Characterization of a Continuous, In-situ, Near-Time Fluorescence Sensor for the Detection of Fecal Contamination in Drinking Water

by

Olivia August Harmon

B.S., The University of Alabama, 2020

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Department of Civil, Environmental and Architectural Engineering

2022

Committee Members:

Evan Thomas, Chair

Julie Korak

Karl Linden

Harmon, Olivia August (M.S., Environmental Engineering)

Characterization of a Continuous, In-situ, Near-Time Fluorescence Sensor for the Detection of Fecal Contamination in Drinking Water

Thesis directed by Prof. Evan Thomas

A highly sensitive, continuous, in-situ, remotely reporting fluorescence sensor coupled with a machine learning model to predict high-risk fecal contamination was evaluated. The sensor behavior was characterized to multiple fluorescence quenching parameters through lab bench top analysis. The sensor's minimum detection limit of powdered tryptophan dissolved in deionized water was found to be 0.05 ppb, its minimum detection limit of *E. coli* present in wastewater effluent was 10 CFU/100 mL, and its minimum detection limit of lab grown *E. coli* in DI water was 33 CFU/100 mL. A correction factor was calculated to correct for the decline in fluorescence response to water temperature. Inner filter effects were shown to have negligible impact in an operational context. It was established that the fluorescence signal increases by approximately 82% with the formation of biofilm, while the sensitivity of the sensor is reduced by approximately 5% with the formation of mineral scaling. Four sensors were installed on Boulder Creek in Colorado and over the course of 88 days 298 ground truth samples were enumerated for *E. coli* through membrane filtration. This data built the training and validation data set for a machine learning model. The performance of this model was improved by incorporating a proxy feature for biofouling that was based on time since cleaning the cuvette. This model has the ability to predict high risk fecal contamination with 83% accuracy, has a sensitivity (true positive rate) of 80%, specificity (true negative rate) of 86%, and can distinguish between all risk categories established by the World Health Organization 64% accuracy. This sensor, combined with the highly skilled machine learning model, has the ability to provide a more consistent informative data set about fecal contamination risk in drinking water to water service providers and individual consumers.

Dedication

For my family, I wouldn't be the person I am today without your constant encouragement, guidance, and love.

To my parents who showed me how to have a career that can make a real difference in people's lives. Mom, for always being there to listen to me talk through my worries, accomplishments, and everything in between. Even while you were teaching, tutoring, and getting your own Master's Degree, you always made the time to support me. Thank you for reading over all of my papers (including this one) for grammar and spelling mistakes. Dad, for encouraging me to cultivate my love of math and science and for pushing me to challenge my comfort zone. Thank you for sharing with me your love of the outdoors and 90s hip-hop (I wouldn't have made it through this without A Tribe Called Quest).

To my siblings, for fully understanding why I am the way I am. My sister Caroline, for teaching me to stand up for myself and love others fiercely. My brother Noah, for teaching me to be kind and look on the bright side of things. My brother Eli, for showing me that it is inspiring to be unironically enthusiastic about the things that you love. My sister Eve, for teaching me that love is unconditional and it is okay to treat every event as your own personal performance.

Acknowledgements

First, I want to thank Dr. Emily Bedell. I have learned more from you than I can put into words, and I am so grateful to have you as a role model and friend. It has been such an honor to work with you on this project. I know it is a cliche, but I TRULY could not have done any of this without you.

I would like to thank all of the teachers and professors I had who shared their wisdom and supported me from the very beginning. Mrs. Elizabeth Mayfield Pruitt, thank you for teaching me how to see the good in everyone. Mr. Tim Hurry, thank you for challenging me and showing me how creative math can be. Mrs. Kelly Reeves, thank you for teaching me that it is good to push yourself, but your health and happiness should take top priority. Mrs. Michele Cooley, thank you for teaching me how to be a writer and citizen scholar. Mrs. Jennifer Ayers, Mr. Terrance Cobb, and Mr. Ron Pence, thank you for teaching me confidence and how to be an effective leader. Dr. Jair Lizarazo-Adarme, thank you for not only encouraging me to pursue higher education but teaching me how to be the best person and engineer I can be. Dr. Laura MacDonald, thank you for supporting me from the very beginning of this journey, you are an incredible researcher, friend, and person. I would also like to thank my thesis committee, Dr. Julie Korak, Dr. Karl Linden, and Dr. Evan Thomas, for their roles in my academic development.

Arthur Ginley, thank you for providing me with a space where I feel listened to and valued. You're support and advice has been invaluable.

Finally I want to thank all of my friends. You have surrounded me with so much love and support that no amount of thanks will ever be enough.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Background Information

Reliable and affordable access to safe drinking water is essential to human health. However, it is estimated that worldwide around two billion people use a drinking water source that has been contaminated with feces, which can cause severe diarrheal infections (86). This contamination is one of the leading causes of disease and death, particularly among children in low- and middle-income communities (LMICs). While more incidences of diarrhea are associated with contaminated drinking water in LMICs, there are still outbreaks in high-income communities (62). A recent study found that nearly half a million households in the United States do not have access to reliable sanitation services, over a thousand community water systems are in serious violation of the Safe Drinking Water Act, and over 20,000 permittees were in significant noncompliance with the Clean Water Act (53). These numbers do not include those experiencing homelessness or the over 40 million households in the United States (U.S.) who use domestic wells, natural water sources, or other types of drinking water supplies that are not regulated by the U.S. Environmental Protection Agency (EPA) (36). Low-income and non-White people, particularly Black and indigenous populations that continue to be plagued by injustices stemming from the legacies of colonialism, are disproportionately affected by these inadequacies in water and sanitation services (52). This inadequate access to water, sanitation, and hygiene (WASH) services is a public health crisis. 30% of all waterborne illnesses in the U.S. were associated with drinking untreated groundwater between 1971 and 2008 (19).

1.1 Water Quality Monitoring

To improve public health, it is important for water service providers and in some cases households to monitor fecal contamination in the drinking water, but that can be expensive and time-consuming. Traditionally, in order to test water for fecal contamination, one must retrieve a sample from the water point, take it to a lab, plate this sample, and then incubate the plated sample for 18-24 hours. Consumers may already be exposed by the time contamination in the water source is detected. These tests on average cost 21 dollars per sample including consumables, equipment, lab access, and logistics (11; 20). Challenges such as these have prompted the United Nations International Children's Emergency Fund (UNICEF) to ask researchers to develop a real-time, in-situ Escherichia coli (E.coli) detection product (81). *E. coli* is the World Health Organization's (WHO) recommended indicator for fecal contamination of drinking water and is the only coliform that is almost exclusively associated with feces (76). UNICEF set minimum performance requirements in the target product profile (TPP) to include the following: being battery-based, requiring minimal processing, not needing reagent mixing or incubation, having qualitative output ranges of fecal contaminations, being able to sample multiple water sources, being able to detect 10 colony forming units (CFUs) per 100 mL with the false positives and negatives below 10%, having a detection time of less than 3 hr, having a minimum life span of two years, and being portable.

1.2 Tryptophan-like Fluorescence

While there are other potential methods to detect *E. coli*, using spectroscopy to measure tryptophan-like fluorescence (TLF) in drinking water systems offers an opportunity to meet the requirements identified for UNICEF's TPP for a real-time, in-situ E.coli detection product. Measuring TLF, or Peak T1 relies on the fluorescence of tryptophan whose presence can indicate it as its own free molecule or in proteins, peptides, humic structures, or in the presence of microbial activity (34; 29; 35). This phenomenon is even more apparent in E.coli, which has the highest TLF per occurrence compared to other bacteria because it produces indole from lactose and tryptophan

(74). TLF has an emission wavelength of around 360 nm and an excitation wavelength of 275 nm (17). It is important to note that there is some uncertainty over what is actually being measured when using this method, but, recently, multiple studies have shown promising data that measuring TLF can be used as a risk assessment tool in place of or in addition to microbial testing in drinking water systems (7)(72). Tryptophan is a free molecule and, when bound to structures where microbial activity is occurring, can be intracellular or extracellular. As intracellular molecules, tryptophan is found in bacteria as structural and functional proteins used in endospore formation; metabolic pathways, and byproducts. If occurring as extracellular molecules, it is found as secreted signaling molecules, exotoxins, metabolic byproducts, and cellular debris (29). Tryptophan contains an amine group, a carboxylic acid group, and a side chain indole group. As microbial activity converts tryptophan to a side chain indole group fluorescence enhances at a 33% greater intensity (29). It is important to note that while indole occurs in many natural phenomena, it is even more apparent in mammalian digestive systems and feces and it can be measured in concentrations as high as 0.5 mM in *E. coli* cultures (46). Previous studies have been conducted that conclude the intracellular or extracellular occurrence of TLF depends on the type of water source such as surface water or ground water. In natural water sources, it is more likely to have extracellular TLF and in lab-grown settings, it is more likely to have intercellular TLF (75)(29).

1.3 Sensor Design

Since 2018, the development, demonstration, and optimization of a low-cost fluorescence sensor that meets the requirements of UNICEF's TPP have been ongoing at the University of Colorado Boulder. These efforts have led to the first publication of this work, as well as a patent application on the design of the prototype and machine learning algorithms that will improve the sensitivity and specificity of the sensor (10). Since the last publication updates have been made to the design of the sensor to be autonomous, highly sensitive to tryptophan concentrations, simple to maintain and clean, able to detect and measure proxies for biofouling through machine learning, and have range and gain levels that are applicable to natural and treated water.

In the design used for the experiments described is shown in Figure 1.1, the sensor used a UV quartz flow-through cuvette from FireFlySci (www.fireflysci.com). This cuvette has a path length of 1 cm. LEDs and the SiPM PCB boards were mounted onto a 3D printed sleeve that slid directly onto the cuvette. A bandpass filter was mounted on the sleeve between the SiPM and the cuvette which was centered around 357 +/-22 nm (23). Rubber tubing connected the cuvette to a peristaltic pump. The main PCB board, cuvette sleeve, and pump were mounted in a polycase waterproof case with a 3D printed mount.

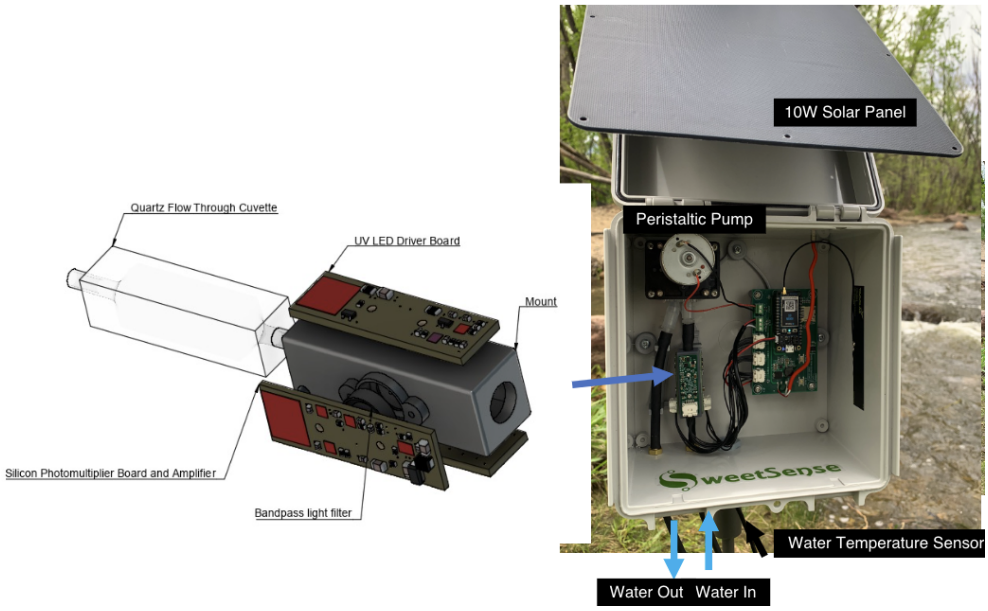

Figure 1.1: The opened enclosure box shows the configuration of the peristaltic pump which pulls water through a flow through cuvette. UV LED and SiPM driver boards are mounted around the cuvette and connected to a microcontroller that controls measurements taken by the SiPM, water temperature sensor, and board temperature sensor. The Particle Boron board then transmits the data via LTE to an online platform.

1.4 Gaps in Knowledge

While the results from this sensor have been promising, it is important to note that challenges and barriers concerning parameters that affect the sensitivity of the signal need to be examined before the sensor can be produced. In this research, I examined how pH, temperature, mineral scaling, and biofouling affect the sensor's sensitivity and signal strength in order to improve on the

current low-cost fluorescence sensor.

1.4.1 pH

Previous studies show that pH affects the intensity of humic, fulvic, and tryptophan-like fluorescence which would likely affect the sensor signal sensitivity (58) (7). Literature indicates that the highest loss of TLF occurs in acidic solutions (low pH below 4.5). The most plausible reasoning for this is due to the deprotonation or protonation of acidic or basic functional groups that are chemically bound to fluorophores (45). However, in another recent study, this trend was not observed, and instead, there was a more variable response in fluorescence to pH (7). This difference could be because this study was conducted on heterogeneous natural water samples while the previous studies were conducted on humic substances. This only reinforces the need to test the effects of pH on this sensor to discover the pH range at which this sensor can operate properly, especially since pH will not be monitored in real-time.

1.4.2 Temperature

Temperature has been shown to affect a TLF signal. Higher temperatures increase collisional quenching, which increases the likelihood of an excited electron to return to the ground state energy through a radiationless pathway (6). Collisional quenching is when a fluorophore deactivates through contact with another molecule, leading to a decrease in fluorescence activity without a chemical reaction occurring (78). A recent study on how temperature affects tryptophan-like fluorometers shows that TLF is negatively correlated to temperature (38). These results emphasize the importance of adding a correction model based on temperature to improve the accuracy of the results. While the trends of temperature affecting TLF have been fairly well documented to be reproducible, it is still important that a correction model be sensor-specific. Our sensor configuration includes the addition of a thermocouple to monitor temperature and correct measurements in real time. It is also important to test the effects of temperature on this specific sensor to set an optimal temperature range for operation.

1.4.3 Turbidity

Turbidity is the measurement of the clarity of a liquid and is an common test of water quality. Clarity is influenced by suspended particles in the water which can cause increased scattering and excitation of light (38). A study of in-situ Chromophoric Dissolved Organic Matter fluorometers indicated that at greater than 400 NTU (the units in which turbidity is measured) the fluorescence signal is reduced by 80% when the excitation wavelength is 370 +/-10 nm (22). While this finding is significant to consider, it does not indicate anything about TLF. Studies concerning how turbidity influences TLF sensors are uncommon; however, a recent study shows an increase in TLF intensity at low to moderate turbidity (70). These results do not agree with other findings which reported more fluorescence intensity at low and high turbidity (22). However, they were pronounced, non-linear, and repeatable between tryptophan concentrations (38). These trends were more apparent when comparing sediment types as opposed to sensor types examined in this study. This discrepancy could be due to increased stray light reaching the fluorometer photodiode in silt when compared to clay, or the fact that clay might have higher attenuance than silt. Furthermore, it could be due to the fact that in this study, the sediment was treated to remove organic matter, which could have increased the ratio of soft to hard scatters, thus reducing absorption. After examining previous studies, it was discovered that there is a clear need to monitor turbidity for this sensor specifically and recognize a turbidity range for optimal operation.

1.4.4 Mineral Scaling

Mineral scaling is the deposition of minerals on the interior surfaces of water lines and containers. This most often occurs when water contains carbonates or bicarbonates and the calcium, sulfates, and magnesium in those components are heated. There have not been many studies on how mineral scaling affects fluorescence sensor signals directly but several studies teach how to quantify or monitor $CaCO_3$ scale formation using evanescent field spectroscopy (61) (12) (55). Another study focused on the development of a novel fiber optic sensor for real-time sensing of silica scale

formation (57). Each of these studies has a component where water with a known concentration of chemicals is pumped through the sensor to promote mineral scaling. To quantify this scaling, the transmittance responses were measured. These responses indicated that the transmittances decreased to a constant value over time (57) (55). With these results, one can make a reasonable assumption that as mineral scaling occurs, the transmittance will decrease causing fluorescence to decrease as well. When less light is able to pass through the cuvette for this sensor, there is less light to fluoresce.

