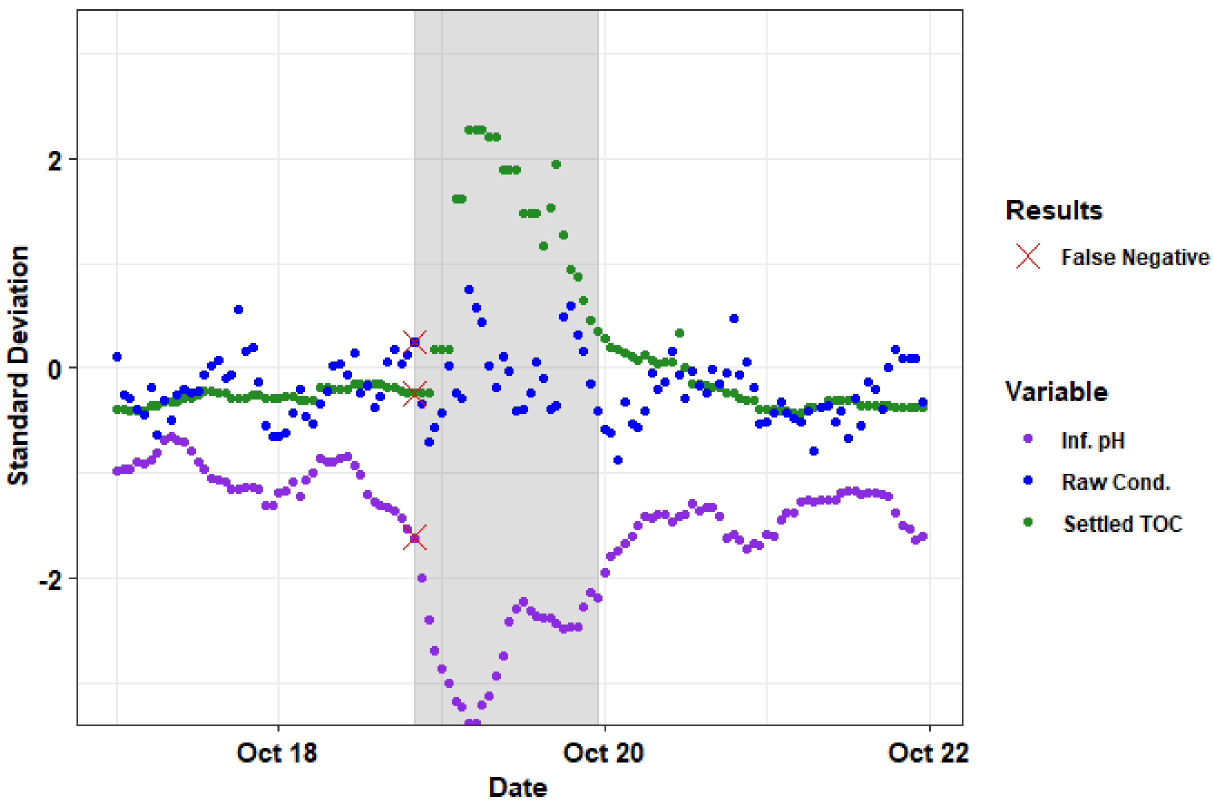


Figure B-13. Testing Set Results of bstTree (n=0.1, mstop=150, maxdepth=3).
Trained on the raw data of thirteen variables (raw conductivity, influent nitrate, influent conductivity, influent UVT, settled NH4+, settled ORP, settled TOC, biofilter influent total Cl2, biofilter influent pH). The shaded gray area represents the Event. Red X's indicate false negatives. The three variables whose omission would have resulted in greatest loss in testing set accuracy (influent pH, raw conductivity, and settled TOC), are shown.

B.3.3 Actual Thresholds

A simpler approach than SML for alerts is to set a fixed threshold on a single variable (e.g., settled TOC over 10 mg/L triggers an alert). This is a common approach in current practice at wastewater and drinking water facilities, so it was applied to this dataset as a reference against which to benchmark the performance of the SML models.

This section shows the time until detection using alert thresholds values that were in place at the SWIFT RC. These alerts were set conservatively lower than corresponding alarms, which were based on ensuring the public health and regulatory compliance. Alerts were in place on six of the variables shared for this study (Table B-8). Only three were triggered during the testing set event: secondary effluent turbidity, settled water total chlorine, and ozone dose. Secondary effluent turbidity triggered soonest during the event, after just two hours. However, there were also two alerts for effluent turbidity within the five-day testing set not associated with the industrial event.

Table B-8. Actual Threshold-Based Alerts in Place at SWIFT RC and Their Performance Detecting the Event in the Testing Set.

Accuracy, false positives, and time until first detection are all for the testing set.

Location	Variable	Unit	Actual Alert Threshold	Accuracy	False Positives	Time Until 1st Detection (hr)
Secondary Effluent / SWIFT RC Influent	Total Inorganic Nitrogen	mg/L	4	76.7%	0	Never
	Conductivity	mS/cm	1500	76.7%	0	Never
	Turbidity	NTU	3.5	95.8%	2	2
Settled Water (Post Floc/Sed)	Monochloramine	mg/L	2	76.7%	0	Never
	Total chlorine	mg/L	2	77.5%	0	5
Ozonation System	Ozone Dose	mg/L	7	89.2%	0	5

B.3.4 Data-Driven Thresholds

Current fixed-threshold-based alerts at the SWIFT RC are based on safety factors, critical control points, and ensuring the public health or regulatory compliance. However, another approach would be to set alert thresholds based on the maximum (or a high percentile) of the data considered normal. This approach would be SML-like, in that it would be data-driven, and thresholds could be trained, tested, and refined over time. However, compared the SML methods described above, this approach would be much simpler since it would be monovariate. In this section, alerts were set based on maximum or minimum normal datapoint for each variable in the training set. Alerts set this way are herein called “data-driven thresholds.” For pH, UVT, and disinfectant residuals, the data-driven threshold was set to the minimum normal datapoint in the training set (Figure B14). Otherwise, the data-driven threshold was set to the maximum normal datapoint in the training set.

For most variables, a data-driven threshold would not have predicted any events in the testing set (equivalent to the NIR) (see Table B8). However, a threshold based influent UVT would have achieved 98.3 percent testing set accuracy with zero false positives. This testing set accuracy would be equal or better than all but one of the SML models evaluated.

However, greater sample size would be expected to improve the relative performance of the SML models. In contrast, greater sample size might not improve the data-driven threshold results. The use of minima or maxima to set thresholds like was done here would become increasingly conservative (i.e., fewer false positives, more false negatives) with greater sample size because it would allow more time for non-industrial outliers in the Normal training data. This could be counteracted somewhat by setting the threshold based on a specified percentile that strikes the desired balanced between false positives and false negatives.

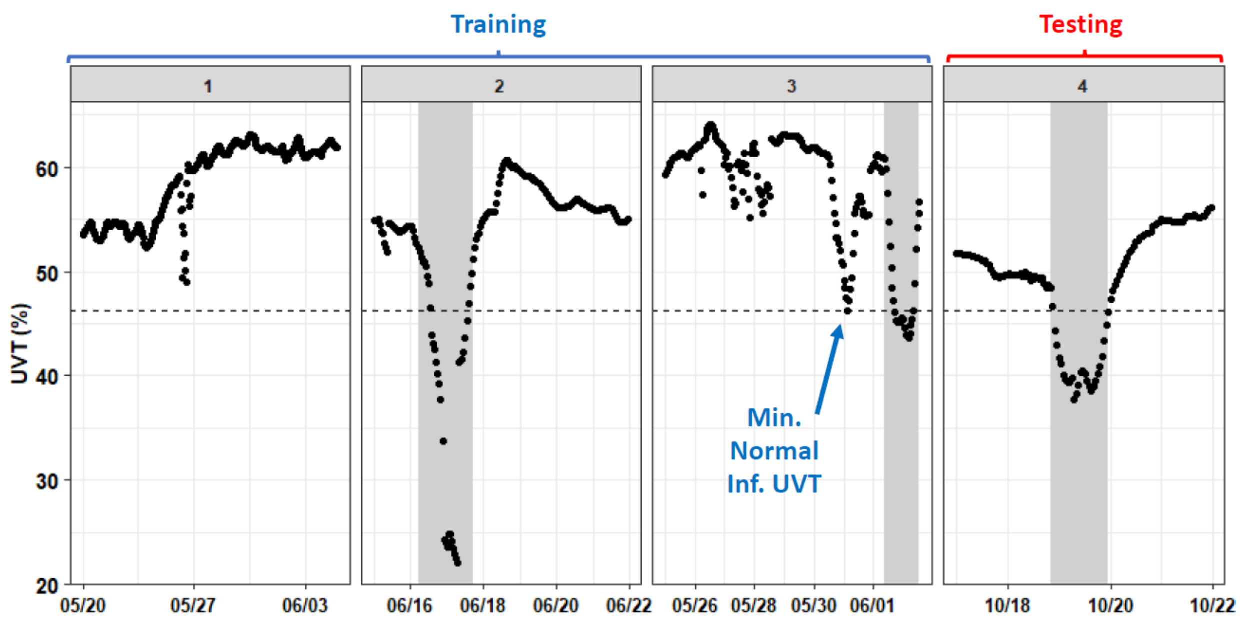


Figure B-14: Data-Driven Threshold Example Using Influent UVT.
The dashed black line represents the threshold. The blue arrow indicates the Normal, training set datapoint on which it was based. Shaded grey areas indicate events.

Table B-9. Performance of a Fixed Threshold Approach Based on each of the 30 Variables.
Since the thresholds were set to the maximum Normal value of each variable in the training set, all thresholds would have resulted in zero training set false positives.

Location	Variable	Training set Event Sensitivity	Testing set Accuracy	Testing set p-value	Testing set False Positives	Testing set Event Sensitivity
Raw Wastewater Influent	Conductivity	0%	76.7%	0.55	0	0%
Secondary Wastewater Effluent	Flow	0%	76.7%	0.55	0	0%
	Total Nitrogen	0%	76.7%	0.55	0	0%
	Total Inorganic Nitrogen	0%	76.7%	0.55	0	0%
	Total Organic Carbon	63.6%	82.5%	0.077	1	29%
	Nitrite	0%	76.7%	0.55	0	0%
	Nitrogen Oxides	0%	76.7%	0.55	0	0%
	Nitrate	0%	76.7%	0.55	0	0%
	Ammonia	0%	71.7%	0.92	8	7%
	Conductivity	0%	76.7%	0.55	0	0%
	UV Transmittance	60.6%	98.3%	9.9×10^{-12}	0	93%
	Turbidity	48.5%	95.8%	8.2×10^{-9}	3	93%
	pH	27.3%	82.5%	0.077	0	25%
	Temperature	0%	76.7%	0.55	0	0%
Settled Water	UV Transmittance	40.9%	79.2%	0.30	0	11%
	Monochloramine	18.2%	76.7%	0.55	0	0%
	Ammonium	0%	76.7%	0.55	0	0%
	Total chlorine	0%	76.7%	0.55	0	0%
	Redox potential	0%	76.7%	0.55	0	0%
	Total Organic Carbon	75.8%	93.3%	1.1×10^{-6}	0	71%
	Total Nitrogen	0%	76.7%	0.55	0	0%
	Free Ammonia	0%	76.7%	0.55	0	0%
Ozonation System	Ozone Dose	31.8%	80%	0.23	0	14%
	Ozone Sidestream Flow	0%	76.7%	0.55	0	0%
	Ozone Residual Setpoint	0%	70%	0.96	8	0%
	Ozone Residual	0%	76.7%	0.55	0	0%
Biofiltration Influent	UV Transmittance	63.6%	93.3%	1.1×10^{-6}	0	71%
	Total Chlorine	0%	76.7%	0.55	0	0%
	Redox potential	0%	23.3%	1	92	100%
	pH	21.2%	85.8%	0.0090	0	39%

B.4 Discussion

Testing set accuracy has limitations as a metric of success for SML models. For “unbalanced” data [e.g., data with many more of one class than the other(s)], such as used here, models could achieve over 70 percent accuracy by always assuming datapoints were Normal, or by randomly guessing Random vs Event based solely on their proportion in the training set. Furthermore, using only testing set accuracy, the success of models cannot be directly compared across studies, since the accuracy would depend in part on the proportion of classes in the respective datasets.

