

# Geostatistical Exploration of Some Water Quality Clusters in Laguna Lake, Philippines

Pantolla, H. G.,<sup>1,2\*</sup> Malay, C. A.,<sup>3</sup> Leonisa, M. E. M.,<sup>3</sup> Reyes Jr., R. M.<sup>3</sup> and Legaspi, M. J. S.<sup>4,5</sup>

<sup>1</sup>Mathematics and Statistics Unit, Institute of Applied Sciences, Kolehiyo ng Lungsod ng Dasmariñas, City of Dasmariñas, Cavite, Philippines, E-mail: hgpantolla@kld.edu.ph\*

<sup>2</sup>Department of Physical Sciences and Mathematics, University of the Philippines Manila, Metro Manila, Philippines

<sup>3</sup>Graduate Studies of Science and Computing, De La Salle University – Dasmariñas, City of Dasmariñas, Cavite, Philippines

E-mail: camalay@dlsud.edu.ph, memleonisa@dlsud.edu.ph, rmreyes@dlsud.edu.ph

<sup>4</sup>National Service Training Program Unit, Institute of Human Potential and Development, Kolehiyo ng Lungsod ng Dasmariñas, City of Dasmariñas, Cavite, Philippines

E-mail: mslegaspi@kld.edu.ph

<sup>5</sup>Distance, Open, and Transnational University, Central Luzon State University, Science City of Muñoz, Nueva Ecija, Philippines

\*Corresponding Author

DOI: <https://doi.org/10.52939/ijg.v20i4.3143>

## Abstract

*This paper underscores the benefits of employing data visualization and geostatistical analyses in monitoring the water quality parameters of Laguna Lake in the Philippines, particularly emphasizing the influence of quarterly seasonality based on the Philippine climate. Preliminary post-COVID-19 lockdown analyses support and extend the literature in highlighting the persistent risks posed by population growth and industrialization of Metro Manila, also known as the National Capital Region, to the ecosystem and economic benefits of the lake. While findings corroborate existing studies, this paper proposes novel avenues for future research, pinpointing locations with elevated water quality parameters and exploring the impact of different seasonal quarters in the Philippines. The findings also offer potential contributions to Sustainable Development Goal 6 and proactive climate change responses. Targeting stakeholders, economic managers, and health officials overseeing Laguna Lake, the study provides a foundation for informed decision-making and anticipatory measures crucial for addressing potential fish population declines, especially during warmer seasons. Contiguity in Metro Manila raises concerns about water quality, necessitating decisive interventions to mitigate long-term human and fish health risks from contaminant accumulation. Integrating the Geographical Information System enhances interpretation by revealing clusters of high or low water quality parameter values, empowering effective resource management. Further, identifying the hot spots of the water quality parameters by seasonal quarter may equip authorities with essential insights for safeguarding the sustainability, economic vitality, and well-being of the dependent populations of the lake on ways forward after the COVID-19 lockdowns were fully lifted.*

**Keywords:** Geographical Information System, Laguna Lake, Spatial Analysis, Water Quality Parameters

## 1. Introduction

### 1.1 Water Quality and Human Health

Even before the onset of the COVID-19 pandemic, the world was already off track to achieve Sustainable Development Goal 6 (SDG 6), which aims to guarantee universal access to clean water and sanitation by 2030. A significant number of individuals across the globe continue to lack access to safe drinking water and sanitation.

However, both services have been recognized as fundamental human rights for a considerable period. Numerous water sources are experiencing depletion, increased contamination, or a combination of both. Water-intensive sectors such as industry, agriculture, and energy production are experiencing growth to fulfill the demands of a growing population [1].

According to the World Health Organization (WHO), the health consequences of water quality are substantial, regardless of its use for drinking, household purposes, food production, or recreational purposes. Poor water quality can result in the occurrence of disease outbreaks and contribute to the prevalence of diseases occurring at different time intervals. Efforts to ensure water safety enhance public health and promote socioeconomic advancement and overall welfare [2].

Moreover, water quality plays a crucial role in the growth and maturation of fish. Farmers will suffer significant financial losses as a result of the deterioration in water quality [3][4][5] and [6]. Furthermore, pollutants eventually settle in aquatic ecosystems. Urbanization, industrialization, and agricultural practices are all causes of anthropogenic water pollution. In due course, aquatic environments become polluted due to excessive use of pesticides, fertilizers, and waste from residential and industrial areas. Water quality degradation can spread infectious diseases such as dysentery, diarrhea, and jaundice [7].

Aquatic contamination is a prominent type of pollution that leads to significant negative health impacts and loss of life. The inherent capacity of water to counteract contamination can be compromised if contamination is not regulated [7]. Fluctuations in water quality have diverse economic repercussions, affecting human health, ecosystem health, agricultural and fisheries productivity, and recreational and amenity uses. While certain effects can be measured, many others cannot and require non-market valuation techniques to be quantified in monetary terms. Valuation necessitates prior scientific monitoring and comprehension of intricate biophysical relationships. For instance, to distinguish between pollution originating from agricultural activities and pollution from non-agricultural sources or to track the movement of dispersed pollution through intricate hydrological systems. The latter aspect is significant because combining physical distance and lags in cause-and-effect relationships introduces intricacy in assessing and comparing monetary values [8]. Therefore, it is crucial to conduct regular monitoring and control the discharge of pollutants into the nearby aquatic ecosystems [7]. This study would attempt to contribute to monitoring some of the most critical water quality parameters that are being observed are the biochemical oxygen demand (BOD), dissolved oxygen (DO), and fecal coliforms (FC) in the largest lake in the Philippines, specifically in the early periods after the consecutive COVID-19 lockdowns.