1.4.5 Biofouling

Biofouling, or the accumulation of biofilm, is the undesired deposition of microorganisms and sticky extracellular polymeric substances on surfaces (26). Biofilms are made in 5 stages: 1) planktonic cells attach to the surface, 2) the attached cells divide and extracellular polymeric substances are formed, 3) then the extracellular polymeric substance expands, 4) the biofilm matures, and 5) the mature biofilm releases planktonic bacteria back into the environment (27). The degree to which biofouling occurs varies as a function of the environment. Anti-biofouling mechanisms are put in place to make sure the surfaces of the tools that are growing biofilms are working properly. Research shows that as biofilms accumulate on the surface of the cuvette in a sensor, a higher rate of false positives and signal drift will occur (74).

Biofouling has been shown to affect fluorescence sensor signals in the field and interferes with the otherwise reliable sensors (24). One potential method toward addressing these challenges without regular maintenance and calibration is through machine learning based synthetic calibration that has been validated with ground truth data that has been collected manually (37).

1.5 Experimental Setup

Each experiment done in this study was conducted with the sensor that is described in Figure 1.1. However, different configurations of the sensor's setup were developed for context of the experiment being conducted. Three different setups were designed for field testing, lab testing,

and biofouling and mineral scaling testing and are described below.

1.5.1 Field Testing

Figure 1.2 depicts the setup for one of the sensors installed on Boulder Creek near 55th Street. Adaptations made to the setup for field testing include a camera stake with adjustable screws that the enclosure of the sensor is mounted on to easily and securely stake it into the ground near the testing site. Also, the sensor's inlet and outlet tubes were secured to a ceramic brick so they would remain submerged under the water, but above the creek floor.

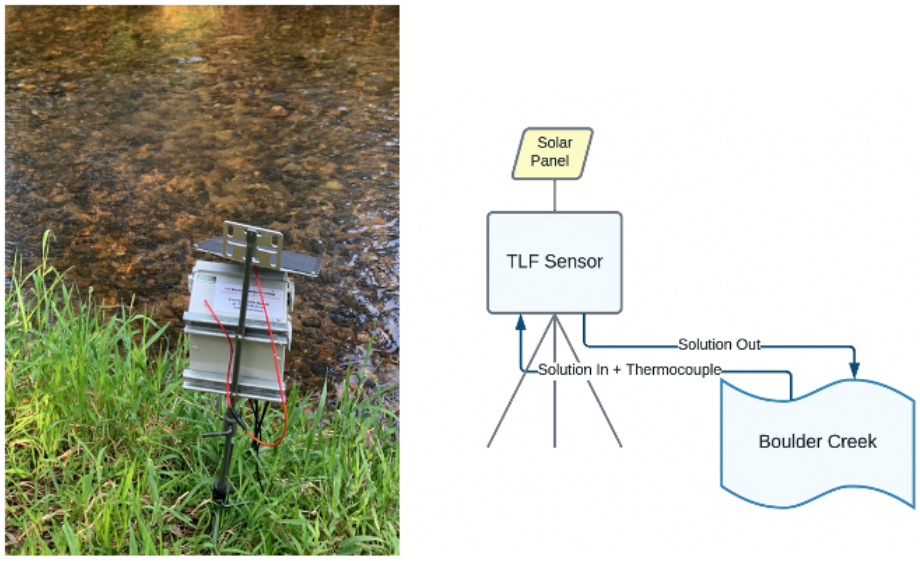

Figure 1.2: Photograph and Flowchart of the Field Testing Setup 9315 on Boulder Creek near 55th Street

The three other sensors were placed in various other locations, but the setup for each of the field sensors are fairly similar.

1.5.2 Lab Testing

Figure 1.3 depicts the setup used for all the tests conducted in chapter 2 (tryptophan, wastewater, and *E. coli* sensitivity and the viable cell test) and chapter 3 (pH, turbidity, and temperature). The box containing the sensor rested on a lap bench top stand while in use. The testing bottle

containing the solution being pumped through the sensor at that particular test was set on a combination heat and stir plate. A laptop was set up to run Particle, which was the program used to call on the sensor to pump and/or measure each sample.

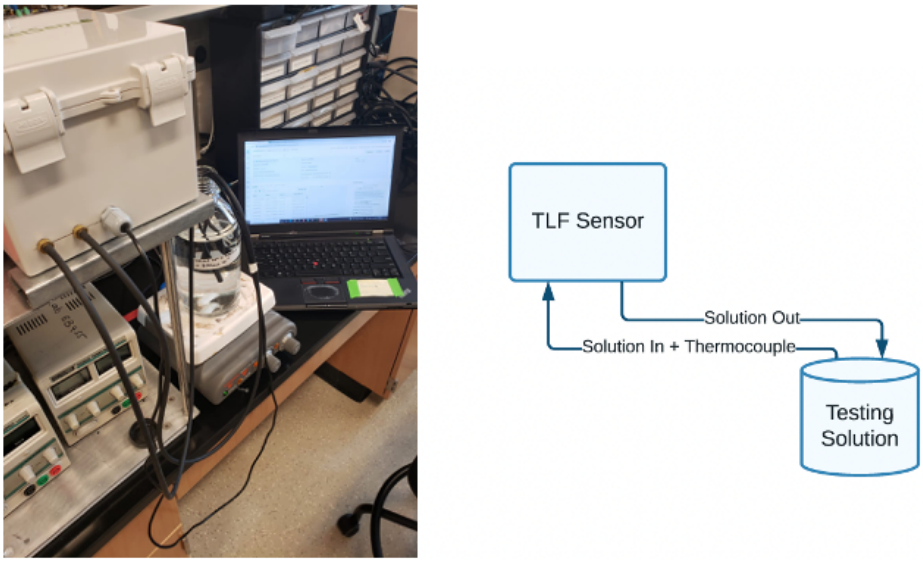

Figure 1.3: Photograph and Flowchart of the Lab Testing Setup

Many of the in-lab experiments required other equipment such as a pH meter that are not pictured in Figure 1.3 but other than that the setup remained constant. Details of the analytical methods will be discussed in detail in chapter 2 and 3.

1.5.3 Biofouling and Mineral Scaling Testing

Figure 1.4 depicts the setup used for the biofouling and mineral scaling experiments. Since the senors used in these experiments were the ones pulled from the field, the box containing the sensor remained on the camera stake with adjustable screws. A 5 gallon bucket was fitted with a cooling jacket to better mimic temperature changes in typical drinking water systems. The thermocouple, inlet, and outlet tubes were placed inside the 5 gallon bucket.

The contents of the bucket would vary based on the experiment, but the setup remained the same nonetheless.

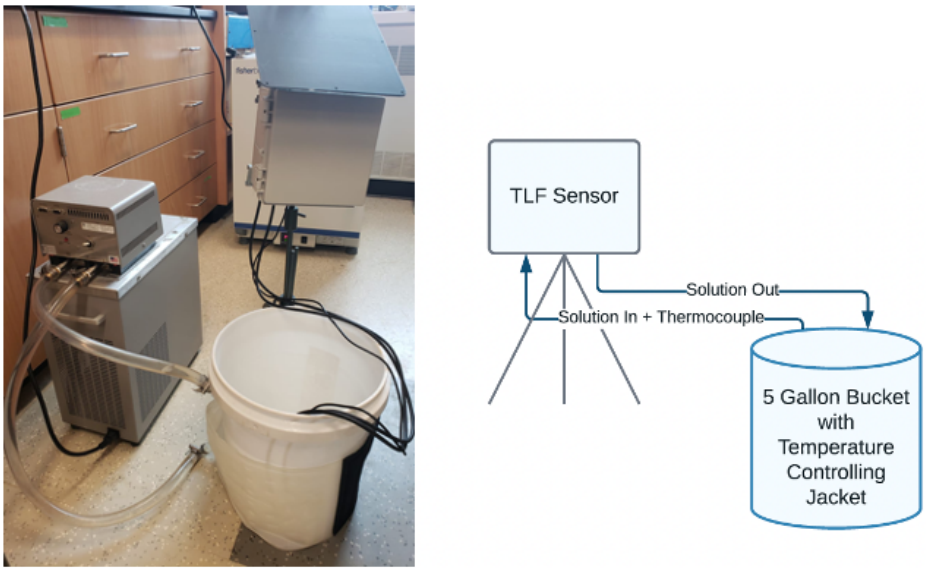

Figure 1.4: Photograph and Flowchart of the Biofouling and Mineral Scaling Lab Testing Setup

1.6 Contributions

Successful research is inherently collaborative, and my thesis is no exception, as I was able to work closely with Emily Bedell and Katie Fankhauser on this project. While there was a lot of overlap in the contents of our work, as seen in Figure 1.5, it is important to give credit to everybody's individual contributions.

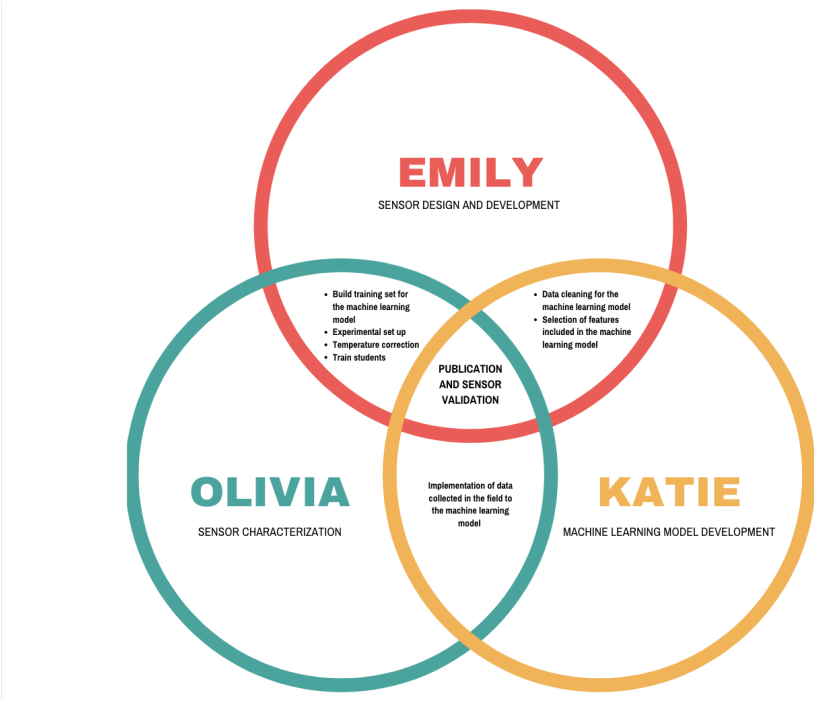

Figure 1.5: Venn Diagram of the Contributions from Emily Bedell, Olivia Harmon, and Katie Fankhauser

I took the lead on the characterization of this sensor which led me to ask and explore the following research questions:

(1) What is the detection limit of *E. coli* and total coliforms in terms of tryptophan, lab grown *E. coli*, and wastewater effluent? Explored in chapter 2.

(2) How will the fluorescence output be affected by pH, temperature, and turbidity? What operational limits for pH and turbidity and correction factors for temperature will need to be set to account for this difference? Explored in chapter 3.

(3) How does the formation of biofilms and mineral scaling inside the flow through cuvette impact the sensor's signal? Explored in chapter 4.

(4) Can a machine learning model be designed and implemented to attenuate the effects of biofilms and increase the specificity and sensitivity of the sensor? Explored in chapter 5.

In order to explore research questions 1 and 2, I conducted literature reviews to become familiar with the background of TLF, the current knowledge of the available TLF sensors, and how pH, temperature, and turbidity affect the fluorescence output. Using the knowledge I gained conducting the literature reviews, I figured out how to run each experiment to add to the knowledge available to answer research questions 1 and 2 using our sensor. After I conducted each experiment and analyzed the data in RStudio, I made small adjustments to the experimental design before writing the standard operating procedure for each experiment. Details about these experiments can be found in the methods section of chapters 2 and 3. I then trained graduate and undergraduate students on how to successfully conduct each experiment to increase the amount of data that could be collected. Using an R code that Emily Bedell and I wrote, I analyzed all the data collected and made recommendations for future work based on those results.

In order to explore research question 3, I conducted literature reviews to become familiar with the background of how biofilm and mineral scaling forms, how it can impact fluorescence, and how to enumerate it. I organized meetings with other graduate students, postdoctoral researchers, and faculty. Using the knowledge I gained from these meetings and conducting the literature reviews, I planned how to run each experiment that would form biofilm or scaling on the inside of the flow through cuvette. I worked with Emily to measure how the formation of biofilm or scaling affects the sensor's signal. I also figured out how to enumerate the biofilm and scaling that formed on the inside of the flow through cuvette. Details about these experiments can be found in the methods section of chapter 4. I trained graduate and undergraduate students on how to conduct the experiment to aid in the formation of biofilm and monitor contamination events in order to increase the amount of data that could be collected. I conducted all of the experiments to enumerate the biofilm and

mineral scaling through membrane filtration and spectral analysis. I designed and 3D printed a cuvette holder and caps so the cuvette could be filled with DI water and fit into each spectroscopy machine. Using this data and the R package staRdom I created excitation-emission matrices to further investigate how the formation of biofilm and mineral scaling affects fluorescence.

For research question 4, I assisted in the formation of the data for the training set for the machine learning model, which is a set of data that is used to train the model and set hidden figures and patterns that are in the data. I worked with Emily Bedell to set up four sensors to collect data for this training set on Boulder Creek. To collect ground truth samples, Emily, 8 other graduate and undergraduate students, and I took and enumerated 298 samples from the creek where each sensor was set up. I trained these students on how to properly collect, record, plate, and enumerate samples from the creek. Details about these methods can be found in the methods section of chapter 5. I organized weekly meetings with this team to coordinate sample collection.

Chapter 2

Sensor Signal Sensitivity to Changes in Water Composition

2.1 Introduction

Sensitivity of the sensor signal to parameter changes in water was conducted. All of the results from the Tryptophan, wastewater, and *E. coli* validation sensitivity experiments were analyzed in R version 4.0.5. Analysis of Variance (ANOVA) and t-tests were conducted on the outputs of these experiments to examine the statistical significance.

2.2 Methods

2.2.1 Tryptophan Sensitivity

Standard L-tryptophan solutions (Sigma-Aldrich reagent grade L-tryptophan) were made by mixing 1000 mL of deionized (DI) water and 0.1 g powdered tryptophan for 30 min to create a solution of 100 ppm tryptophan. This 100 ppm tryptophan solution was used to prepare 0.05, 0.1, 0.5, 1, 3, 10, 30, 70 ppb L-tryptophan standards. Each solution was tested within 72 hrs after being made. Before testing, the sensor was rinsed by pumping DI water through for 60 s. To collect data for each solution, starting with the lowest concentration (DI water) and ending with the highest concentration (100 ppb), the inlet tube was placed in the solution and the outlet tube placed in a waste container. For each dilution ten TLF measurements, with the average of 80 samples per measurement, were taken from the sensor. The sensor was rinsed between each solution by running DI water through the tubes and cuvette for 30 s.

2.2.2 Wastewater Sensitivity

Wastewater effluent was collected from the Boulder Wastewater Treatment Facility and stored in a walk in fridge at 4°C for a maximum of 5 days until it was used for testing. Standard dilutions were made by mixing DI water and wastewater effluent to prepare 10%, 25%, 40%, 55%, 70%, 85%, and 100% dilutions of wastewater effluent. Each solution was made right before the beginning of testing and kept for a maximum of 48 hr refrigerating the dilutions that were not being tested at the time. Sensor data for each dilution, starting with the lowest concentration (DI) and ending with the highest concentration (100% wastewater effluent dilutions), was collected and analyzed by the same method used to collect the tryptophan data. The pH of DI and each wastewater effluent dilution was measured since DI water exposed to air has the potential to be acidic (65). *E. coli* and total coliforms (TC) present in each dilution were enumerated through membrane filtration by plating a filter with m-ColiBlue24 broth (EPA Approved Hach Co.: 10029 method). m-ColiBlue24 broth indicates *E. coli* colonies by blue coloration resulting from specific activity of β -glucuronidase and TC by red coloration resulting from specific activity of β -galactosidase (76). Samples were plated in triplicates and incubated at 35°C for 18-24 hr.

