Journal of Research High School

Volume 2022 (2) 85 - 90
ISSN 2688-3651 [online]

Correlation of the *E. coli* Concentration with Other Water Quality Parameters in an Impaired Local Creek

Christina D. Hwang[1 *]

[1]Liberal Arts and Science Academy, Austin, TX USA

Received May 16, 2022; Revised August 12, 2022; Accepted, August 29, 2022

Abstract

This study was to determine a correlation of *E. coli* concentrations with other typical water quality parameters of pH, dissolved oxygen, temperature, turbidity, electrical conductivity, suspended solids, chloride, sulfate, nitrate, phosphorus, and total hardness as well as flow rate. The recent water quality data from Plum Creek, Texas from September 2020 to June 2021 were utilized. The Pearson's correlation coefficients (r) were calculated between the *E. coli* concentration and the aforementioned water quality parameters. Results showed that the *E. coli* concentration correlates very well with the suspended solids concentration (r = 0.57) and flow rate (r = 0.629). Therefore, the suspended solids and flow rate that are relatively easy and quick can serve as indicators of the *E. coli* concentration. It is construed that stormwater runon to Plum Creek during rainstorms significantly increased the *E. coli* concentration in the Plum Creek.

Keywords: E. coli, Pearson's correlation, water quality

1. Introduction

Disease-causing microorganisms (pathogens) contaminate recreational or drinking water and cause waterborne illness (Gelting et al., 2011; Liu et al., 2018). Diarrhea is the most common waterborne illness and, following pneumonia, it is the second leading cause of death for children under the age of five. In the United States, it is estimated that one in 44 people gets sick from waterborne illness each year and that hospitalizations and emergency department visits due to waterborne illness caused 3.33 billion dollars in direct healthcare costs in 2014 (CDC, 2020).

Escherichia coliform (*E. coli*) is an indicator microorganism of the presence of pathogens in water. That is, when large concentrations of *E. coli* are present in the water, there is a high probability that fecal contamination has occurred, and therefore, other pathogens are also present in water. The United States Environmental Protection Agency (USEPA) set the level for *E. coli* in recreational freshwater at 126 Colony Forming Units (CFU) per 100 mL of water sample (USEPA, 2012). *E. coli* prevalence was also an effective predictor of foodborne pathogen presence in pastured poultry farms (Xu et al., 2022). On the other hand, several research have shown that the concentration of *E. coli* does not necessarily correlate with the concentration or presence of pathogens (Dorner et al., 2007).

Plum Creek begins north of the City of Kyle in Hays County, Texas, and runs 52 miles southeast through Caldwell County before meeting the San Marcos River. Plum Creek and its tributaries drain an area of 397 mile^2. Plum Creek has played an important role as a reliable water source for settlers and their livestock and an area with recreation opportunities for families. In 2004, Plum Creek was

* Corresponding Author
christina.d.hwang@gmail.com

Advisor: Christina W. Lopez
christina.lopez@txstate.edu

listed by the State of Texas as having *E. coli* levels that impaired contact recreation uses of the stream. In 2018, a Bacterial Source Tracking study showed that the Plum Creek was substantially impaired by *E. coli* due to wildlife (feral hogs), failing septic tanks, and cattle wastes (TWRI, 2018). The land use in the Plum Creek Watershed currently consists of 10.4% urban, 11% cultivated crops, 15% forest, 33.8% pasture/hay, and 25.8% grasslands (PCWP, 2022a).

To ensure water quality, routine and targeted monitoring is conducted by the Guadalupe-Blanco River Authority through grant funding from the Texas State Soil and Water Conservation Board. *E. coli* is typically quantified by either multiple tube/multiple well (e.g., Colilert) or membrane filtration (e.g., mColiblue-24) technique, and both require incubation for 24 hours and are labor and material intensive. As the presence of *E. coli* indicates the presence of pathogens responsible for water-borne diseases, early identification of potential presence of pathogens will ensure health and safety of the users of the stream. This study, therefore, aims to discover correlation of the *E. coli* concentrations with other water quality parameters that are easily measured and readily available for public use.

2. Materials and Methods

2.1. Water Quality Data

The recent water quality data from September 2020 to June 2021 were extracted from monthly routine monitoring at three monitoring sites (PCWP, 2022b) that were chosen based on their locations in the watershed (upper, middle, and lower sections in the Plum Creek, Fig. 1). Each site had 10 monitoring events during the aforementioned time period and, therefore, a total of 30 data per water quality parameter was used in the current study. Water quality parameters included in this study are pH, dissolved oxygen, temperature, turbidity, electrical conductivity (EC), suspended solids, chloride, sulfate, nitrate, phosphorus, total hardness, and *E.* coli as well as flow rate.

