Spring 2022

Escherichia coli in Lake Erie: Impact on Lake Health

Mary Grmek
The University of Akron, mmg121@uakron.edu

***Escherichia coli* in Lake Erie: Impact on Lake Health**

3100:497:002

Author: Mary Grmek

Sponsor: Dr. Randy Mitchell

Advisor: Dr. Brian Bagatto

Readers: Dr. Hazel Barton and Elizabeth Militante-Advey, REHS

Date: April 22, 2022

Abstract

Lake Erie is a vital resource for its surrounding area, and one of its most important purposes is as a recreational body of water. In determining whether Lake Erie's water quality meets U.S. EPA standards, daily surveys and water samples were taken on behalf of the Lake County General Health District, as well as *E. coli* samples every other day. This paper quantified the *E. coli* count for both Mentor Headlands Beach and Fairport Harbor Beach in Lake County over the swimming seasons for 2020 and 2021, as well as evaluated which variables collected in daily surveys impact *E. coli* count. Both wave height and 24-hour precipitation (collected through USGS gage 04212100 Grand River near Painesville, OH) were shown to be associated with *E. coli* count. The overall health of Lake Erie, according to U.S. EPA standards, is decent, but the Harmful Algal Bloom severity index suggests otherwise.

Keywords: *Escherichia coli*, Lake Erie, Environmental Health, Water Quality, Fecal Coliform

Introduction

Despite the widely held belief that Lake Erie has grown cleaner and cleaner over the years, in actuality its *E. coli* levels have increased significantly as of 2014 (Weiskerger and Whitman 1239). This has a destructive impact on public health. Lake Erie is utilized in many ways, with recreational use being one of the most popular. With higher *E. coli* levels impacting swimming safety, beachgoers are less likely to go swimming, which can result in the loss of millions of dollars, specifically in tourism (Wolf et al. 468). This is therefore a problem that affects not only the overall environment and ecosystem of the lake, but also an industry reliant on it.

Escherichia coli is a rod-shaped, Gram-negative bacteria that *has* long been utilized as an indicator of water quality. It is widely studied and is typically referred to as a fecal indicator bacterium (FIB) (Jang et al. 570). To be considered a fecal indicator bacterium, a bacterium must fulfill certain requirements; these include presence in intestinal tracts, similar survival to pathogens, inability to multiply, and non-pathogenic, among others (Ishii and Sadowsky 101). The presence of FIB in water is a good indicator that the body of water has not only been polluted by fecal material but could also potentially house other pathogens (An et al. 771). *E. coli* can originate from a variety of sources, including wastewater discharges, runoff (both storm and agricultural), and human and animal waste (An et al. 771). *E. coli* is measured in both MPN (most probable number) and cfu/100 mL (colony forming units per 100 mL), which both refer to the estimated concentration within a solution. For the remainder of this paper, *E. coli* will be referred to in MPN. MPN and cfu/100 mL are freely interchangeable and refer to the same amount of *E. coli*.

One of the most prominent threats to the overall health of both Lake Erie and the people who utilize it is Harmful Algal Blooms (HABs). Though occurring elsewhere in the world, HABs are especially prolific in Lake Erie, specifically the Western basin near Toledo. HABs are defined by the National Oceanic and Atmospheric Administration (NOAA) as "when colonies of algae grow out of control and produce toxic or harmful effects on people and [marine life]" (NOAA, 2016). They can have a variety of effects, with some of the more debilitating including hypoxia and dead zones, which can in turn cause eutrophication. Other HABs are incredibly toxic to the surrounding environment and can induce life-threatening illnesses to a variety of organisms. The main contributing factor to HABs is thought to be an increase in the amount of rate-limiting nutrients, specifically Nitrogen and Phosphorus (Anderson et al. 705). These nutrients are typically found in stream runoff and can cause a sharp increase in these HABs. Though no link has been found between *E. coli* and HABs specifically, HABs play a vital role in determining water quality. In 2021, the severity index of HABs spotted in Lake Erie was a 6.0 out of 10, which is classified as moderately severe (NOAA Western…2021). This is in sharp contrast to the severity index of 2020, which came in at only a 3.0 (NOAA Western…2020).

