Optical properties of water for prediction of wastewater contamination in surface water

Steven R. Corsi, Laura A. De Cicco, Angela Hanson, Peter Lenaker, Brian Bergamaschi, Brian Pellerin

# Abstract

# Introduction

Sewage contamination from illicit discharges and leaking sewer infrastructure in the Great Lakes is a substantial source of pollution in tributaries and nearshore waters. Mistakes during construction of sewage infrastructure lead to misconnections into the storm sewer system, and many metropolitan areas have an aging sanitary sewer infrastructure with failures in the system resulting in sewage exfiltration. Contaminants such as nutrients, pharmaceuticals, hormones, toxic compounds, and pathogens that are found in sewage can have a substantial effect on the aquatic ecosystem {Nilson et al, 2019, Thrupp et al, 2018, Lenaker et al., 2017, Sauer, 2011}. There are more than 1400 wastewater treatment facilities in the U.S. and Canada that discharge 18 million m3 (4.8 billion gallons) of treated effluent to the Great Lakes each day {Arvai et al. 2014}. However, a large volume of sewage never makes it to wastewater treatment plants. One United States Environmental Protection Agency study reported between 12% and 49% of sewage flows are lost due to leaking infrastructure {Rutsch, Franz, and Krebs 2007}. At the low end of these estimates (12%), basin-wide leakages in the Great Lakes would be more than 1.9 million m3 (500 million gallons) per day.

Multiple factors influence the level of sewage contamination in a receiving stream at any given time including the number of sources in a drainage basin, the dynamic nature of urban hydrology, the efficiency of urban stormwater conveyance systems, and the level of infiltration and inflow (I & I) that stress the sanitary sewer systems {kaushal, 2012; Templar, 2016; Sauer et al., 2011}. These factors make defining the quantity, the timing, and the location of sewage contamination a challenging task to accomplish in a time-, labor- and cost-efficient manner for those who administer Illicit Discharge Detection and Elimination (IDDE) programs.

Multiple methods for quantifying sewage contamination in waterways exist, but each of them has challenges involved. Some are not specific to sewage, including chemical analysis such as ammonia, or fecal indicator bacteria such as fecal coliform or Escherichia coli (E. coli) {Panasiuk et al., 2015, Brown, Caraco, and Pitt, 2004}. Other methods that are more specific to sewage include human bacteria markers, pharmaceuticals, and some personal care products {Ahmed et al, 2009, Feng et al, 2018, Tran et al, 2014}. These options provide more definitive information on sewage contamination, but analytical costs can be limiting and reporting results takes weeks to months. Given the dynamic nature of contamination in urban surface waters, techniques to define the results on a relatively fine time scale (minutes to hours) would be valuable to allow for accurate sewage load computation to facilitate comparison of contributions from different areas of a watershed. This would allow resource managers to direct mitigation efforts to areas with the greatest levels of contamination.

Human sewage has distinct optical properties of water that are different than those typically observed in natural waters {Baker et al, 2001}. Optical property analysis includes measurement of fluorescence and absorbance spectra that serve to characterize the composition of dissolved organic matter (DOM) in water. There are many sources of DOM in natural waters that influence optical properties, but previous research efforts have identified useful signals that predict sewage presence in natural waters {Cao et al., 2009; Dickerson et al., 2007}. These studies, however, have typically been limited in scope to single study areas. Technology does exist to incorporate measurement of select optical signals into field sensors {Kraus, 2017, Bergamaschi, 2011; pellerin snowmelt, 2012 }, but for these sensors to be effective more broadly, research is needed to determine the degree of geographic transferability and the level of consistency over small to large watershed scales. Practitioners responsible for IDDE programs would benefit from development of new tools such these for rapid, accurate, and cost-effective evaluation of sewage contamination.

The overall objective of this research is to identify optical signals that could be incorporated into a real time sensor system for detection of sewage in surface waters. Specific objectives to achieve this are: 1) To characterize the optical properties in surface water samples in a diverse set of geographic and watershed settings, 2) to define sewage presence and magnitude by concurrent analysis of human-associated bacteria, 3) To identify the optical signals that best serve as surrogates to predict sewage contamination for development of field sensors, 4) to conduct this research in multiple settings including: variable land use, variable hydrologic conditions, through different seasons, and at multiple watershed scales, to define limitations of such signals, and 5) to provide stakeholders sufficient information to implement monitoring systems that can predict sewage presence in real time.

# Methods

Sources of wastewater contamination are diverse in nature with large variations in composition. Wastewater can include waste from residential areas, industrial discharges, and commercial applications. These sources contain various degrees of human waste, detergents from a wide variety of applications, food waste, and an untold variety of other materials that are introduced into the wastewater system. This diversity in sources can result in highly variable chemical and microbial composition leading to heterogeneity in wastewater depending on time and location in the system. In designing a sensor system to detect wastewater, this heterogeneity must be considered to determine the potential transferability and relative accuracy considering contributions from multiple and unique source signatures. The current study was designed to examine these sources of variability by including a range of three spatial scales from very small to large as the wastewater signal changed from a high degree of local influence that could be dominated by very few sources to a mixture of many potential sources.

## Site selection

Samples were collected from areas around the Great Lakes at three different drainage basin scales (Figure 1, (Figure 1, Table 1), including: 1). Watersheds: Eight tributaries of the Great Lakes with a gradient of urban to agricultural land use, 2). Subwatersheds: Three locations were sampled on the Menomonee River in the Milwaukee Metropolitan area, and 3). Small scale: Multiple storm sewers and open channels (213 sites) were sampled in three subwatersheds within the Great Lakes Basin including the Middle Branch of the Clinton River in Macomb County, Michigan (65 sites), Red Creek in Monroe County, New York (88 sites), and the Kinnickinnic River in Milwaukee (60 sites).