One alternative metric is “balanced accuracy,” or what the accuracy would be if there were equal percentages of each class in the dataset. Balanced accuracy is more intercomparable across studies and cannot be increased by increasing the proportion of a specific class. However, in the context of alert systems for the water or wastewater industry, false positives would be a more important error type than false negatives. False positives (i.e., Normal datapoints incorrectly predicted as Event) would waste resources and eventually lead to a boy-who-cried-wolf scenario in which the alert system is disregarded or discontinued. For hourly data frequency, even a 1 percent false positive rate would lead to false alerts roughly twice per week, which would be plainly unacceptable to utility operators. In contrast, many of the datapoints labeled and predicted as an Event in this study could be considered to occur at low levels that would not yet pose an immediate threat to the operation or treatment goals of the facility. Thus, a higher false negative rate could be considered tolerable compared to the acceptable false positive rate. Considering the above, for a dataset with 75 percent Normal data, a model with 0 percent false positive rate, 80 percent false negative rate, and 80 percent accuracy and 60 percent balanced accuracy (Figure B-15 Example A) would be considered far preferable to a model with 40 percent false positive rate, 0 percent false negative rate, and 70 percent accuracy and 80 percent balanced accuracy (Figure B-15 Example B). Balanced accuracies for the optimized versions of the models selected for in-depth evaluation are shown in Table B-7. Except for plr, the optimized versions of all models selected for in-depth evaluation had balanced accuracy over 88 percent. bstTree had the highest test set balanced accuracy at 98.2 percent.

Example A

		Reference	
		Normal	Event
Prediction	Normal	750	200
	Event	0	50

n	1000
Sensitivity	20%
Balanced Accuracy	60%
Accuracy	80%
FPR	0%
FNR	80%

Example B

		Reference	
		Normal	Event
Prediction	Normal	450	0
	Event	300	250

n	1000
Sensitivity	100%
Balanced Accuracy	80%
Accuracy	70%
FPR	40%
FNR	0%

Figure B-15. Hypothetical Model Result Examples with Contrasting Accuracy and Balanced Accuracy.
FPR is false positive rate. FNR is false negative rate.

Another alternative to accuracy is Cohen's Kappa. Cohen's Kappa compares the agreement between the true classifications and the model classifications to the agreement that could occur due to random allocation (Cohen 1960). The formula for Cohen's Kappa with two classes is:

Equation B-1. Adapted from Chicco et al. (2021).

$$\kappa = \frac{2 \cdot (TP \cdot TN - FP \cdot FN)}{(TP+FP)\cdot(FP+TN)+(TP+FN)\cdot(FN+TN)}$$

Where TP is true positives, FP is false positives, FN is false negatives, and TN is true negatives. One of the limitations with Cohen's Kappa is that there is not a universally agreed magnitude considered adequate (i.e., less consensus compared to the typically acceptable p-value threshold of less than 0.05). One highly cited guideline is that Cohen's Kappa above 0.81 is almost perfect agreement (Landis and Koch 1977). Except plr, the optimized versions of all models selected for in-depth evaluation exceeded this threshold. bstTree had the highest Cohen's Kappa at 0.976.

B.5 Conclusions

- The model bstTree had the highest testing set accuracy for this dataset, 99.2 percent. bstTree would have detected the event in about an hour with zero false positives over the 5-day testing set. bstTree also the highest balanced accuracy at 98.2 percent and Cohen's Kappa at 0.976. Thus, bstTree would have been selected for future monitoring and alerts among the SML models investigated in this study.
- A data-driven fixed threshold based on influent UVT would have resulted in a testing set accuracy of 98.3 percent, below that of bstTree but only by about 1 percent. Based on training set results and regulatory considerations, a threshold based on settled TOC would have been more likely chosen in practice. A settled TOC-based threshold would have resulted in a testing set accuracy of only 93.3 percent. This data-driven threshold for UVT or the actual threshold for secondary effluent turbidity would have detected the event in about two hours, one hour slower than bstTree.
- The most beneficial preprocessing method differed among the SML model types. Two models performed best without preprocessing, one with PCA, and three with raw data and differences from the rolling median.
- In many cases, some variables could be omitted to decrease training time without loss in accuracy. However, the optimal selection of variables depended on the model.
- Certain SML model types from within the random forest family (e.g. ORFsvm, rfRules) had testing set accuracies that depended on the seed to the random number generator. Thus, the accuracy of these models would be more uncertain in full-scale applications, even with appropriate validation and testing procedures.

Looking to the future, the team would make following recommendations:

- As next steps to engineer an accurate, practical, SML-based alert system at HRSD, the team would recommend repeating the above analyses but with greater sample size, including multiple instances of the events in the testing set. This would provide greater confidence about the relative performance of the models, particularly whether the highest-performing model would be best for detecting all events of this type, not just the individual event in this testing set. After that, a small number of high-performing SML models could be piloted in real-time, until an additional event occurs. The time until first detection of the SML models could then be compared in the field against human monitoring and other alert approaches.
- Since Event and Normal datapoints in this dataset were distinguished based on human judgement, the best the models could possibly do would be to match—not exceed—human judgement. On the other hand, a human monitoring the data in real-time might not have concluded that an event was occurring as soon as a human evaluating the whole dataset retrospectively. In future research on machine learning for wastewater or reuse alert systems, this could be achieved by simulating industrial discharges in a pilot or flume like the one at Clean Water Services (CWS) (see Section 3.1). Alternatively, real full-scale

industrial events could be labelled objectively if the industrial source is known and keeps records of discharge flow (e.g., the landfill that discharges limited quantities of leachate to the WWTP that feeds SWIFT RC) (Gonzalez et al. 2021; Nading et al. 2022).

- A limitation of SML-based alert systems is that they are designed to detect events of a known, previously documented type. If a new type of industrial discharge were to occur associated with a different response from the online instrumentation, this may or may not trigger an SML-based alert. Changes in the water quality pattern at the AWTF during industrial discharge events could also occur due to changes in the treatment operation response at the WWTP. So, a strategic solution would be to employ both SML-based and threshold-based alerts or alarms. This would combine the sensitivity of SML with the generalizability of thresholds. These additional thresholds could be set based on training set data, health-based goals, or operational considerations. Advanced multivariate statistical methods for fault or outlier detection other than SML also merit further research in the context of wastewater and reuse (Klanderman et al. 2020).

APPENDIX C

Clean Water Services Machine Learning Case Study

C.1 Introduction

Clean Water Services (CWS) in Hillsboro, Oregon collected data for ten water quality variables from six sensors in a flume of real wastewater influent. Two types of simulated industrial discharge were spiked into the flume on twenty total occasions. The two spike types were (1) a blend of bleach, NaCl, and NaOH and (2) humic acid. The data was originally collected for the purpose of testing sensor containment devices and comparing the instruments under various conditions such as ragging, FOG, and different velocities (Section 3.1 and Appendix A). However, due to (1) high sample size, (2) numerous variables, and (3) known labels for when spikes were occurring, the dataset was also suitable for testing the ability of SML to detect industrial discharges.

In a previous study, (Appendix B), SML performed only marginally better than fixed thresholds to detect real industrial discharges in a demonstration-scale AWTF. However, that studied labelled data into just two classes (Normal and Event). Since the CWS had two types of simulated industrial discharges, datapoints were labelled into three classes (Normal, Spike 1, and Spike 2), so a fixed threshold on a single variable would not be applicable. Furthermore, the CWS dataset has more than three instances of both spike types, which enables multiple instances of each spike type in chronological split training and testing sets. This in turn facilitates a more robust evaluation. Since the spikes were added at known times, there is less risk that datapoints may have been mislabeled. The greater noisiness of the CWS dataset due to its location in primary effluent could be expected to lead to more benefit from preprocessing, or greater difficulty for detecting the events through human visual monitoring. The CWS dataset has greater total sample size which should in theory improve the accuracy of SML models, all else equal. On the other hand, the CWS dataset contains a smaller sample size of datapoints that are true positives, which limits the representativeness of the training set and robust evaluation with the testing set.

C.2 Methods

C.2.1 Construction of the Flume

See Appendix A for the construction of the flume.

C.2.2 Sensors Installed in the Flume

The installed sensors are summarized briefly in Table C1. and described in detail in Appendix A. BOD, COD, TSS, and nitrate were estimated based their site-specific correlations with UV absorbance at various wavelengths measured by the s::can Spectro::lyser.

Table C-1. Water quality variables used for supervised machine learning.

Water Quality Variable	Brand	Instrument	Unit
pH	Yosemitech	Y532-A	
Redox Potential (ORP)	ECD	ORP Pt Cap peek, two-tang probe	mV
pH	ECD	Extended Life pH Electrode RADEL body	
BOD	s::can	Spectro::lyser	mg/L as O_2
COD	s::can	Spectro::lyser	mg/L as O_2
TSS	s::can	Spectro::lyser	mg/L
UV254	s::can	Spectro::lyser	1/m
Nitrate	s::can	Spectro::lyser	mg/L as N
Conductivity	s::can	condu::lyser	µS/cm
pH	s::can	pH::lyser	

C.2.3 Experiments Conducted

Experiments were run with the original goal of comparing the sensors' reliability under various conditions, particularly different velocities, which were hypothesized to impact fouling. The experimental phases are summarized in Table C-2 and described in detail in Appendix A. While both types of simulated industrial discharge were conducted in each experiment phase, in some cases, data for one or more sensors was not available at the time of the spike. Table C-3 shows which Experiments had data for all sensors for at least one datapoint of each spike.

Table C-2. Summary of Flume Experiments and Industrial Spikes.