Determining water quality and the overall health of a water body can be difficult, as so many factors have an effect on water bodies as big as Lake Erie. To evaluate the water quality of Lake Erie, we turn to the U.S. EPA. Lake Erie is classified as a recreational body of freshwater, and thus *E. coli* is the ideal FIB, as declared by the U.S. EPA in 1986. Additionally, they recommended that water contact is not advised for freshwater beaches when *E. coli* levels surpass 235 cfu/100 mL (colony forming units per 100 mL of water) or "the geometric mean of *E. coli* counts of at least 5 samples, equally spread over a 30-day period, exceeds 126 colonies per 100 mL water" (U.S. EPA, 1986). The first guideline was the one most utilized in data collection for the Lake County General Health District. Though Xu et al. (2001) proposes a multi-pronged assessment technique with three distinct kinds of indicators (structural, functional, and system-level) that may be more useful, it is not utilized within this study.

The Lake County General Health District sought to better understand the factors affecting *E. coli* count and HABs with regard to lake health at Mentor Headlands Beach and Fairport Harbor Beach. This report concerns the possible influence of several causes, including USGS wind speed, USGS wind direction, wave height, and USGS 24-hour rainfall. This data collection occurred through the Lake County General Health District as the position of Water Quality Intern. The purpose of this data collection was to predict the *E. coli* count at each beach every day and determine if the water quality was safe enough to swim, depending on the U.S. EPA criteria. As a result, data collection only occurred during the swimming season, which typically falls from late May to early September, ending on Labor Day. Though no data was collected for HABs in this study, I will still touch on their impact and how they could relate to the results of this study.

Methods

Study Area

The *E. coli* levels were quantified from water samples collected every other day from late May to early September in both 2020 and 2021. These samples were collected by me and measured 100 mL each. They were collected in waist deep water at two different sites, Mentor Headlands Beach and Fairport Harbor Beach.

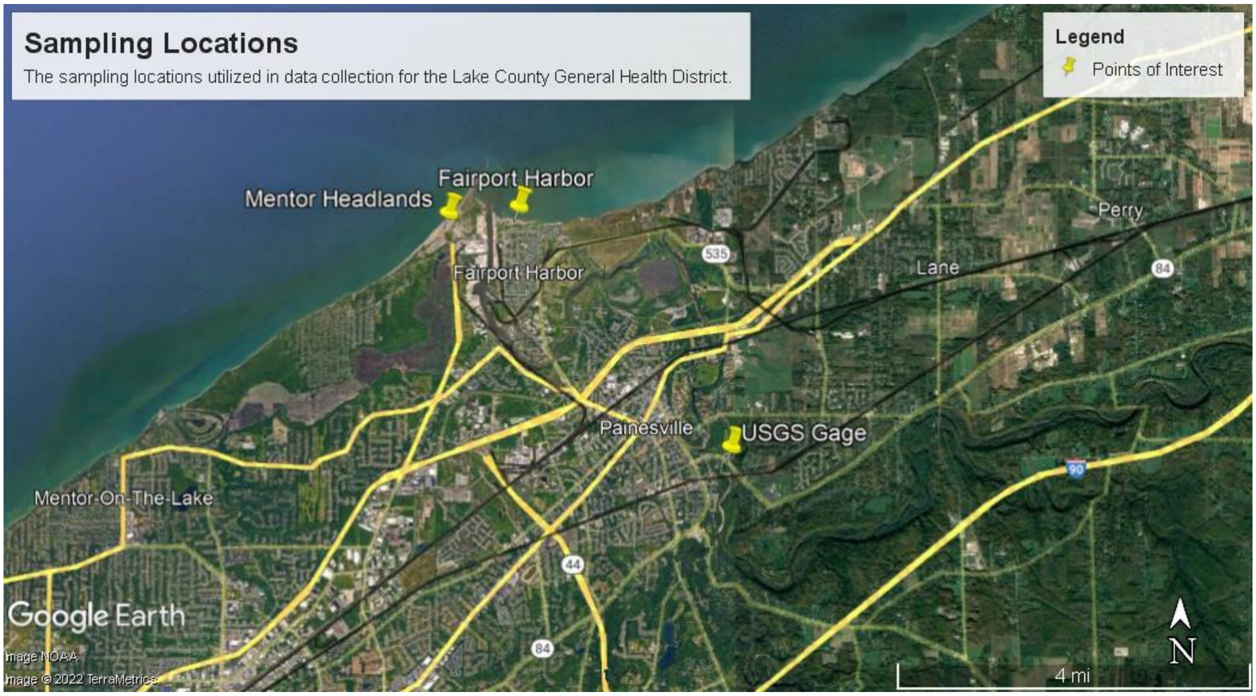


Figure 1. A photo taken from Google Earth Pro that highlights the two sampling beaches (Mentor Headlands and Fairport Harbor) and the location of the two relative to each other and the USGS gage used for data collection.