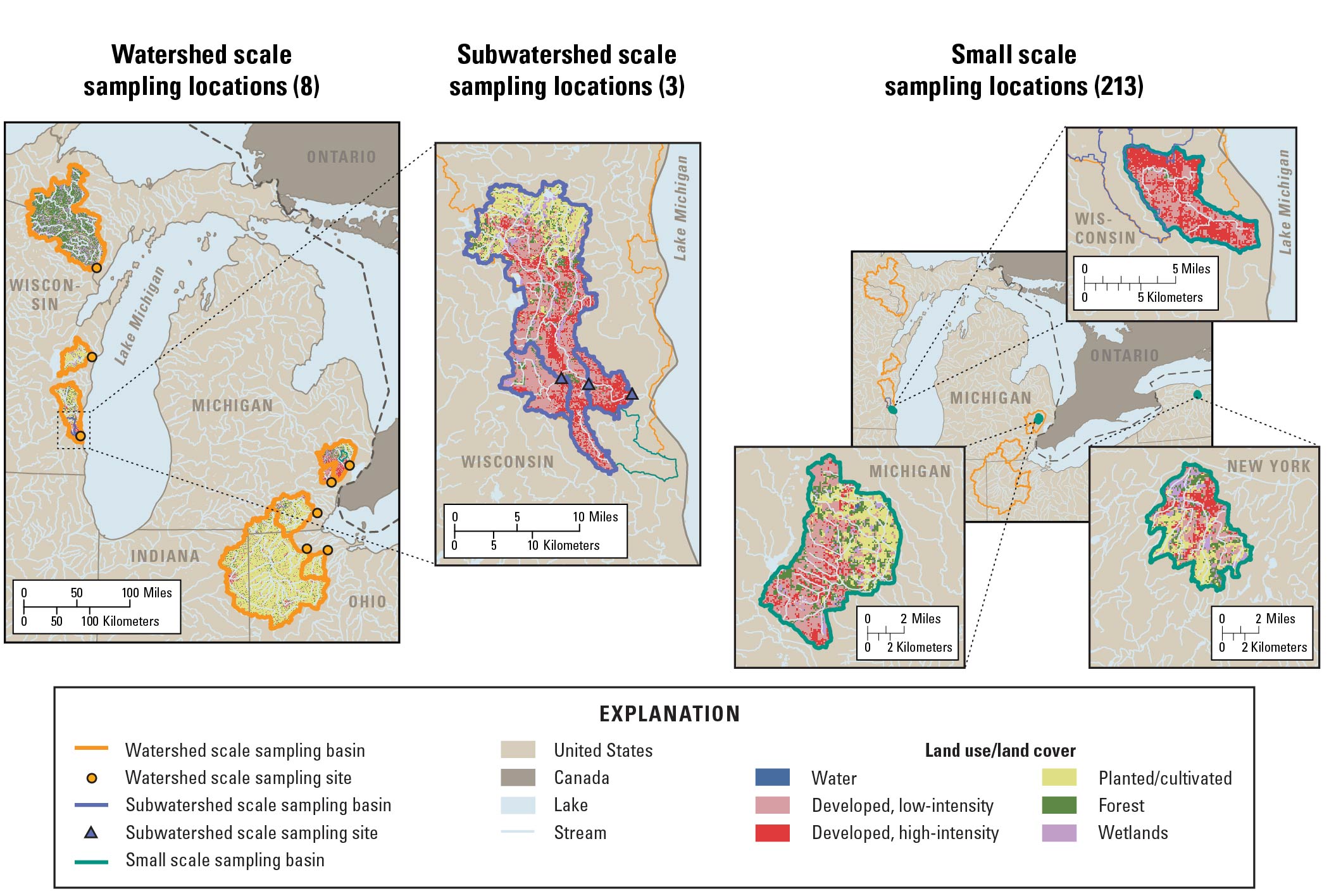


Figure 1: Location of sampling sites including small storm sewers and open channels (small scale), subwatersheds of the Milwaukee River (subwatershed scale), and tributaries of the Great Lakes (watershed scale).

Table 1. Land cover for sampling sites including small storm sewers and open channels (small scale), subwatersheds of the Milwaukee River (subwatershed scale), and tributaries of the Great Lakes (watershed scale)

Table: Table 1: Land cover for sampling sites including small storm sewers and open channels (small scale), subwatersheds of the Milwaukee River (subwatershed scale), and tributaries of the Great Lakes (watershed scale) 

Sampling design

Samples at all three drainage basin scales were collected throughout each season of the year to capture seasonal variations, and samples were collected during low-flow periods and runoff-event periods (rainfall and snowmelt) to capture variable hydrologic conditions. Watershed-scale locations were sampled over a two-year period from 2011-2013 for a total of 236 samples. Subwatersheds were sampled over a four-year period from 2011-2014 for a total of 127 samples. Small-scale sites were sampled from 2014-2016 for a total of 593 samples.

At watershed- and subwatershed-scale locations, samples were collected over a 24-hour duration for low-flow periods, and throughout the duration of increased streamflow for runoff-event periods. An individual sample included multiple subsamples that were collected using automatic samplers. Subsample collection was triggered by accumulation of streamflow to a pre-determined threshold volume that was chosen based on anticipated streamflow magnitude for any given sampling period. Analysis of these samples resulted in flow-weighted mean concentrations for the sampling period. A detailed description of the sample collection procedure as well as a diagram of the automatic sampler design has been previously published (Lenaker et al., 2017).

At the small-scale locations, discrete grab samples were collected by direct bottle submersion into the predominant flow of water or by peristaltic pump. Direct bottle submersion used 500 mL autoclaved polypropylene bottles for bacteria markers and a 500 mL baked amber glass bottles for optical properties and dissolved organic carbon (DOC). The pump systems included a peristaltic pump (Cole-Parmer, Vernon Hills, IL), a stainless-steel sample strainer (Teledyne ISCO, Lincoln, NE) and a combination of Masterflex PharmaPure pump tubing (Cole-Parmer, Vernon Hills, IL), polytetrafluoroethylene and polyvinyl chloride tubing (Fluorotherm polymers, Parsippany, NJ; Thermo ScientificTM, NalgeneTM, Rochester, NY), and a stainless-steel manual ball valve with Yor-Lok fittings (McMaster-Carr, Elmhurst, IL) to direct flow into the sample bottles. To characterize sources of wastewater, samples were periodically collected from sanitary sewers with small and large contributing areas. Small sanitary sewer sampling locations included residential and commercial areas near small-scale environmental sampling locations. Large sanitary sewer sampling points were located at the influent to Jones Island and South Shore water reclamation facilities in Milwaukee, WI and Van Lare Treatment plant in Rochester, NY.

## Laboratory analyses

Samples were analyzed for optical properties of water (fluorescence and absorbance spectra measured in a laboratory setting), DOC, two human-associated bacteria markers, two fecal indicator bacteria, and field-measured turbidity. Samples from the large watershed and subwatershed scales were also analyzed for human-specific viruses. Human virus and human-associated bacteria data have been previously published for the watershed scale (Lenaker et al., 2018) and in part for the subwatershed scale (Lenaker et al., 2019). Data that have not been published elsewhere are published as a data release (De Cicco et al., 2020 (in progress)). For convenience all data used in the current study have been compiled and provided in the supporting information (Tables SI-1-3).

### Bacteria

Bacteria analyses were performed at the University off Wisconsin-Milwaukee School of Freshwater Sciences. For human-associated and indicator bacteria analyses, a volume of 200 mL was filtered onto 0.22 μm pore-size mixed cellulose ester filters (47 mm diameter; Millipore, Billerica, MA), placed in 2 mL screw-cap vials, and stored at -80°C prior to DNA isolation. The frozen filters were broken into small fragments using a sterile metal spatula. DNA was extracted from the fragments using the MPBIO FastDNA® SPIN Kit for Soil (MP Biomedicals, Santa Anna, CA) according to manufacturer’s instructions, with the exception of the lysis step in which a bead beater (BioSpec, Bartlesville, OK) was used for 2 min., and the final step in which DNA was eluted using 150 ul of DES. Samples were analyzed six to nine months after collection.