2.2.3 *E. coli* Sensitivity

A working culture was prepared using *E. coli* K-12, the strain of *E. coli* that researchers know the most about, stored in individual tubes in a freezer at -80°C (59). To prepare the working culture, *E. coli* K-12 was taken from the freezer and placed in a sterilized fume hood to warm up. While the sample was defrosting, tape, a permanent marker, 25 mL pipette tip, motorized pipette, 25-100 μL pipette tip, 10-100 μL pipette, three Erlenmeyer flasks, and tin foil used to cover the flask were gathered, sterilized, and placed in the fume hood. Two of the Erlenmeyer flasks were labeled, one nutrient broth and the other nutrient broth + *E. coli*, with tape. The third Erlenmeyer flask was filled with 75 mL of nutrient broth. Nutrient broth was made once a month using 7.2 g of Gather DifcoTM Nutrient Broth powder and 900 mL of DI water mixed for 30 min then Autoclaved

at 121°C. Using the 25 mL pipette tip and motorized pipette, 25 mL of the nutrient broth were transferred to each of the labeled Erlenmeyer flasks. Next, using the 10-100 μL pipette, 25 μL of *E. coli* K-12 were transferred from the defrosted sample tube to the Erlenmeyer flask labeled nutrient broth + *E. coli*. The top of each Erlenmeyer flask was immediately covered with tin foil and incubated in the benchtop shaking incubator for at least 15 hr at 37°C and 121 rpm. After 15 hr, both Erlenmeyer flasks were compared to ensure the nutrient broth had not been contaminated as seen in Figure 2.1.

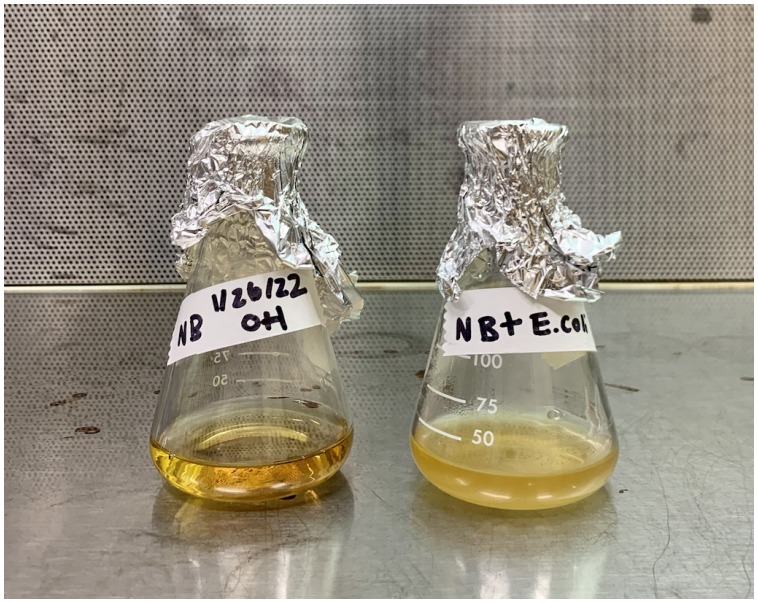

Figure 2.1: The Erlenmeyer flask on the left contains only Nutrient Broth and the Erlenmeyer flask on the right contains Nutrient Broth and *E. coli* K-12. This distinction is made clear when you compare the cloudiness of the Nutrient Broth on the right to the clearness of the Nutrient Broth on the left.

Once it was confirmed there was no contamination, a sterilized fume hood was turned on, and the nutrient broth + *E. coli* Erlenmeyer flask, an Erlenmeyer flask of 75 mL of PBS, 25 mL pipette tip, motorized pipette, and one 50 mL centrifuge tube labeled "WC *E. coli* K-12 Date, Time" were placed inside to clean the working culture. The working culture (continents of nutrient broth + *E. coli* Erlenmeyer flask) was poured in the labeled 50 mL centrifuge tube then placed in the balanced centrifuge set at 3000 rpm for 10 min. After 10 min in the centrifuge, the working

culture was returned to the fume hood and the liquid continents of the centrifuge tube was poured back into the flask labeled nutrient broth + *E. coli*, and 25 mL of PBS was added to the centrifuge tube then shaken until all the particles are dissolved. After using this method to clean the working culture two more times, the liquid content of the centrifuge tube was transferred to the flask labeled nutrient broth + *E. coli* for the last time. Then, 10 mL of PBS was added to the centrifuge tube and shaken until all the particles were dissolved creating the final WC of 10^9 CFU of *E. coli*/mL. Figure 2.2 depicts a flowchart that describes this process.

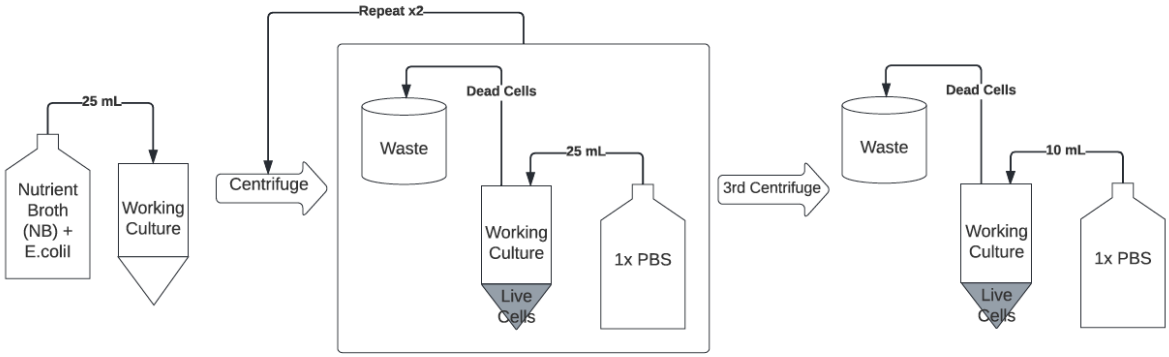

Figure 2.2: Flowchart describing the process for preparing the working culture

This WC was used to prepare 10^7, 10^6, 10^5,10^4, 10^3, 10^2, 10^1, 10^0, and 10^{-1} CFU of *E. coli*/mL dilutions. These values were validated by plating each dilution 3 times, incubating the plates at 37°C for 20 hr, then counting the colonies that grew on each plate. The *E. coli* solutions were made right before each test and kept for a maximum of 48 hr. Sensor data for each solution, starting with the lowest concentration (DI) and ending with the highest concentration (10^7 CFU of *E. coli*/mL dilutions), was collected and analyzed using the same method that was used to collect the tryptophan data.

2.2.4 Chlorination Impact Test

To determine the impact of chlorine on the sensor's output 0.24 mL of bleach was added to 1000 mL wastewater effluent, a 50 ppb tryptophan solution, and DI water. Then, 0.24 mL of bleach was added to each dilution and mixed for 30 min. The pH of DI water, 50 ppb tryptophan, and wastewater effluent was measured before bleach was added and after bleach was mixed in each solution for 30 minutes. Total chlorine present was measured at minute 1 and 30 and free chlorine was measured at minute 30 using a spectrophotometer. 10 measurements with the sensor were taken for each solution (wastewater effluent, wastewater effluent + bleach, 50 ppb tryptophan, 50 ppb tryptophan + bleach, DI water, DI + bleach). The concentration of *E. coli* and TC before and no more than 60 minutes after the addition of bleach was enumerated by membrane filtration using the method described above.

2.3 Results and Discussion

2.3.1 Tryptophan Sensitivity

At $p < 0.01$ (according to EPA MDL Procedure) this sensor is able to show a statistically significant difference between DI water and 0.05 ppb tryptophan at all four current inputs to the LEDs (28). As seen in Figure 2.3, the design goal of 1 ppb tryptophan (signifying high risk contamination) was met. As the current increases, the sensitivity at low tryptophan concentrations increases; as the current decreases, a higher range of tryptophan concentrations can be detected.

2.3.2 Wastewater Sensitivity

The results from examining the sensor response to various dilutions of wastewater effluent from the Boulder Wastewater Treatment Facility are shown in Figure 2.4 below. From this data, it is apparent that the sensor can detect 2 CFU/100 mL *E. coli* in the presence of other microbes and debris found in wastewater effluent. These findings further support the sensitivity results of the tryptophan and *E. coli* testing.

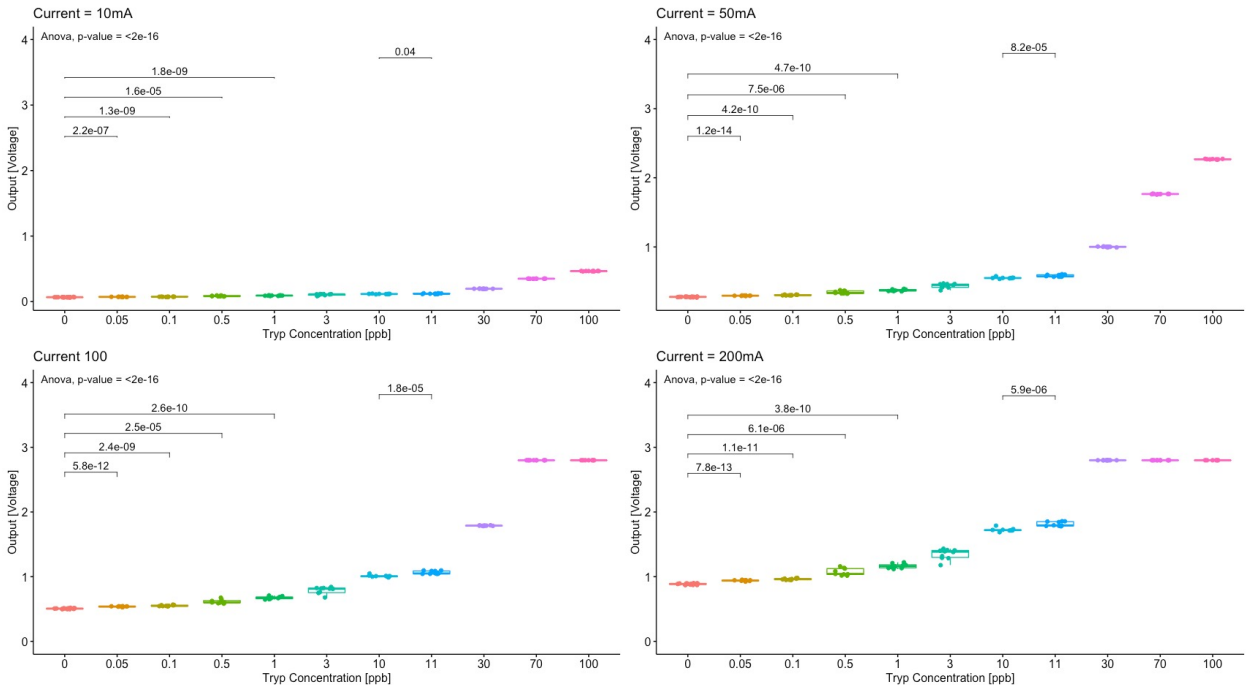

Figure 2.3: Sensor response from tryptophan dissolved in DI water at four different currents powering the LEDs, indicated at the top of each graph. Analysis of Variance (ANOVA) p-value shows the difference across all concentrations.

The sensor was able to statistically significantly detect ($p < 0.01$) *E. coli* concentrations in wastewater effluent above 10 CFU/100mL, which signifies intermediate risk contamination. These results are seen in the figure below. The R^2 between *E. coli* present in the wastewater and sensor output was 0.93. There was a drop in sensor output in the range of 1000 CFU/100mL. The drop in sensor output can be attributed partly to inner filter effect (IFE), but could also be a result of light scatter from particles. The absorbance data collected on the UV-VIS spectrophotometer shows increasing absorbance as the concentrations increase. The calculated corrected fluorescence due to IFE increases the R^2 between *E. coli* and sensor output to 0.95 as seen in Table 2.1.

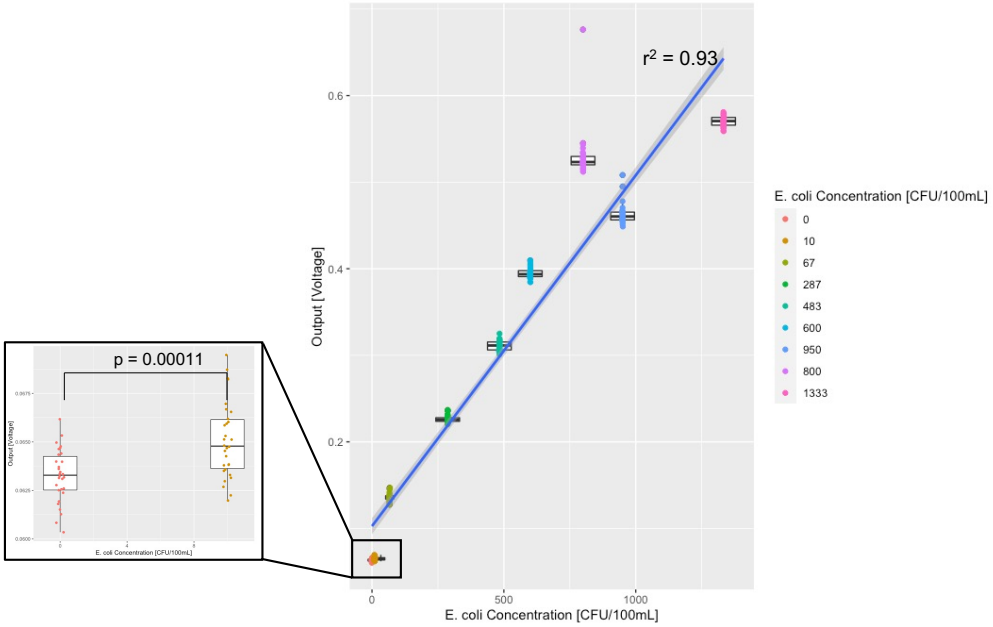

Figure 2.4: Sensor response from wastewater effluent dilutions graphed continuously. The bar between 0 and 10 CFU/100mL show the significant differences between indicated concentrations calculated by a student's t-test.

There was not a significant change in pH in the various wastewater dilutions as seen in Table 2.2.

Table 2.1: Corrected Fluorescence Based on Inner Filter Effects

E.coli Concentration [CFU/100mL]	Measured Fluorescence [V]	$A_{\lambda ex}$	$A_{\lambda em}$	Corrected Fluorescence [V]
0	0.06	0.00	0.00	0.06
10	0.07	0.01	0.00	0.07
67	0.14	0.01	0.00	0.14
287	0.23	0.03	0.01	0.24
483	0.31	0.05	0.02	0.34
600	0.40	0.07	0.03	0.44
800	0.53	0.10	0.04	0.62
950	0.46	0.11	0.04	0.55
1333	0.57	0.13	0.05	0.70

Table 2.2: pH of DI Water and Wastewater Effluent Dilutions

Percentage of Wastewater Effluent	Measured pH
0%	6.67
10%	6.29
12.5%	6.50
25%	6.75
50%	6.96
100%	7.02

2.3.3 *E. coli* Sensitivity

The results from examining the sensor's response to lab grown *E. coli* K-12 are shown in Figure 2.5 below. This test showed the sensor response at 33 CFU/100mL at its lowest. Concentrations that are higher than 33 CFU/100mL are significantly different from the sensor response to DI water but average voltage output varies until the *E. coli* concentration reaches 1533 CFU/100mL. These varying readings could be due to inconsistencies in concentrations for each dilution measurement or IFE.