Data Collection

In addition to *E. coli* count, I recorded several other variables onsite every day at the sample time, including wave height (inches), water temperature (Celsius and Fahrenheit), air temperature (Celsius and Fahrenheit), swimmer count, bird count, and beach-goer count, as well as the presence of any dead fish. All of this data was collected by me, the Lake County water quality intern. I also pulled other relevant data from the Nowcast website, which reports daily recreational water quality across Ohio, Michigan, and Pennsylvania, and exported as a csv file (U.S. EPA, 2022). NowCast pulls data from the USGS gage 04212100 Grand River near Painesville, OH. This data included current lake level, lake discharge at 0800, rainfall in the past 24, 48, and 72 hours, wind speed, and wind direction. The daily change in lake level was calculated in Excel by subtracting the current day's level from the previous day's level. NowCast also pulled data from the NWS site for Burke Lakefront Airport in Cleveland, OH and included the rainfall for the past 24, 48, and 72 hours (as well as a weighted calculation for 48 and 72 hours), wind speed, and wind direction. This process for data collection was conducted for both summers that samples were taken.

E. coli **Assessment**

The exact amount of *E. coli* in each sample was measured in cfu/100 mL and assessed by using Colilert-24, a single-use capsule reagent which identifies *E. coli* without the need for media (IDEXX). Once the capsule of Colilert-24 was added to each sample bottle, the bottle was shaken until the Colilert was fully dissolved. The bottles were left out for several minutes to let any bubbles dissipate. The samples were then loaded into IDEXX Quanti-Tray 2000s, which contain 97 separate wells (49 large, 48 small) and can quantify up to 2,419 cfu/100 mL in a given 100 mL sample (Weber-Scientific 2022). Once loaded, the trays were sealed and left to incubate at 95 degrees Fahrenheit. After incubating for 24 hours, the *E. coli* levels were determined utilizing a blacklight – if *E. coli* were present in a given well, it appeared to glow under the blacklight. Once all the black lit wells were counted, *E. coli* counts were measured using a table provided by IDEXX. This table detailed the predicted amount of *E. coli* based on the number of large and small wells that lit up under blacklight – the more wells that lit up, the more *E. coli* there were determined to be in the sample.

Analysis

Data analysis was conducted using the GraphPad InStat software. Each day's data (*E. coli* sample count, weather, lake conditions) was treated as a separate data point. Multiple linear regressions between *E. coli* (MPN) and wave height (inches), USGS wind speed (MPH), USGS wind direction (degrees), and USGS 24-hour rainfall (inches) at each beach for each year. The intention was to discover the impact, if any, that the latter four variables had on *E. coli* count.

Results

E. coli counts varied widely amongst both beaches from sample day to sample day. To quantify the amount of *E.coli*, the mean MPN of *E. coli* was calculated for each month's data was collected at each site, and can be viewed in Tables 1 and 2. The mean *E. coli* in MPN for Fairport Harbor was 146.04 in 2021, and 45.77 in 2020. The mean *E. coli* in MPN for Mentor Headlands was 88.84 in 2021, and 83.26 in 2020. It should be noted that for the months of May and September in both years that data collection occurred significantly less than in the months of June, July, and August.The *E. coli* count over the months is visualized in Figures 3-6.

E. coli counts were sometimes related to measured factors, and sometimes not, and the important factors differed among years and sites. Daily wave height was associated with *E. coli* count at Mentor Headlands in 2020 and 2021. Likewise, USGS 24-hour rainfall was associated with *E. coli* count at Fairport Harbor in 2021 and Mentor Headlands in 2021. No variables were associated with the *E. coli* count at Fairport Harbor in 2020. These multiple linear regressions were run for each beach for each year, with the intention of showing how *E. coli* count changed between the years. Of these variables, only two were found to have an association with *E. coli* count according to the multiple linear regressions. For Mentor Headlands in 2020, this was the multiple linear regression equation used:

```
[A:E. coli (MPN)] = 16.095 + 73.126*[B:24h rain (in)] + 1.853*[C:Wind sp. (MPH)]
     + 0.04023*[D:Wind dir. (deg)] + 4.914*[E:Wave ht (in)]
```

Wave height was found to impact *E. coli* count with a p value of 0.0495. For Mentor Headlands in 2021, this was the multiple linear regression equation used:

```
[A:E.coli (MPN)] =  -22.111 + 28.518*[B:24h rain (in)] + 1.821*[C:Wind sp (mph)]
     + 0.2587*[D:Wind dir (deg)] + 3.572*[E:Wave ht (in)]
```