Samples were analyzed by qPCR for the human Bacteroides (HB), human Lachnospiraceae (Lachno2), E. coli (EC) and enterococci (ENT) assays as described in detail previously (Templar et al., 2016; Dila et al., 2018; Sauer et al., 2011; Ludwig and Schleifer, 2000). Briefly, assays for HB, Lachno2, EC and ENT were carried out by Applied Beiosystems StepOne Plus ™ Real-Time PCR System Thermal Cycling Block (Applied Biosystems; Foster City, CA) with Taqman hydrolysis probe chemistry, and assay conditions according to manufacturer’s instructions. Reactions were carried out in volumes of 25 μL, with 5 μL of sample added as template. The lower limit of quantification was determined to be 15 copy number (CN) per reaction, which is equivalent to 225 CN/100 mL when a sample volume of 200 mL was filtered or 112.5 CN/100 mL when a sample volume of 400 mL was filtered. Signals below 35 cycles (and not within the quantifiable range) were considered detectable but not quantifiable. All qPCR runs included two previously analyzed environmental samples as controls. All no-DNA template controls were negative. All assays were performed in duplicate and compared to values in standard curves. Previously published primers and probes were used for all analyses and are listed in supporting information (Table SI-1). Standard curves (slope and intercept values) are reported in supporting information (Table SI-2).

Periodic testing for extraction efficiency and inhibition followed procedures described previously (McLellan et al., 2018), where 0.2 µg of salmon testes DNA was added to 1 mL DNA extraction buffer (efficiency), or a final concentration of 0.2 ng µL–1 salmon testes DNA was added to extracted environmental samples (inhibition). The extraction efficiency was 19.8% (±6%) and was similar to a previously determined efficiency (Sauer et al., 2011) of 15.3% (± 2.7%). There was no indication of inhibition during qPCR in river water (Lenaker et al,, 2018, McLellan et al., 2018) or stormwater (Sauer et al., 2011). A subset of samples analyzed for the current study were consistent with these results. Additional details for DNA extraction and qPCR assays have been published (Lenaker et al., 2018).

**UWM add enterococci culture methods here**

### Viruses

Human virus analyses were performed at the U.S. Department of Agriculture and U.S. Geological Survey joint Laboratory for Infectious Disease and the Environment. Methods for analysis of human viruses have been previously published (Lenaker et al., 2017; Lenaker et al., 2019), and additional information is provided below. A portion of the human virus sample results from subwatershed scale sites have also been published (Lenaker et al., 2019). All human virus results are available in the U.S. Geological Survey National Water Information System (<https://nwis.waterdata.usgs.gov/usa/nwis/qwdata>) and are also provided in the supporting information (SI-Table 3).

Following elution using 3% beef extract with 0.05 M glycine (pH 9.5) as previously described (Lenaker et al. 2018; Millen et al. 2012), eluate from paired glass wool filters and pre-filters was adjusted to pH 7.0, combined, and concentrated by polyethylene glycol precipitation (Lambertini et al. 2008). For wastewater influent, the entire 1-liter sample was concentrated by polyethylene glycol precipitation. Briefly, polyethylene glycol 8000 (8% m/v) and NaCl (0.2 M) were added to the eluate from filters or to 1-L influent sample, which was then stirred for a minimum of 2 hours at 4° C and incubated overnight (4° C). Samples were centrifuged at 4700 x g for 45 min (4° C), and the pellet was resuspended in Tris-EDTA buffer. Sample concentrate (volume ranged 1.3 – 28.5 mL) was frozen at -80 °C.

Nucleic acids were extracted from 280 µL of the concentrated water and wastewater samples using QIAamp DNA blood mini kit and buffer AVL (Qiagen, Valencia, CA), producing 200 µL of extracted nucleic acid. Virus RNA was reverse transcribed using random hexamers (ProMega, Madison, WI) and SuperScript® II reverse transcriptase (Invitrogen Life Technologies, Rockville, MD).

Following extraction, virus RNA was reverse-transcribed by adding 15.48 µL of the extracted nucleic acids to 15.48 µL nuclease-free water and 1.26 µL random hexamers (ProMega, Madison, WI). After heating for 5 min at 95 °C, 57.78 µL of reverse transcription master mix was added to the mixture. The reverse transcription master mix consisted of the following components, reported as final concentrations: 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl2, 0.6 nM dithiothreitol, 70 µM of each deoxynucleoside triphosphate (ProMega), 0.72 U RNAsin® (ProMega), 50 U SuperScript® II reverse transcriptase (Invitrogen Life Technologies, Rockville, MD). The reaction was incubated for 60 min at 42 °C followed by 5 min at 95 °C.

Samples were analyzed for 18 viral and bacterial pathogen gene targets by quantitative polymerase chain reaction (qPCR) using the LightCycler 480 (Roche Diagnostics, Mannheim Germany) and hydrolysis probes (Probes Master kit; Roche Diagnostics); primers, probes, and assay references are listed in Table XX. Six µL extracted DNA or cDNA from reverse transcription were added to 14 µL master mix. Thermocycling began at 95 °C for 5 min followed by 45 cycles of 10 s at 95 °C and 1 min at 60 °C. Cycle of quantification (Cq) values were calculated using the second derivative maximum method. qPCR was performed in duplicate, and the average of positive replicates was reported. Standard curves were created using synthesized DNA oligos (Integrated DNA Technologies; Coralville, IA); standard curve performance metrics are reported in Table XX.

Reverse transcription-qPCR inhibition was evaluated for all samples using hepatitis G virus RNA following Gibson et al. (2012). Three samples were inhibited; inhibition was mitigated by dilution (1:5) with AE buffer.

Human enterovirus was used for an extraction and reverse transcription positive control and was evaluated qualitatively. DNA oligos (Integrated DNA technologies) were used as qPCR positive controls for each assay and must yield Cq values within 1 cycle of the expected value.

Negative controls must exhibit no fluorescence above the baseline (i.e., no Cq value) to be acceptable and were included with all nucleic acid extraction, reverse transcription, and qPCR batches. Equipment blanks (sterile phosphate buffered saline solution pumped through the sampling apparatus and processed as above) were analyzed throughout the study (5 per site) and were negative for all microbial targets. Recovery controls (n = 20) for glass wool filtration were analyzed as described in Lambertini et al. (2008) by seeding adenovirus 41, poliovirus Sabin 3, and enterohemorrhagic E. coli O175:H7 into water from each site (2 baseflow and 2 runoff events per site).

### Optical properties of water

Absorbance, fluorescence and DOC measurements were performed by the Organic Matter Research Laboratory (OMRL) at the California Water Science Center (US Geological Survey, Sacramento, CA) over a five-year period between February 2011 and June 2016.

DOC concentrations were measured by a total organic carbon analyzer (TOC-VCSH, Shimadzu Scientific Instruments, Columbia, Maryland) using high-temperature catalytic combustion according to a modified version of USEPA method 415.3 (Potter and Wimsatt, 2012). The accuracy and precision of these measurements were within data quality objectives as indicated by an internal laboratory standard (caffeine), laboratory replicates, and matrix spikes. The laboratory reporting limit for DOC concentration was 0.30 mg C L−1 based on three times the standard deviation of a low concentration standard measured over an annual cycle.

Absorbance spectra and fluorescence excitation-emission matrices (EEMs) were measured on filtered (0.45-μm nominal pore size syringe filter) water samples at room temperature (21 °C) in an acid-cleaned 1 cm quartz cuvette (Starna Cells, Inc., CA, USA, parts 1-Q-10, 3-Q-10). Over the five-year period, measurements of absorbance spectra and fluorescence matrices were made using similar analytical methods, but three different instruments as described below.