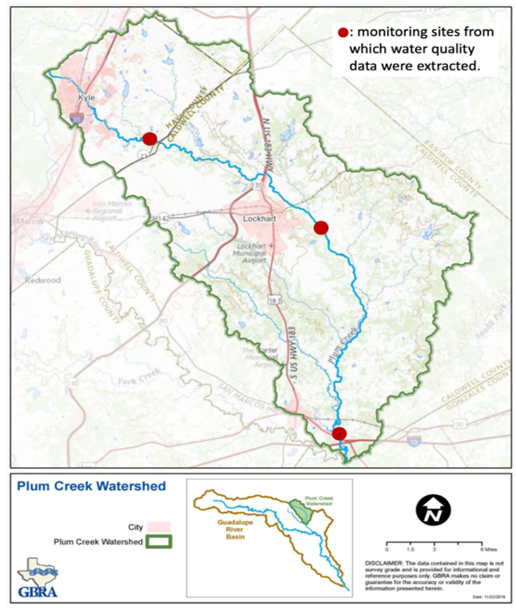

Figure 1. Three locations of monitoring sites in the main stream of the Plum Creek where water quality data were obtained.

2.2. Statistical Analysis

Statistical correlation analyses were performed using the Pearson's correlation coefficient formula (Eq. 1).

$$r = \frac{n(\Sigma xy)-(\Sigma x)(\Sigma y)}{\sqrt{\left[n\Sigma x^2-(\Sigma x)^2\right]\left[n\Sigma y^2-(\Sigma y)^2\right]}} \quad (1)$$

Where, r = Pearson's correlation coefficient, n = number of the variable pairs to be compared, and x and y = variables.

Eq. 1 provides a Pearson's correlation coefficient between (-)1 and (+)1, where (-1), zero, and (+1) indicates a strong negative correlation, no relationship, and a strong positive correlation, respectively. The larger the coefficient, the stronger the relationship between two variables (Table 1).

Table 1. The scale of Pearson's correlation coefficient, r

Absolute value of an "r" value	Extent of correlation
0.2	Weak
0.5	Moderate
0.8	Strong
1.0	Perfect

Pearson's correlation has been widely used in pattern recognition between two variables. For example, Mohanty et al. (2022, in press) used the Pearson's correlation to assess interrelationship between water quality parameters and Deng (2019) examined correlations between water quality parameters and the structure and connectivity of the river networks.

3. Results and Discussion

The *E. coli* concentration inversely correlated with pH, dissolved oxygen, EC, chloride, sulfate, nitrate, phosphorus, and total hardness concentrations. That is, the greater their concentrations, the lower the *E. coli* concentration was. In comparison, the *E. coli* concentration correlated positively with flow rate, temperature, turbidity, and suspended solids concentrations (Table 2 and Fig. 2). The flow rate showed the greatest correlation (r = 0.63) of all to the *E. coli* concentration, followed by the suspended solids concentration (r = 0.57). The *E. coli* concentration had a weak correlation with pH and temperature and, therefore, no further discussion about them was made in this article.

The concentrations of nitrate and phosphorus moderately correlated with the *E. coli* concentration with a value of r being -0.30 and -0.21, respectively. Nitrate and phosphorus are essential nutrients for growth and reproduction of living organisms, but in excess amount they can cause significant human health and water quality problems. In this study, Plum Creek had a nitrate concentration in the range of 0.71 mg/L to 12.4 mg/L with an average being 5.0 mg/L. Although Plum Creek is not a drinking water source for human consumption, the public can be exposed to a high level of nitrate concentration in Plum Creek during recreational activities. Chronic ingestion of drinking water high in the nitrate concentration can cause cancer and birth defects (Ward et al., 2018). In this regard, the USEPA set a maximum contamination level at 10 mg/L (as N) for nitrate (USEPA, 2022). The over-enrichment of aquatic ecosystems with phosphorus (and nitrate) contributes to algal bloom (i.e., eutrophication) resulting in decreased levels of dissolved oxygen and increased levels of algal toxins (Schindler, 2018). The phosphorus concentration in Plum Creek ranged between 0.32 mg P/L and 3.79 mg P/L with an average being 1.3 mg P/L. The USEPA has set a recommended limit of 0.05 mg/L for total phosphates in streams that enter lakes and 0.1 mg/L for total phosphorus in flowing waters to control eutrophication (USEPA, 1986). Therefore, the monitoring data indicate that Plum Creek is impaired with nitrate and phosphorus concentrations and needs to control input of these nutrients from wastewater treatment discharges or agricultural farm wash-off during rainstorms.