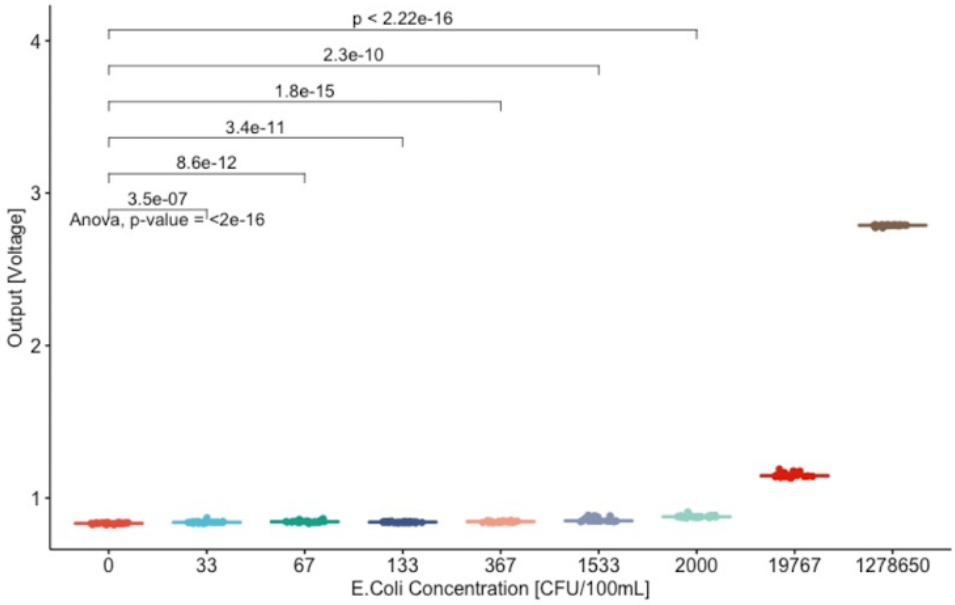

Figure 2.5: Sensor response to lab grown *E. coli* K-12. The bar between 0 and 10 CFU/100mL show the significant differences between indicated concentrations calculated by a student's t-test.

2.3.4 Impact of Chlorination

Adding bleach, a strong oxidant, to solutions of DI water, 50 ppb of tryptophan, and wastewater effluent significantly lowered the signal as seen in Figure 2.6. The pH of DI water, 50 ppb tryptophan, and wastewater effluent was 6.67, 6.82 and 7.02 before bleach was added and 8.07, 8.21, and 7.32 30 minutes after bleach was added, respectively. The free chlorine present after 30 min in each solution was 5.9, 5.5, and 5.6 mg/L for the DI, 50 ppb tryptophan, and wastewater

effluent, respectively.

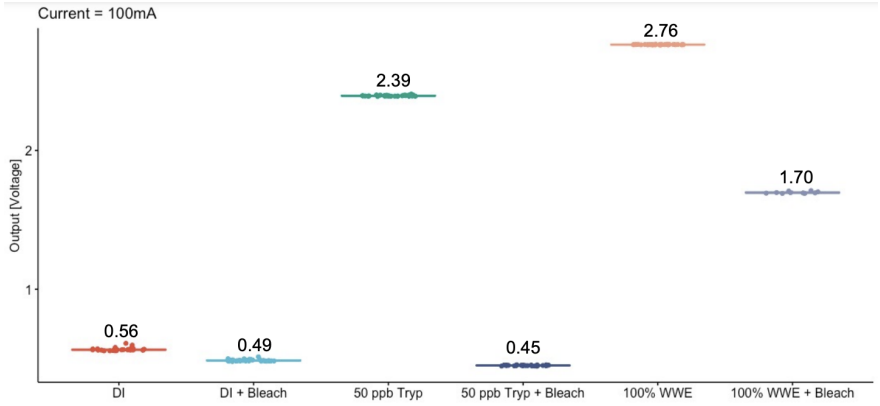

Figure 2.6: Sensor response adding bleach to a DI, a tryptophan solution, and wastewater effluent.

When bleach was added to the tryptophan solutions all of the tryptophan thoroughly degraded to the point that the signal from the sensor was lower than the signal observed using DI. Similar studies have concluded that 0.025% bleach concentration was able to completely degrade the tryptophan in *E. coli* that was lab grown through the destruction of the indole ring, cell protein, and amino acids (3). Pairing these conclusions with the results seen in our experiment one can conclude that the signal remains once bleach is added to wastewater effluent is due to extracellular material and not tryptophan.

Chapter 3

Operational Limitations for pH and Temperature

3.1 Introduction

Testing the sensor in the lab continued by assessing the operational limitations for pH, turbidity, and temperature. By conducting the experiments laid out below, the degree to which each parameter affects the fluorescence output of the sensor was measured. Using the data collected, the operational limits for pH and turbidity were set and a correction factor for temperature was set.

3.2 Methods

3.2.1 Operational Limitation for pH

The methods described in the sensitivity chapter were followed to make tryptophan dilutions; the only exception for this experiment was that tryptophan dilutions were made using DI and tap water and 50 ppb was the only tryptophan dilution tested. A 0.1 M solution of HNO_3 and a 0.1 M solution of NaOH were created with DI water to vary the pH of DI, 50 ppb made with DI, tap water, and 50 ppb made with tap water from 3 to 11 in 0.5 increments. Both the 0.1 M solution of HNO_3 and a 0.1 M solution of NaOH were tested on the lab fluorometer and no inherent fluoresence was measured. The inlet tube was placed in a separate bottle of DI used for flushing out the tubes and cuvette and the outlet tube in a separate bottle for waste, then DI water was pumped through for 60 s before data is recorded. Increments of 1 mL of HNO_3 or NaOH were added to each dilution and mixed on a stir plate for 4 min until the pH-adjusted solution reached

equilibrium. After 4 min, a pH meter (Sealed, gel-filled, epoxy body, Ag/AgCl pH Sensor from Vernier Software Technology) was used to record the pH before pumping the solution for 10 s to flush out the tubes and cuvette with the new acidic or basic solution. After 10 sec, the inlet tube, thermocouple, and outlet tube were placed in the bottle containing the solution being tested, and the sensor recorded 10 measurements of voltage outputs at four current levels to the LEDs: 10, 50, 100, and 200 mA. The sensor data collected was evaluated in RStudio to create a line graph that was used to show the difference between each tryptophan concentration as the pH changes for each current.

3.2.2 Operational Limitation for Turbidity

The sediment chosen for this experiment was Fuller's Earth (D50 = 11.9 µm), a clay. All sediment used in this experiment was treated with hydrogen peroxide to remove organic matter then rinsed with DI water and dried in an oven at 65 °C for 24 hrs following the method outlined by Gray et al (31). Turbidity was evaluated for DI water and a 50 ppb tryptophan standard solution. The treated sediment was weighed and added incrementally to each tryptophan standard to collect data for 0, 20, 50, 100, 250, 500, and 1000 NTUs. Before testing, the inlet tube was placed in a separate bottle of DI used for flushing out the tubes and cuvette and the outlet tube in a separate bottle for waste, then DI water was pumped through for 60 s before data is recorded. After 60 s, pre-measured treated sediment was added to the tryptophan dilution and mixed for 5 min. After 5 min, the mixed solution was pumped for 10 s to flush out the tubes and cuvette with the new turbidity. Then 50 mL of the dilution was collected in a sample tube to measure turbidity on the turbidimeter (Hach 2100N turbidimeter). The inlet tube, thermocouple, and outlet tube were placed in the testing dilution and the sensor recorded 10 measurements of voltage outputs at four current levels to the LEDs: 10, 50, 100, and 200 mA. The sensor data collected was evaluated in RStudio to create a line graph that was used to show the difference between each tryptophan concentration as turbidity changes for each of the 4 currents.

3.2.3 Temperature Variation and Correction

Tryptophan solutions were made following the methods described in the sensitivity chapter. Each solution (DI water, 1, 3, 10, 30, 70, 100, 200 ppb) was refrigerated overnight (for at least 12 hr) in the walk in refrigerator at 4°C. Solutions were removed from the refrigerator one at a time in ascending concentration, starting with DI water, and tested. Once removed from the refrigerator, the solution was pumped through the sensor to flush out the tubes and cuvette from the previous test. Then solutions were placed on a hotplate stirrer and stirred during the entirety of the test; the heating element was applied when the solution began to reach room temperature. The temperature of the water was monitored from 7°C to 35°C using a thermocouple. Measurements were taken at least three times per degree Celsius by the sensor. The sensor recorded 10 measurements of voltage outputs at four current levels to the LEDs: 10, 50, 100, and 200 mA. The sensor data collected was evaluated in RStudio to create a line graph that was used to show the difference between each tryptophan concentration as turbidity changes for each of the 4 currents.

3.3 Results and Discussion

3.3.1 pH Sensitivity

pH was shown to significantly impact the sensor's signal outside of the range pH = 6 to 9. As seen in Figure 3.1, the sensor's signal varies over 1 volt when the pH ranges from 3 to 6 and a little less than 1 volt when the pH ranges from 9 to 11. In contrast, when the pH = 6 to 9 the sensor's signal varies less than 0.5 volts, which is not significant or outside of natural ranges and thus would not impact the machine learning output.

Similar effects of pH on the sensor's signal were observed in varying tryptophan dilutions in Figure 3.2. For all dilutions but especially 100 ppb the sensor's signal increases below pH of 6 and above a pH or 9 until the signal saturates at a pH of 9.5.

The increase of the sensor's signal with varying pH could be due to the many different potential sources of TLF in water samples and the direct changes in the behavior of the fluorophore

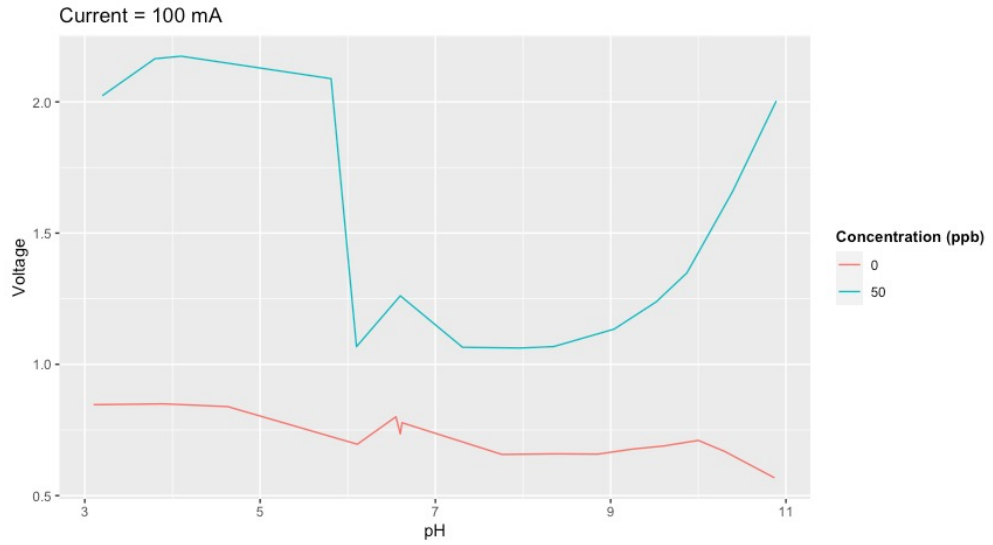

Figure 3.1: Sensor response to varying pH at increasing tryptophan concentrations

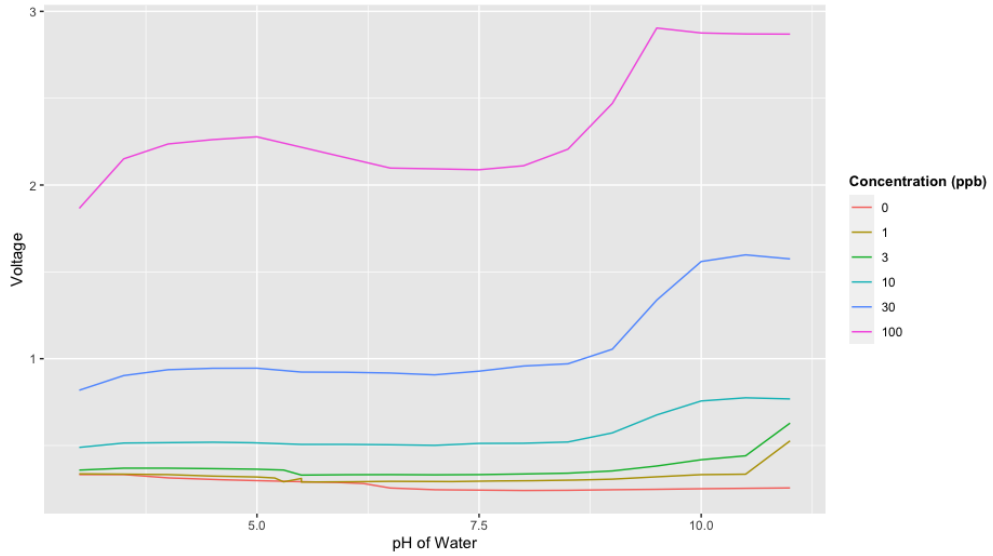

Figure 3.2: Sensor response to varying pH at increasing tryptophan concentrations

associated with the folding and unfolding of proteins in acidic solutions (45). The increase in signal below a pH of 6 seen in Figure 3.2 could be due to changes in protonation states based on the suggestions made by Chen et al. (44). Similarly the increase in signal above a pH of 9 could be due to ionization.

3.3.2 Turbidity Sensitivity

Turbidity was shown to significantly impact the sensor's signal in the range or 0 to 100 NTUs as seen in Figure 3.3. In fact, the sensor's signal increases for DI water and 50 ppb tryptophan.

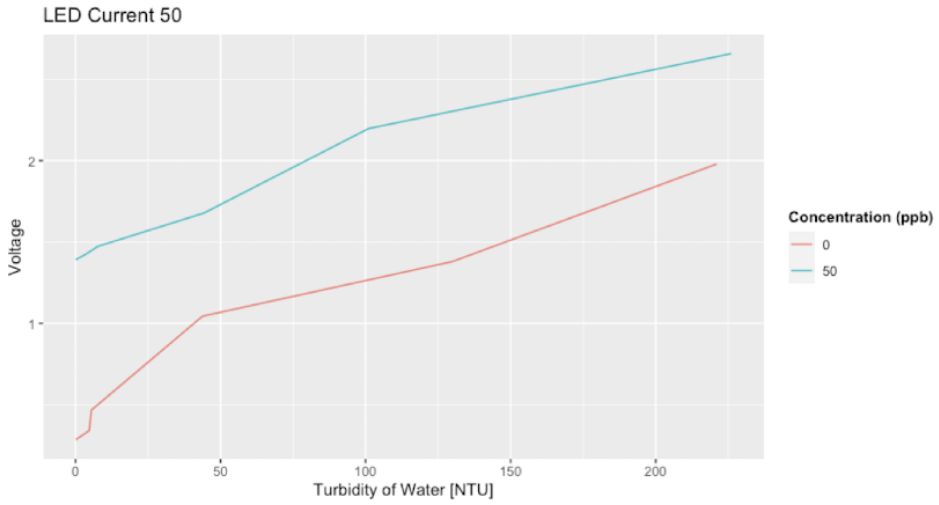

Figure 3.3: Sensor response to varying Turbidity with DI water and 50 ppb Tryptophan

The results are not in line with previous findings presented in the background but these differences can be explained. Based on the literature review, it was predicted that as turbidity increased the sensor's signal would decrease. However, it is possible that the clay that was used to vary the turbidity to could be contributing to the increase in signal due scattering that could occur during pulsing. Hydrogen peroxide was used to remove organic matter from the clay before testing. While this is a procedure used in other turbidity tests, hydrogen peroxide has been known to cause clumping in the clay which makes it harder for the clay to be evenly distributed through a solution and raises the signal.To further characterize this sensor future turbidity test with other

types of sediment and water sources should be conducted.