Samples collected between February 2011 and July 2014 (n=406) were analyzed for absorbance in 1 nm increments between 200 and 750 nm using a CARY-300® spectrophotometer equipped with a photodiode (Agilent Technologies, Santa Clara, CA, USA). Fluorescence was measured using a Fluoromax-4® spectrofluorometer (Horiba Instruments, NJ, USA) equipped with a 150 W xenon arc lamp. Fluorescence intensity was measured at excitation wavelengths of 240 nm to 440 nm in 5 nm increments and emission wavelengths of 290 nm to 600 nm in 2 nm increments.

Absorbance spectra and fluorescence EEMs for samples collected between August 2014 and June 2016 (n=612) were simultaneously measured using an Aqualog® equipped with a photodiode and charge-coupled device (CCD) (Horiba Instruments, NJ, USA) according to the method described by Hansen et al., 2018. Briefly, excitation and absorbance spectra were measured using a 150 W xenon lamp, a 5 nm bandpass, and a 1 s integration time at wavelengths of 240–600 nm. Emission spectra were collected with a CCD at approximately 1.64 nm (4 pixel) intervals at wavelengths of 250–600 nm. To reduce ultraviolet light exposure of the sample and limit the effects of photobleaching, excitation and absorbance wavelengths were scanned from low to high energy (i.e., visible to ultraviolet range) during analysis.

For comparability of fluorescence measurements between the different instruments, Fluoromax-4 measurements (counts per second) were converted to Aqualog measurements (microvolts) by applying a conversion matrix created from 36 environmental samples ranging in DOM concentration and composition that were analyzed on both instruments on 11 analytical dates. An average of the ratio of the 36 samples analyzed on both instruments was used to calculate the conversion matrix. Fluoromax-4 EEMs were resized by linear interpolation and divided by the conversion matrix to render them comparable to EEMs generated by the Aqualog.

Correction procedures for optical data included instrument-specific excitation and emission corrections, baseline subtraction, normalization to the daily water Raman peak area (Murphy, 2010), and the removal of Rayleigh scatter lines. Concentration-related inner filter effects were corrected as described by Ohno (2002). Absorbance (Aλ) data are reported as absorbance units (AU), obtained directly from the instrument. Fluorescence data are expressed in Raman-normalized intensity units (RU). High concentration samples with A254 > 3.0 AU were diluted and then reanalyzed to ensure linearity in the wavelengths of interest.

1. Field QA results

* Blanks
* Replicates

## Regression modeling

Regression modeling was used to explore relations between bacteria concentrations (response variables) and optical signals (explanatory variables). In all cases, response variables were log base 10 transformed. From a practical standpoint for potential field application, explanatory variable selection for these regressions was done in two stages: First, optical signals measured in the laboratory that represent signals available in current field sensors were considered (sensor signals). Next, explanatory variables were expanded to use of additional optical signals and variables that were derived from optical signals (alternative signals). Only those variables that were not correlated to the field sensor signals (with a )correlation coefficient less than 0.95) were considered. These signals included direct fluorescence or absorbance signals at specific wavelengths, means of fluorescence or absorbance signals over specified bandwidths, ratios of these signals, spectral slopes determined as the slope of the absorbance curve over specified wavelengths in exponential space (**add reference here**), and several optical indices that have been used in previous research (**add references here**). Given the potential for seasonality in DOC composition in streams (@kaiser\_seasonal\_2001, @coble\_temporal\_2016), seasonal varaibles (sine and cosine of julian day/(2)) were used to develop interaction terms with optical signals. Optical signals used as sensor signal variables included signals F (often referred to as fDOM or CDOM) and T (often referred to as tryptophan-like fluorescence) as well as turbidity. Final explanatory variables used in this analysis are presented in the supporting information (Table SI-xx).

**For small-scale sites, it was not uncommon to find a model that appeared to be a good fit overall but did not provide a good fit for data from individual sampling events.**

The model selection process included several steps with the ultimate goal of minimizing error in prediction while choosing a model that has a reasonable fit for all sites in the data set for watershed and subwatershed site. For watershed and subwatershed models, it was not uncommon to find a model that appeared to be a good fit overall but did not provide a good fit for data from one or more of the individual sites. For this reason, exploration of potential regression models was conducted using data sets that included multiple sites as well as one single-site data set (Milwaukee River). The Milwaukee River was modeled independently from other sites because the nature of organic matter influence was substantially different than other sites: The moniotoring station at this site was near Lake Michigan and a seiche effect is commonly present, posing the potential for Lake Michigan water influence. In addition, sites with similar land cover were grouped into one data set (e.g. urban, agrucultural). For the single-site data set, ordinary least squares regression (OLS) was used. For data sets with more than one site included, linear heirarchical mixed effects models (LME) were used with sites included as random variables to account for potential shifts in the statistical relationships among sites (**add references for base r and lme4**). For both modeling techniques, 5-fold cross validation repeated 50 times was conducted to estimate predictive accuracy. The median root mean squared residuals (RMSE) for the 50 model fits was used to compare among models with different explanatory variable combinations. For models that included only one site (Milwaukee River), the model with the lowest median sum of RMSE was chosen. For models that included multiple sites, model options within 3% of the of the lowest median sum of RMSE were evaluated further by considering the quality of fit for each individual site: The RMSE was computed for individual sites, and the maximum and minimum individual site RMSE values for each model were identified. The final model was chosen based on the smallest difference between the maximum and minimum individual site RMSE. This minimized the chances of choosing a multiple-site model with a poor fit for individual sites. For small-scale sites, model development attempts often appeared to be a good fit overall, but in all cases, models did not provide a consistently good fit for data from individual sampling events. the Wilcoxen rank sum test was used to test for differences between seasons and hydrologic conditions. Statistical analyses were done in the R statistical programming language including base R functionality for correlation, the Wilcoxen rank sum test, and ordinary least squares regression (R core team, 2020), and the lmer4 package for linear hierarchical mixed effects regression (Bates et al., 2015).

# Results

Throughout the study, 996 environmental samples were collected, including 590 samples from small scale sites, 170 samples from subwatershed scale sites, and 236 samples from watershed scale sites. The sum of the two human indicator bacteria markers (sHM) varied from below the reporting level to a maximum of 8.95 x 107 cn/100 ml (Figure 2).