Experimental Phase	Water Velocity (m/s)	SML Modeling Dataset	Start	End	All Data Available for Spike 1	All Data Available for Spike 2
Experiment 1	0.4	Training	2021-03-29 09:00	2021-04-05 08:20	✔	✔
Experiment 2	0.24	Training	2021-04-05 09:10	2021-04-12 07:40	✔	
Experiment 3	0.17	Training	2021-04-12 08:30	2021-04-19 08:30		✔
Experiment 4	0.4	Testing	2021-04-27 08:10	2021-05-04 08:40	✔	✔
Experiment 5	0.24	Testing	2021-05-04 09:40	2021-05-11 10:00	✔	
Experiment 6	0.17	Testing	2021-05-11 10:30	2021-05-18 08:50	✔	✔

C.2.4 Data Collection and Preparation

Data was recorded for ten water quality variables (Table C-1) using the sensors described in Appendix A. Data was collected every two minutes for the s::can sensors, every 5 minutes for the ECD sensors, and every 10 minutes for the Yosemitech sensor. Data was analyzed at 10-minute intervals, with data from the more frequent instruments averaged over the preceding 10-minute period. Known cleaning events were omitted. Observations were omitted if any of the independent variables included had missing data.

C.2.5 Supervised Machine Learning

Supervised machine learning was conducted in R version 4.1.0 using the caret package (Kuhn 2008). Observations were labelled Normal, Spike 1, or Spike 2. Data from 11:20 a.m., March 27th, 2021 through the end of Experiment 3 (8:20am, April 19th, 2021) was used as a training set. Three instances each of Spike 1 and Spike 2 occurred over the time range of the training set. However, data for one or more variables was not available during Experiment 2 Spike 2 and Experiment 3 Spike 1 (Figure C-1). So, there were effectively two instances each of both spike types when modeling with all variables. Data immediately after the end of Experiment 3 (8:30 a.m., April 19th, 2021) through the end of Experiment 6 (8:50 a.m., May 18th, 2021) was used as the testing set. Thus, there was roughly a 50:50 train:test split. The testing set time range had three instances each of Spike 1 and Spike 2. However, one or more variables was unavailable for all of Experiment 5 Spike 2 and parts of Experiment 5 Spike 2, Experiment 6 Spike 1 and Experiment 6 Spike 2. Since observations were only included when data was available for all variables, the sample size depended on the variables included (Table C-3).

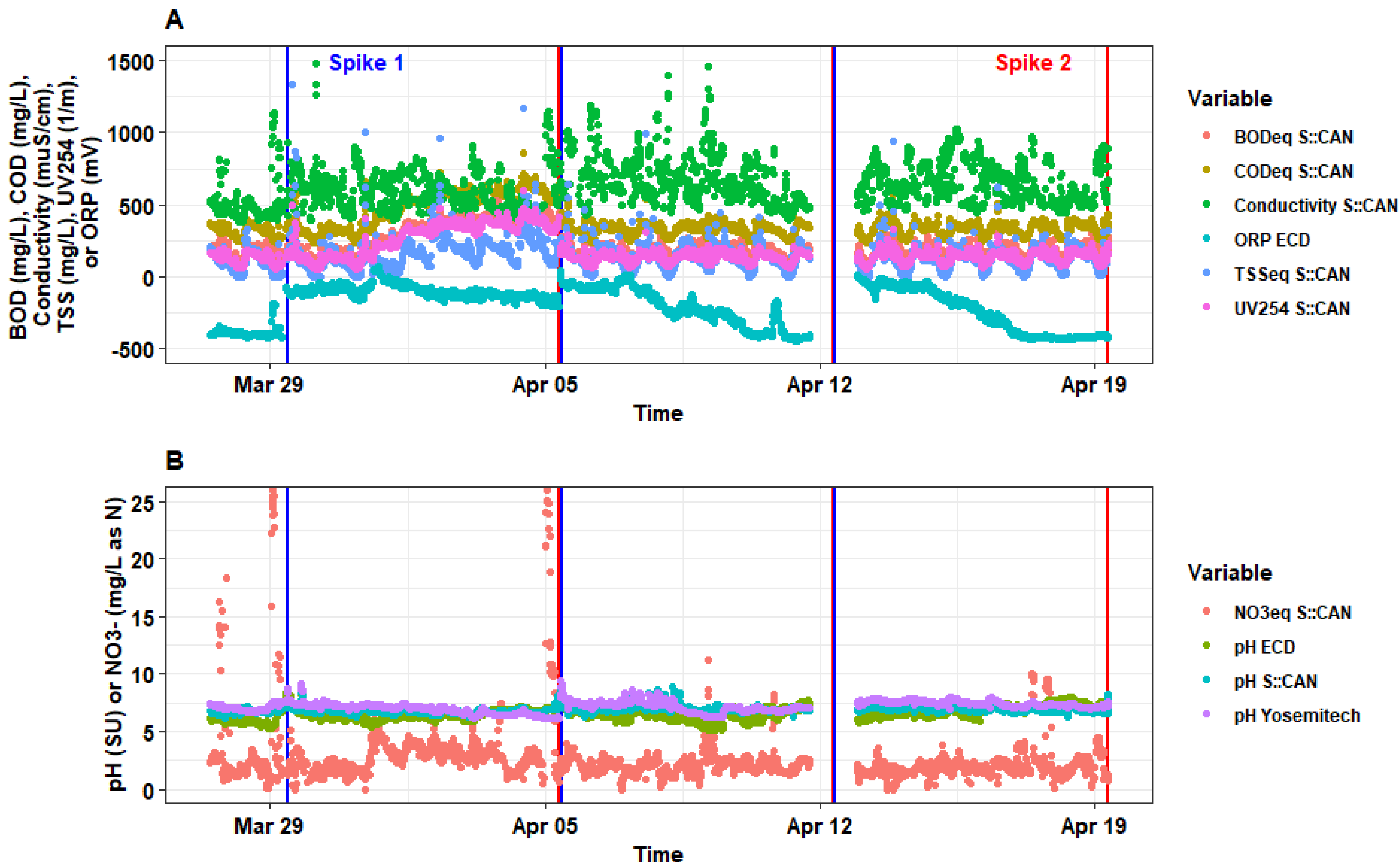


Figure C-1. Training Set Water Quality Data for (A) Variables with Mean Absolute Values Over 100 and (B) Variables with Mean Absolute Values Below 100. Blue and red solid vertical lines indicate Spike 1 and Spike 2 events, respectively.

Table C-3. Sample Sizes With and Without Different Variables.
(Nitrate as measured by the s::can spectro::lyser, pH as measured by the Yosemitech probe, and ORP as measured by the ECD probe)

Variables	Training Set			Testing Set		
	Total	Spike 1	Spike 2	Total	Spike 1	Spike 2
All	2473	9	9	2802	8	6
No s::can Nitrate	2473	9	9	2802	8	6
No Yosemitech pH	2836	10	9	3808	10	10
No ECD ORP	2756	12	11	2944	8	6

Twenty-eight models were considered for screening based on achieving at least 95 percent testing set accuracy detecting stormwater or wastewater effluent in surface water with raw data for nine water quality variables in Thompson and Dickenson (2021). However, two of those models—Random Forest Rule-Based Model and Partial Least Squares with a kernel algorithm for wide datasets—were omitting for requiring over 10 minutes per training iteration. So, 26 models (Table C-4) were trained and tested on raw data for all ten variables. Among these, the three with highest testing set accuracy were selected for in-depth evaluation.

Models were screened on raw data (i.e., no preprocessing) using default tuning parameters in the caret package. The following metrics were recorded for each model: training set accuracy; testing set accuracy; testing set specificity (accuracy of the model when the true classification was not Normal); total false alerts (i.e., false positives, Normal observations incorrectly predicted as either Spike 1 or Spike 2); total false negatives (Spike 1 or Spike 2 incorrectly predicted as Normal); total misclassifications (Spike 1 incorrectly predicted as Spike 2 or vice versa); and p-value that the testing set accuracy exceeds the no information rate (NIR). NIR is the accuracy that could be achieved by always assuming the most common label, which in this case was Normal. The NIR was 99.50 percent since the vast majority of the data was Normal. The training set accuracy was internally cross-validated with 25 bootstraps (Kuhn 2008). This means 25 random samples were selected with replacement (i.e., with the possibility of the same datapoint being selected twice) with the same sample size as the original training set. These random samples were then split 75:25 into training and testing sets; the average accuracy on these random testing sets selected from within the training set was considered the training set accuracy. This bootstrapped training set accuracy was used for selecting tuning parameters.

Table C-4. List of 26 SML Models Screened.

Model	Abb.
Bagged MARS	bagEarth
C5.0	C5.0
Single C5.0 Ruleset	C5.0Rules
Single C5.0 Tree	C5.0Tree
Partial Least Squares	kernelpls
Weighted k-Nearest Neighbors	kknn
Linear Discriminant Analysis	lda
Linear Discriminant Analysis with Number of Discriminant Functions Tuning Parameter	lda2
Localized Linear Discriminant Analysis	loclda
Least Squares Support Vector Machine with Radial Basis Function Kernel	lssvmRadial
Mixture Discriminant Analysis	mda
Neural Networks with Feature Extraction	pcaNNet
Penalized Discriminant Analysis	pda
Penalized Linear Discriminant Analysis	PenalizedLDA
Partial Least Squares	pls
Random Forest	Rborist
Shrinkage Discriminant Analysis	sda
Partial Least Squares	simpls
Sparse Partial Least Squares	spls
Linear Discriminant Analysis with Stepwise Feature Selection	stepLDA
Support Vector Machines with Linear Kernel (kernlab package)	svmLinear
Support Vector Machines with Linear Kernel (e1071 package)	svmLinear2
Support Vector Machines with Radial Basis Function Kernel	svmRadial
Support Vector Machines with Radial Basis Function Kernel with Cost Tuning Parameter	svmRadialCost
Support Vector Machines with Radial Basis Function Kernel with Sigma and Cost Tuning Parameters	svmRadialSigma
Support Vector Machines with Radial Basis Function Kernel with Sigma, Cost, and Weight Tuning Parameters	svmRadialWeights

The three models selected for in-depth evaluation were trained and tested with thirty or more distinct seeds to check whether their performance was subject to random chance (see Appendix B for discussion on why this is a best practice). The three selected models were then trained and tested with the nitrate variable omitted, since this variable was not fully calibrated during Experiments 1 and 2. Next, the models were trained and tested with Yosemitech pH or ECD ORP data omitted, since these variables had missing data. Yosemitech pH had 416 and 1193 observations missing from the training and testing sets, respectively. ECD ORP had 341 and 306 observations missing from the training and testing sets, respectively. So, omitting these variables enabled greater samples sizes in both the training and testing sets. The three models were also trained and tested with full-scale wastewater influent flow as an eleventh variable.

C.2.6 Preprocessing

Four preprocessing methods were assessed to enhance model accuracy: (A) principal component analysis (PCA); (B) difference from the 24-hour rolling median; (C) difference from the training set median for each hour of the day (1:00 a.m. to 1:50 a.m., 2:00 to 2:50 a.m., etc.); and (D) B then C. PCA was conducted to promote diversity among the variables, considering that each principal component is perpendicular (non-correlated) with the others. PCA has previously been applied as a preprocessing technique for SML (Rodriguez, Kuncheva, and

Alonso 2006). The PCA model was constructed based on the training set and then the scores for each principal component were then also calculated on the testing set. The difference between each observation and the median of the past day (i.e., 144 observations at 10-minute intervals) was calculated to correct for drift due to fouling. The difference from the rolling median was provided to the models both in addition to and instead of the raw data, in case both the total value and the recent relative change of the variables were useful for event detection. The difference between each observation and the median for that hour of the day within the training set was calculated to correct for diurnal patterns.