3.3.3 Temperature Sensitivity

The sensor output was negatively correlated to water temperature. It was also observed that the impact of water temperature on the signal was higher at higher tryptophan concentrations. Similarly to Watras et al. 2011, fluorescence declined exponentially with water temperature at all concentrations tested, thus the methods for temperature correction of a fluorescence sensor described in that study were followed (85). The data was fitted to the relationship:

$$TLF_m = TLF_r e^{\rho(T_m - T_r)} \tag{3.1}$$

Where T is temperature (°C), the subscripts r and m represent the reference and measured values, and ρ is the temperature coefficient (C^{-1}). Equation 3.1 was fit to each concentration and the mean ρ was found to be -0.03. Using -0.03 for the value of ρ in Equation 3.1, the effect of temperature can almost completely be removed from the raw data. For the lower concentrations, correcting the data causes a small increase of the sensor output with temperature as seen in Figure 3.4.

Looking at Figure 3.4, it is clear that increasing water temperature attenuates the sensor's signal. These impacts are due to an increase in temperature, resulting in an increase in collisional quenching, which increases the likelihood of an electron to return to the ground state energy through a radiationless pathway. In other words, collisional quenching is when the fluorophore deactivates through contact with another molecule, leading to a decrease in fluorescence activity without a chemical reaction occurring (85). A correction factor was established through an exponential fit to the experimental data; this correction factor was incorporated to correct data with water temperature values to 20°C.

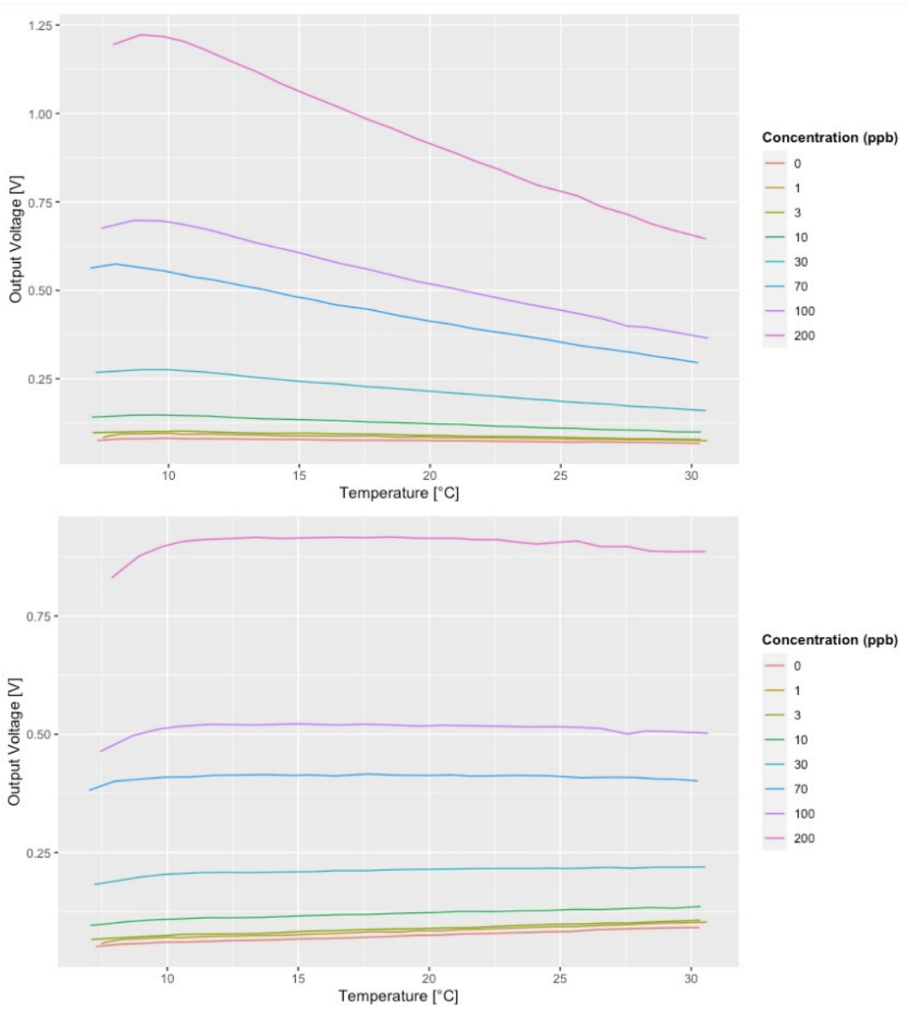

Figure 3.4: The top figure is the sensor response to increasing temperature at increasing tryptophan concentrations. The bottom figure is data corrected to temperature at 20°C.

Chapter 4

Evaluation of the Effects Biofouling and Mineral Scaling have on the Sensor

4.1 Introduction

To further improve the accuracy of the sensor, the effects of biofouling and mineral scaling on the sensor's signal was evaluated. By conducting the experiments laid out below, biofouling and mineral scaling was quantified and the degree to which each affects the fluorescence output of the sensor was measured. Using the data collected, the results were incorporated into a machine learning model that corrects for these differences in signal.

4.2 Methods

4.2.1 Biofilm Growth

To monitor biofilm growth, the sensors were set up in the lab to automatically collect data once an hour on the hour. To simulate treated drinking water sources that have contamination events, these sensors ran with tap water from the lab that was occasionally spiked with wastewater effluent. Ground truth samples were collected at 10:00 am, 12:00 pm, and 3:00 pm to be plated, incubated for 24 hr, and then *E. coli* and total coliforms were counted. Once a month, a flow-through cuvette from one of the sensors was removed and cleaned.

4.2.2 Estimation of Biofouling and Mineral Scaling Through Spectral Analysis

Before cleaning, the biofilm was enumerated through spectral analysis to create excitation-emission matrices (EEMs) and by plate counts. For spectral analysis, a cuvette holder and plugs

were designed, and 3D printed so that the flow-through cuvette fit in each machine and retain DI water to mimic the conditions that occur when the sensor is pulling a sample. To design the holder and caps, measurements of the flow-through cuvette and the cuvette holder from the lab were taken and adaptations were made. With these measurements, a new cuvette holder was drawn using Autodesk Fusion 360 shown in Figure 4.1 below.

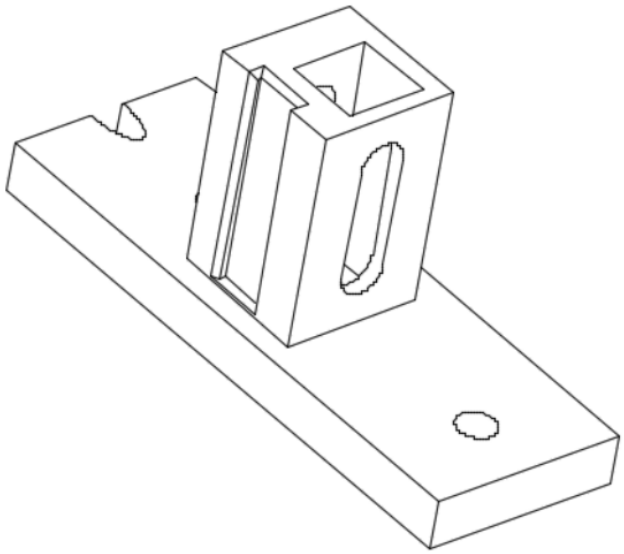

Figure 4.1: Design of Cuvette Holder with Adaptations for a Flow-Through Cuvette

Since the cuvettes will be examined by the spectroscopy instruments filled with DI water, caps are needed to plug one end of the flow-through cuvette. Using the measurements of the flow-through cuvettes in the sensors, caps were drawn in Autodesk Fusion 360 as seen in Figure 4.2.

Once the design of the cuvette holder and caps were finalized, they were saved as Mesh in order to open in PrusaSlicer. In PrusaSlicer, the final adjustments were made for 3D printing and saved to an SDCX card. The SDCX card was inserted into the 3D printer and the cuvette holder and caps were created.

Fluorescence data were collected using the lab fluorescence spectrophotometer (Fluoromax-

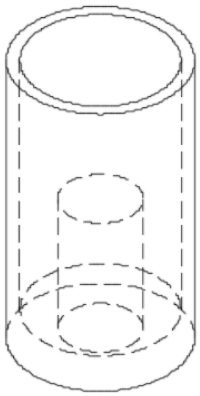

Figure 4.2: Design of a Flow-Through Cuvette Cap

4). Emission wavelengths were measured from 300 to 400 nm in 2 nm increments, excitation wavelengths from 300 to 240 nm in 10 nm increments, and the excitation and emission bandpass was set to 5 nm and a 0.25 sec integration time. Absorbance data were collected using an ultraviolet-visible spectrophotometer (Cary 4000) which has a maximum absorbance of 1 and a scan rate of 600 nm/min. Absorbance spectra were measured from a wavelength of 200 to 800 nm at 1 nm intervals. Once fluorescence and absorbance data was collected, the excitation-emission matrix (EEM) were corrected using the staRdom (R package version 1.1.14) (33). The R package staRdom corrects inner filtering, removes Raman scattering by subtracting a blank EEM (or the data from the clean cuvette) from the sample EEM (or data from the fouled or scaled cuvette) the raw data using a blank, and removes Rayleigh scattering.

This code corrects for primary and secondary inner filter effects by using Equation 4.1 Where F_{corr} is the sensor's corrected fluorescence signal, F_{obs} is the sensor's observed fluorescence signal, $A_{\lambda ex}$ is the absorbance at the excitation wavelength, and $A_{\lambda em}$ is the fluorescence observed at the emission wavelength. Primary inner filtering occurs when the excitation's beam light intensity is lost entering a sample in the cuvette (80). Secondary inner filtering happens when light that is emitted as fluorescence is re-absorbed by a sample in the cuvette (80).

$$F_{corr} = F_{obs} * 10^{(A_{\lambda ex} + A_{\lambda em})/2} \tag{4.1}$$

Since there is a correction factor for each specific excitation and emission wavelength, pairing a matrix of correction factors is multiplied in an element-wise manner with the sample EEM matrix in the R code.

1st and 2nd order Rayleigh scattering is another correction made in this R code. 1st order Rayleigh scattering seen in EEMs as a diagonal line of intensities in which the emission and excitation wavelength are equal and 2nd order Rayleigh scattering seen as a diagonal line of intensities where the emission wavelength is twice the wavelength of excitation (91). This R code masks 1st and 2nd order Rayleigh scattering by replacing the values for intensities in these regions with zero since that data does not represent the true fluorescence of a sample.

Furthermore, the fluorescence data from the cuvettes that have not been cleaned for at least one month, "dirty" cuvette, EEMs were adjusted by subtracting the blank EEMs after inner filter corrections were made (18). The blank EEMs were made by conducting this spectral analysis on cuvettes that were cleaned before the sensors were set up for the lab testing. Correcting with the blank removes Raman scattering and ensures that any fluorescence signal is not due to issues with the instrument.

4.2.3 Platting Biofilm Through Membrane Filtration

After the spectral analysis was finished, the biofilm that accumulated in the cuvette was plated. The cuvette was filled with a solution of PBS and DI water, plugges at either end, shaken for 30 secs to detach biofilm, and then the solution with biofilm was poured into a clean sample tube. This process was performed 15 times per cuvette. 10 mL of this solution was used for membrane filtration. This technique was used to enumerate *E. coli* and total coliforms (TC) present in the biofilm by plating a filter with m-ColiBlue24 broth (EPA Approved Hach Co.: 10029 method). Samples were plated in triplicates and incubated at 35°C for 18-24 hr.

Table 4.1: Composition of the Scaling Model Solution

	Scaling Model Solution
$CaCl_2$ (M)	1.67×10^{-2}
$MgSO_4$ (M)	1.05×10^{-2}
Na_2SO_4 (M)	1.45×10^{-2}

4.2.4 Mineral Scaling Formation

Mineral scaling experiments were conducted using a scaling model solution at a concentration factor of eight compared to the composition of the Colorado River consisting of calcium chloride dihydrate, magnesium sulfate, and sodium sulfate (66) the details of which are found in Table 4.1. The gypsum saturation index (SI) was calculated in terms of Equation 4.2 with visual MINTEQ to find the degree of supersaturation of the scaling model solution (32).

$$SI = \frac{(Ca^{2+})(SO_4^{2-})}{K_{sp,Ca_2SO_4}} \tag{4.2}$$

Where (Ca^{2+}) and (SO_4^{2-}) are the activities of each of the ions and K_{sp} is the solubility constant for gypsum. The saturation index was found to be 1.01 making it supersaturated. T

Before each experiment, DI water was recycled through the cuvette in the sensor for 2 hr. A portion of the scaling solution ($MgSO_4$ and Na_2SO_4 solutions) was added to the DI water and recycled for another 2 hr. Then the rest of the scaling model solution ($CaCl_2$ solution) was added and recycled through the cuvette for 24 hr to form scaling. To test the effects that mineral scaling has on the sensor's signal, 50 ppb tryptophan dilution and DI water alternated being pumped through the sensor in 30-min intervals. The impact on the sensor's signal will be examined by comparing the results from the same experiment conducted with a clean cuvette.

4.3 Results and Discussion

4.3.1 Biofouling Sensitivity

As wastewater effluent was introduced to the sensor for extended periods of time, an increase in the sensor's signal was observed even when only sampling DI water as seen in Figure 4.3 below.

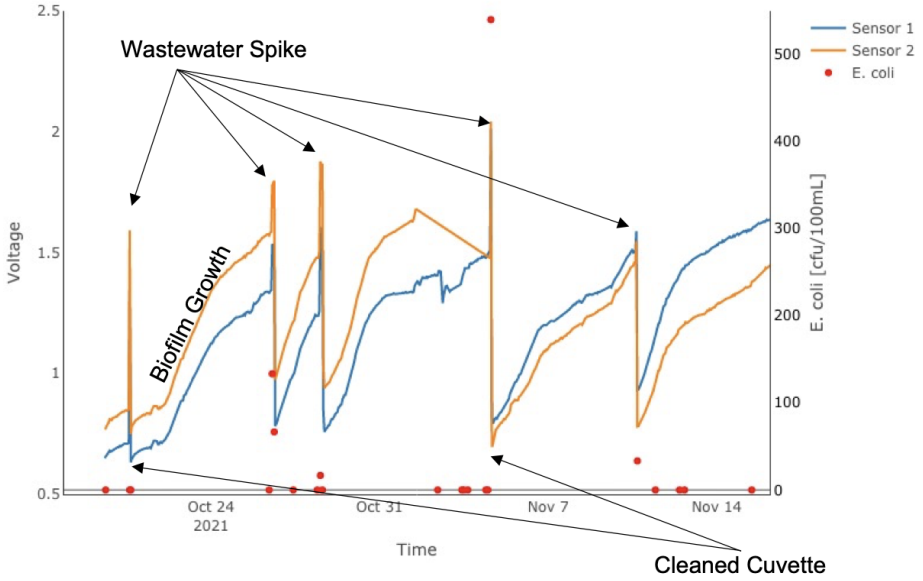

Figure 4.3: Data from two sensors sampling tap water spiked with wastewater effluent five times over the course of one month, combined with the *E. coli* data from samples taken before and after each spike.

There was an 82% increase calculated for the average amount of signal growth between wastewater spikes. This signal increase is predicted to be due to biofilm growth on the inside of the cuvette. This hypothesis is further investigated through spectral analysis and plate counts as described in the methods section. WWE contamination spikes can be observed in the sensor signal through signal increased caused by biofouling. For example, contamination events that have 17 CFU/100mL show a significant increase in signal, even with a fouled cuvette. There is an instance of sensor signal spike with no *E. coli* present in the wastewater effluent which may be characterized as a false positive. This could also indicate that there was a contamination event at a previous time, but all live fecal indicating bacteria had died off. The EEMs for the "dirty" cuvette using

the clean cuvette as a blank show an increase in signal at the excitation wavelength of 275 nm and emission wavelength of 340 nm which is the conditions the sensor operates at Figure 4.4.