### 1.2 Water Quality Parameters

The BOD refers to the quantity of oxygen that organisms utilize to metabolize oxidizable organic substances within a specified timeframe [9] and [10]. BOD serves as a measure of organic pollution in freshwater bodies [10][11] and [12], and it is closely associated with microbiological contamination. Elevated levels of BOD decrease the amount of oxygen present, leading to the deterioration of aquatic habitats and the loss of biodiversity [10] and [13]. Additionally, it hampers the usability of water. The primary contributors to high BOD levels in freshwater systems are human activities, including the discharge of domestic and livestock waste, industrial emissions, and combined sewer overflows. During transportation through the stream network, the levels of BOD are diminished through microbial degradation (river self-purification) and dilution before reaching the seas.

The concentration of dissolved oxygen (DO) is crucial for maintaining water quality in aquaculture [5][14] and [15]. Sufficient oxygen levels are essential for intensive fish farming [15] and [16]. Estimating fish metabolism can be done indirectly by measuring oxygen consumption [15] and [17]. In addition to factors such as feed and temperature, DO plays a crucial role in controlling growth. Maintaining a consistent DO level above a critical threshold is essential for optimal feed consumption, growth, and feed conversion efficiency [15] and [18]. Moreover, changes in hematological, hormonal, biochemical, and osmoregulatory processes in the blood and plasma can be influenced by the availability of DO [19].

Coliforms, including FC, are bacteria that can survive with or without oxygen, have a negative reaction to the Gram stain, ferment lactose, and do not form spores. They are commonly found in the environment and the intestines of humans and other animals in significant quantities [20] and [21]. The coliform group of bacteria, which includes both pathogenic and non-pathogenic forms, is commonly used to indicate microbial contamination [21] and [22]. Total coliforms, including bacteria widely found in soil, plants, and animals, including FC and *Escherichia coli*, also known as *E. coli*, respond to the natural environment and treatment processes like pathogens. Examining the coliform bacteria approximates the quantity or density of harmful bacteria in the sample [21] and [23]. In 2012, the United States Environmental Protection Agency (EPA) advised using thermo-tolerant coliforms, specifically FC, as an indicator of water quality standards due to their association with fecal matter.

Although FC may not directly result in illness, their abundance indicates an increased likelihood of encountering disease-causing bacteria and viruses, such as pathogenic *E. coli* [21]. The EPA advises the assessment of *E. coli* or Enterococci to ascertain potential health hazards further. *E. coli* is a benign bacterium that naturally resides in the intestines of warm-blooded animals, including humans [21].

### 1.3 Objectives of the Study

The main objective of the study is to generate some pieces of Geographic Information System (GIS) that visually displays the potential detection of spatial clusters and statistically significant high values (hot spots) and low values (cold spots) of BOD, DO, and FC in the lockdowns imposed due to the COVID-19 in 2021, and the year after. The results aim to offer insights into the distribution pattern of these common water quality parameters obtained from samples collected at the selected lake as the study setting and its tributaries. Should spatial clusters of hot spots or cold spots exist, the maps will be overlaid onto maps depicting hot spots or cold spots and spatial interpolations of the BOD, DO, and FC. Policymakers and experts may use such maps to provide policies and recommendations, especially in (a) conservation efforts to the lake and (b) valuable advice to the fisherfolk that is economically reliant on the selected study setting. The geostatistical analyses applied herein may also contribute to the scientific literature about monitoring water quality, such as in chemical properties in an aquatic environment and, ultimately, identifying clusters of water quality parameters [24] and [25].

## 2. Methods

### 2.1 Study Setting

The selected study setting for this study is Laguna Lake, also known as *Laguna de Bay* in the Philippines. The lake is located at 14°22'59.99" N and 121°14'60.00" E [26]. As stated in the 2013 Laguna de Bay Ecosystems Health Report Card led by the Laguna Lake Development Authority (LLDA), the lake is the most expansive body of water located within the borders of the Philippines. It ranks as the third largest in the Southeast Asian region [27]. It spans a surface area of 900 square kilometers, with an average depth of 2.5 meters and an elevation approximately 1 meter higher than sea level. The lake is surrounded by the provinces of Laguna to the east, west, and southwest, Rizal to the north to the northeast, and Metro Manila, also known as the National Capital Region (NCR), to the northwest. The lake comprises three separate bays: the West Bay, Central Bay, and East Bay, which meet at the

South Bay. The West Bay watershed is characterized by its high population density and extensive urbanization, primarily due to its inclusion of a portion of NCR. In contrast, the East Bay watershed exhibits the lowest population and development levels. Talim Island is the geographical barrier separating the West and Central Bays, making it the largest and most densely inhabited island in the Lake. The sole discharge point of the Lake is the Napindan Channel, which is linked to Manila Bay through the Pasig River [27].

The lake serves multiple purposes, including flood control, power generation, recreation, industrial cooling, irrigation, and waste disposal, and it is a source of potable water. Its primary use is for fisheries and aquaculture, which form the fishing industry of the Locality. Approximately 13,000 fishermen are purportedly financially dependent on the lake. According to the LLDA, the lake produces an annual yield of approximately 80,000 to 90% metric tons of fish [28].