Both wave height and USGS 24-hour rainfall were originally associated with *E. coli* count, with respective p values of 0.0082 and <0.0001. There was a significant outlier present in the regression performed for Mentor Headlands in 2021 – this was shown to have an influence on the regression results. When removed, the regression was not statistically significant for USGS 24-hour rainfall (p-value of 0.6951), but remained significant for wave height. Fairport Harbor in 2020 did not have any variables that indicated statistical significance in association with *E. coli* count. For Fairport Harbor in 2021, this was the multiple linear regression equation used:

```
[A:E. coli (MPN)] = 48.674 + 978.12*[B:24h rain (in)] + 10.814*[C:Wind sp (mph)]
     - 0.2404*[D:Wind dir (deg)] + 4.022*[E:Wave ht (in)]
```

USGS 24-hour rainfall was found to impact *E. coli* count with a p value of <0.0001. Regardless, the only two variables that were statistically significant predictors of E. coli count were wave height and USGS 24-hour rainfall. USGS wind speed and direction therefore had no bearing on *E. coli* count in either year. For each statistically significant variable, a scatter plot was generated comparing the given variable against the *E. coli* count in MPN. This data is visualized in Figures 7-11.

Table 1. Table displaying the mean amount of *E. coli* in MPN, the standard deviation of *E. coli*, and the number of *E. coli* samples taken per month in both 2020 and 2021 for Mentor Headlands Beach.

Mentor Headlands	Mean *E. coli* in MPN per month	SD *E. coli* per month	N *E. coli* per month
2021			
May	104.94	64.83	5
June	88.20	227.97	16
July	111.61	218.92	17
August	64.51	73.64	14
September	74.94	36.72	5
2020			
May	36.40	48.37	6
June	70.18	100.85	16
July	63.54	68.20	16
August	96.73	110.83	16
September	149.43	209.69	4

Fairport Harbor	Mean *E. coli* in MPN per month	SD *E. coli* per month	N *E. coli* per month
2021			
May	278.23	173.03	4
June	53.65	84.54	16
July	159.12	293.32	19
August	206.15	459.28	16
September	33.05	19.18	4
2020			
May	83.97	98.62	6
June	30.49	33.95	14
July	46.57	65.40	18
August	54.36	105.45	16
September	13.47	12.61	3

Table 2. Table displaying the mean amount of *E. coli* in MPN, the standard deviation of *E. coli*, and the number of *E. coli* samples taken per month in both 2020 and 2021 for Fairport Harbor Beach.

MENTOR HEADLANDS		
2021		
	P-Value	Parameter Estimate
Overall	<0.0001	-22.111
USGS 24-hour rainfall (in)	<0.0001	28.518
USGS wind speed (mph)	0.6056	1.821
direction (degrees)	0.0603	0.2587
Wave height (in)	0.0082	3.572
2020		
Overall	0.0063	16.095
USGS 24-hour rainfall (in)	0.4073	73.126
USGS wind speed (mph)	0.7587	1.853
USGS wind direction	0.8291	0.04023
Wave height (in)	0.0495	2.43

Table 3. Table displaying the overall p-value and parameter estimate for the multiple regression analysis of data from both years at Mentor Headlands Beach, as well as the individual p-values and parameter estimate of each variable.

FAIRPORT HARBOR		
2021		
	P-Value	Parameter Estimate
Overall	<0.0001	48.674
USGS 24-hour rainfall (in)	<0.0001	978.12
USGS wind speed (mph)	0.3951	10.814
USGS wind direction (degrees)	0.6232	-0.2404
Wave height (in)	0.7429	4.022
2020		
Overall	0.097	22.379
USGS 24-hour rainfall (in)	0.1043	136.76
USGS wind speed (mph)	0.445	-3.842
USGS wind direction (degrees)	0.4394	0.1477
Wave height (in)	0.4394	2.517

Table 4. Table displaying the overall p-value and parameter estimate for multiple regression analysis of the data from both years at Fairport Harbor Beach, as well as the individual p-values and parameter estimate of each variable.