A moderately negative correlation of the concentrations of chloride (r = -0.32) and sulfate (r = -0.38) with the *E. coli* concentration was observed from the monitoring data. It is known that the chloride and sulfate concentrations in the water typically originate from anthropogenic sources such as wastewater treatment plants, septic systems, and industrial discharges (USEPA, 2022). The concentrations ranged between 48.6 mg/L and 226 mg/L for chloride and between 37.1 mg/L and 106 mg/L for sulfate. Nevertheless, these chloride and sulfate concentrations were below the non-enforceable Secondary Standard of 250 mg/L that the USEPA National Secondary Drinking Water Regulations set (USEPA, 2022).

Total hardness concentration also showed a moderately negative relationship (r = -0.33) with the *E. coli* concentrations. It ranged between 163 mg $CaCO_3$/L and 317 mg $CaCO_3$/L, showing a "very hard" characteristic of Plum Creek water. However, these high total hardness concentrations may be attributed to regional geological characteristics. In fact, the central Texas regions generally have a "hard" or "very hard" drinking water, with the total hardness at about 200 mg $CaCO_3$/L (HydroFlow, n.d.).

The monitoring data also showed that the *E. coli* concentration moderately correlated with EC (r = -0.34). EC increases with an increase in the concentrations of dissolved salts and other inorganic chemicals that conduct electrical current. Therefore, significant changes of conductivity may indicate a discharge of inorganic conductive pollutants to the aquatic resource. Plum Creek had an EC in the range of 540 /cm to 1290 /cm. The USEPA does not set either a Primary or a Secondary Standard for EC, but it recommends a value of less than 1,000 /cm for drinking water. EC is related to the total dissolved solids concentration (TDS) as follows:

$$TDS(mg/L) = k \bullet EC(\mu\Omega/cm) \quad (2)$$

where the value of *k* increases with an increase of ionic strength in water, and it ranges between 0.5 and 0.75 (Rusydi, 2018). Based on Eq. 2, the TDS concentration in Plum Creek is estimated to be in the range of 300 mg/L and 900 mg/L. The USEPA National Secondary Drinking Water Regulations set a non-enforceable Secondary Standard of 500 mg/L for TDS. Therefore, it can be said that Plum Creek is occasionally impaired with EC and TDS concentration.

As one of two physical water quality parameters that positively correlated with the *E. coli* concentration, turbidity of Plum Creek ranged between 2.7 and 70 NTU. Higher turbidity levels are often associated with higher levels of pathogenic microorganisms. However, turbidity was moderately correlated with the *E. coli* concentration (r = 0.37). Materials that cause water to be turbid (opaque or cloudy) include clay, silt, inorganic and organic matters, and microscopic organisms which are in general washed into rivers or creeks during a rainstorm. Also, solids on the bottom of rivers or creeks can be suspended during high flows.

The suspended solids concentration of Plum Creek ranged between 1.5 mg/L and 84.7 mg/L and was correlated very well (r = 0.57) with the *E. coli* concentration. The flow rate fluctuated from 1.9 ft^3/sec to 125 ft^3/sec and is also correlated very well with the *E. coli* concentration (r = 0.63). In fact, the suspended solids concentration and flow rate had a high correlation each other (r = 0.80). Therefore, the suspended solids concentration and flow rate can be served as indicators of the *E. coli* concentration in Plum Creek. In comparison to the *E. coli* enumeration that is costly, needs special equipment and apparatus, and takes 24 hours or longer for completion, suspended solids can be relatively easily and quickly quantified by a membrane filtration technique that takes approximately 2 hours and flow rate can be measured real-time with a flowmeter that measures water velocity which is multiplied by a cross-sectional area of the stream to calculate the flowrate.