The western region is most profitable for capturing fisheries and two types of aquacultures. The constant influx of saline water creates a favorable environment for milkfish to thrive in brackish conditions. Furthermore, the lake sustains fisheries. However, pollution contamination presents a significant risk. Moreover, invasive species threaten not only the lake's ecological stability but also the well-being of native species and their economic productivity [28].

### 2.2 Data Used

Overall, 52 monitoring stations are covering the lake. There are 15 Laguna Lake monitoring stations and 37 monitoring stations from its tributaries. However, the analysis dropped Station 22b (Tanay River – Midstream) and Station 30 (Binangonan River) due to multiple missing values within the covered period. The data of BOD in mg/L, DO in mg/L, and FO in MPN/100ml, which are openly available, were collected from the website of the LLDA E-Library [29]. Figure 1 shows the map of the study setting with the locations of the monitoring stations. Panel A shows the location of Laguna Lake in the Philippines, Panel B shows Laguna Lake with its tributaries in blue lines, and Panel C shows the location of each monitoring station where the data of every water quality parameter was collected. The averaged data of December-January-February (DJF) from December 2021 to February 2022, as well as the quarters of March-April-May (MAM), June-July-August (JJA), and September-October-November (SON) in 2022, will be utilized.

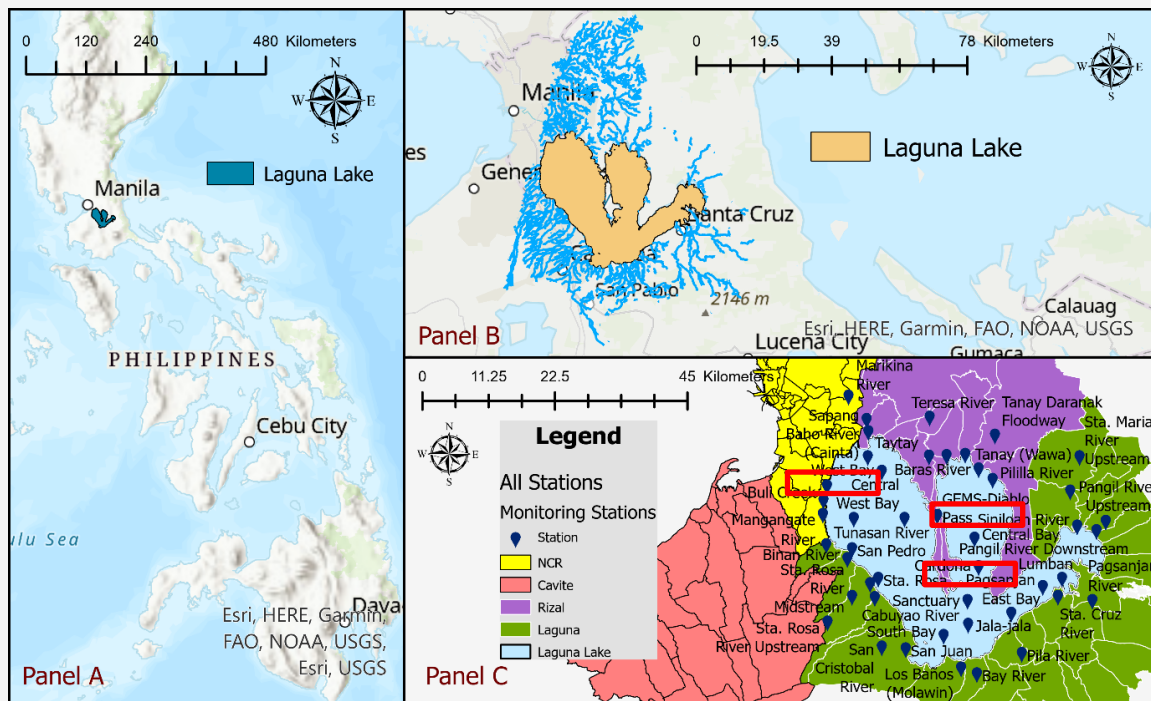


Figure 1: Laguna Lake with monitoring stations

This was done to potentially accommodate variations in the environment caused by seasonal changes relative to the climatic system of the Philippines. Necessary imputations will be applied. The *mean imputation* of the two months covering missing data, the *last observation carried forward*, or the *next observation carried backward* imputations were applied necessarily.

### 2.3 Spatial Autocorrelation

The application of statistical techniques in a spatial context presents numerous obstacles. Waldo Tobler, a geographer and statistician, succinctly expressed a fundamental factor that influences the analysis of spatially referenced data in his well-known and often paraphrased first law of geography, stating that *everything is related to everything else, but near things are more related than far things* [30]. This law defines the statistical concept of positive spatial autocorrelation, wherein observations taken nearby exhibit greater similarity than those taken at greater distances [31]. For this paper, the researchers want to determine if the water quality parameters per seasonal quarter in the monitoring stations near each other are spatially autocorrelated.