C.2.7 Hyperparameter Tuning

Lastly, the three selected models, with their optimal variable selection and optimal preprocessing technique(s), were trained and tested with a wider array of settings for their hyperparameters (a.k.a. tuning parameters). Training and testing set accuracies were calculated for each combination of hyperparameter settings. This revealed whether selecting the hyperparameter settings based on the highest training set accuracy would have led to the hyperparameter settings with highest testing set accuracy. However, as per machine learning validation protocol, the testing set was used for evaluation purposes only, and was not used to inform the hyperparameter settings in the final model recommendation(s).

C.3 Results

C.3.1 Screening Results with Raw Data

Among the 26 models evaluated on raw data, the three with highest testing set accuracy were Neural Networks with Feature Extraction (pcaNNet), Bagged Multivariate Additive Regression Spline (bagEarth), and Single C5.0 Ruleset (C5.0Rules) (Table C-5). **So, these three models were selected for more in-depth evaluation.** The testing set accuracies of pcaNNet, bagEarth, and C5.0Rules using the raw data were 99.75 percent, 99.75 percent, and 99.68 percent, respectively. For all three models, there were zero false positives and at least one detection of each Spike type. So, even without preprocessing or hyperparameter tuning, these three models could be considered satisfactory and useful. The testing set accuracies of pcaNNet and bagEarth had p-values of 0.031 relative to the NIR and C5.0Rules had a p-value of 0.11. Based on the conventional criterion of p-value < 0.05, pcaNNet and bagEarth could be considered significantly better than the NIR. However, considering the number of models screened, it is conceivable the success of these models was in part due to random chance.

Table C-5. Screening Results for 26 Models with Default Tuning Parameters and Raw Data for Ten Water Quality Variables.

*Select for in-depth evaluation based on testing set accuracy.

Abb.	Training Set	Testing Set					
	Accuracy	Accuracy	p-value Acc. > NIR	Specificity	Total False Alerts	Total False Negatives	Mis-classifications
pcaNNet*	99.42%	99.75%	0.031	50%	0	7	0
bagEarth*	99.31%	99.75%	0.031	50%	0	7	0
C5.0Rules*	99.23%	99.68%	0.109	43%	0	8	1
C5.0	99.38%	99.64%	0.175	57%	2	6	2
lssvmRadial	99.35%	99.61%	0.259	21%	0	11	0
lda	98.51%	99.61%	0.259	43%	3	8	0
lda2	98.57%	99.61%	0.259	43%	0	8	3
pda	98.51%	99.61%	0.259	43%	3	8	0
Rborist	99.35%	99.57%	0.358	14%	0	12	0
PenalizedLDA	96.58%	99.54%	0.464	36%	4	9	0
svmRadial	99.30%	99.50%	0.570	0%	0	14	0
svmRadialSigma	99.30%	99.50%	0.570	0%	0	14	0
spls	99.29%	99.50%	0.570	0%	0	14	0
kernelpls	99.29%	99.50%	0.570	0%	0	14	0
svmRadialCost	99.30%	99.50%	0.570	0%	0	14	0
svmRadialWeights	NA	99.50%	0.570	0%	0	14	0
pls	99.29%	99.50%	0.570	0%	0	14	0
simpls	99.29%	99.50%	0.570	0%	0	14	0
stepLDA	99.23%	99.46%	0.670	0%	1	14	0
sda	99.03%	99.39%	0.828	0%	3	14	0
C5.0Tree	99.15%	99.39%	0.828	50%	8	7	2
kknn	99.33%	99.25%	0.971	50%	14	7	0
mda	98.75%	99.11%	0.997	43%	16	8	1
loclda	99.29%	99.00%	1.000	29%	18	10	0
svmLinear	99.48%	98.93%	1.000	50%	23	7	0
svmLinear2	99.48%	98.93%	1.000	50%	23	7	0

pcaNNet is a feed-forward, single-layer neural network with PCA preprocessing for feature extraction (Venables and Ripley 2002). pcaNNet has two hyperparameters: size and decay. Size is the number of units in the hidden layer. Decay is a regularization penalty applied to the sum of squares of the weights of the units in the hidden layer.

Open-source implementations of multivariate adaptive regression splines are abbreviated "Earth" because "MARS" has been trademarked. Earth uses weighted sums of basic functions, which can be constants, hinge functions, or products of hinge functions (Friedman 1991). Bagging is bootstrap aggregation, or taking the average of models trained on random samples of the data with replacement with the same sample size as the original data (Irizarry 2019).

bagEarth is the bagged version of Earth. bagEarth has two tuning parameters: degree and nprune. Degree is the maximum degree of interaction, i.e., the maximum number of hinge functions multiplied together. Nprune is the maximum number of terms in the Earth models.

C5.0Rules generates a single set of if-then rules using the C5.0 system (Kuhn et al. 2020). As implemented in the caret package, C5.0Rules does not having any tunable hyperparameters.

C.3.2 Checking for Randomness

Out of 100 trials, pcaNNet had a testing set accuracy of 99.75 percent 88 times and 99.79 percent 12 times. Thus, there was an element of randomness in the performance of this model, but its median accuracy was still equal or greater than the other models tested. Similarly, in 30 trials, bagEarth had a testing set accuracy ranging from 99.71 percent to 99.75 percent but would still have ranked first or second among the models tested on raw data. C5.0Rules testing set accuracy did not depend on the seed. So, multiple seed iterations were not tested in the remainder of this study.

C.3.3 Omitting and Adding Variables

Omitting s::can nitrate reduced testing set accuracy for pcaNNet, bagEarth, and C5.0Rules from 99.75 percent, 99.75 percent, and 99.68 percent respectively to 99.61 percent, 99.64 percent, and 99.57 percent respectively. So, it was concluded that the nitrate variable conveyed useful information for classification despite not being calibrated during Experiments 1 and 2.

Omitting Yosemitech pH reduced testing set accuracy for pcaNNet, bagEarth, and C5.0Rules from 99.75 percent, 99.75 percent, and 99.68 percent respectively to 98.20 percent, 99.53 percent, and 99.37 percent respectively. This loss of accuracy occurred even though the training set sample size increased from 2473 to 2836. Furthermore, all three models went from having zero false positives to at least one false positive. So, the Yosemitech pH was a useful enough variable to include in the models despite its relatively frequent data gaps.

Omitting ECD ORP reduced testing set accuracy for pcaNNet from 99.75 percent to 96.84 percent despite the training set sample size increasing from 2473 to 2756. However, for bagEarth, the testing set accuracy increased slightly from 99.75 percent to 99.76 percent, with the same number of true positives and false negatives but with more true negatives from the expanded sample size. For C5.0Rules, the testing set accuracy improved from 99.68 percent to 99.73 percent (or 99.71 percent on the original testing set) with zero false positives and two misclassifications. So, it was decided to omit ECD ORP from the training sets for bagEarth and C5.0Rules models but not pcaNNet. However, to ensure that comparisons among the models were fair, all three models were compared on the testing set observations for which data was available for all ten variables (n=2802) in the analyses below.

Including flow as an eleventh variable decreased pcaNNet, bagEarth, and C5.0Rules testing set accuracy from 99.75 percent, 99.75 percent, and 99.73 percent respectively to 99.71 percent, 99.75 percent, and 99.71 percent respectively.

C.3.4 Preprocessing

None of the preprocessing techniques evaluated improved the accuracy of the three models. PCA preprocessing reduced the bagEarth and C5.0Rules testing set accuracies from 99.75 percent and 99.71 percent to 99.50 percent and 99.61 percent, respectively. PCA preprocessing was not tested with pcaNNet because that model already automatically includes PCA preprocessing. An example of the other preprocessing techniques is demonstrated in Figure C-2 and Figure C-3 for the UV254 variable. Visually, it appears that these preprocessing techniques succeeded in reducing the variance of the data due to drift or diurnal patterns. Nevertheless. taking the difference from the rolling median of the previous 24-hours decreased pcaNNet, bagEarth, and C5.0Rules testing set accuracy from 99.75 percent, 99.75 percent, and 99.71 percent respectively to 99.14 percent, 99.00 percent, and 99.46 percent, respectively. pcaNNet was equally accurate after subtracting the hourly median from each variable; however, BagEarth and C5.0Rules testing set accuracy were reduced from 99.75 percent and 99.71 percent respectively to 99.71 percent and 99.50 percent respectively. After subtracting both the 24-hour rolling median and the median for each hour of the day, pcaNNet, bagEarth, and C5.0Rule testing set accuracy decreased from 99.75 percent, 99.75 percent, and 99.71 percent respectively to 99.50 percent, 99.50 percent, and 99.54 percent respectively.

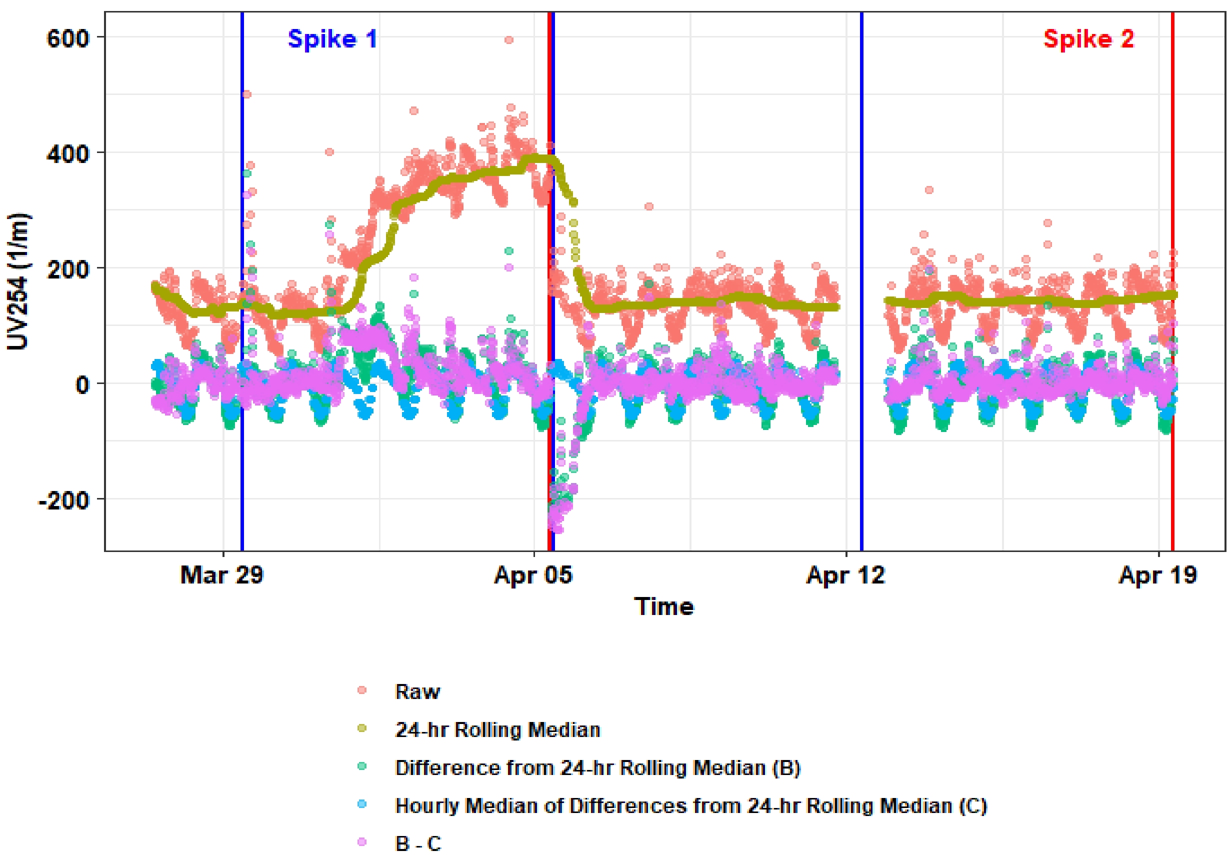


Figure C-2. Raw Data (red); the 24-hr Rolling Median (yellow); Difference Between Raw Data and the 24-hr Rolling Median (green); Hourly Median of Differences Between the Raw Data and the 24-hr Rolling Median (blue); and the Difference Between the Difference from the 24-hr Rolling Median and Its Hourly Medians (purple) for s::can UV254 (1/m) in the training set.