Table 2. Pearson's correlation coefficients between *E. coli* concentration (CFU/100 mL) and other water quality parameters

Water Quality Parameter	Pearson's Correlation Coefficient	Correlation Direction	Correlation Extent
pH	-0.07	Negatively	None to Weak
Dissolved Oxygen (mg/L)	-0.17		
Phosphorus (mg P/L)	-0.21		Weak to Moderate
Nitrate (mg N/L)	-0.30		
Chloride (mg/L)	-0.32		
Total Hardness (mg $CaCO_3$/L)	-0.33		
Electrical Conductivity (/cm)	-0.34		
Sulfate (mg/L)	-0.38		
Temperature (°C)	0.31	Positively	Weak to Moderate
Turbidity (NTU)	0.37		
Suspended Solids (mg/L)	0.57		
Flow Rate (ft^3/sec)	0.63		

4. Conclusion

Based on high correlations of the *E. coli* concentration with the suspended solids concentration and flow rate, it is construed that a rainstorm that causes stormwater runon to Plum Creek watershed significantly increased the *E. coli* concentration in Plum Creek. This finding is in good agreement with the results from the previous Bacterial Source Tracking study that attributed the impairment of Plum Creek with *E. coli* to wildlife (feral hogs), failing septic tanks, and cattle wastes.

This study indicates that a high concentration of *E. coli* in Plum Creek can be warned to public when those easily measured surrogate water quality parameters of suspended solids concentration and flow rate are high. It should be, however, noted that there were a few outliers in water quality data (Fig. 2) which could have produced different results if they had been eliminated through a statistical evaluation such as Grubbs' test. In addition, further studies are warranted to elucidate the levels of suspended solids and flow rate that indicate the threshold impairment concentration of *E. coli*.