The Global Indicator of Spatial Autocorrelation attempts to ascertain the presence of clustering of outliers in the lake. The widely utilized Global Moran's  $I$  statistic will be employed to determine if

the clustering of the monitoring stations in terms of specific water quality parameters by quarter are random events or not [32], i.e., statistically significant. The null hypothesis for this statistic posits that the observed clustering results from random chance. Conversely, the alternative hypothesis posits that the clustering that may have taken place is not random. The Global Moran's  $I$  exhibit a distinct attribute that deviates from the conventional correlation coefficient. It quantifies the resemblance between adjacent regions. The formula for the Global Moran's  $I$  [31] and [33], plugging in the concepts, is given in Equation 1.

$$I = \frac{\sum_{u=1}^A \sum_{v=1}^A c_{u,v} (G_u - \bar{G})(G_v - \bar{G})}{S^2 \sum_{u=1}^A \sum_{v=1}^A c_{u,v}} \quad \text{Equation 1}$$

where the distance  $c_{u,v}$  between the centroids of monitoring stations  $u$  and  $v$ , and the minimum distance required for each monitoring station to have at least one neighbor.  $\bar{G}$  and  $S^2$  can be determined from Equation 2 and 3, respectively

$$\bar{G} = \frac{1}{A} \sum_{u=1}^A G_u \quad \text{Equation 2}$$

$$S^2 = \frac{1}{A} \sum_{u=1}^A (G_u - \bar{G})^2 \quad \text{Equation 3}$$

The total number of monitoring stations is denoted by  $A$  in Equation 1 and thus,  $A=50$ . An explicitly defined water quality parameter, given a seasonal quarter, on the other hand, is represented by the  $G$  statistic for every monitoring station  $u$  and  $v$ , whereas  $\bar{G}$  and  $S^2$ , in succession, are the mean and variance of each water quality parameter per quarter. The Local Indicator of Spatial Autocorrelation (LISA) was employed to identify potential clusters. The statistic specifically used was the standardized Local Moran's  $I$  [31] and [33]. The indicator represented as  $I_u^z$  with  $s$  as the standard deviation of poverty incidence across the region, is calculated and is given in Equation 4 as follows:

$$I_u^z = \frac{1}{s} (G_u - \bar{G}) \left[ \sum_v \frac{c_{u,v}}{s} (G_v - \bar{G}) \right] \quad \text{Equation 4}$$

#### 2.4 Hot Spots and Cold Spots

The High/Low Clustering (Getis-Ord General  $G$ ) tool measures the concentration of high or low values for a given study area. It is given by Equation 5:

$$G = \frac{\sum_{u=1}^A \sum_{v=1}^A \omega_{u,v} G_u G_v}{\sum_{u=1}^A \sum_{v=1}^A G_u G_v}, \forall u \neq v. \quad \text{Equation 5}$$

In equation 5,  $G_u$  and  $G_v$  are attribute values for features  $u$  and  $v$ , and  $\omega_{u,v}$  is the spatial weight between  $u$  and  $v$ . The number of monitoring stations is denoted by  $A$  in the dataset, and  $\forall u \neq v$  guarantees that  $u$  and  $v$  are pairwise distinct [34]. Furthermore, the score statistic  $z_G$  is calculated in Equation 6 through:

$$z_G = \frac{G - \text{Exp}[G]}{\sqrt{\text{Var}[G]}} \quad \text{Equation 6}$$

$$\text{Where: } \text{Exp}[G] = \frac{\sum_{u=1}^A \sum_{v=1}^A \omega_{u,v}}{A(A-1)}, \forall u \neq v$$

$$\text{and } \text{Var}[G] = \text{Exp}[G^2] - \text{Exp}[G]^2.$$

The study will also apply Optimized Hot Spot Analysis. The implementation will specifically focus on the Getis-Ord  $G_i^*$  (G-I star) Analysis [35]. The  $G_i^*$  statistic is a means of quantifying the level of clustering of an attribute in different locations [36].

This statistic identifies monitoring stations as either hot spots or cold spots with varying confidence levels. In the context of this study, a *hot spot* of a water quality parameter is a monitoring station with extremely high values given a quarter that is clustered. In contrast, a *cold spot* of a water quality parameter is defined as a monitoring station with extremely low values. Furthermore, a hot spot is empirically identified when the observed values surpass the expected values. In contrast, a cold spot is statistically detected when the observed values fall below expected [36]. A statistically significant hot spot or cold spot is dependent and relative to the values of the neighboring monitoring stations. Equation 5 provides a standardized value for the statistic [33] and [35].  $G_i^*$  is defined in Equation 7.

$$G_i^* = \frac{\sum_{v=1}^A \omega_{u,v} G_v - \bar{G} \sum_{v=1}^A \omega_{u,v}}{\sigma_G \sqrt{\frac{1}{A-1} \left[ \left( \sum_{v=1}^A \omega_{u,v}^2 \right) - \left( \sum_{v=1}^A \omega_{u,v} \right)^2 \right]}}. \quad \text{Equation 7}$$

In this paper,  $G_v$  is the value of the water quality parameter of interest in a monitoring station  $v$  with its corresponding spatial weight given by  $\omega_{u,v}$  between monitoring stations  $u$  and  $v$ .  $A$  denotes the total number of monitoring stations. The mean value of a water quality ( $\bar{G}$ ) is defined in Equation 8.

$$\bar{G} = \frac{1}{A} \sum_{u=1}^A G_u. \quad \text{Equation 8}$$

On the other hand,  $\sigma_G$  denotes the standard deviation, which is given by Equation 9.

$$\sigma_G = \sqrt{\left( \frac{1}{A} \sum_{v=1}^A G_v^2 \right) - \bar{G}^2} \quad \text{Equation 9}$$

#### 2.4 Kriging Method to be Applied

To estimate the spatial distribution of BOD, DO, and FO per quarter, this study will specifically use Empirical Bayesian Kriging (EBK). EBK is a spatial interpolation technique that necessitates minimal interactive modeling and is distinguished by its robustness and simplicity. This geostatistical interpolation method allows for automated kriging processing. The EBK automatically calculates the parameters using simulation and sub-setting instead of manually adjusting the settings to produce *accurate* results. EBK has a distinct advantage over other kriging techniques in its capacity to calibrate errors resulting from the estimation of the underlying semi-variogram [38].