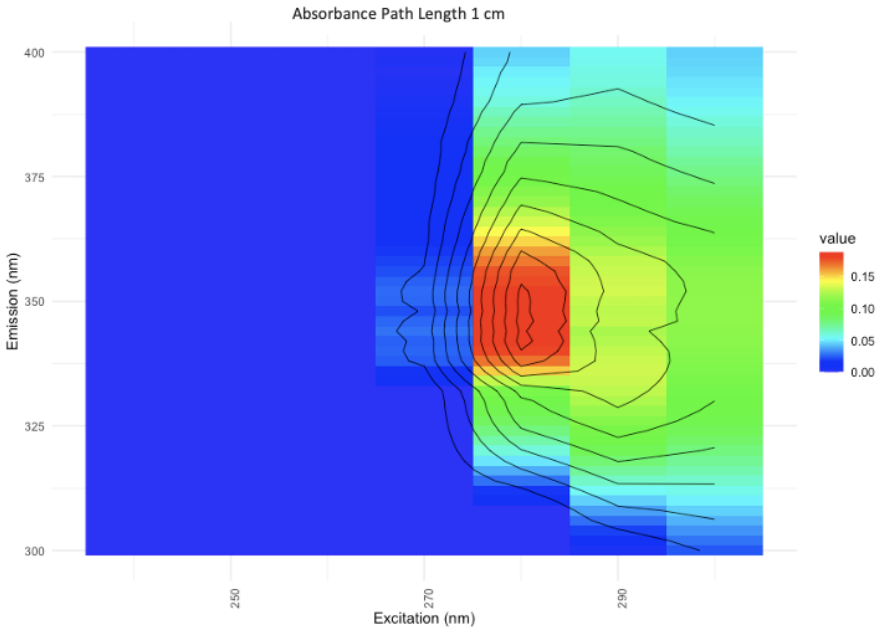

Figure 4.4: EEM of a cuvette that had been in a sensor after two weeks of sampling with wastewater effluent spikes in tap water with an absorbance path length of 1 cm.

Monitoring absorbance, with the UV-VIS, in a cuvette with biofilms present will impact the outcome of the Beer-Lambert law, as it relies on path length and the concentration of a solution inside the cuvette (47). Also, since the cuvettes contained only DI water, the increase in fluorescence at the excitation wavelength of 275 nm and emission wavelength of 340 nm is assumed to be due to a biofilm that has formed. Examining the results from the plate counts from this cuvette, which show that there is 10 CFU/100 mL of *E. coli*, further leads to the conclusion that biofilm growth is contributing to the increasing signal.

In short, when spiking the sensor with wastewater effluent, an increase in the sensor's signal occurred over time, even when there was not a contamination event. This increase in signal is predicted to be due to the formation of biofilm on the faces of the cuvette, further demonstrated by an increase in fluorescence of a cuvette filled with DI water observed by a benchtop flourimeter

and UV-VIS. This, paired with the results from platting, suggests that the fluorescence present is because of organic matter present on the lenses, not in the solution pumped through cuvette.

4.3.2 Mineral Scaling Sensitivity

After the mineral scaling solution was recycled through the sensor, a 50 ppb tryptophan dilution was pumped through the sensor and compared to data collected when the same dilution was run through a clean cuvette. When this data was compared it was concluded that a reduced sensitivity of the sensor by approximately 5% was observed. This decrease in sensitivity is predicted to be due to mineral scaling forming on the inside of the cuvette which increases attenuation. The 5% reduction in sensitivity was calculated by taking the difference in mean of the sensor's signal with a clean cuvette and a cuvette with mineral scaling between 50 ppb Tryptophan dilution and DI. This hypothesis is further investigated through spectral analysis through the EEM seen in Figure 4.5.

There is zero fluorescence at the excitation wavelength of 275 nm and emission wavelength of 340 nm in the EEM of the cuvette pictured in Figure 4.5. This, paired with the raw absorbance data Figure 4.6, led to the conclusion that when mineral scaling occurs there will be a decrease in sensitivity. Mineral scaling on the inside of the cuvette showed a 42.4% increase in absorbance at a wavelength of 275 nm and 38.1% increase at 360 nm. Percent increase was calculated using absorbance data collected on the UV-VIS before and after scaling occurred.

It is important to note that the EEMs for biofilm and mineral scaling should be seen as qualitative data. The intention of these EEMs was to examine the impact of biofouling and mineral scaling inside the cuvettes had on fluorescence measured by a standard instrument, then compare those results to the results observed by our sensor. These EEMs should not be used as a measurement to quantify what is present in biofilm or mineral scaling since neither of those are dissolved organic materials.

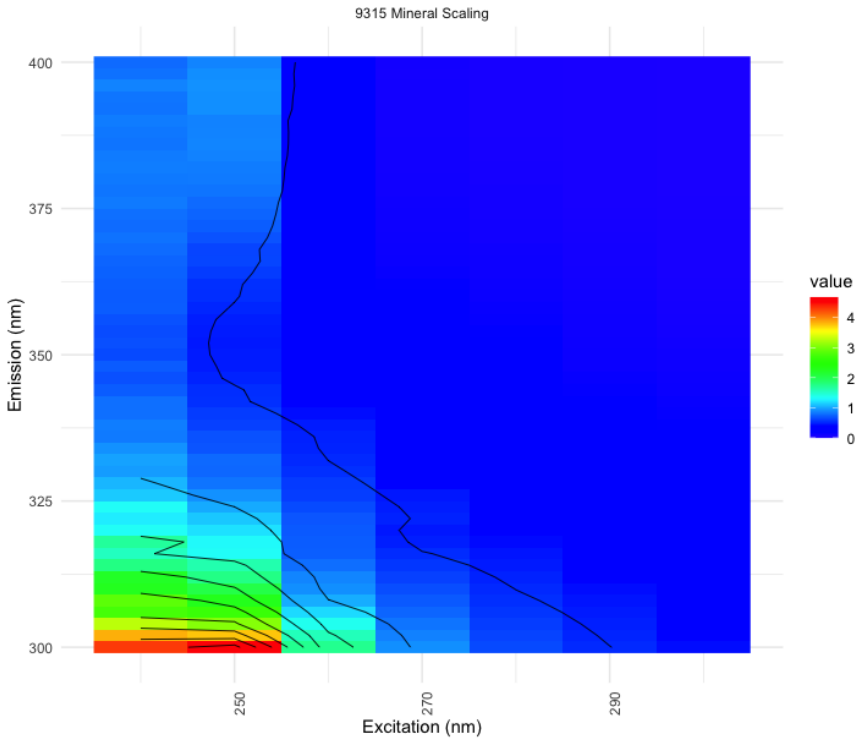

Figure 4.5: EEMs of a cuvette that had been in a sensor after one week of sampling with mineral scaling solution and DI water with an absorbance path length of 1 cm

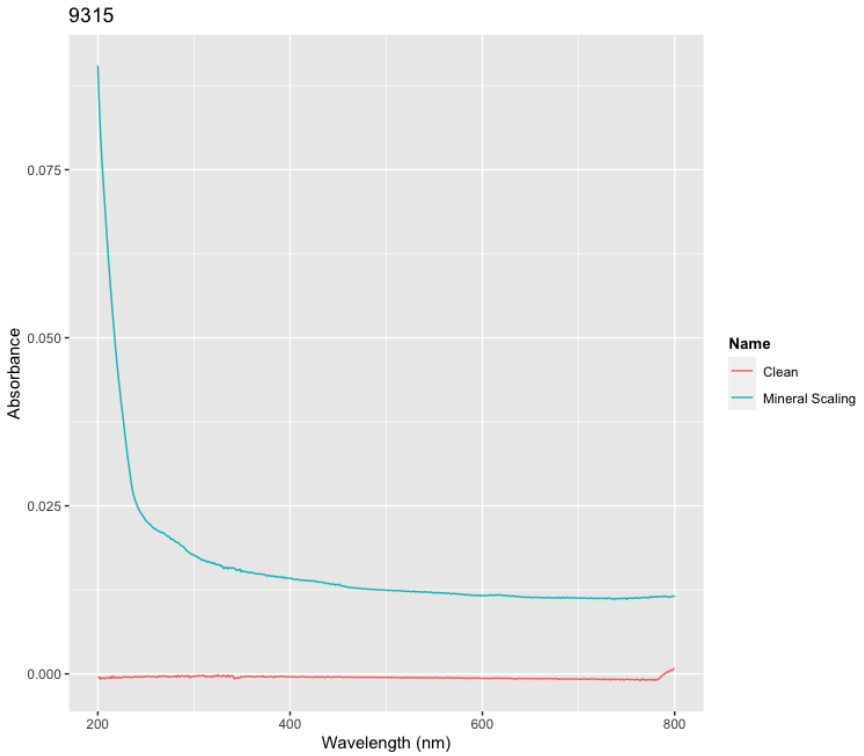

Figure 4.6: Raw absorbance data for a cuvette that had been in a sensor after one week of sampling with mineral scaling solution and DI water

Chapter 5

Field Validation

5.1 Introduction

In the summer of 2021, four sensors were placed on the Boulder Creek in Boulder, Colorado for 88 days to validate the sensor's functionality in the field.

5.2 Methods

5.2.0.1 Study Area

Boulder Creek is east of the Continental Divide in the Front Range of the Colorado Rocky Mountains. It flows out of the foothills of these Mountains and through Boulder, Colorado. Its flow is primarily derived from snow melt and minor springs west of the city of Boulder. Boulder Creek is an important water source for drinking-water supply, recreation, agriculture, and aquatic life. Boulder, Colorado has a semi-arid climate with an average rainfall of 21 inches annually (54). Figure 5.1 depicts where four sensors were placed on Boulder Creek to monitor the fecal contamination at sites upstream, within, and downstream of the city.

5.2.0.2 Sensor Sample Collection

Once in the field, each sensor was set up to sample the water in Boulder Creek every 10 min. The typical setup for each field sensor is described and pictures in Figure 1.2 in chapter 1. Every 10 min the sensors started the sampling sequence as laid out below:

Figure 5.1: Sensor and sampling locations along Boulder Creek in Boulder, Colorado, United States

Creek water was pumped through the sensor for 20 s to flush the cuvette and tubes. A 2 s wait time was set after pumping for 20 s so air bubbles were able to disperse. When taking a measurement, the LEDs were pulsed on for approximately 1000 ms, and 80 readings taken from the SiPM were averaged. With each measurement one data point was recorded at each current level. Creek water was then pumped through for another 5 s, as another measurement was taken. This sequence was repeated for three times every 10 min.

5.2.0.3 Ground Truth Sample Enumeration

To build and validate a machine learning model with the raw sensor data, a training and validation data set of laboratory enumerated microbial contamination was collected. Creek water samples were collected at each of the four sensor sites 13 times per week on average for 13 weeks. Each of these samples were collected each sensor's site within 1 min of the sensor collecting a fluorescence reading to appropriately match ground truth samples to sensor measurements. mWater, a phone based survey tool (www.mWater.co) was used to record and organize sampling data. Samples were collected in labeled 50 mL sample tubes, put into a cooler containing an ice pack, and transported to the lab for processing within two hours of collection.

Membrane filtration was used to enumerate *E. coli* and total coliforms (TC) present in the biofilm by plating a filter with m-ColiBlue24 broth (EPA Approved Hach Co.: 10029 method). m-ColiBlue24 broth indicates *E. coli* colonies by blue coloration resulting from specific activity of β -glucuronidase and TC by red coloration resulting from specific activity of β -galactosidase (76). Samples were plated in triplicates and incubated at 35°C for 18-24 hr. 10 mL of sampled creek water was filtered through a 0.45 micrometer filter then incubated using methods described above.

Plates were enumerated by counting the *E. coli* coliforms and TC was present after incubation and that data was recorded in mWater so it could be properly matched with the sample collection time. Any major event or change to each sensor was also recorded in mWater, such as the date and time of sensor installation, removal, replacement, and cleaning.

5.3 Results and Discussion

5.3.1 Field Validation

Ground truth measurements of *E. coli* in Boulder Creek ranged from 0 CFUs/100 mL to 9580 CFUs/100 mL with an average of 120 CFUs/100 mL over 13 weeks of 298 independent observations (between 69 and 82 from each sampling site). This data was sorted between WHO risk categories of very low, low, intermediate, high, and very high risk, which are detailed in Table 5.1. There were few observations made at the very low and very high risk categories. The highest *E. coli* measurement, 9580 CFUs/100 mL, which at the sensor location cloest to main campus.

Table 5.1: Prevalence and distribution of fecal contamination risk categories observed during field experimentation.

WHO Risk Category	*E. coli* CFUs/100 mL	# in Sample	% in Sample
Very low	0 - 1	19	6.4%
Low	1 - 10	18	6.0%
Intermediate	11 - 100	103	35%
High	101 - 1000	145	49%
Very high	1000+	13	4.4%
Total		298	100%

The time the plated creek water samples that contained *E. coli* were matched with the sensor measurement taken at that same time as logged in mWater. This subset formed the training and testing data-sets for the machine learning model. Due to the fact that *E. coli* presence in the creek water was the desired response variable for this study, several inputs were hypothesized to explain whether or not contamination was present at the time of testing. Some of these inputs were continuous TLF measurements from the sensors that were present in the creek.

As mentioned before, the sensor recorded voltage outputs at four current levels to the LEDs: 10, 50, 100, and 200 mA at each reading. However, in the context of field validation, not one of the current input levels performed better than any of the the others. Due to this trend, the voltage output data was normalized by each current level and was averaged so it could be summarized at one characteristic voltage per *E. coli* enumeration. Before being normalized, the voltage data from

the sensor was standardized to 20°C using water temperature measurements.

A binary model was developed to predict if each sample was ≥ 100 CFUs *E. coli* / 100 mL vs. <100 CFUs / 100 mL, which is at least at the WHO high risk level. Another model for multinomial classification investigated performance at predicting the correct of five WHO risk categories described in Table 5.1. Thus, continuous *E. coli* measurements were characterized by the number of CFUs per 100 mL of water at very low risk, low risk, intermediate risk, high risk, and very high risk (86).

The training set of data was paired with features, which are other data that can be predictive of a certain outcome worked as the foundation of this machine learning model. In the case of this study the outcome is the ability for this sensor to predict WHO fecal contamination risk categories. These features, in order of relative importance included: fluorescence emission (sensor voltage), hours since cuvette was cleaned, relative voltage (percentile), temperature of the sensor, relative voltage (difference from a 7 day average), day of the season, bias volatage, relative voltage (z-score), rainfall, and temperature of water.

The feature fluorescence emission is the retrieved voltage from the in-situ sensor, once it was standardized and normalized. Relative voltage was explained by the z-score, difference from rolling 7-day average, and percentile. As seen in Chapter 4, biofouling and mineral scaling almost certainly impact the sensor signal. This is further seen in the field validation data when the number of hours since cleaning of the cuvette for each sensor was included as one of the features in the machine learning model as a proxy for fouling (17). An estimate of daily municipal rainfall included as a feature was collected from NOAA Physical Sciences Laboratory. It is well known that fecal contamination varies seasonally with rainfall (39). Water temperature corrected to 20°C using Equation (3.1) was used as a feature. The temperature of the SiPM will impact its output; for this reason, temperature inside the sensor enclosure was included as a feature (43).

5.3.2 Machine Learning Model

The machine learning model was developed by Katie Fankhauser and Emily Bedell using TLF from the sensors along Boulder Creek as the principal feature were able to identify high risk of fecal contamination in natural waters with outstanding accuracy. The ROC curve (Figure 5.2) demonstrates that both high sensitivity, true positive rate, and specificity, true negative rate, is achievable. At the chosen predicted probability threshold, the sensitivity and specificity were respectively 80% and 86%. It also shows, from the area under the curve, the model had a probability of 86% of accurately discriminating high contamination risk.