Blue and red solid vertical lines indicate Spike 1 and Spike 2 events, respectively.

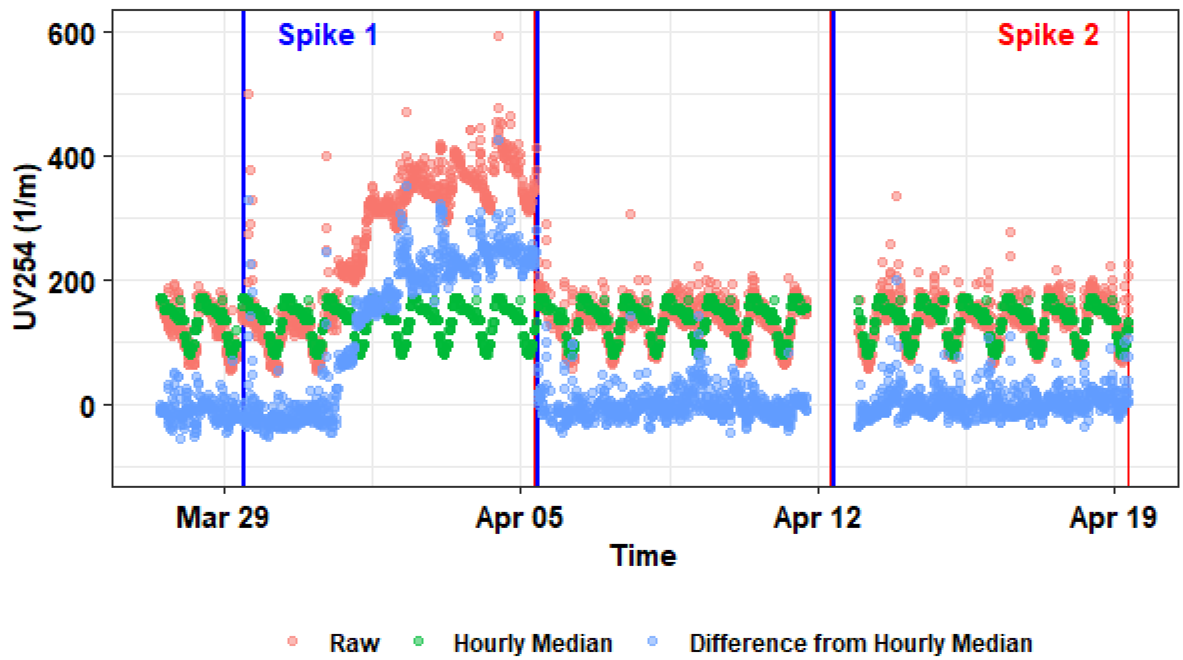


Figure C-3. Raw Data (red); the Median for Each Hour of the Day (green).
The difference from the median of each hour of the day (blue) for s::can UV254 (1/m) in the training set. Blue and red solid vertical lines indicate Spike 1 and Spike 2 events, respectively.

C.3.5 Hyperparameter Tuning

bagEarth had higher testing set accuracy with degree equal to one but higher training set accuracy with degree equal to two, indicating that higher degree led to overfitting for this model and dataset (Figure C4). bagEarth had highest testing set accuracy with degree equal to one and nprune equal or greater to 9, which includes the default settings used above.

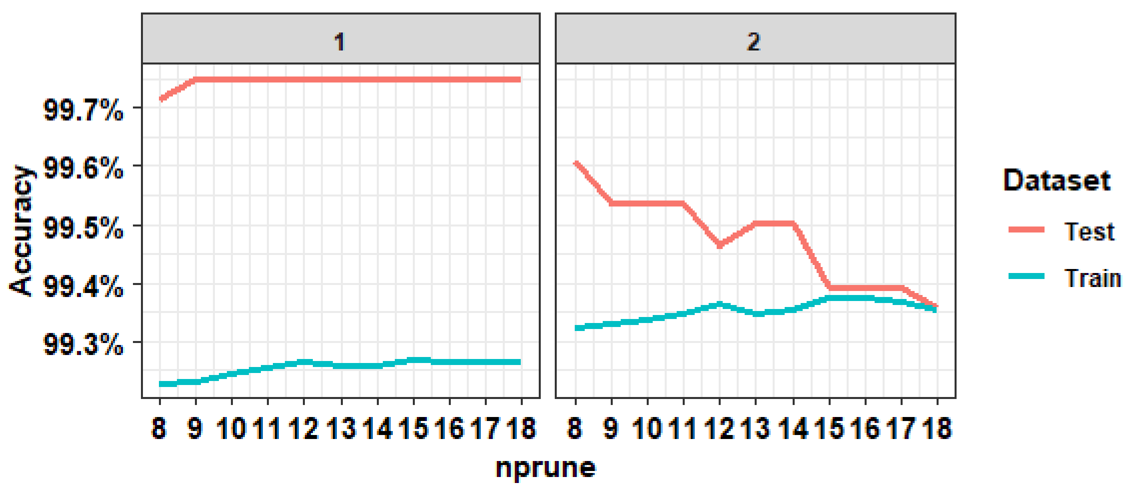


Figure C-4. Train and Testing Set Accuracy of bagEarth with Degree 1 or 2 and nprune from 8 to 18.
Panel heading indicates degree.

Ranging size from one to ten and decay from 0.1×2^{-4} to 0.1×2^{6}, pcaNNet train and testing set accuracies were both generally highest with size equal to or greater than four (Figure C-5).

However, training set accuracy was highest with decay of 0.00625, but testing set accuracy was highest with decay equal to 0.1. The highest pcaNNet testing set accuracy was 99.79 percent with decay equal to 0.1 and size in the range four to ten, or decay equal to four and decay equal to 0.05. This testing set accuracy was the highest yet in this study with seed set to one. However, pcaNNet with default tuning (size=5 and decay=0.1) had also achieved this testing set accuracy in 12 percent of replicates with different seeds. So, it was plausible that these tune settings had the highest testing set accuracies due in part to random chance.

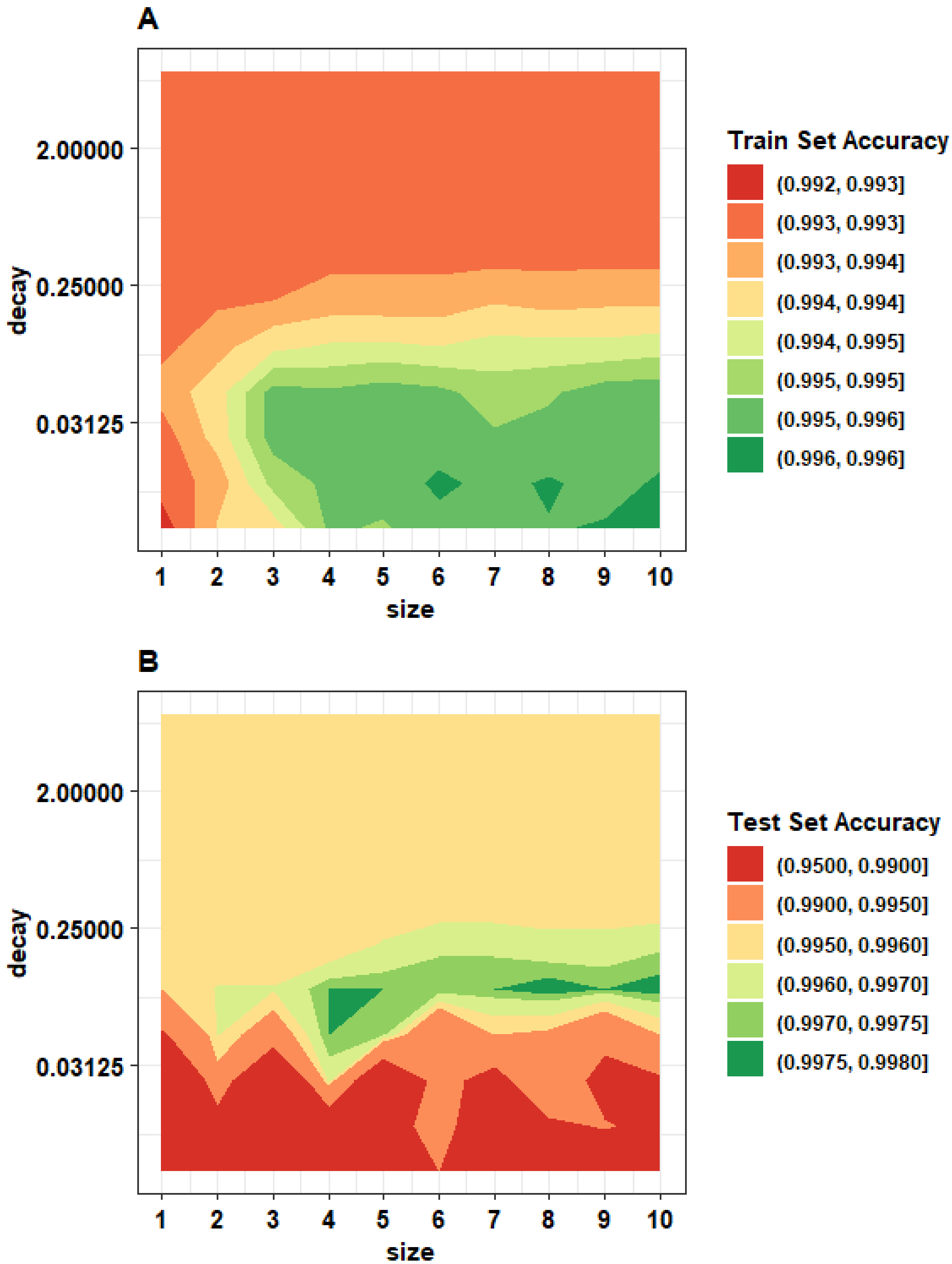


Figure C-5. pcaNNet (A) Train and (B) Testing Set Accuracies with Size from 1 to 10 and Decay from 0.1×2^{-4} to 0.1×2^{6}.