Furthermore, EBK can predict the errors linked to the generated predicted values and the values of regions that have not been sampled. The variogram parameter is simulated numerous times.

Subsequently, the variogram models are computed based on the simulated data. This procedure enhances the accuracy of EBK prediction compared to traditional kriging methods [39] and [40]. Furthermore, empirical evidence has shown that EBK can produce precise forecasts for nonstationary and non-Gaussian data. This includes scenarios where data variations across different locations are not uniform, establishing EBK as a dependable automated interpolation technique [40] and [41]. Other kriging methods may struggle to provide accurate forecasts due to the limited number of weather stations and the vast geographical area. Therefore, the EBK will be implemented accordingly. The EBK framework [39] is presented in Equation 10.

$$\sum_{a=1}^A \pi_a^k \Theta(\lambda_a) + \sum_{a=1}^A \omega_a I_m(\lambda_a) = P_m(\lambda_0) < M_m$$

Equation 10

In equation 10,  $M$  is the interpolated value,  $m$  is the parameter, and  $\lambda_0$  is the critical value of the specified water quality parameter given a target quarter. A monitoring station in this study is represented by  $\lambda_a$  and the total number of locations is denoted by  $A$ . The weight provided to a known monitoring station is  $\omega_a$ . On the other hand, the  $I_m$  is a multiplier could be 0 if it is higher than, or 1 if lower than  $P_m$ . The  $\pi_a^k$  component is the kriging weight represented which is estimated based on the cross-variogram between components in Equation 11 and Equation 12 given below:

$$\Theta(\lambda) = \phi/a$$

Equation 11

$$I_m(\lambda, m) = \begin{cases} 0 & \text{if } \lambda(\lambda) \geq P_m \\ 1 & \text{otherwise} \end{cases}$$

Equation 12

In Equation 5,  $\phi$  denotes the  $\phi^h$  order statistics for a water quality parameter observed at the location  $\lambda$  [39] and [40]. This makes the estimation process for the regions without values simple [39] and [42]. The process begins with an initial estimation of a semi-variogram using existing data. This is followed by simulating a new value based on the calculated variogram. Subsequently, a new semi-variogram is recalculated using the recently generated data, and

the updated weight of the variogram is determined based on Bayes' Rule [39] and [42].

### 2.5 Procedures

All maps and analyses shall be generated in ArcGIS Pro 3.1.2. The (a) Global Moran's  $I$ , (b) Getis-Ord  $G_i^*$ , and EBK shall be generated separately for each quality parameter by quarter. The p-value of a resulting Global Moran's  $I$  that is less than the default 0.05 significance level indicates sufficient proof of spatial autocorrelation. Obtaining an empirically proven global index of spatial autocorrelation summarizes an evidence-based interpretation that the entirety of the study area has a spatial similarity between neighboring observations [31]. This shall entail running Local Moran's  $I$  analyses with maps that will be overlaid with the corresponding EBK maps. The outputs of the Local Moran's  $I$ , which water monitoring stations are spatially autocorrelated, may show how clustered the monitoring stations are and how each monitoring station responds to the fluctuations in the water quality parameters of its nearest neighboring monitoring stations. On the other hand, significant Getis-Ord  $G_i^*$  results using the Optimized Hot Spot Analysis at the 0.01, 0.05, and 0.10 levels of significance shall also be compared to the next quarter. i.e., comparisons of the DJF results with the MAM results, the MAM results with the JJA results, and the JJA results with the SON results shall be presented visually. The significant Getis-Ord  $G_i^*$  results would reveal how if there are high, low, or random clustering. The comparisons shall include confidence levels of 90%, 95%, and 99%. In summary, GIS will be provided to visualize (a) the analysis of spatial autocorrelation, which may show the significance of clustering per water quality parameter, and (b) the analysis of hot spots (or cold spots) of water quality parameters.

The K-Nearest Neighbors (KNN) algorithm will be used to train the data from the monitoring stations. This was applied since the water monitoring stations are point features. In the context of spatial data analysis, the KNN is a nonparametric supervised learning classifier that accounts for the Euclidean proximity of the observations in generating predictions or classifications about the grouping of each observation [43]. Necessarily, the automatic weighting generated from the KNN algorithm was applied to all the geostatistical analyses and the EBK. All applicable analyses shall be replicated by 9,999 to achieve robustness of estimates.