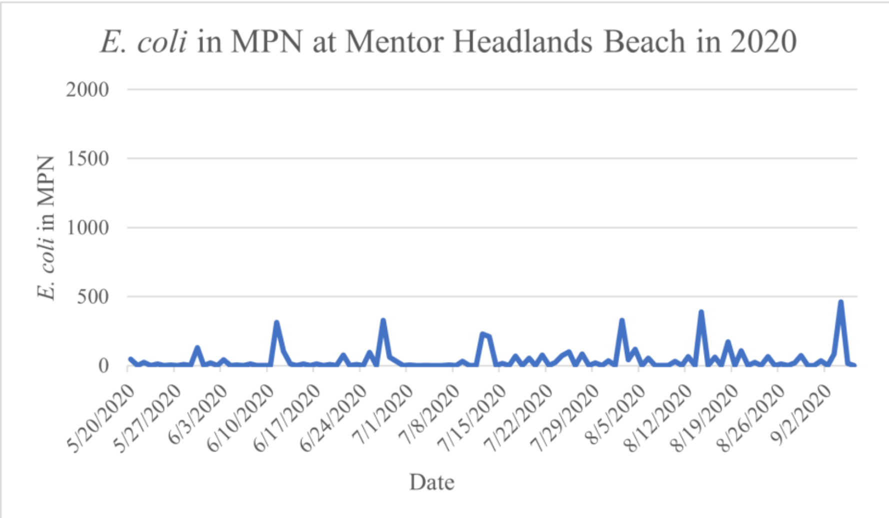
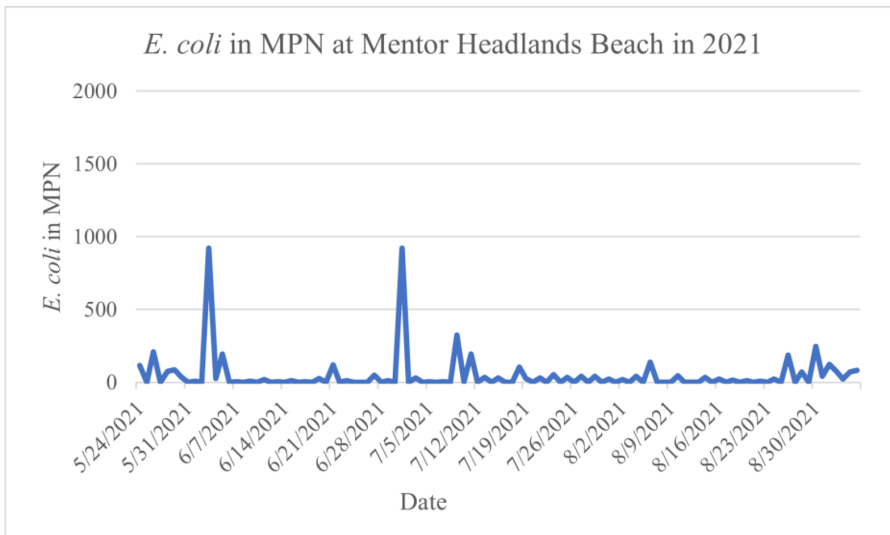


Figure 3. A line graph displaying the amount of *E. coli* over time at Mentor Headlands Beach in 2020.

Figure 4. A line graph displaying the amount of *E. coli* over time at Mentor Headlands Beach in 2021.



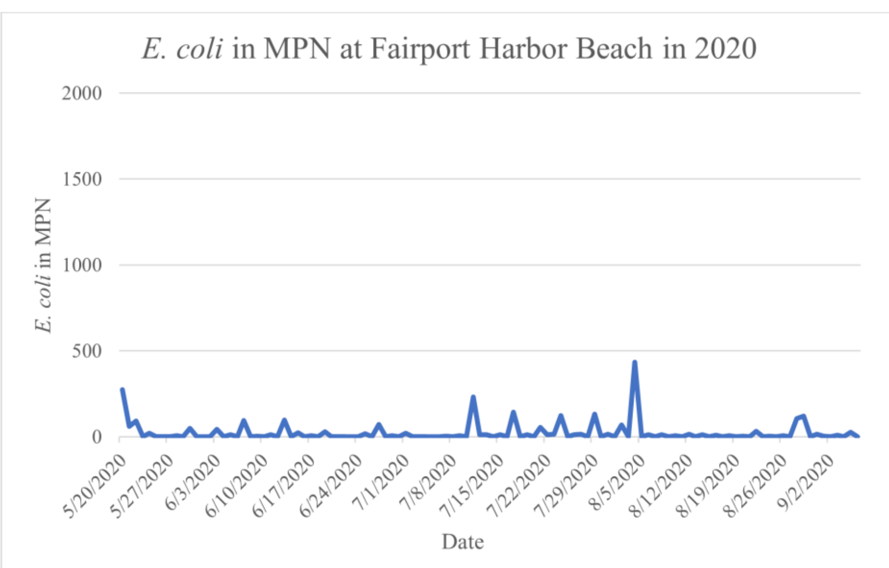


Figure 5. A line graph displaying the amount of *E. coli* over time at Fairport Harbor Beach in 2020.