Seed was set to one. The decay axis is $\log_2$ transformed.

So, this pcaNNet tuning was repeated with 30 replications with 30 distinct seeds. This procedure produced a smoother contour plot (Figure C-6). It also produced a unique maximum testing set accuracy: 99.79 percent with size of 10 and decay of 0.1. This testing set accuracy was the highest in this study. However, it should be noted that these hyperparameter tune settings would not have been selected based on the training set bootstrapping cross-validation procedure with a large array of potential tune settings. Higher decays are associated with underfitting and lower decays are associated with overfitting (Venables and Ripley 2002). Decays in the range 0.0001 to 0.1 are generally recommended, and both the training set accuracy optimal decay and the testing set accuracy optimal decay would fall in this range (Venables and Ripley 2002). Nevertheless, for this particular combination of model and dataset, it appears that the bootstrapping cross-validation procedure would not recommend the optimal range of decay for truly new timeseries data. Since standard machine learning procedure is to use testing sets for evaluation purposes only and not hyperparameter selection, pcaNNet with size of 10 and decay of 0.1 should not be recommended for implementation at CWS without further validation and testing.

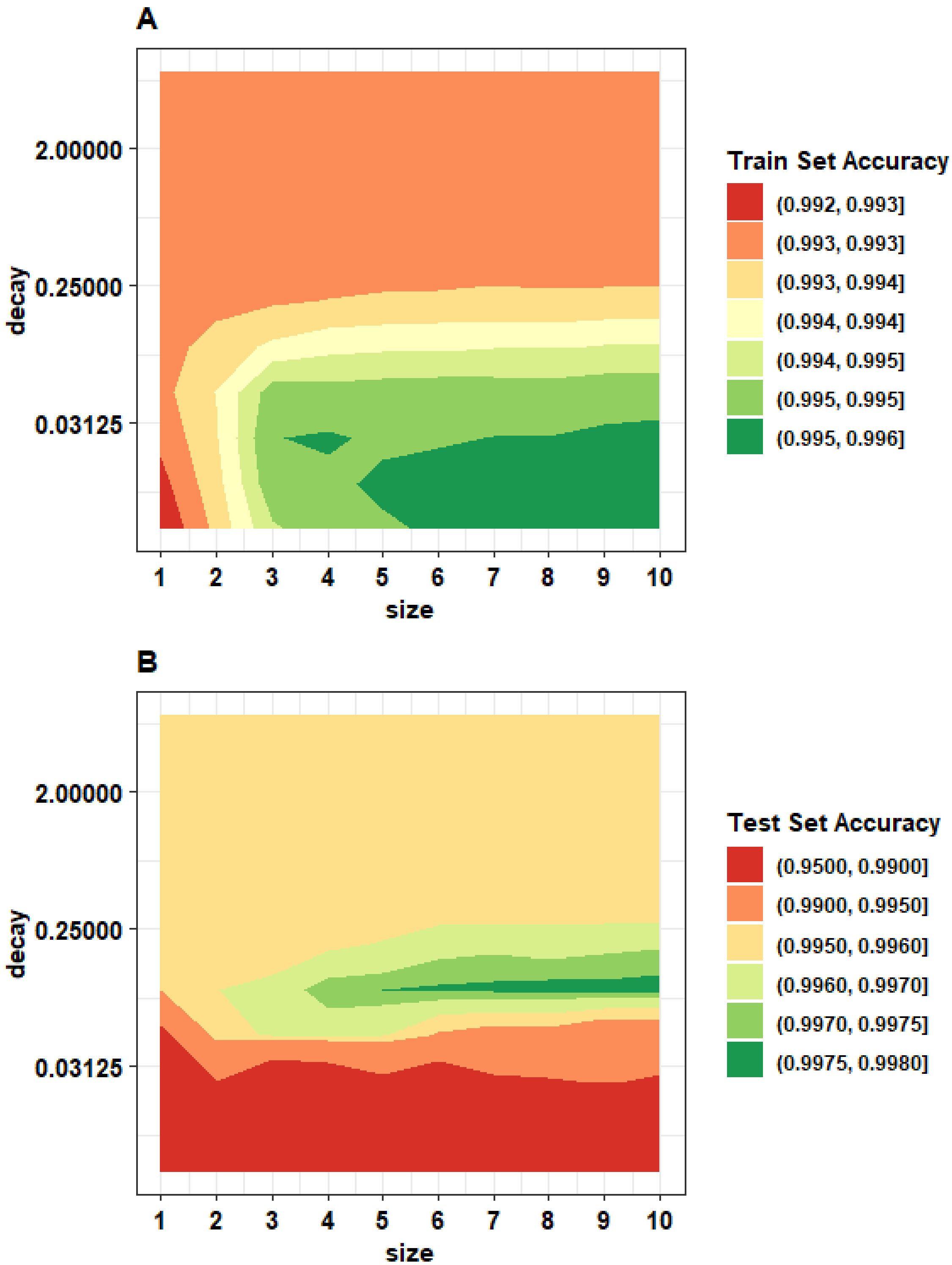


Figure C-6. pcaNNet (A) Train and (B) Testing Set Accuracies with Size from 1 to 10 and Decay from $0.1{\times}2^{-4}$ to $0.1{\times}2^{6}$.

Accuracy is mean of 30 distinct seeds. The decay axis is $\log_2$ transformed.

C.4 Discussion

As discussed in the HRSD case study (Appendix B), accuracy is arguably a misleading metric of success. Models tested on unbalanced data could achieve high-sounding accuracy simply by usually predicting the most common class. This is especially true for the CWS data used in this section, since the percentage of Normal data in the testing set was 99.5 percent, as opposed to 76.7 percent in the HRSD dataset. pcaNNet and bagEarth had testing set balanced accuracy of 66.7 percent (100 percent accuracy among Normal datapoints and 50 percent accuracy among both spike types) (Table C-6). Treating the two spike types as a single class would increase the balanced accuracy to 75 percent, still below the balanced accuracies of the optimized models for the HRSD dataset. C5.0Rules had a slightly lower balanced accuracy of 62.5 percent.

Table C-6. Performance Metrics of the Three Models Selected for In-Depth Evaluation.

Abb.		pcaNNet	bagEarth	C5.0Rules
Variables Excluded		None	ECD ORP	ECD ORP
Training Set	Accuracy	99.42%	99.31%	99.23%
	Training Time	34	480	4
Testing Set	Accuracy	99.75%	99.75%	99.68%
	Balanced Accuracy	66.70%	66.70%	62.50%
	Cohen's Kappa	0.666	0.666	0.635
	False Positives	0	0	0
	False Negatives	7	7	6
	Misclassifications	0	0	2

Another alternative to accuracy is Cohen's Kappa. The formula for Cohen's Kappa with three or more classes is:

Equation C-1. Adapted from Chicco et al. (2021).

$$K = \frac{c \times s - \sum_k^K p_k \times t_k}{s^2 - \sum_k^K p_k \times t_k}$$

Where *c* is the total number of observations correctly predicted, *s* is the total number of observations, p_k is the number of times class *k* was predicted, and t_k is the number of times class *k* truly occurs. The three optimized models had Cohen's Kappas ranging 0.635-0.666, which could be considered good or substantial (Table C-6) (Landis and Koch 1977; Fleiss 1981).

Not all errors in wastewater influent monitoring are equally important. Considering the high frequency of data collection, a very low false positive rate is required. For data analyzed every ten minutes, a 1 percent false positive rate would equate to false alerts more than once per day; 0.1 percent would mean false alerts approximately weekly. Too many false alerts would initially lead to poor allocation of resources and eventually to complacency or discontinued use of the model. Furthermore, the spikes simulated in this study could be considered non-extreme, i.e., would not exceed the equalization and treatment capability of the WWTP. More extreme industrial events would eventually be noticed by human monitoring or fixed thresholds. Furthermore, even one true positive within the sequential set of two to five 10-minute observations within each spiking event could be considered a true detection of the spiking event as a whole. So, a few false negatives within datasets of the size used here could be relatively acceptable. Considering this contextual information, false positives would be considered a worse error type than false negatives. Misclassifications of Spike 1 as Spike 2 or vice versa would be the most minor type of error since either prediction would instigate human investigation and corrective action.

Considering the above, a weight of 4 was applied to each false positive, 2 to each false negative, and 1 to each spike misclassification. bagEarth and pcaNNet both had 7 false negatives, while C5.0Rules had 6 false negatives and 2 misclassifications (Table C-6). With this weighting scheme, all three models would have a score of 14 (or 0.0050 normalized by dividing by the testing set sample size). So, all three models could be considered equally good on this testing set, despite the slightly lower unweighted accuracy for C5.0Rules.

The importance of errors for classification of timeseries data depends not only on the types of errors but also the locations of errors. For example, at least one detection within each consecutive spiking event would be more useful than complete detection of one spiking event but all false negatives for the others, even if the overall accuracy were the same. However, the pcaNNet, bagEarth, and C5.0Rules were equivalent by this criterion too with at least one true positive in three out of five consecutive spike events with available data in the testing set (Figure C-7 and Figure C-8). In fact, pcaNNet (with all water quality variables) and bagEarth (omitting ECD ORP) were exactly equivalent in terms of their errors and true predictions.