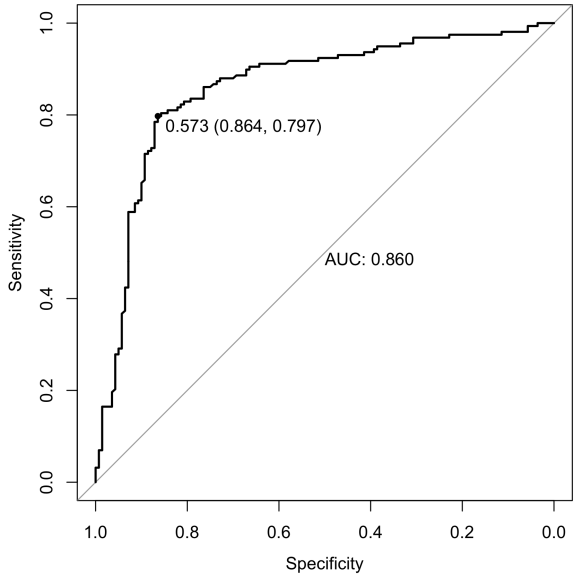

Figure 5.2: Receiver Operating Characteristic (ROC) curve. The point on the curve indicate the probability threshold used to potentially categorize high fecal contamination risk and the test specificity and sensitivity. The area under the curve, a measure of test discrimination, is stated on the graph.

This model had a 64% accuracy in distinguishing between the WHO establish risk categiories that are detailed in Table 5.1.

Chapter 6

Applications and Future Work

6.1 Drinking Water Quality

As of 2019 it is estimated that at least 2 billion people rely on a drinking water source that is contaminated with fecal matter (86). Exposure to these pathogens can cause diarrhea, which is the eighth leading cause of death for all ages and the third leading cause of death for children under 5 years old (87; 88).

The research presented here establishes that a continuous, in-situ, near-time fluorescence sensor coupled with a machine learning model can detect fecal contamination in water sources with high accuracy. This TLF sensor is able to reliably predict high risk fecal contamination in surface water and treated tap water that has been spiked with wastewater effluent, simulating a contamination event. The device and machine learning model presented display a contribution to existing research and knowledge through a novel, in-situ, remotely reporting sensor that when coupled with a machine learning model is highly reliable and accurate in performance.

6.2 Applications of this Sensor in Drinking Water Services

Households globally are served by varying water technologies and service levels. High-tech systems refer to infrastructure such as centralized pipes where municipalities or private companies are responsible for the funding and technical support, and certified operators conduct operation and maintenance, leaving less responsibility to homeowners. Mid-tech refers to solutions that provide advances in access to water and sanitation services, while making moderate use of resources,

materials, and technology when compared to high-tech systems. Low-tech systems offer users a minimum level of service that may not provide all of the health benefits seen when using high or mid-tech solutions (50).

Over the years, many environmental health and engineering interventions have been introduced to improve health outcomes in communities along this spectrum, but these interventions often fall short of expectations, lacking resilience as measured by sustained delivery, cost-effectiveness, positive health outcomes, or other intended impacts (30), (71). There are a plethora of examples of technology that have failed due to preventable issues, such as lack of operation and maintenance budgets, limited supplies and expertise, lack of supply chains for replacement parts, and conflicting management priorities (13). Electronic sensors have been developed and applied within environmental health programs to support health studies and provisioning of basic services (4), (79). Sensors are most certainly not a comprehensive or sufficient measurement tool on their own, but they have provided insights that can help monitor and refine interventions, especially when paired with citizen science (8), (69).

As previously mentioned, it is important for the health of every community for water service providers or even homeowners to monitor fecal contamination in drinking water, which is traditionally time consuming and expensive (11). The sensor, coupled with a machine learning model, proposed and studied in this research showed promising results and could be used as an alternative to traditional microbiological testing in high-, mid-, and low-tech contexts.

This sensor has the potential to be used anywhere in the world; however, the examples I will primarily focus on moving forward in my thesis will be communities in the United States. This focus is due to the wide range of access to clean and reliable drinking water across the nation and the fact that I have the most experience working and living in the U.S. and would simply feel uncomfortable speaking for countries where I am unfamiliar with the language, culture, and people and offering uninformed solutions to challenges of which I do not fully understand the context.

6.2.1 High-Tech

Those who are fortunate enough to have in-home access to clean and reliable drinking water in the United States more than likely rely on piped services provided by the local drinking water treatment plant. Most drinking water is treated by a similar process, but since the testing of this sensor occurred in Boulder, Colorado, the focus will be on the exact process the City of Boulder uses to treat drinking water. As raw water comes in, chlorine is added as pre-disinfection at different locations in the treatment process to kill microorganisms and prevent new ones from growing. Next, coagulants are added to water to form large collections of suspended particles referred to as "flocs". As heavier flocs settle at the bottom of the basins, Dissolved Air Flotation units are deployed to produce micro-bubbles that attach to the flocs and float them to the surface so they can form a removable "sludge blanket." After the sludge blanket is removed, water flows to filters that remove smaller floc particles. After filtration more chlorine is added in post-disinfection to ensure the water will remain clean in the piping and storage systems. This is also the time that additives such as fluoride and corrosion controlling chemicals are added (16).

The City of Boulder routinely monitors for contaminants in drinking water following state and national regulations. However, even with this extensive treatment process and regular water quality monitoring, there was still 1 sample, of 124 samples, that tested positive for Total Coliform Bacteria according to the 2021 Drinking Water Quality Report (15). While this is an extremely low occurrence, it does highlight the importance of monitoring drinking water quality even in high-tech systems. Another fact to note is that there were 124 samples collected for this annual report. If this system incorporated the sensor, then there could be a nearly continuous monitoring of the water quality which would alert operators of a contamination event as soon as it happens.

6.2.2 Mid- and Low-Tech

There is still a significant portion of the population living in the United States that does not have basic access to safe drinking water and sanitation. In fact, 17% of people living in rural areas

report having issues with access to safe drinking water. This is even more staggering for Native American households, which are 19 times more likely to lack complete plumbing when compared to white households. This statistic stresses that race is the strongest predictor of water and sanitation access in the United States, a so-called "developed" or "first world" country (52).

In response to these shortcomings, mid- and low-tech systems have been introduced in households and communities that lack piped water.

6.2.2.1 Alaska Native Tribal Health Consortium - Practicum

In rural Alaska there are over 3,000 households rely on low-tech solutions with no access to piped water or sewer services whatsoever (1). This inadequate access to WASH services is a public health crisis and as previously mentioned has led to the introduction of mid- and low-tech water treatment and sanitation systems at the house hold level.

One example of this is the Portable Alternative Sanitation System (PASS) pictured in Figure 6.1 (2). This system was was engineered by The Alaska Native Tribal Health Consortium (ANTHC), where completed my practicum in the summer of 2021.

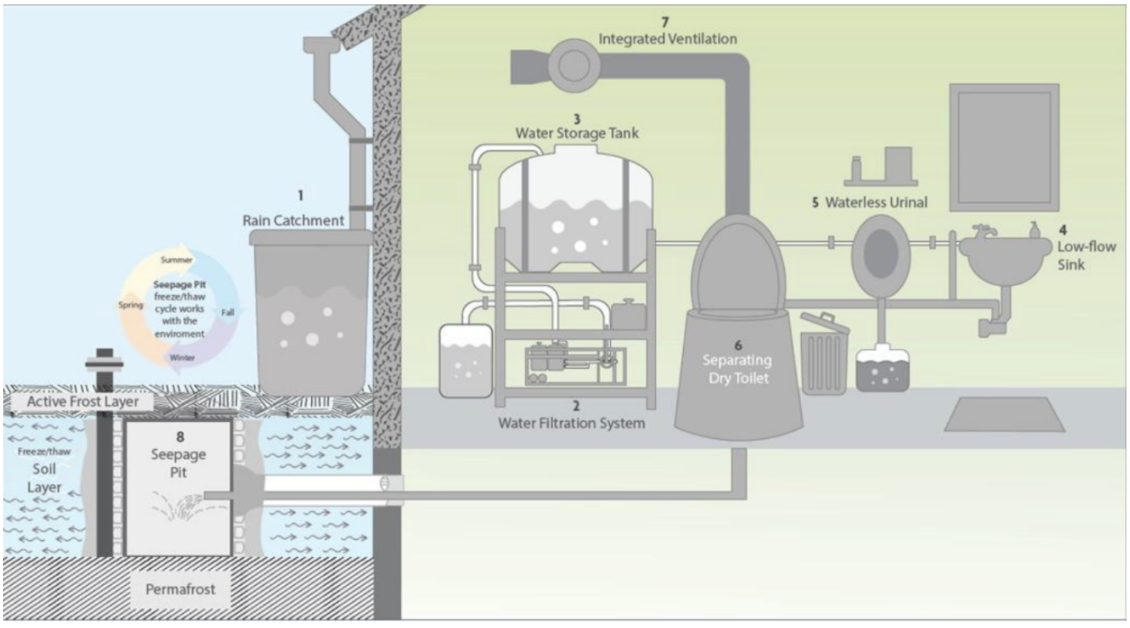

Figure 6.1: Diagram of the Portable Alternative Sanitation System

This system includes a small water treatment system that incorporates two cryptosporidium-rated filters and disinfection using a chlorine (bleach) injector. Water is pumped through the treatment system then into the water storage tank elevated on top of the water treatment system. When needed, this water flows from the water storage tank to the sink by gravity. The PASS toilet can operate in 3 different modes: urine-diverting drying toilet (UDDT), container mode, or ventilated honey bucket. If the system is operated as a UDDT, urine is diverted outside of the home to an underground seepage pit, and feces is disposed of in a honey bucket located inside the separating dry toilet. If the system is operated in container mode, urine is diverted into a urine storage container and feces are disposed of in a ventilated honey bucket. When the system is in ventilated honey bucket mode, both the urine and feces will be disposed of in a honey bucket located in the separating dry toilet, and, as in all the other modes, the vent is open and a fan is turned on, eliminating the need for chemicals to mask odor. The seepage pit is a vertically oriented perforated pipe located 4 feet below the ground surface designed to infiltrate the surrounding soil with only small amounts of greywater or urine. Recently, mini-PASS was developed. It is essentially the same as full PASS but only offers ventilated honey bucket mode, so it does not require a seepage pit.

During my practicum, I collaborated with the research and engineering departments at ANTHC developing materials for homeowners, ANTHC employees, and donors and made improvements on the PASS and mini-PASS. This work led me to a position with ANTHC as a Research Associate. Where I currently work collaboratively with other ANTHC research and clinical staff to ensure the successful completion of research projects examining the public health effects of PASS and mini-PASS.

While Figure 6.1 depicts a rain catchment system, it has been my experience working with ANTHC that many communities that use PASS rely on many different water sources to fill their water storage tank. This water may be hauled from a community water treatment plant, rain catchment systems, local rivers, streams, or from melted ice. Hauling water limits the amount of water community members use each day to between 3 and 5.4 gallons per capita per day (gpcpd), due to the time, effort, limited storage, and cost of hauling (51). When water is collected from a

watering point the water might be considered safe but can be contaminated when hauling or during storage. By adding the sensor discussed in this study, homeowners would be able to monitor the quality of water in their PASS system, which could lower the rates of illnesses associated with fecal contamination in drinking water.

6.2.2.2 Well Water Monitoring in the United States

The rise in exposure to various contaminates in drinking water have led to the increased popularity of point-of-use (POU) drinking water treatment in the US. Notably over 43 million people living in the US use private wells as their main drinking water source (89). These domestic wells are not typically regulated by federal or state laws and homeowners are the ones responsible for testing their wells for contaminants. Due to this many private wells are not regularly tested which can lead to many different health risks. With addition of this sensor once adapted could be installed in wells and the main water pipes in homes to alert homeowners of *E. coli* contamination events.

Due to the formations of biofilms and mineral scaling a cleaning procedure will need to be added to the current and any future designs. The machine learning aspects of this model should be able to account for some of the biofouling and scaling but overtime a cleaning function will be necessary. One way the cuvette could be cleaned is by the user injecting a cleaning solution through the sensor. This would work similarly to the chlorine injection used in PASS.

6.3 Looking Forward

Before these applications can be made, there are further studies that should be conducted. For example, there is substantial background fluorescence noise in surface water due to natural organic matter present. The study presented in Boulder, Creek is an example of one of these more difficult environments, especially when compared to groundwater or treated tap water (84). Future work would be testing the sensor and machine learning model's ability to predict various contamination in other types of drinking water environments.

Bibliography

[1] Alaska Native Tribal Health Consortium. Healthy Alaskans — Alaska Native Tribal Health Consortium.

[2] Alaska Native Tribal Health Consortium. Portable Alternative Sanitation System connects in-home sanitation systems where it was impossible before, 2019.

[3] Alexandra Alimova, A Katz, Masood Siddique, Glenn Minko, Howard E Savage, Menhendra K Shah, Richard B Rosen, and Robert R Alfano. Native Fluorescence Changes Induced by Bactericidal Agents. IEEE SENSORS JOURNAL, 5(4), 2005.

[4] Luis Andres, Kwasi Boateng, Christian Borja-Vega, and Evan Thomas. A review of in-situ and remote sensing technologies to monitor water and sanitation interventions. Water (Switzerland), 10(6), 6 2018.

[5] Robert Bain, Ryan Cronk, Jim Wright, Hong Yang, Tom Slaymaker, and Jamie Bartram. Fecal Contamination of Drinking-Water in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis. PLoS Medicine, 11(5), 2014.

[6] Andy Baker. Thermal fluorescence quenching properties of dissolved organic matter. Water Research, 39(18):4405–4412, 11 2005.

[7] Andy Baker, Sarah Elliott, and Jamie R. Lead. Effects of filtration and pH perturbation on freshwater organic matter fluorescence. Chemosphere, 67(10):2035–2043, 2007.

[8] Timothy M Barzyk, Hongtai Huang, Ronald Williams, Amanda Kaufman, and Jonathan Essoka. Advice and Frequently Asked Questions (FAQs) for Citizen-Science Environmental Health Assessments. 2018.

[9] Emily Bedell, Karl Linden, Julie Korak, Claire Monteleoni, and Joe Brown. Design, optimization, and implementation of a near-time reporting in-situ fecal contamination risk alarm sensor for drinking water Background/Problem Statement. Technical report.

[10] Emily Bedell, Taylor Sharpe, Timothy Purvis, Joe Brown, and Evan Thomas. Demonstration of tryptophan-like fluorescence sensor concepts for fecal exposure detection in drinkingwater in remote and resource constrained settings. Sustainability (Switzerland), 12(9), 5 2020.

[11] Marie Claude Besner, Michèle Prévost, and Stig Regli. Assessing the public health risk of microbial intrusion events in distribution systems: Conceptual model, available data, and challenges, 2011.

[12] Martijn Boerkamp, David W Lamb, Peter G Lye, Christopher M Fellows, Ali Al-Hamzah, and Andrew D Wallace. Detecting and Monitoring Industrial Scale Formation Using an Intrinsic Exposed-Core Optical Fiber Sensor.

[13] Hilary Schaffer Boudet, Dilanka Chinthana Jayasundera, and Jennifer Davis. Drivers of Conflict in Developing Country Infrastructure Projects: Experience from the Water and Pipeline Sectors. Journal of Construction Engineering and Management, 137(7):498–511, 7 2011.

[14] City of Boulder. Water Treatment — City of Boulder.

[15] City of Boulder. Drinking Water Quality Report. Technical report, 2021.

[16] City of Boulder. Water Treatment — City of Boulder, 2021.

[17] Paula Coble, Jamie Lead, Andy Baker, Darren Reynolds, and Robert Spencer. Aquatic Organic Matter Fluorescence. Cambridge University Press, Cambridge, 2014.

[18] Paula G. Coble. Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Marine Chemistry, 51(4):325–346, 1 1996.

[19] Gunther F. Craun, Joan M. Brunkard, Jonathan S. Yoder, Virginia A. Roberts, Joe Carpenter, Tim Wade, Rebecca L. Calderon, Jacquelin M. Roberts, Michael J. Beach, and Sharon L. Roy. Causes of outbreaks associated with drinking water in the United States from 1971 to 2006, 2010.

[20] Caroline Delaire, Rachel Peletz, Emily Kumpel, Joyce Kisiangani, Robert Bain, and Ranjiv Khush. How Much Will It Cost to Monitor Microbial Drinking Water Quality in Sub-Saharan Africa? Environmental Science and Technology, 51(11):5869–5878, 6 2017.