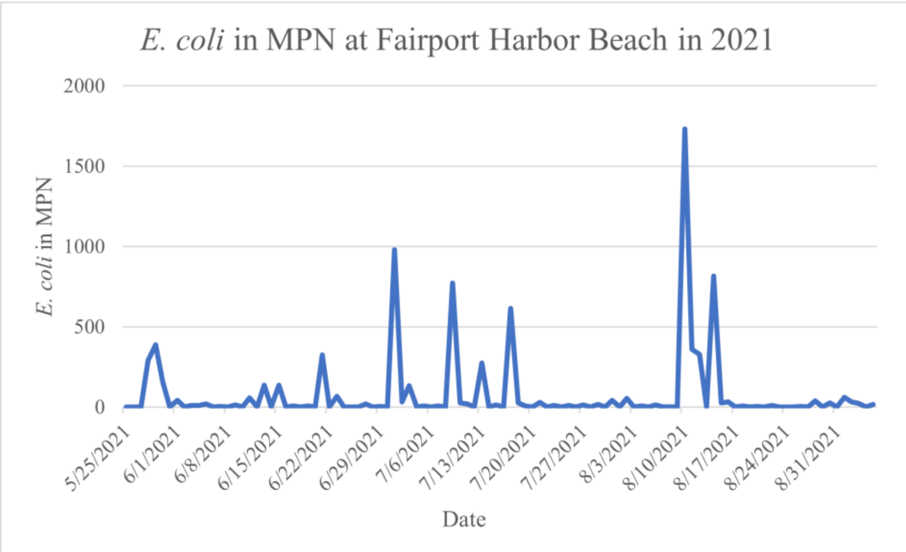
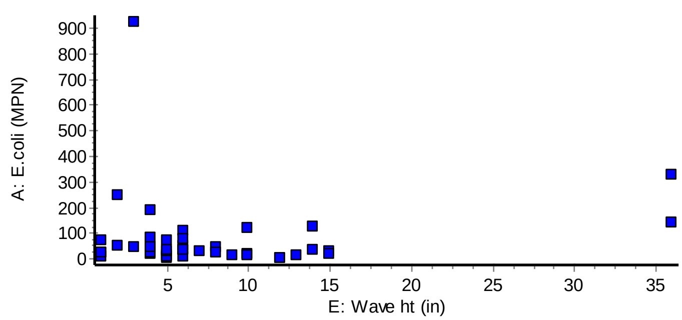


Figure 6. A line graph displaying the amount of *E. coli* over time at Fairport Harbor Beach in 2021.

Figure 7. A scatter plot visualizing the *E. coli* in MPN versus the wave height in inches for Mentor Headlands Beach in 2020. The multiple linear regression for this comparison was statistically significant.



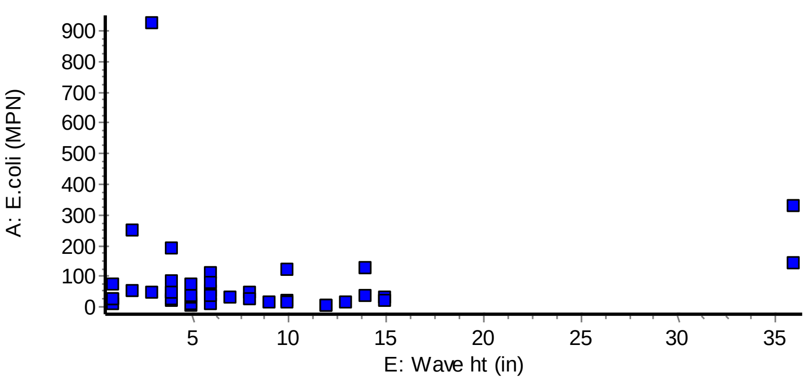
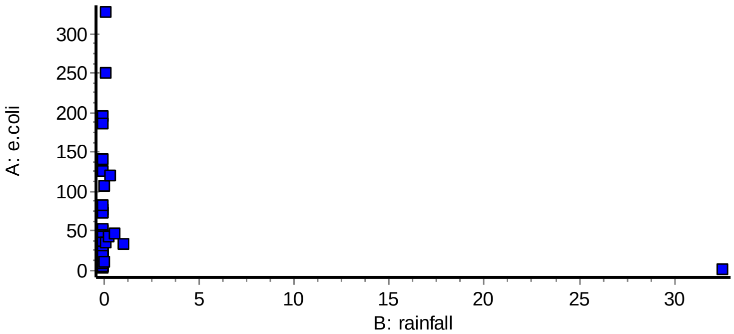
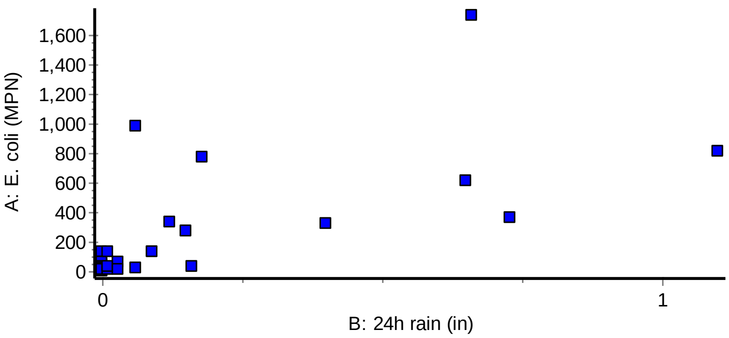