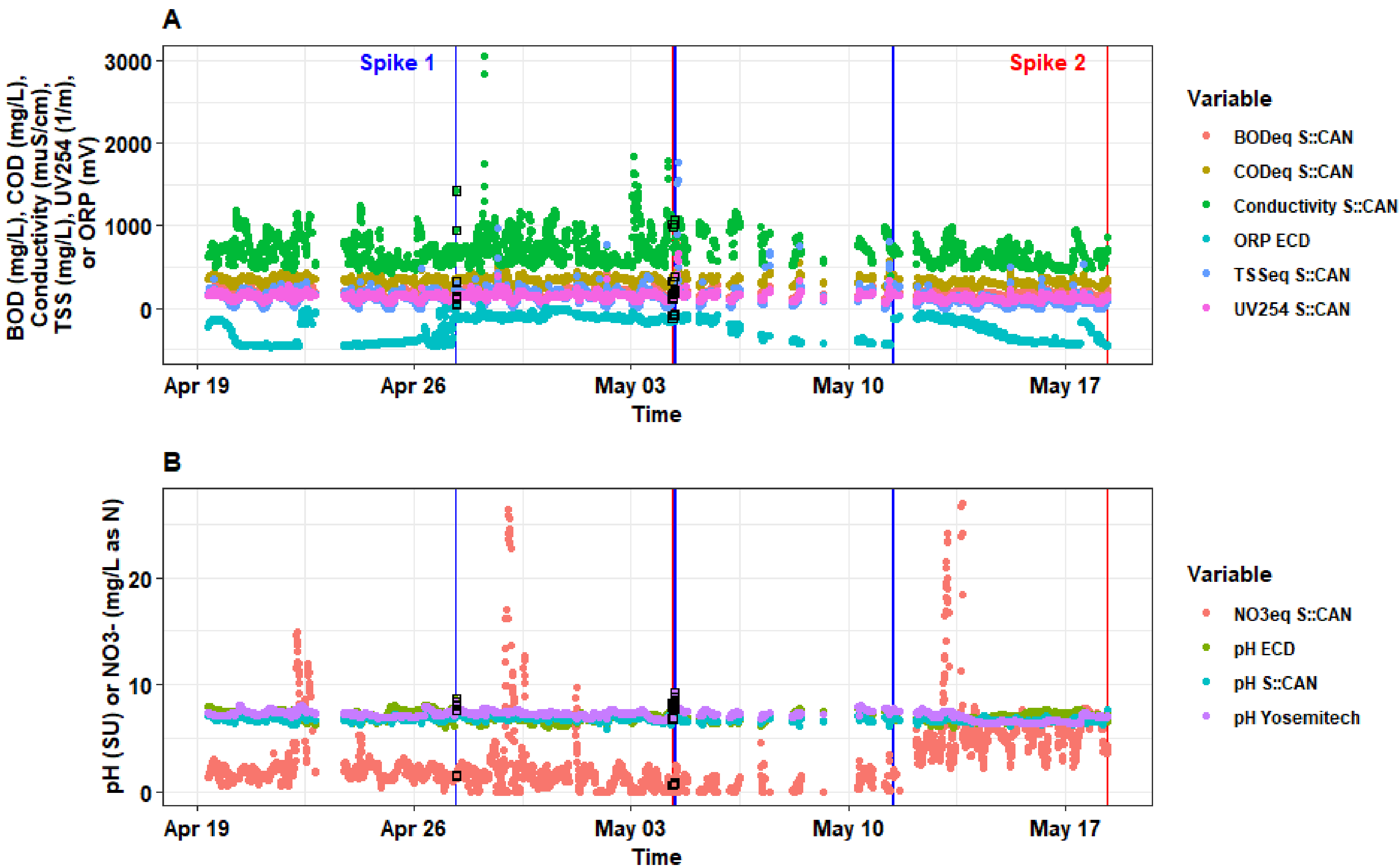


Figure C-7. Testing Set Results for pcaNNet with All Water Quality Variables or bagEarth Omitting ECD ORP, Which Produced the Same Predictions. Black squares indicate true positives. (A) variables with mean absolute values over 100 and (B) variables with mean absolute values below 100. Blue and red solid vertical lines indicate Spike 1 and Spike 2 events, respectively.

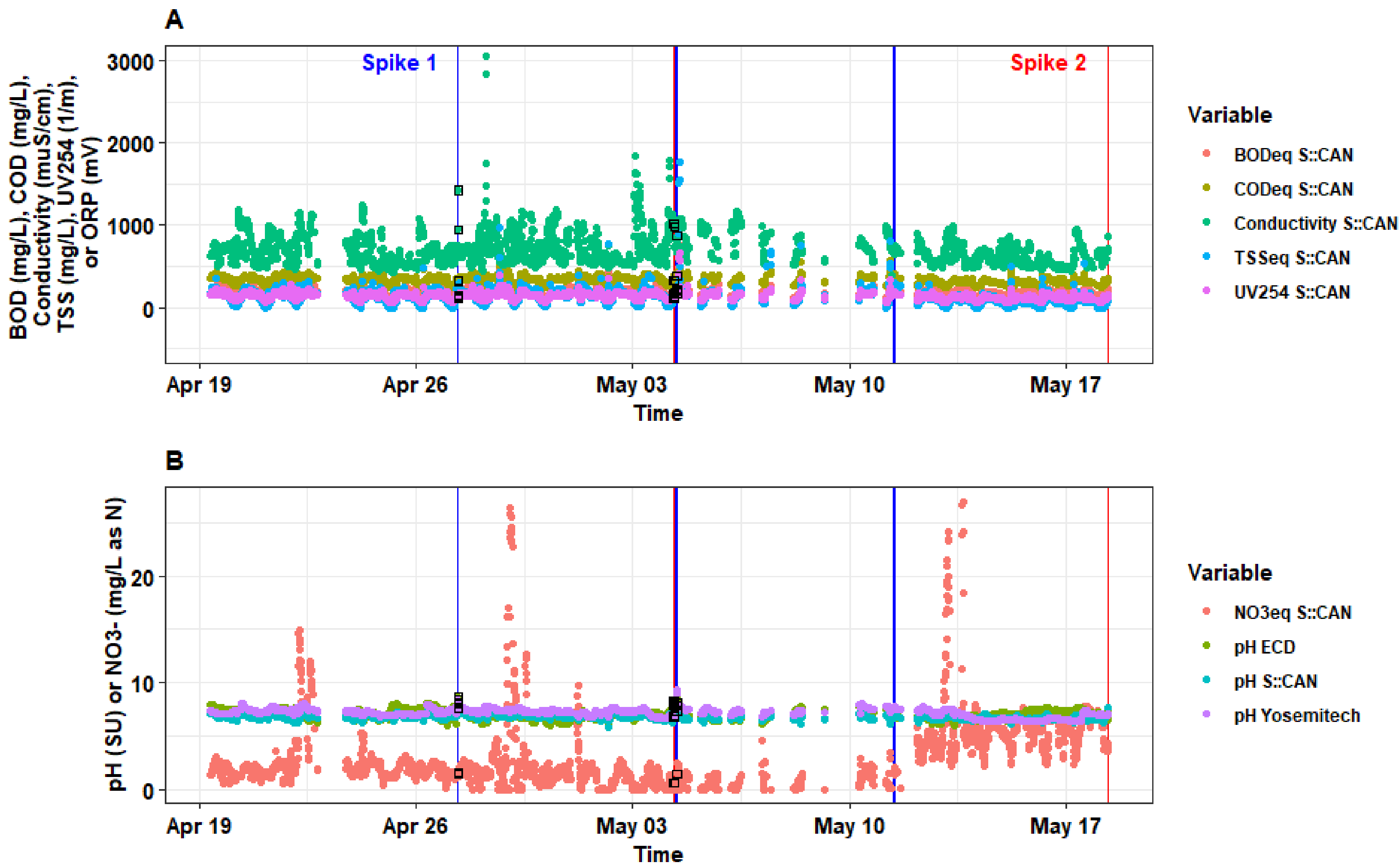


Figure C-8. Testing Set Results for C5.0Rules with All Water Quality Variables.

Black squares indicate true positives. (A) variables with mean absolute values over 100 and (B) variables with mean absolute values below 100. Blue and red solid vertical lines indicate Spike 1 and Spike 2 events, respectively.

Another potential criterion for SML model selection is training time. A quicker training time enables training and testing on more iterations (e.g. hyperparameter tunings or random replicates), facilitating ongoing optimization. Training time shorter than the data recording interval would be ideal since this would enable the model to be updated after each new datapoint in real-time. C5.0Rules, pcaNNet, and bagEarth training required 4 seconds, 34 seconds, and 480 seconds, respectively, on a laptop with 11th Gen Intel(R) Core(TM) i7-1185G7 @ 3.00GHz and 32 GB of RAM. In practice, these training times could increase over time with increasing training set sample size or the addition of new water quality sensors. So, bagEarth could be considered borderline too slow, while C5.0Rules and pcaNNet would be sufficiently fast.

Interpretability is another potential criterion for SML model selection, especially in fields such as drinking water and wastewater treatment where clear communication with regulators and the public is a priority. A neural network model is "black box" by nature. In contrast, C5.0Rules generates a single, interpretable ruleset. In this case, C5.0Rules generated five rules using just three variables (ECD pH, s::can pH, and s::can nitrate) (Table C-7). An additional benefit of this model in this case would be that the Yosemitech pH variable was not used and so could be discontinued, saving operations and maintenance cost.

Table C-7. Rules for C5.0Rules using raw data for water quality variables excluding ECD ORP.

Rule	Output
s::can NO_3 eq > 1.92	Normal
s::can pH ≤ 7.21	Normal
ECD pH ≤ 7.85	Normal
s::can NO_3 eq > 1.62 and s::can NO_3 eq ≤ 1.92 and ECD pH > 7.85 and s::can pH > 7.21	Spike 1
s::can NO_3 eq ≤ 1.16 and ECD pH > 7.85 and s::can pH > 7.21	Spike 2

It could be considered surprising that the rule for Spike 2 was based in part on nitrate, since Spike 2 was addition of humic acid, not nitrate. However, considering the nitrate variable is an absorbance-based estimate, it appears the humic acid interfered with the absorbance wavelengths used to estimate nitrate. Counterintuitively, pH increased during the Spike 2 humic acid addition. Possibly, the humic acid solution had higher pH than the wastewater influent due to higher pH in the tap water used for the dilution or the presence of a buffer in the humic acid standard.

It is notable that in this study, testing set accuracies were almost always higher than training set accuracies. Such a trend is unusual, though not impossible. In this case, it may have been due to the major fouling incident occurred on the s::can spectro::lyser instrument (which measured BOD, COD, TSS, UV254, and nitrate) that occurred during Experiment 1 in the training set (Figure C-1). Such a drastic fouling event did not occur during the testing set (Figure C-7).

C.5 Conclusions

- pcaNNet, bagEarth, and C5.0Rules were the most accurate models using raw data for ten water quality variables with testing set accuracies of 99.75 percent, 99.75 percent, and 99.68 percent, respectively.
- pcaNNet, bagEarth, and C5.0Rules had zero false positives and at least one detection of both spike types in the testing set.
- Omitting ECD ORP to increase the effective sample size of the other variables did not decrease bagEarth accuracy and increased C5.0Rules testing set accuracy to 99.71 percent.
- Preprocessing with PCA or median-based methods to remove diurnal patterns and drift did not increase the testing set accuracy of pcaNNet, bagEarth, or C5.0Rules.
- Considering the relative importance of types and locations of errors, pcaNNet, bagEarth, and C5.0Rules were equally successful.
- However, considering training time and interpretability, the team would recommend C5.0Rules for this application.

References

Atla, A., R. Tada, V. Sheng, and N. Singireddy. 2011. “Sensitivity of Different Machine Learning Algorithms to Noise.” *Journal of Computing Sciences in Colleges,* 26 (5): 96–103.

Broeke, J. V. D. 2014. *Compendium of Sensors and Monitors and Their Use in the Global Water Industry*. Project 4428. Denver, CO: The Water Research Foundation.

Chicco, D., M. J. Warrens, and G. Jurman. 2021. “The Matthews Correlation Coefficient (MCC) Is More Informative than Cohen’s Kappa and Brier Score in Binary Classification Assessment.” *IEEE Access,* 9: 78368–81. https://doi.org/10.1109/ACCESS.2021.3084050.

Clopper, C. J., and Pearson, E. S. 1934. "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial." *Biometrika,* 26: 404–413. https://doi.org/https://doi.org/10.2307/2331986

Code of Federal Regulations (CFR). 2023a. “Title 40, Chapter I, Subchapter N, Part 469.” https://www.ecfr.gov/current/title-40/chapter-I/subchapter-N/part-469

Code of Federal Regulations (CFR). 2023b. "Title 40, Chapter I, Subchapter N, Part 437." https://www.ecfr.gov/current/title-40/chapter-I/subchapter-N/part-437

Cohen, J. 1960. “A Coefficient of Agreement for Nominal Scales.” *Educational and Psychological Measurement,* 20 (1): 37–46.