[21] Isabel Domínguez, Edgar Ricardo Oviedo-Ocaña, Karen Hurtado, Andrés Barón, and Ralph P. Hall. Assessing sustainability in rural water supply systems in developing countries using a novel tool based on multi-criteria analysis. Sustainability (Switzerland), 11(19), 10 2019.

[22] Bryan D. Downing, Brian A. Pellerin, Brian A. Bergamaschi, John Franco Saraceno, and Tamara E.C. Kraus. Seeing the light: The effects of particles, dissolved materials, and temperature on in situ measurements of DOM fluorescence in rivers and streams. Limnology and Oceanography: Methods, 10(OCTOBER):767–775, 2012.

[23] Edmund. Bandpass Filter.

[24] Matthias Fischer, Martin Wahl, and Gernot Friedrichs. Design and field application of a UV-LED based optical fiber biofilm sensor. Biosensors and Bioelectronics, 33(1):172–178, 3 2012.

[25] Matthias Fischer, Martin Wahl, and Gernot Friedrichs. Design and field application of a UV-LED based optical fiber biofilm sensor. Biosensors and Bioelectronics, 33(1):172–178, 3 2012.

[26] H Flemming. Biofouling in water systems - cases, causes and conutermeasures.

[27] Hans Curt Flemming, Thomas R. Neu, and Daniel J. Wozniak. The EPS matrix: The "House of Biofilm Cells", 11 2007.

[28] William T. Foreman, Teresa L. Williams, Edward T. Furlong, Dawn M. Hemmerle, Sarah J. Stetson, Virendra K. Jha, Mary C. Noriega, Jessica A. Decess, Carmen Reed-Parker, and Mark W. Sandstrom. Comparison of detection limits estimated using single- and multi-concentration spike-based and blank-based procedures. Talanta, 228:122139, 6 2021.

[29] B. G. Fox, R. M.S. Thorn, A. M. Anesio, and D. M. Reynolds. The in situ bacterial production of fluorescent organic matter; an investigation at a species level. Water Research, 125:350–359, 2017.

[30] Joshua V. Garn, Gloria D. Sclar, Matthew C. Freeman, Gauthami Penakalapati, Kelly T. Alexander, Patrick Brooks, Eva A. Rehfuess, Sophie Boisson, Kate O. Medlicott, and Thomas F. Clasen. The impact of sanitation interventions on latrine coverage and latrine use: A systematic review and meta-analysis. International Journal of Hygiene and Environmental Health, 220(2):329–340, 4 2017.

[31] Andrew B. Gray, Gregory B. Pasternack, and Elizabeth B. Watson. Hydrogen peroxide treatment effects on the particle size distribution of alluvial and marsh sediments. Holocene, 20(2):293–301, 3 2010.

[32] Jon Gustafsson. Visual MINTEQ – Visual MINTEQ – a free equilibrium speciation model, 2013.

[33] Richard A. Harshman. PARAFAC: An "Explanatory" Factor Analysis Procedure. The Journal of the Acoustical Society of America, 50(1A):117–117, 7 1971.

[34] R. K. Henderson, A. Baker, K. R. Murphy, A. Hambly, R. M. Stuetz, and S. J. Khan. Fluorescence as a potential monitoring tool for recycled water systems: A review, 2009.

[35] Naomi Hudson, Andy Baker, and Darren Reynolds. Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters - A review, 7 2007.

[36] Tyler D. Johnson, Kenneth Belitz, and Melissa A. Lombard. Estimating domestic well locations and populations served in the contiguous U.S. for years 2000 and 2010. Science of the Total Environment, 687:1261–1273, 10 2019.

[37] Kathleen Joslyn and John Lipor. A Supervised Learning Approach to Water Quality Parameter Prediction and Fault Detection. In Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018, pages 2511–2514. Institute of Electrical and Electronics Engineers Inc., 1 2019.

[38] K. Khamis, J. P.R. Sorensen, C. Bradley, D. M. Hannah, D. J. Lapworth, and R. Stevens. In situ tryptophan-like fluorometers: Assessing turbidity and temperature effects for freshwater applications. Environmental Sciences: Processes and Impacts, 17(4):740–752, 4 2015.

[39] Caroline Kostyla, Rob Bain, Ryan Cronk, and Jamie Bartram. Seasonal variation of fecal contamination in drinking water sources in developing countries: A systematic review, 5 2015.

[40] Dolly Kothawala, Kathleen Murphy, Colin Stedmon, Gesa Weyhenmeyer, and Lars Tranvik. Inner filter correction of dissolved organic matter fluorescence. 2013.

[41] P. Kuhnert, J. Nicolet, and J. Frey. Rapid and accurate identification of Escherichia coli K-12 strains. Applied and Environmental Microbiology, 61(11):4135–4139, 1995.

[42] Peter Kuhnert, Jacques Nicolet, and Joachim Frey. Rapid and Accurate Identification of Escherichia coli K-12 Strains. Technical Report 11, 1995.

[43] Evgeny Kuznetsov. Temperature-compensated silicon photomultiplier. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 912:226–230, 12 2018.

[44] Y A L E E N H E E R, § A N D K A R L B O O K S H —. Fluorescence Excitation-Emission Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter. 2003.

[45] Joseph R Lakowicz. Principles of Fluorescence Spectroscopy Third Edition. Technical report.

[46] Jin Hyung Lee, Thomas K. Wood, and Jintae Lee. Roles of indole as an interspecies and interkingdom signaling molecule, 2015.

[47] J. H.T. Luong, K. A. Mahmoud, and K. B. Male. Instrumentation and Analytical Methods. Comprehensive Biotechnology, Second Edition, 2:829–838, 9 2011.

[48] J. H.T. Luong, K. A. Mahmoud, and K. B. Male. Instrumentation and Analytical Methods. Comprehensive Biotechnology, Second Edition, 2:829–838, 9 2011.

[49] Jitka Macadam and Simon A Parsons. Calcium carbonate scale formation and control. Technical report.

[50] Kaitlin Mattos, John Warren, Laura Eichelberger, Jessica Kaminsky, and Karl G. Linden. Pathways to the successful function and use of mid-tech household water and sanitation systems. Journal of Water Sanitation and Hygiene for Development, 11(6):994–1005, 11 2021.

[51] Kaitlin J Mattos, Laura Eichelberger, John Warren, Aaron Dotson, Millie Hawley, and Karl G Linden. Household Water, Sanitation, and Hygiene Practices Impact Pathogen Exposure in Remote, Rural, Unpiped Communities.

[52] George Mcgraw. Closing the Water Access Gap in the United States Chief Executive Officer DigDeep Radhika Fox Chief Executive Officer US Water Alliance 4 Closing the Water Access Gap in the United States: A National Action Plan ACKNOWLEDGMENTS. Technical report.

[53] J. Tom Mueller and Stephen Gasteyer. The widespread and unjust drinking water and clean water crisis in the United States. Nature Communications, 12(1), 12 2021.

[54] Sheila F. Murphy. State of the watershed : water quality of Boulder Creek, Colorado. U.S. Dept. of the Interior, U.S. Geological Survey, 2006.

[55] Takuya Okazaki. Fundamental Study on the Development of Fiber Optic Sensor for Real-time Sensin g of CaCO3 Scale Formation in Geothermal Water.

[56] Takuya Okazaki, Kenichiro Imai, Shin Tan, Yun Yong, Faidz Rahman, Noriko Hata, Shigeru Taguchi, Ueda Akira, and Hideki Kuramitz. Fundamental Study on the Development of Fiber Optic Sensor f or Real-time Sensin g of CaCO3 S cale Formation in Geothermal Water. Technical report, 2015.

[57] Takuya Okazaki, Tatsuya Orii, Akira Ueda, Akiko Ozawa, and Hideki Kuramitz. Fiber Optic Sensor for Real-Time Sensing of Silica Scale Formation in Geothermal Water. Scientific Reports, 7(1), 12 2017.

[58] N. Patel-Sorrentino, S. Mounier, Y. Lucas, and J. Y. Benaim. Effects of UV-visible irradiation on natural organic matter from the Amazon basin. Science of the Total Environment, 321(1-3):231–239, 4 2004.

[59] Nicole T. Perna, Jeremy D. Glasner, Valerie Burland, and Guy Plunkett. The Genomes of Escherichia coli K-12 and Pathogenic E. coli. Escherichia Coli, pages 3–53, 2002.

[60] Amanda Pezzuto and Emily Sarver. A lab study of mineral scale buildup on lined and traditional PE water pipes for acid mine drainage treatment applications. Journal of Sustainable Mining, 19(1):33–45, 2020.

[61] R Philip-Chandy, P J Scully, and D Thomas. A novel technique for on-line measurement of scaling using a multimode optical ®bre sensor for industrial applications. Technical report.

[62] Annette Prüss-Ustün, Jennyfer Wolf, Jamie Bartram, Thomas Clasen, Oliver Cumming, Matthew C. Freeman, Bruce Gordon, Paul R. Hunter, Kate Medlicott, and Richard Johnston. Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: An updated analysis with a focus on low- and middle-income countries. International Journal of Hygiene and Environmental Health, 222(5):765–777, 6 2019.

[63] Annette Pr€ Uss-Ust€ Un, Jamie Bartram, Thomas Clasen, John M Colford, Oliver Cumming, Valerie Curtis, Sophie Bonjour, Alan D Dangour, Jennifer De France, Lorna Fewtrell, Matthew C Freeman, Bruce Gordon, Paul R Hunter, Richard B Johnston, Colin Mathers, Daniel M€ Ausezahl, Kate Medlicott, Maria Neira, Meredith Stocks, Jennyfer Wolf, and Sandy Cairncross. Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries. 2014.

[64] Matthias Pucher, Urban Wünsch, Gabriele Weigelhofer, Kathleen Murphy, Thomas Hein, and Daniel Graeber. StaRdom: Versatile software for analyzing spectroscopic data of dissolved organic matter in R. Water (Switzerland), 11(11), 11 2019.

[65] PureTec. The Relationship Between ph an Deionized Water. 2010.

[66] Anditya Rahardianto, Wen Yi Shih, Ron Wai Lee, and Yoram Cohen. Diagnostic characterization of gypsum scale formation and control in RO membrane desalination of brackish water. Journal of Membrane Science, 279(1-2):655–668, 8 2006.

[67] Tara Randall, Karl G Linden, Joann Silverstein, and Cresten Mansfeldt. HIGH IMPACT UVLED APPLICATIONS: RECLAIMED WASTEWATER EFFLUENT IRRIGATION. Technical report.

[68] D. M. Reynolds. Rapid and direct determination of tryptophan in water using synchronous fluorescence spectroscopy. Water Research, 37(13):3055–3060, 2003.

[69] A. Ripoll, M. Viana, M. Padrosa, X. Querol, A. Minutolo, K. M. Hou, J. M. Barcelo-Ordinas, and J. Garcia-Vidal. Testing the performance of sensors for ozone pollution monitoring in a citizen science approach. Science of The Total Environment, 651:1166–1179, 2 2019.

[70] John Franco Saraceno, Brian A. Pellerin, Bryan D. Downing, Emmanuel Boss, Philip A.M. Bachand, and Brian A. Bergamaschi. High-frequency in situ optical measurements during

a storm event: Assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes. Journal of Geophysical Research: Biogeosciences, 114(4), 2009.

[71] Ryan W. Schweitzer and James R. Mihelcic. Assessing sustainability of community management of rural water systems in the developing world. Journal of Water Sanitation and Hygiene for Development, 2(1):20–30, 2012.

[72] J. P.R. Sorensen, A. Sadhu, G. Sampath, S. Sugden, S. Dutta Gupta, D. J. Lapworth, B. P. Marchant, and S. Pedley. Are sanitation interventions a threat to drinking water supplies in rural India? An application of tryptophan-like fluorescence. Water Research, 88:923–932, 1 2016.

[73] James P R Sorensen, Andy Baker, Susan A Cumberland, Dan J Lapworth, Alan M Macdonald, Steve Pedley, Richard G Taylor, and Jade S T Ward. Real-time detection of faecally contaminated drinking water with tryptophan-like fluorescence: defining threshold values. Technical report.

[74] James P.R. Sorensen, Andy Baker, Susan A. Cumberland, Dan J. Lapworth, Alan M. MacDonald, Steve Pedley, Richard G. Taylor, and Jade S.T. Ward. Real-time detection of faecally contaminated drinking water with tryptophan-like fluorescence: defining threshold values. Science of the Total Environment, 622-623:1250–1257, 5 2018.

[75] James P.R. Sorensen, Mor Talla Diaw, Abdoulaye Pouye, Raphaëlle Roffo, Djim M.L. Diongue, Seynabou C. Faye, Cheikh B. Gaye, Bethany G. Fox, Timothy Goodall, Daniel J. Lapworth, Alan M. MacDonald, Daniel S. Read, Lena Ciric, and Richard G. Taylor. In-situ fluorescence spectroscopy indicates total bacterial abundance and dissolved organic carbon. Science of the Total Environment, 738, 10 2020.

[76] Pam Tallon, Brenda Magajna, Cassandra Lofranco, and Kam Tin Leung. MICROBIAL INDICATORS OF FAECAL CONTAMINATION IN WATER: A CURRENT PERSPECTIVE. Technical report.

[77] Hsiou kun Tan, Willis B. Wheeler, and Cheng i. Wei. Reaction of chlorine dioxide with amino acids and peptides: Kinetics and mutagenicity studies. Mutation Research/Genetic Toxicology, 188(4):259–266, 8 1987.

[78] Scott D. Tanner, Vladimir I. Baranov, and Dmitry R. Bandura. Reaction cells and collision cells for ICP-MS: a tutorial review. Spectrochimica Acta Part B: Atomic Spectroscopy, 57(9):1361–1452, 9 2002.

[79] Evan Thomas, Luis Alberto Andrés, Christian Borja-Vega, and Germán Sturzenegger. Innovations in WASH Impact Measures Water and Sanitation Measurement Technologies and Practices to Inform the Sustainable Development Goals D I R E C T I O N S I N D E V E L O P M E N T Infrastructure. Technical report.

[80] Sheryl A Tucker, Vicki L Amszi, and William E Acree. The Modern Student laboratory: Fluorescence Spectroscopy Primary and Secondary Inner Filtering Effect of K2Cr207 on Fluorescence Emission Intensities of Quinine Sulfate. Technical report.

[81] UNICEF. UNICEF Target Product Profile Rapid E. coli Detection.

[82] Brett Walton. Infographic: Household Wells in the United States - Circle of Blue, 2018.

[83] Jade S.T. Ward, Daniel J. Lapworth, Daniel S. Read, Steve Pedley, Sembeyawo T. Banda, Maurice Monjerezi, Gloria Gwengweya, and Alan M. MacDonald. Large-scale survey of seasonal drinking water quality in Malawi using in situ tryptophan-like fluorescence and conventional water quality indicators. Science of the Total Environment, 744, 11 2020.

[84] Joseph Wasswa and Natalie Mladenov. Improved Temperature Compensation for in Situ Humic-Like and Tryptophan-Like Fluorescence Acquisition in Diverse Water Types. Environmental Engineering Science, 35(9):971–977, 9 2018.

[85] C.J. Watras, P.C. Hanson, T.L. Stacy, K.M. Morrison, J. Mather, Y.-H. Hu, and P. Milewski. A temperature compensation method for CDOM fluorescence sensors in freshwater. 2011.

[86] WHO. Drinking-water Fact Sheet. Technical report, 2019.

[87] World Health Organization. Child Mortality and Causes of Death, 2020.

[88] World Health Organization. The top 10 causes of death, 2020.

[89] Jishan Wu, Miao Cao, Draco Tong, Zach Finkelstein, and Eric M V Hoek. A critical review of point-of-use drinking water treatment in the United States.

[90] Shang Tian Yang, Xiaoguang Liu, and Yali Zhang. Metabolic Engineering-Applications, Methods, and Challenges. Bioprocessing for Value-Added Products from Renewable Resources, pages 73–118, 2007.

[91] Richard G. Zepp, Wade M. Sheldon, and Mary Ann Moran. Dissolved organic fluorophores in southeastern US coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in excitation–emission matrices. Marine Chemistry, 89(1-4):15–36, 10 2004.