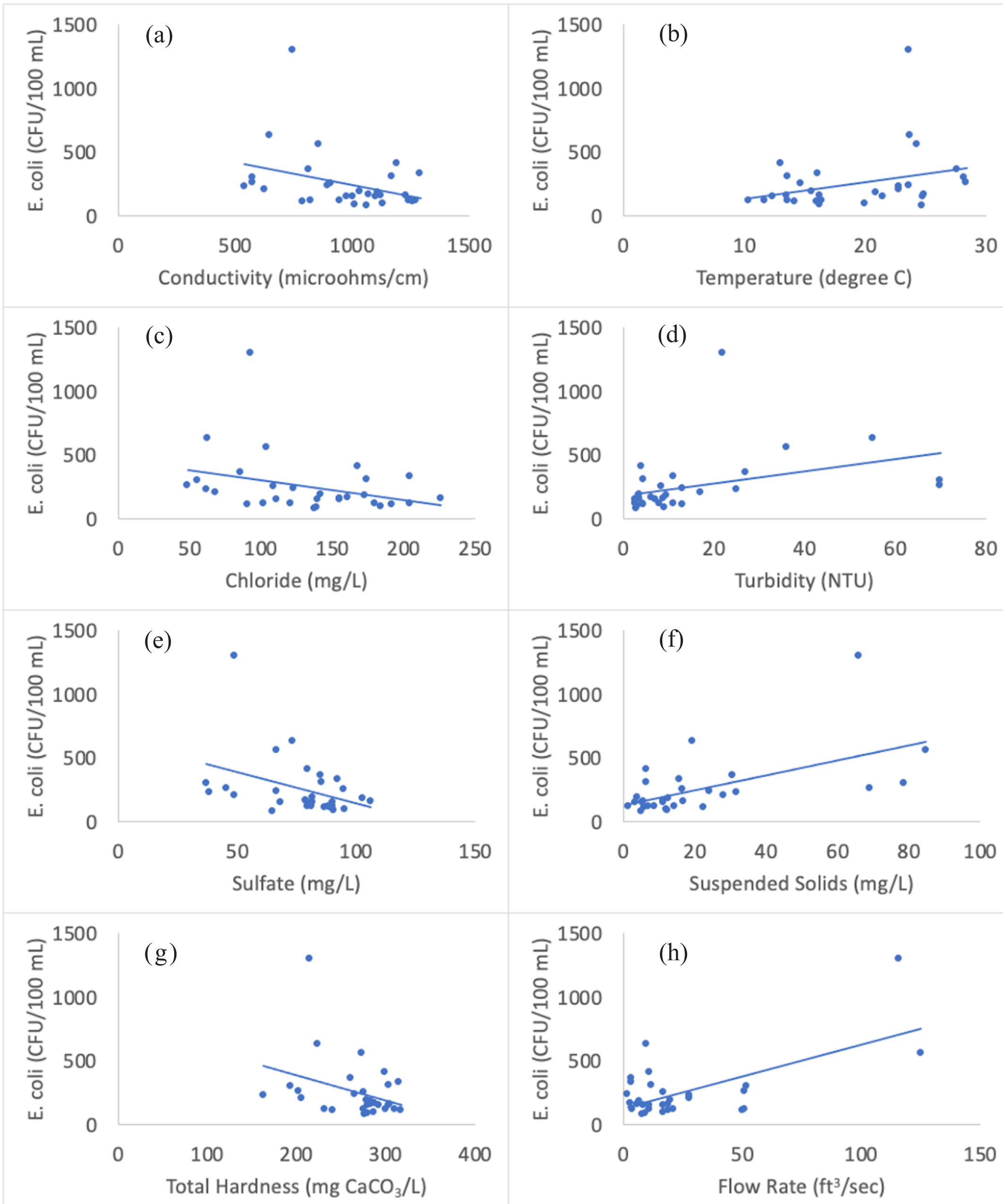

Figure 2. Correlation of the *E. coli* concentration with other water quality parameters. (a) *E. coli* vs. conductivity, (b) *E. coli* vs. temperature, (c) *E. coli* vs. chloride, (d) *E. coli* vs. turbidity, (e) *E. coli* vs. sulfate, (f) *E. coli* vs. suspended solids, (g) *E. coli* vs. total hardness, and (h) *E. coli* vs. flow rate. Presented are those correlations with the Pearson's correlation coefficients (r) being greater than ±0.3.

Acknowledgment

The author is grateful to the Guadalupe-Blanco River Authority for sharing the water quality data which are critical part of this research. The author also appreciates her parents for encouraging her to pursue this journey of scientific investigation. Special thanks extend to research advisor, Christina W. Lopez, for her guidance throughout this endeavor.

References

CDC. (2020). *Waterborne Disease in the United States*. Centers for Disease Control and Prevention. Retrieved May 8, 2022, from https://www.cdc.gov/healthywater/surveillance/burden/findings.html

Deng, X. (2019). Correlations between water quality and the structure and connectivity of the river network in the Southern Jiangsu Plain, Eastern China. *Science of the Total Environment*, 664, 583-594.

Dorner, S.M., et al. (2007). Pathogen and indicator variability in a heavily impacted watershed. *Journal of Water and Health*, 5, 241-257.

Gelting, R.J. et al. (2011). Irrigation water issues potentially related to the 2006 multistate E. coli O157:H7 outbreak associated with spinach. *Agricultural Water Management*, 98, 1395-1402.

HydroFlow. (n.d.). *Texas*. Retrieved March 16, 2022, from https://www.hydroflow-usa.com/texas-water-hardness

Liu, Z. et al. (2018). Association between floods and infectious diarrhea and their effect modifiers in Hunan province, China: A two-stage model. *Science of the Total Environment*, 626, 630-637.

Mohanty, L., Panigrahi, S.K., Patnaik, M. (2022, in press). Assessment of water quality of Mahanadi basin using statistical and wavelet techniques. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2022.03.381

PCWP. (2022a). 2022 Update to the Plum Creek Watershed Protection Plan. Retrieved August 11, 2022, from https://plumcreekwatershed.org/wp-content/uploads/pcwpp-2022.pdf.pdf

PCWP. (2022b). *Monitoring*. Plum Creek Watershed Partnership. Retrieved March 16, 2022, from https://plumcreekwatershed.org/water-quality/monitoring

Rusydi, A.F. (2018). Correlation between conductivity and total dissolved solid in various type of water: A review. IOP Conf. Series: *Earth and Environmental Science*, 118, 012019. https://doi:10.1088/1755-1315/118/1/012019

Schindler, D.W. et al. (2018). Reducing Phosphorus to Curb Lake Eutrophication is a Success. *Environmental Science & Technology*, 17, 8923-8929.

TWRI. (2018). *Texas Bacterial Source Tracking Program Application, Expansion and Marker Evaluation (FY16-FY17), TR-508*. Texas Water Resources Institute.

USEPA. (2022). *Drinking Water Regulations and Contaminants*. Retrieved March 16, 2022, from https://www.epa.gov/sdwa/drinking-water-regulations-and-contaminants

USEPA. (2012). *Recreational Water Quality Criteria*. Office of Water 820-F-12-058, US Environmental Protection Agency, Washington, D.C.

USEPA. (1986). Quality criteria for water. Office of Water 440/5-86-001, US Environmental Protection Agency, Washington, D.C.

Ward, M.H. et al. (2018). Drinking Water Nitrate and Human Health: An Updated Review. *International Journal of Environmental Research and Public Health*, 15(7), 1557

Xu, X. et al. (2022). Using E. coli population to predict foodborne pathogens in pastured poultry farms. *Food Microbiology*, 108, 104092.