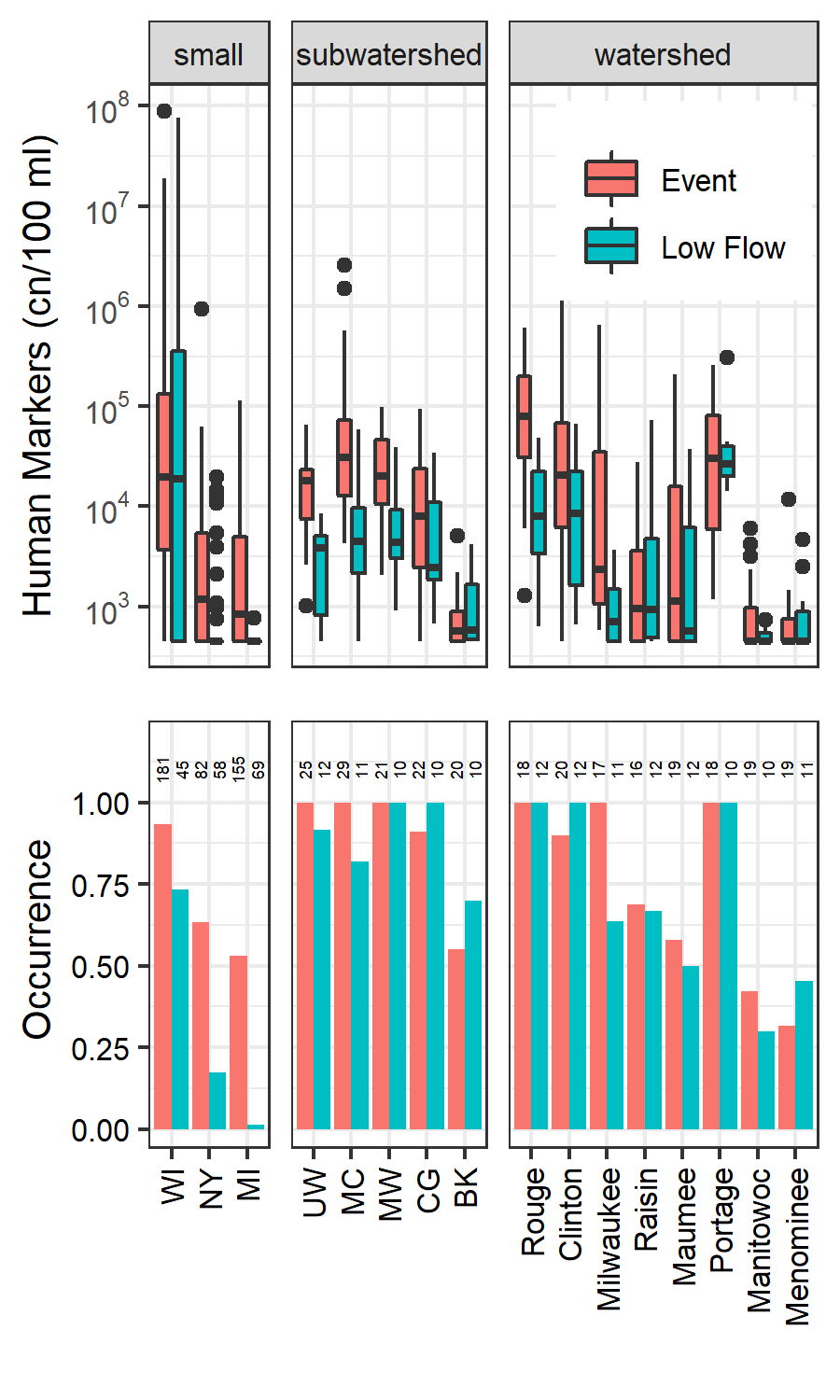


Figure 2: The sum of human indicator bacteria markers A. concentrations and B. Occurrence proportion in samples from small storm sewers and open channels (small scale), subwatersheds of the Milwaukee River (subwatershed scale), and tributaries of the Great Lakes (watershed scale). For each scale, sites are ordered by most to least urban land use percentage from left to right. Number of observations in each concentration bin are provided above the bars.

Remove CG and BK?

There were differences in sHM concentrations among sites and by hydrologic condition (Figure 2). At small-scale sites, median concentrations of sHM were 43 times greater in samples from the Wisconsin sites collected in the Kinnickinic River system than those from the other states (p < 0.05). In samples from small-scale sites in Michigan (Middle Branch Clinton River) and New York (Red Creek), results indicated greater sHM concentrations during runoff events than during periods of low flow (p < 0.05), but results from the small scale sites in Wisconsin indicated no significant difference in sHM concentrations by hydrologic condition.

Samples collected at sites with greater than 25% urban land cover at the watershed and subwatershed scales had greater sHM concentrations during runoff event periods than low-flow periods (p<0.05). Concentrations of sHM at sites with less than 25% urban land cover at these two scales did not vary significantly by hydrologic condition.

Human viruses were analyzed for the watershed- and subwatershed-scale sites concurrently with the human-associated and the general indicator bacteria, providing a means to assess potential exposure to human pathogens from sampled water. The fraction of samples for which human viruses were present increased with increasing sHM concentration, but occurrence of human viruses did not increase with increasing concentration of the non-specific fecal indicators E. coli and enterococci (Figure 3). These results reinforce the the assumption that human-associated bacteria markers represent presence of sewage and an increased risk of exposure to human pathogens with increasing human-associated bacteria concentrations.