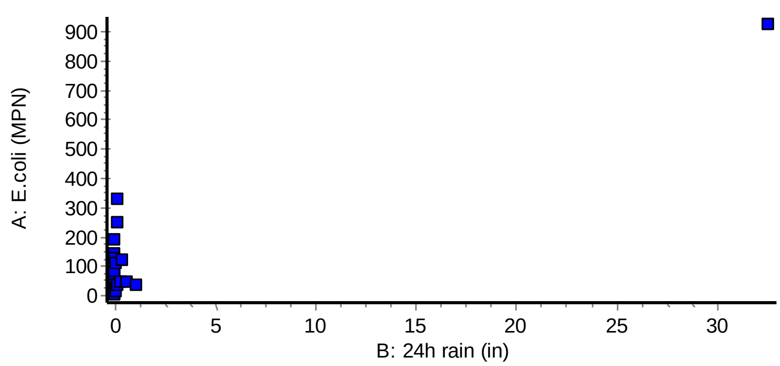
Figure 8. A scatter plot visualizing the *E. coli* in MPN versus the wave height in inches for Mentor Headlands Beach in 2021. The multiple linear regression for this comparison was statistically significant.





Figures 9 and 10. Scatterplots showcasing the *E. coli* in MPN versus the USGS 24-hour rainfall in inches. Figure 9 is the data without the two days of outliers, while Figure 10 is the data with the two days of outliers. These outliers are the sole reason for statistical significance between these two variables. The multiple linear regression for Figure 10 was found to be statistically significant, while the multiple linear regression for Figure 9 was not found to be statistically significant.

Figure 11. A scatter plot visualizing the *E. coli* in MPN versus the USGS 24-hour rainfall in inches for Fairport Harbor Beach in 2021. The multiple linear regression for this comparison was statistically significant.



Discussion

The mean *E. coli* count at each beach displays a wide amount of variation. Though the average amount of *E. coli* at Mentor Headlands was fairly consistent in both years, Fairport Harbor displayed an exceptionally substantial increase in *E. coli* count over the year, jumping from 45.77 MPN in 2020 to 146.04 MPN in 2021. Something to note with the *E. coli* count in 2021 for both beaches is that the USGS 24-hour rainfall had an impact, though no statistical significance was found for Mentor Headlands in 2021 when the larger outliers (i.e., the two days with an *E. coli* count of 920.8) were removed. Regardless, this could potentially be the reasoning behind this increase in average *E. coli* at both beaches. Another interesting comparison to be made is the spike that occurs in August 2021 at Fairport Harbor. This spike occurs around August 10 and reaches past 1800 MPN. Interestingly enough, Mentor Headlands displays no spike in or around this date, suggesting that USGS 24-hour rainfall, though impactful, is not the only factor that impacts *E. coli* count. If it truly were the only variable, the *E. coli* count at both beaches would read the same, which is not the case. Additionally, wave height was found to be a significant variable that affects *E. coli* count at Mentor Headlands for both 2020 and 2021. Fairport Harbor, adversely, was not significantly affected by wave height.

According to U.S. EPA standards, the water quality of these specific beaches off of Lake Erie is acceptable for its purpose as a recreational body of water. As stated earlier, water contact is not advised for freshwater beaches when *E. coli* levels surpass 235 cfu/100 mL. The average *E. coli* in MPN for each beach, even monthly, does not surpass this number, so as a whole Lake Erie would be determined safe. There were certain days in every season that surpassed this number, but more often than not Lake Erie was safe to swim in. Another aspect to consider with overall lake health and water quality is the amount of HABs in Lake Erie of each year. In 2020, Lake Erie's severity index for HABs was a 3.0, while the severity index for Lake Erie in 2021 was a 6.0 – a sharp increase from one year to the next. This does not bode well for the health of the lake; HABs can have negative consequences for all organisms who utilize the lake in their lives. It is important to note that while this data spans only two years, literature has indicated that *E. coli* counts as recently as 2014 have also been higher than in past years (Weiskerger and Whitman 1239). While no literature explores the possible connection between *E. coli* and HABs, it is interesting to note that both beaches experienced an increase in *E. coli* from 2020 to 2021, similar to how the severity index increased for Lake Erie. Even if these two variables are unlinked, it is possible there may be affecting factors they share (i.e., precipitation, wind speed, lake level, etc). This highlights a research direction for the future.