### 3. Results and Discussion














#### 3.1 Exploratory Data Analysis

Figure 2, Figure 3, and Figure 4 shows the GIS of the classified levels of water quality parameters. The BOD was leveled into 5, the DO was classified into 3, and the FC was leveled into 5. Logarithmic transformation was done to the FC data to minimize the variance and make visualization easier. Irrespective of the monitoring station class, the BOD and DO levels were based on the Department of Environment and Natural Resources Administrative Order (DENR) Administrative Order Number 2016 – 08. On the other hand, the non-transformed levels of FC were based on DENR Water Quality Guidelines Number 2021 – 19. Table 1 shows the ranges of the data.

#### 3.2. Spatial Autocorrelation, Hot Spots, and Spatial Interpolation

The Global Moran's  $I$  and Getis-Ord General  $G$  results are shown in Table 2 and Table 3, respectively. In both analyses, all p-values are less than 0.05 except for the FC of SON. Table 2 provides statistical proof that for all four seasonal quarters, the water quality parameters are spatially autocorrelated aside from the FC of SON. This indicates that the water quality parameters for all seasonal quarters, BOD, DO, and FC sans SON, are spatially autocorrelated statistically. This implies that some monitoring stations are significantly affected by fluctuations in the levels of water quality parameters in some neighboring monitoring stations.

**Table 1:** Colors used for the symbology of ranges for each water quality parameter

BOD		DO		FC (log-transformed)	
Color	Range	Color	Range	Color	Range
	At most 3		At most 2		2.92 to at most 3.91
	More than 3 to less than 5		More than 2 to less than 5		More than 3.91 to less than 4.61
	More than 5 to less than 7		More than 5		More than 4.61 to less than 5.30
	More than 7 to less than 15	-	-		More than 5.30 to less than 5.99
	More than 15	-	-		More than 5.99 to less than 20

**Table 2:** Global Moran's  $I$  analyses for global spatial autocorrelation of the water quality parameters by quarter

Water Quality Parameter	Quarter	p-value	Spatial Autocorrelation
BOD	DJF	0.0032*	Clustered
	MAM	<0.0001*	Clustered
	JJA	0.0038*	Clustered
	SON	0.0001*	Clustered
DO	DJF	0.0001*	Clustered
	MAM	<0.0001*	Clustered
	JJA	<0.0001*	Clustered
	SON	<0.0001*	Clustered
FC	DJF	<0.0001*	Clustered
	MAM	0.0006*	Clustered
	JJA	0.0034*	Clustered
	SON	0.2178	Random














\*Significant at the 0.05 level

**Table 3:** Getis-Ord General  $G$  analyses for clustering of the water quality parameters by quarter

Water Quality Parameter	Quarter	p-value	Spatial Autocorrelation
BOD	DJF	0.0037*	High
	MAM	<0.0001*	High
	JJA	0.0052*	High
	SON	<0.0001*	High
DO	DJF	0.0011*	High
	MAM	0.0025*	High
	JJA	<0.0001*	High
	SON	0.0003*	High
FC	DJF	<0.0001*	High
	MAM	0.0008*	High
	JJA	0.0045*	High
	SON	0.2114	Random

\* Significant at the 0.05 level

**Table 4:** Colors used for the symbology of ranges for each water quality parameter

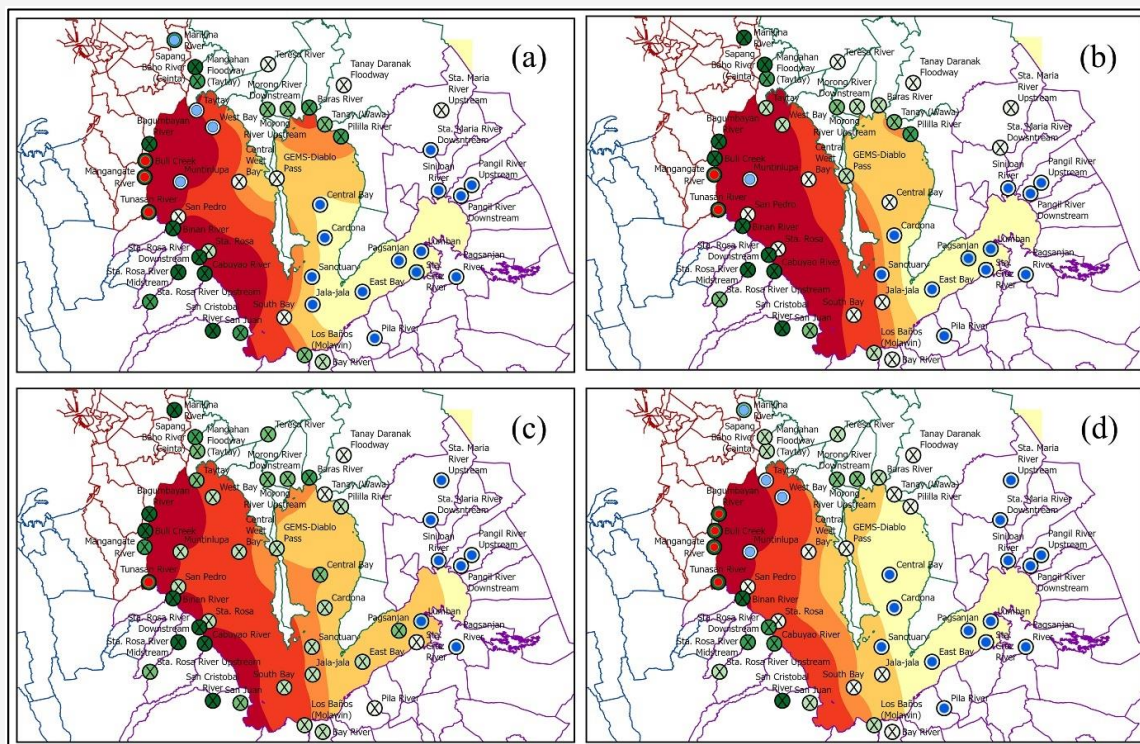
BOD		DO		FC (log-transformed)	
Color	Range (mg/L)	Color	Range (mg/L)	Color	Range (MPN/100ML)
	At most 3		At most 2		2.92 to at most 3.91
	More than 3 to less than 5		More than 2 to less than 5		More than 3.91 to less than 4.61
	More than 5 to less than 7		More than 5		More than 4.61 to less than 5.30
	More than 7 to less than 15				More than 5.30 to less than 5.99
	More than 15				More than 5.99 to less than 20