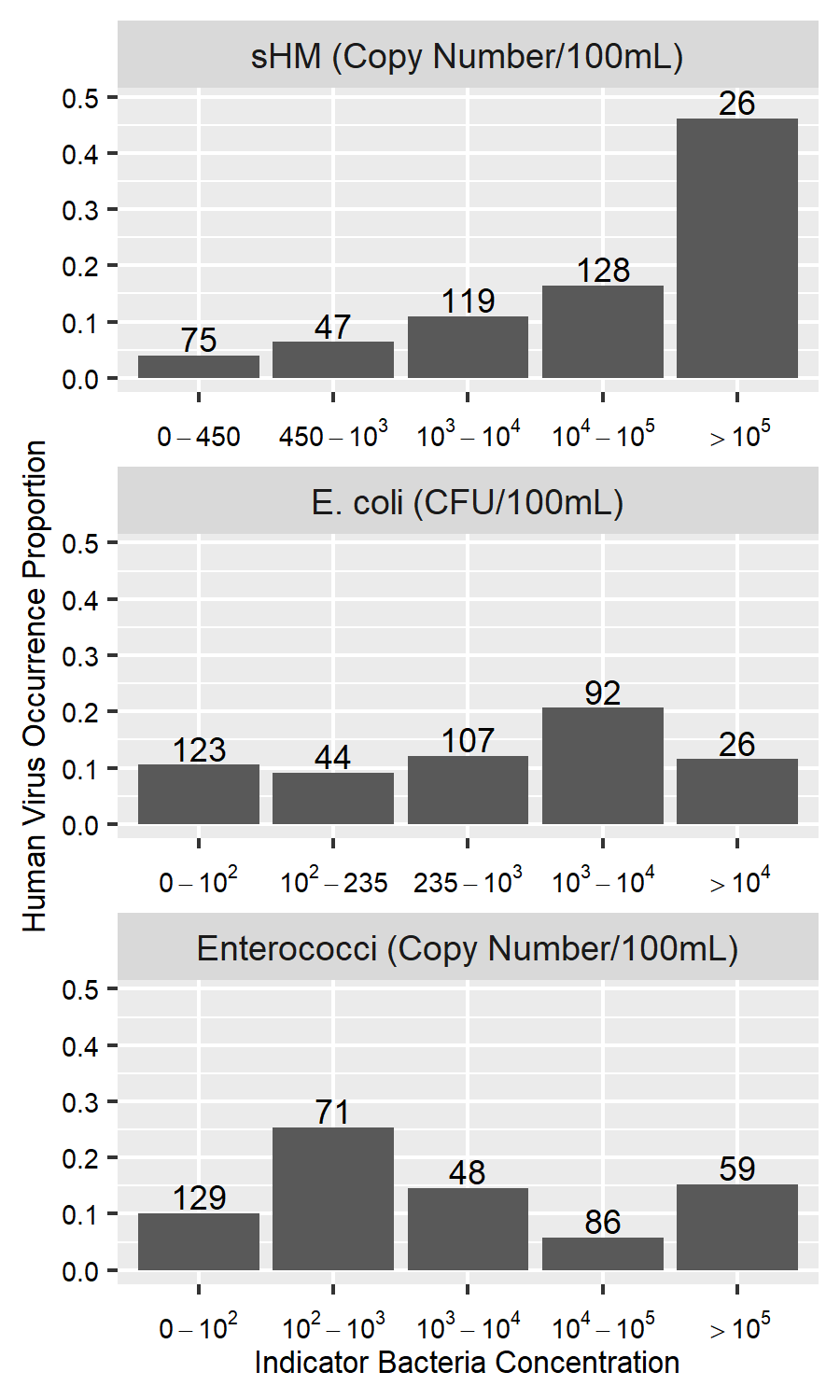


Figure 3: Fraction of human virus occurrence as compared to concentration bins for A. the sum of human associated bacteria, B. Escherichia coli, and C. Enterococci in samples from subwatersheds of the Milwaukee River (subwatershed scale), and tributaries of the Great Lakes (watershed scale). Number of observations in each concentration bin are provided above the bars. sHM concentration of 450 Copy Number/100mL represents the analytical detection level. E. coli concentration of 235 CFU/100mL represents a common recreational water quality criteria.

## Regression Modeling

Regression models were explored for describing variability of human-associated bacteria using optical signals. The first consideration in regression modeling was the choice of sites to include in common models. Watershed and subwatershed sites were initially grouped together for a unified model. These efforts resulted in poor cross-validated predictive capability, and no adequate models were discovered. Next, categories from watershed and subwatershed sites were chosen for similarity in land cover and potential diversity of water quality influences as well as geographic proximity. These groups were chosen to minimize differences in background DOC that could lead to confounding signals in regressions, resulting in four site categories: Watershed-scale sites were separated into three categories for model development including Agricultural (Maumee, Portage, and Raisin), Urban (Clinton and Rouge), and single site (Milwaukee), and the three subwatershed-scale sites were included as the fourth category. Separating sites into these four categories resulted in improved cross-validated predictive capability for resulting models over the watershed- and subwatershed-scale unified model, so the four categories were selected for inclusion in final model development (Table 2).

Model development attempts for the small-scale sites also included four data groupings: first, data from all sites and samples within the three regions were used to develop one unified model, and second, separate data sets from each of the three regions was used to develop individual models. Resulting models for small-scale sites were all considered to be inadequate for the intended purpose: To meet sensor development needs for tracing contamination back to the source, regressions must provide increasing predictions with increasing human bacteria concentrations for individual events within a storm sewer system or small drainage basin. Attempts to develop models for the small-scale sites did not result in models that consistently met this need.

Table: Table 2:

|  | | Human Bacteroides | | Lachnospiraceae | | Enterococci | | Enterococci Culture | | E. coli | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sites | Category | Variables | RMSE | Variables | RMSE | Variables | RMSE | Variables | RMSE | Variables | RMSE |
| Watershed: Agriculture | sensors | Turbidity\_F\_T | 0.56 | Turb\_F\_T | 0.42 | Turb\_F | 0.55 | Turb\_F | 0.67 | Turb\_F | 0.51 |
| non-cor | rMrange.25\_N\_Turbidity | 0.55 | rMrange.25\_N\_Turb | 0.39 | rMrange.25\_F\_Turb | 0.53 | rA280\_A295\_Turb | 0.63 | F\_Turb | 0.51 |
| Watershed: Urban | sensors | Turbidity | 0.78 | Turb | 0.89 | Turb\_F2 | 0.48 | Turb | 0.49 | Turb\_F2 | 0.44 |
| non-cor | rA290\_A295 | 0.76 | rA290\_A295 | 0.86 | rA290\_A295 | 0.48 | rA290\_A295\_Turb | 0.43 | rA290\_A295 | 0.41 |
| Milwaukee River | sensors | Turbidity\_T | 0.60 | Turb\_T | 0.66 | Turb\_T | 0.58 | Turb\_T | 0.82 | Turb\_T | 0.79 |
| non-cor | T\_Turbidity | 0.60 | rS2.25\_S1.25\_Turb | 0.62 | FI\_2005\_Turb | 0.54 | FI\_2005\_Turb | 0.73 | rN\_S3.25\_Turb | 0.73 |
| Subwatersheds | sensors | Turbidity\_T2 | 0.51 | Turb\_F\_T2 | 0.49 | Turb\_F\_T2 | 0.39 |  | NA | Turb\_F\_T2 | 0.44 |
| non-cor | Turbidity | 0.54 | rMrange.25\_N | 0.52 | rMrange.25\_N | 0.48 |  | NA | rMrange.25\_N | 0.49 |

Models were developed with two sets of optical variables: First, using optical signals to represent those used in commercially available field sensors (sensor models), and second, using additional optical signals that were not highly correlated with at least one of the sensor signals (alternate signal models). Two important attributes of the selected models included interaction of the optical variables with a seasonal term, and, for multi-site models, use of individual sites as random effects in the linear mixed effect models. The seasonal term was used to control for variable DOC presence as vegetation cover changed throughout the year and as soil contact of runoff varied with snow cover. Individual sites were used as random effects to help control for variable background DOC in watersheds resulting from different geologic environments and land cover. Model exploration with and without the seasonal and random effects terms verified that inclusion of these two model attributes resulted in improved model fit.

The sensor models included different combinations of fluorescence signal “T” (often referred to as tryptophan-like fluorescence), fluorescence signal “F” (often referred to as “CDOM” or “fDOM”), and turbidity. Signals T and F were commonly included in final models as interaction terms with seasonality. All final sensor models included turbidity, indicating that turbidity may be the most valuable signal for inclusion in these surrogate models with additional optical signals and seasonal terms to enhance accuracy of these.

There was not a substanial and consistent improvement in cross-validation error (RMSE) in the alternate signal models compared to the sensor models (Table 2). The enterococci culture models for the urban large watershed sites and for the Milwaukee River site were the only two models that had 10% or greater decrease in RMSE by using models developed with the alternate signal variables. In addition to T and F, variables in the alternate signal models included two fluorescence signals (M and N), ratios of fluorescence signals, the fluorescence index, and ratios of absorbance signals (Table SI-XX). Collectively, results suggested that there may be minor improvements in model performance by use of the alternate signal variables for a few selected site-organism combinations, but that the optical signals selected for the alternate signal models are not consistent.

The proportion of virus occurrence increased with the sum of fitted sHB from the watershed- and subwatershed-scale models (Figure 4). For instances when fitted model values were below 1000 copy number/100mL, no viruses were detected. Virus detection increased with increasing concentration up to 40% virus detection when sHB fitted values were greater than 100000 copy number/100mL. Because human viruses were measured independently from human bacteria and not used in model development, this result provides independent verification that the sensor models have the capability of predicting the presence of sewage and potential risk from human pathogens.