City of Morro Bay. 2020. Morro Bay. City of Morro Bay. Accessed. March 2, 2023. https://morrobaywrf.com/).

Deng, H., G. Runger, E. Tuv, and W. Bannister. 2014. “CBC: An Associative Classifier with a Small Number of Rules.” *Decision Support Systems,* 59 (1): 163–70. https://doi.org/10.1016/j.dss.2013.11.004.

Drewes, J., P. Anderson, N. Denslow, W. Jakubowski, A. Olivieri, D. Schlenk, and S. Snyder. 2018. *Monitoring Strategies for Constituents of Emerging Concern (CECs) in Recycled Water: Recommendations of a Science Advisory Panel (SCCWRF Technical Report 1032)*. Sacramento, CA, USA: State Water Resources Control Board.

Faiz, F., G. Baxter, S. Collins, F. Sidiroglou, and M. Cran. 2020. “Polyvinylidene Fluoride Coated Optical Fibre for Detecting Perfluorinated Chemicals.” *Sensors and Actuators B: Chemical,* 312 (January). https://doi.org/10.1016/j.snb.2020.128006.

Fleiss, J. L. 1981. *Statistical Methods for Rates and Proportions*. 2nd ed. New York City, New York, USA: John Wiley.

Friedman, J. H. 1991. “Multivariate Adaptive Regression Splines.” *Annals of Statistics,* 19 (1): 1–141.

Fujioka, T., T. Tanisue, S. L. Roback, M. H. Plumlee, K. P. Ishida, and H. Kodamatani. 2017. “Near Real-Time N-Nitrosodimethylamine Monitoring in Potable Water Reuse via Online High-Performance Liquid Chromatography-Photochemical Reaction-Chemiluminescence.”

Environmental Science: Water Research & Technology, 3: 1032–36. https://doi.org/10.1039/C7EW00296C.

Gonzalez, D., K. Thompson, O. Quiñones, E. Dickenson, and C. Bott. 2021. "Granular Activated Carbon-Based Treatment and Mobility of per- and Polyfluoroalkyl Substances in Potable Reuse for Aquifer Recharge." *AWWA Water Science,* 3 (5): e1247. https://doi.org/10.1002/aws2.1247.

Irizarry, R. A. 2019. "Data Analysis and Prediction Algorithms with R." Chapman and Hall/CRC. 2019.

Kadiyala, R., and C. Macintosh. 2018. *Leveraging Other Industries - Big Data Management (Phase I)* Project 4836. Denver, CO: The Water Research Foundation.

Klanderman, M. C., K. B. Newhart, T. Y. Cath, and A. S. Hering. 2020. "Fault Isolation for a Complex Decentralized Waste Water Treatment Facility." *Applied Statistics,* 69 (4): 931–51.

Kodamatani, H., S. L. Roback, M. H. Plumlee, K. P. Ishida, H. Masunaga, N. Maruyama, and T. Fujioka. 2018. "An Inline Ion-Exchange System in a Chemiluminescence-Based Analyzer for Direct Analysis of N -Nitrosamines in Treated Wastewater." *Journal of Chromatography A,* 1553: 51–56.

Kuhn, M. 2008. "Building Predictive Models in R Using the Caret Package." *Journal of Statistical Software,* 28 (5): 1–26. https://doi.org/10.18637/jss.v028.i05.

———. 2019. "The Caret Package. Chapter 6: Available Models." Github. 2019. https://topepo.github.io/caret/available-models.html.

Kuhn, M., S. Weston, M. Culp, N. Coulter, and R. Quinlan. 2020. "C5.0 Decision Trees and Rule-Based Models." CRAN. 2020.

Landis, J. R., and G. G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics,* 33 (1): 159–74.

Law, C. S., J. Wang, S. Gunenthiran, S. Y. Lim, A. D. Abell, L. Ahrens, T. Kumeria, A. Santos, and N. H. Voelcker. 2021. "Real-Time Detection of Per-Fluoroalkyl Substance (PFAS) Self-Assembled Monolayers in Nanoporous Interferometers." *Sensors and Actuators B: Chemical,* 355 (October 2021): 131340. https://doi.org/10.1016/j.snb.2021.131340.

Liggett, J., C. Macintosh, and K. Thompson. 2018. *Designing Sensor Networks and Locations on an Urban Sewershed Scale*. Project 4835. Denver, CO: The Water Research Foundation.

Lindgren, F., P. Geladi, and S. Wold. 1993. "The Kernel Algorithm for PLS." *Journal of Chemometrics,* 7 (1): 45–59.

Livaniou, I., L. Heinrich, A. Michalski, T. Motchan, S. Schuster, T. Stone, and B. Uku. 2020. *Water Quality Sensors Global Horizon Scan*. London, UK: Isle Utilities.

Matsumoto, M., and T. Nishimura. 1998. "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator." *ACM Transactions on Modeling and Computer Simulation,* 8 (1): 3–30. https://doi.org/10.1145/272991.272995.

Menze, B. H., B. M. Kelm, D. N. Splitthoff, U. Koethe, and F. A. Hamprecht. 2011. "On Oblique

Random Forests." In *Machine Learning and Knowledge Discovery in Databases*. Edited by D. Gunopulos, T. Hofmann, D. Malerba, and M. Vazirgiannis. Part II, 453–69. Athens, Greece: Springer.

Nading, T., E.Dickenson, A. Salveson, A. Branch, and L. Schimmoller. 2022. *An Enhanced Source Control Framework for Industrial Contaminants in Potable Reuse*. Alexandria, VA: The Water Research Foundation.

Neemann, J., J. DeCarolis, D. ten Bosch, S. Snyder, I. Pepper, and M. Park. 2019. *Integrated Management of Sensor Data for Real Time Decision Making.* Project 4759. Denver, CO: The Water Research Foundation.

Nolan, J. R. 2002. "Computer Systems That Learn: An Empirical Study of the Effect of Noise on the Performance of Three Classification Methods." *Expert Systems with Applications,* 23 (1): 39–47.

NWRI (National Water Research Institute). 2020. *Independent Advisory Panel Report: METI Report on the Review of Yokogawa Electric Corporation's Rapid Assessment Pathogen Identification (RAPID) Technology for Potable Water Reuse Applications*. Fountain Valley, CA, USA: National Water Research Institute.

Park, M. Y., and T. Hastie. 2008. "Penalized Logistic Regression for Detecting Gene Interactions." *Biostatistics,* 9 (1): 30–50. https://doi.org/10.1093/biostatistics/kxm010.

Peng, J., C. Chen, M. Zhou, X. Xiaohua, Y. Zhou, and C.-H. Luo. 2020. "A Machine-Learning Approach to Forecast Aggravation Risk in Patients with Acute Exacerbation of Chronic Obstructive Pulmonary Disease with Clinical Indicators." *Scientific Reports,* 10: 3118. https://doi.org/10.1038/s41598-020-60042-1.

Pepper, I. L., and S. A. Snyder. 2016. *Monitoring for Reliability and Process Control of Potable Reuse Applications*. Alexandria, VA: Water Environment & Reuse Foundation.

Poona, N., A. van Niekerk, and R. Ismail. 2016. "Investigating the Utility of Oblique Tree-Based Ensembles for the Classification of Hyperspectral Data." *Sensors (Switzerland),* 16 (11). https://doi.org/10.3390/s16111918.

Rännar, S., F. Lindgren, P. Geladi, and S. Wold. 1994. "A PLS Kernel Algorithm for Data Sets with Many Variables and Fewer Objects. Part 1: Theory and Algorithm." *Journal of Chemometrics,* 8 (2): 111–25.

Roback, S. L., H. Kodamatani, T. Fujioka, and M. H. Plumlee. 2020. "Validation of a Novel Direct-Injection Chemiluminescence-Based Method for Recycled Water, Drinking Water, and Wastewater." *Environmental Science: Water Research & Technology,* 6 (4): 1106–15. https://doi.org/10.1039/C9EW00943D.

Rodriguez, J. J., L. I. Kuncheva, and C. J. Alonso. 2006. "Rotation Forest: A New Classifier Ensemble Method." *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 28 (10): 1619–30.

Schölkopf, B., K. K. Sung, C. J. C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik. 1997.

“Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers.” *IEEE Transactions on Signal Processing*, 45 (11): 2758–65. https://doi.org/10.1109/78.650102.

Steinle-Darling, E., P. Carlo, A. Salveson, G. Dorrington, and N. Nye. 2020. *Demonstrating Real-Time Collection System Monitoring for Potable Reuse.* Project 4908. Alexandria, VA: The Water Research Foundation.

SWRCB. 2021. *DPR Framework 2nd Edition Addendum - Early Draft of Anticipated Criteria for Direct Potable Reuse*. Sacramento, CA: California Water Boards.

Thompson, K. in process. *Designing Sensor Networks and Locations on an Urban Sewershed Scale with Big Data Management and Analytics.* Project 4797. Denver, CO: The Water Research Foundation. https://www.waterrf.org/research/projects/designing-sensor-networks-and-locations-urban-sewershed-scale-big-data-management.

Thompson, K. A., and E. R. V. Dickenson. 2021. “Using Machine Learning Classification to Detect Simulated Increases of de Facto Reuse and Urban Stormwater Surges in Surface Water.” *Water Research,* 204: 117556. https://doi.org/10.1016/j.watres.2021.117556.

USEPA (US Environmental Protection Agency). 2021a. National Pretreatment Program. US Environmental Protection Agency. https://www.epa.gov/npdes/national-pretreatment-program.

USEPA. 2021b. “Water Quality Surveillance and Response.” US Environmental Protection Agency. 2021. https://www.epa.gov/waterqualitysurveillance.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. New York, NY, USA: Springer Science. https://doi.org/10.1007/978-1-4757-2719-7_14.

Wang, J., I. Abusallout, R. Marfil-Vega, and D. Hanigan. 2021. “Quantification of Per- and Polyfluoroalkyl Substances with a Modified Total Organic Carbon Analyzer and Ion Chromatography.” *AWWA Water Science,* 3: e1235. https://doi.org/10.1002/aws2.1235.

Wang, Z.. 2018. “Robust Boosting with Truncated Loss Functions.” *Electronic Journal of Statistics,* 12 (1): 599–650. https://doi.org/10.1214/18-EJS1404.

Westerhoff, P., N. Sharma, C. Zeng, T. Karanfil, D. Kim, A. Ghosh, C. Seidel, C. Samson, and A. Eaton. 2022. *Occurrence Survey of Bromide and Iodide in Water Supplies.* Project 4711. Denver, CO: The Water Research Foundation.

THE
Water
Research
FOUNDATION

advancing the science of water®

1199 North Fairfax Street, Suite 900
Alexandria, VA 22314-1445 | 6666 West Quincy Avenue
Denver, CO 80235-3098

www.waterrf.org | info@waterrf.org