**Table 5:** Colors used for the symbology of Local Moran's  $I$ 

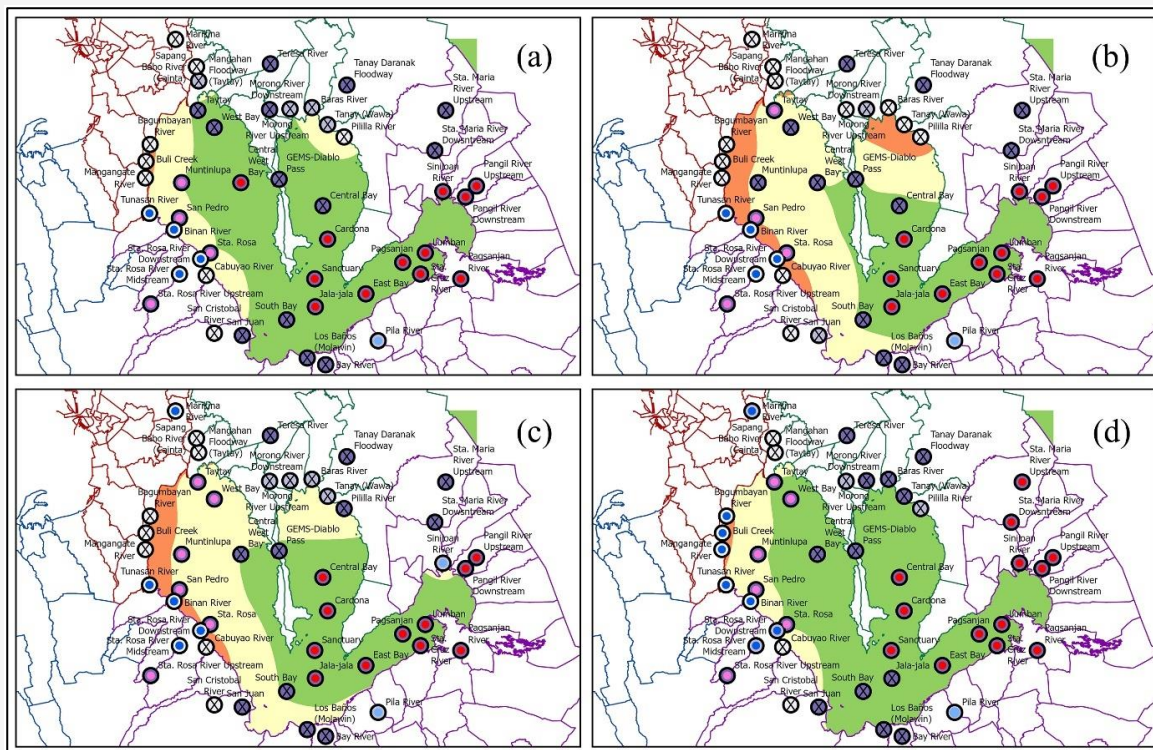
Color	Interpretation
	High-High Clustering
	High-Low Clustering
	Low-High Clustering
	Low-Low Clustering
X	Not Significant

This entails using a LISA, which is the Local Moran's  $I$ . For comparative purposes, the Local Moran's  $I$  was applied to generate maps for all water quality parameters in each seasonal quarter. Table 2 shows the results of the Global Moran's  $I$  analyses. Similarly, Table 3 shows that every water quality parameter per seasonal quarter has significantly high clustering except for FC of SON. There is sufficient evidence that supports the statistically high clustering of the water quality parameters in all the quarters,