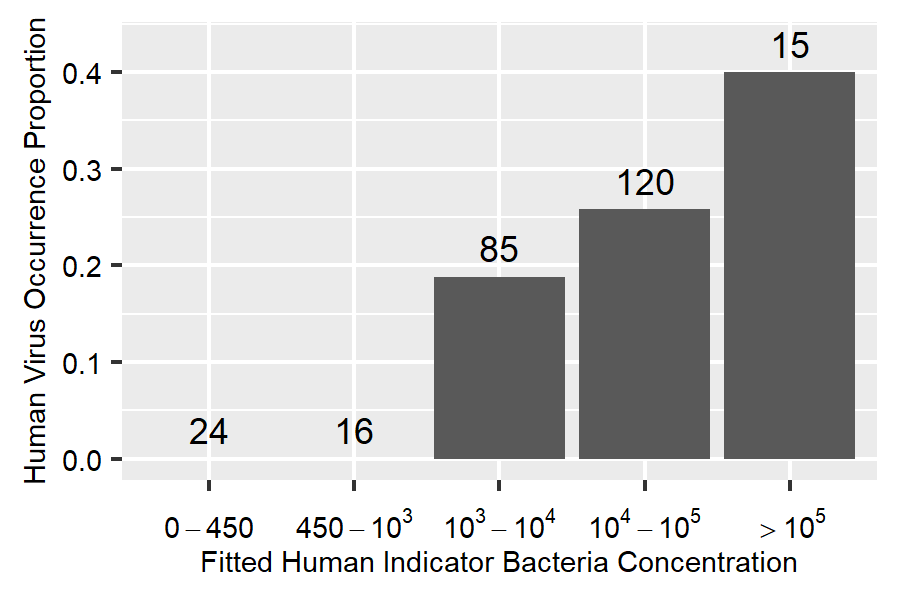


Figure 4: Fraction of human virus occurrence as compared to concentration bins for the sum of human associated bacteria computed from fitted values from regression models developed using laboratory-measured optical signals that represent currently available optical sensors. Number of observations in each concentration bin are provided above the bars. sHM concentration of 450 Copy Number/100mL represents the analytical detection level.

EXAMPLE FOR FONT STYLE

Whole paragraph of colored text

# Discussion

**Microbial Contamination**

The variability of contamination level by watershed was illustrated in result with median concentrations varying from less than the detection limit in samples from watershed scale to greater than 10000 cn/100 ml in several samples from the small-scale sites within the Kinnickinic River watershed. The general increasing human virus occurrence with increasing concentration bins for sHB provided validation that human-associated bacteria markers can be valuable for assessing the presence of sewage and risk of exposure to sewage-contaminated waters. Human viruses, however, are not consistently present in the human waste stream, and when they are present, they are highly variable even in raw wastewater (Lenaker et al, 2018, Fernandez-Cassi et al., 2018; Shrestha et al., 2017). These pathogens are most likely to be present in wastewater when they are present in the human population, and concentrations are likely to vary depending on how pervasive they are in the human population at the time of sampling. Depending on the specific virus, the occurrence and magnitude of virus presence in wastewater may or may not have seasonal patterns (Lenaker et al, 2018, Myrmel et al., 2006). Conversely, presence of human-associated bacteria such as the two measured in samples from the current study are relatively stable in wastewater treatment plant influent over time with concentrations of approximately 108 cn/100 ml an order of magnitude (Lenaker et al., 2019). With a detection level of 225 cn/100 ml, HIB would be detectable even after dilution in natural waters of five to seven orders of magnitude. Therefore, while human-associated bacteria are likely to be present whenever human viruses are present, human viruses are not always present in environmental waters contaminated with human-associated bacteria (Zhang et al., 2020). Depending on monitoring objectives, these two types of parameters can be valuable in different ways. Measurement of human viruses provide more direct information on human health risk while human-associated bacteria provide information on the potential for human health risk (McLellan et al., 2018; Zhang et al., 2020, Boehm et al., 2018) and an additional level of information on wastewater prevalence that can include other contaminants that are potentially harmful to aquatic life regardless of pathogenic presence.

The lack of relation between EC or ENT and virus occurrence is consistent with conclusions from previous work that general fecal indicators are not necessarily good predictors of waterborne pathogens (Boehm et al., 2009; Corsi et al., 2016; Noble and Fuhrman, 2001). These indicators are, however, used commonly as more general indicators of nonspecific fecal contamination from numerous sources (Standridge, 2008; Ercumen et al., 2017; Ishii and Sadowsky, 2008; Kinzelman et al., 2003; Byappanahalli et al., 2012), and have previously been associated with a risk to human health (Wade et al., 2006, Wade et al., 2008; Cabelli, 1989). Concentrations of fecal indicator bacteria in recreational waters are a primary measure currently used for issuing swimming advisories for the protection of human health (Kay et al., 2004; United States Environmental Protection Agency, 2012).

**Site groupings**

The result that model fit improved using selected subsets of the data by site rather than full data sets for each of the large watershed and subwatershed scales indicates the importance of watershed setting. Each subset of data included sites that represented a specific type of watershed setting with potential for distinctly different background DOM compostition that could impact the relation between optical signals and bacteria abundance: The final subsets of data included large watersheds dominated by agricultural influence, large watersheds dominated by urban influence, a mixed land use watershed with seiche effect from Lake Michigan, and subwatersheds dominated by urban influence. The agricultural watersheds were located in the Western Lake Erie drainage basin and all have **> xx%** agricultural influence (Table 1). The urban watersheds were located in the Detroit metropolitan area, and both have active combined sewer overflows at times (Michigan Department of Environmental Quality, 2011, Michigan Department of Environmental Quality, 1995). The Milwaukee River includes a mix of urban and agricultural land use and is influenced by a regular seiche effect that includes transitional influences from the Milwaukee River, the Milwaukee harbor, and Lake Michigan resulting in a unique combination of DOM sources. The subwatershed sites are all within the Menomonee River watershed in the Milwaukee metropolitan area that is dominated by residential, commercial, and industrial land cover, and to a lesser extent, the upper portion of the Menomonee River includes agricultural influence. Each of these groupings have different watershed characteristics unique enough to have influences on the organic matter content, and therefore, different optical properties of water that have potential to change the association with human-associated bacteria.

**Variable selection**

Final variables selected in models could help govern how to design field-level monitoring systems for estimating wastewater contamination. The sensor models indicated that this could potentially be done using currently available T and F fluorescence sensors. These models, however, were developed using data from laboratory analysis. Additional work will be needed to understand the steps necessary to translate these models to a field setting. Analysis of the field sensor response in comparison to the corresponding signals measured by the laboratory instrument will be needed as well as adjustments for the influence of turbidity, temperature, and inner filter effects on field sensor results (Saraceno et al., 2017; Downing et al., 2015; Watras et al., 2014).