There were quite a few difficulties in analysis of this project. Primarily, *E. coli* samples were collected every other day and not every day like the rest of the variables. This results in a smaller sample size per season for *E. coli* compared to variables like wave height or wind speed. Additionally, because so many variables were sourced from outside the Health District, there are gaps of time where no data was collected due to technical and mechanical errors, resulting in an even smaller sample size. The measurements collected from the USGS gage and Burke Lakefront Airport are the same for both beaches, though this may not necessarily be true. The beaches are geographically close, but could have variances in data collected, such as precipitation,

Acknowledgements

I would like to thank the Lake County General Health District for allowing me the opportunity to serve in the Water Quality Intern position, as well as Elizabeth Militante-Advey for her guidance and willingness to act as a reader for this project. I am also thankful to the Department of Biology at the University of Akron, specifically my sponsor Dr. Randall Mitchell, my reader Dr. Hazel Barton, and my advisor Dr. Brian Bagatto. The funding and equipment were provided by the Lake County General Health District.

Works Cited

An, Y.-J., Kampbell, D. H., Peter Breidenbach, G. (2002). Escherichia coli and total coliforms in water and sediments at Lake Marinas. *Environmental Pollution*, *120*(3), 771–778.

Anderson, D. M., Glibert, P. M., Burkholder, J. M. (2002). Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries*, *25*(4), 704–726.

Google. (2022). Google maps. Google; [accessed 2022 March 27]. https://maps.google.com/

IDEXX. (n.d.). Colilert - IDEXX US. IDEXX US; [accessed 2022 April 15]. https://www.idexx.com/en/water/water-products-services/colilert/

Ishii, S., Sadowsky, M. J. (2008). Escherichia coli in the environment: Implications for water quality and human health. *Microbes and Environments*, *23*(2), 101–108.

Jang, J., Hur, H.-G., Sadowsky, M. J., Byappanahalli, M. N., Yan, T., Ishii, S. (2017). Environmental Escherichia coli: Ecology and public health implications - A Review. *Journal of Applied Microbiology*, *123*(3), 570–581.

NOAA Western Lake Erie Harmful Algal Bloom Seasonal Assessment 2020. October 29, 2020. Ohio: NOAA; [accessed 2022 March 23]. https://cdn.coastalscience.noaa.gov/hab-data/bulletins/lake-erie/2020/finalAssessment_2020-09.pdf.

NOAA Western Lake Erie Harmful Algal Bloom Seasonal Assessment 2021. November 1, 2021. Ohio: NOAA; [accessed 2022 March 23].

https://nccospublicstor.blob.core.windows.net/hab-data/bulletins/lake-erie/2021/finalAssessment_2021-08.pdf.

USEPA. (1986). Bacteriological ambient water quality criteria for marine and fresh recreational waters. EPA 440/5-84-002. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH.

USGS. (2022). Great Lakes Nowcast status. USGS; [accessed 2022 April 15]. https://pa.water.usgs.gov/apps/nowcast/

Weber Scientific. (2022). Quanti-Tray® and Quanti-Tray 2000® from Idexx. IDEXX; [accessed 2022 April 15]. https://www.weberscientific.com/quanti-tray-and-quanti-tray-2000-from-idexx

Weiskerger, C.J., Whitman, R.L. (2018). Monitoring E. coli in a changing beachscape. Science of the Total Environment, 619–620:1236–1246.

What is a harmful algal bloom? April 27, 2016. NOAA; [accessed 2022 March 22]. https://www.noaa.gov/what-is-harmful-algal-bloom.

Wolf, D., Chen W., Gopalakrishnan, S., Haab, T., Klaiber, H.A. (2019). The impacts of harmful algal blooms and E. coli on recreational behavior in Lake Erie. Land Economics, 95:455-472.

Xu, F., Tao, S., Dawson, R.W., Li P-G, Cao, J. (2001). Lake Ecosystem Health Assessment: Indicators and methods. *Water Research*, *35*(13), 3157–3167.