Modeling results from this study call into question whether it would be a valuable use of resources to develop additional field senors targeted at optical signals other than T, F, and turbidity for the purpose of wastewater contamination prediction. The lack of substantial improvements in model fit when including alternate signals in the model in combination with the lack of consistency in the alternate signals selected for models among sites and among bacteria modeled would not provide a clear choice on which alternative signals would be of greatest value. Further, each additional sensor added to a monitoring system adds equipment and maintenance cost. Until it becomes clear that alternative sensor development efforts improve accuracy, reliability and reduce maintenance in the field, measurment of signals T, F, and turbidity sensors would likely be the most economical way to predict wastewater contamination using optical field sensors.

**Seasonality**

Inclusion of seasonal variables in the large watershed and subwatershed regression models indicate a seasonal influence in the relation of bacteria to predictor variables. This seasonality could be influenced by several factors: Precipitation and runoff to streams can be exposed to different levels of vegetation over the year, and runoff also has a variable level of exposure to soil depending on the presence of snow and ice (Qui et al., 2016; Shi et al., 2020). In the fall, leachate from leaf litter during precipitation events can be transported to streams. These factors all have influence on the DOM composition and lead to seasonal fluctuations (Mineau et al., 2016; Fork et al., 2020; Meingast et al., 2020). Human bacteria measured in the large watershed scale sites indicated some seasonal differences in concentrations and loadings (Dila et al., 2018). Soil saturation, which is commonly high in the spring in the Great Lakes region, can have a substanial influence on potential for inflow and infiltration into sanitary sewers and the likelihood of exceeding sanitary sewer capacity and increasing the risk of sanitary sewer overflows (Cahoon and Hanke, 2017). Soil saturation level also governs the baseflow level and the severity of precipitation response in streams: increased flow increases the capacity of the stream to carry sediment, thereby increasing turbidity (Mukandan et al., 2013). All of these factors, and likely others, provide potential for seasonal changes in each of the model response and predictor parameters and the overall models themselves.

**Small scale variability**

Small scale sampling included storm sewers and open channels in urban areas. These sites drained small areas, and therefore, have relatively small flows that can be greatly influenced by contamination from individual discharges (e.g misconnections or leaks from single sources). Depending on the nature of the individual contamination source and timing of the discharge in relation to sample collection times, DOC (and therefore optical signals) and human-associated bacteria can vary widely. Wastewater contamination from these individual sites can include sources such as laundry, bathing, and industrial waste that do not necessarily contain human waste, but will have influence on optical signals, and some sources contain human waste that will have great influence on human bacteria as well as optical signals. This type of variability is illustrated in sample results from local sanitary sewers during the sampling events for small scale sites: Concentrations of sHM in local sanitary sewers varied by three orders of magnitude and fluorescence signals T and F varied more than 2 orders of magnitude in these samples (figure SI-xx boxplots). The concentrations of DOC in local sanitary sewers varied from concentrations less than most environmental samples to concentrations greater than the regional WWTP influent samples (2.6-149 mg/L). This type of variability in the targeted contamination source presented a challenge for model development which requires a relatively consistent source signal.

In addition, other non-human sources can influence the optical signals in an unpredictable manner such as, wildlife waste, pet waste, and leaf leachate (Bratt, et al., 2017; Fork et al., 2020). This is likely to be another substantial reason that modeling attempts were unsuccessful for small-scale sites. In contrast, flow contributing to the watershed and subwatershed sites represents water from many diverse, small-scale sources that are homogenized in the main channel, making it more likely to result in a consistent relationship between optical signals and human bacteria to facilitate model development. This difference among different scale sites is a likely reason that development of models for small scale sites was not possible while reasonable models were achievable for the watershed and subwatershed scale sites.

**Surrogate models**

Previous efforts have developed surrogate models for prediction of fecal indicators in recreational waters (Francy et al., 2013, He and He, 2008; Shively et al., 2015; Brooks et al., 2016), and in rivers (Christensen, et al., 2002; Baldwin et al., 2012). These models often use basic water quality parameters such as turbidity and water temperature as well as physical parameters such as wave height, water currents, wind velocities, and stream flow as predictor variables. Some have been used to assist managers of recreational waters for issuing swimming advisories (Heberger et al., 2008; Dada and Hamilton, 2016), and others for helping to understand dynamics of water quality or for computation of stream loadings (McKergow and Davies, 2010; David and Haggard, 2010). The majority of these modeling efforts have focused on prediction of general fecal indicators such as E. coli or fecal coliform that provide general microbial contamination, but do not necessarily indicate presence of wastewater and its associated pathogenic and chemical contamination. Expanding the use of surrogate models for human-associated bacteria in the current study provides opportunity to gain information on the level of wastewater contamination directly associated with breaches in the sanitary conveyance system that IDDE efforts could work to identify and control.

**Application**

Potential exists for valuable applications of optical field sensors, but there remain a number of limitations with this approach that are important to consider: Given that there was no universal relationship for all sites, validation of model effectiveness and calibration of coefficients will be required for each future application. Translation of these concepts to the field setting will need to be validated through long-term monitoring in different hydrologic conditions and throughout all seasons. Accuracy in comparison to laboratory measurements and stability of the sensor signals over time will need to be evaluated through regular quality control measurements. Field applications can be complex with multiple challenging factors to overcome such as biofouling, logistics of in-stream and flow-through system deployments, and assurance that sampling with a sensor at an individual location is representative of water throughout the stream cross section. In this study, models were developed using data from one to two year time periods. Further study would be needed to verify stationarity of these models over time due to changes within a watershed such as runoff management actions that have potential to change the relationship between human-associated bacteria and optical properties of water. In addition, the current study included multiple watersheds with variable influences, but was limited to the Great Lakes region. Validation for additional geological and geographic areas would provide insight into the transferability of these concepts to more diverse watershed settings.

Even with these limitations, the application holds promise for providing real time estimates of general fecal indicator bacteria, human-associated bacteria markers, and wastewater contamination. Successful implementation of a field application could include continuous deployment at stationary sites or discrete sampling during targeted time periods. Continuous deployment would allow for prediction of bacteria concentrations, bacteria loadings when coupled with flow measurement, and estimation of the wastewater content in streams at a fine time scale. This approach would provide advantages over discrete sampling for bacteria directly: Cost of sampling would be substantially less than that from laboratory analysis of numerous samples collected at a fine time scale, results would be available in real time either through telemetry for unattended continuous sampling or on-site for manual sampling operations. Laboratory analyses typically take days to months. Results during continous deployment could also be used to trigger discrete sampling at threshold levels to verify bacteria and wastewater predictions. While the application does not appear to work well at the very small drainage area scale for tracing contamination to the source, deployment in large watersheds and subwatersheds would provide a means to prioritize relative contamination levels in different portions of the watershed, and the time periods in which they occur, allowing resource managers to focus on areas with the largest contamination contributions for remediation.

#References

# Actions needed:

1. Human-associated bacteria markers (CHECK NAME CONSISTENCY)
2. Search for modeling of human-associated bacteria
3. Write abstract

Modeling points to make:

1. MMSD:
   1. There were three CSO events for subwatershed models. All three of these are outliers in the model. The model under-predicts these values for all bacteria modeled. Adding a CSO term aligns these with the rest of the model (adjustment of the intercept), but three observations is not enough to provide confidence in that model term. Additional CSO event samples would be needed to provide sufficient confidence in these results.