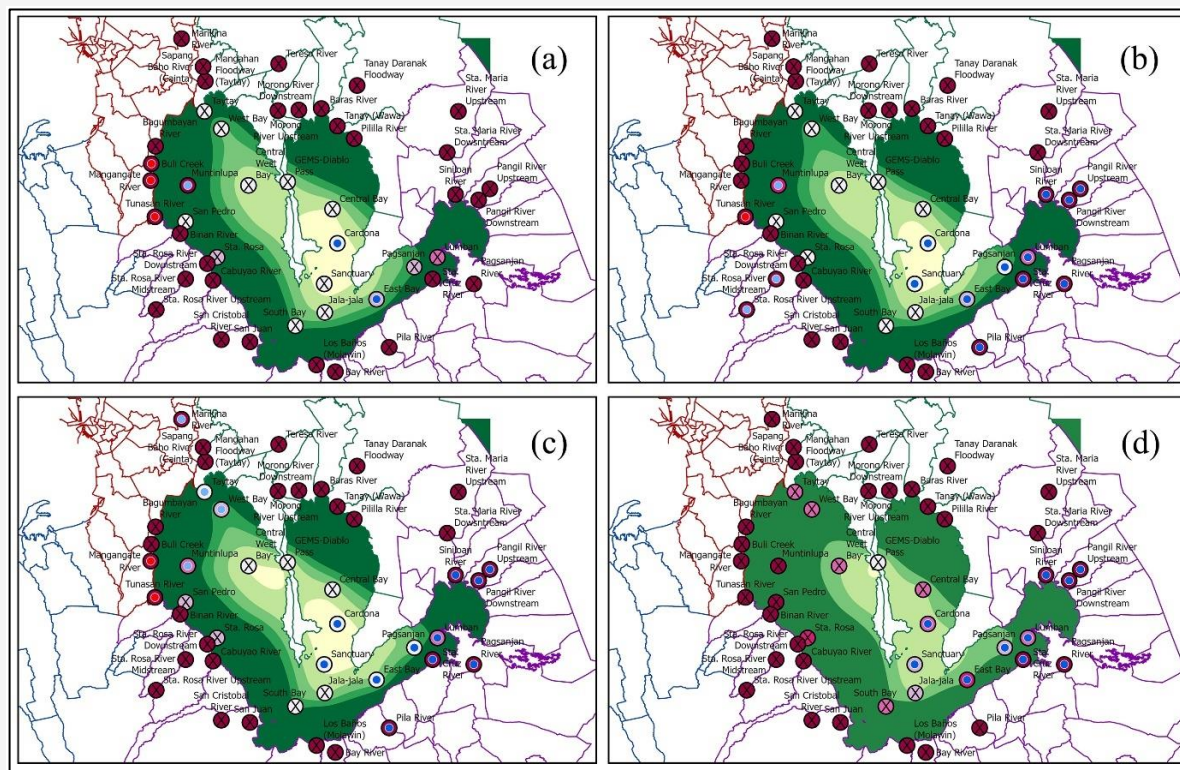
barring FC of SON. Hence, a Getis-Ord  $G_i^*$  will be shown via individual pieces of GIS. Figure 2, Figure 3, and Figure 4 show the overlaid maps of the Local Moran's  $I$  and Getis-Ord  $G_i^*$  with the ranges of water quality parameters given in Table 1. Additionally, these three figures were superimposed on the generated results of the EBK. For easier understanding, the contours of the EBK per water quality parameter are shown in Table 4 and Table 5.



**Figure 2:** BOD clusters and their spatial distribution using EBK for (a) DJF, (b) MAM, (c) JJA, and (d) SON



**Figure 3:** DO clusters and their spatial distribution using EBK for (a) DJF, (b) MAM, (c) JJA, and (d) SON



**Figure 4:** FC clusters and their spatial distribution using EBK for (a) DJF, (b) MAM, (c) JJA, and (d) SON

Figure 2, Figure 3, and Figure 4 show the clustering of the values of the monitoring stations as well. A monitoring station that exhibits a high value of a specific water quality parameter that is neighboring to at least one other monitoring station with a high value of water quality parameter for such quarter, thus giving rise to the term High-High cluster. The same context applies to High-Low, Low-High, and Low-Low clustering. Given that one of the objectives of this study is to identify clusters of monitoring stations with significant clustering of water quality parameters, monitoring stations categorized as High-High and Low-High were prioritized.

Figure 2 illustrates the presence of clustered groups of high values of BOD in the West Bay, which is adjacent to the NCR, in DJF and JJA. Furthermore, the interpolated estimated values through the EBK are consistently higher at the West Bay and South Bay. Higher BOD levels are linked to organic pollution, specifically sewage [27]. The West Bay watershed is characterized by its high population density and extensive urbanization, primarily due to its inclusion of a portion of the NCR [27]. In contrast, the East Bay watershed has the lowest population and development levels [27]. The high BOD in the monitoring stations near the NCR is a glaring

problem. Figure 2 also shows that except for the MAM quarter, the EBK in the maps shows the concentration of high BOD in the West Bay except for the warmer months. These consistently elevated levels of BOD may further deteriorate since all sewage effluent from the city's waterways eventually flows into Laguna Lake or Manila Bay [44].

Figure 3 shows an interesting pattern for the DO, at least in the duration of the time frame of the used data. In the colder seasons, the DJF and SON primarily have clusters of monitoring stations with high dissolved oxygen levels in the Central and East Bay. This aligns with the assertion that warm water has a lower capacity to retain DO than cold water [45] relative to the climatic seasonality of the NCR. The NCR is classified under the Köppen climate classification as having a tropical wet and dry climate and tropical monsoon climate. The dry season in Metro Manila spans from January to May, while the wet season extends from June to December [46]. Conversely, the monitoring stations in the West Bay near the NCR have high DO levels in the MAM, generally in the warmer months, and JJA, where seasonal rains are prevalent. The EBK also shows that the DO levels exceed 5 mg/l in almost Laguna Lake on DJF and SON.

The East Bay region has a greater density of fishermen operating within a relatively small fishing area, with a fishing ground allocation of only 1 fisherman per 28 hectares. Additionally, it is characterized by the highest concentration of the invasive clown knife fish. This species was introduced into the Lake via the East Bay and likely proliferated more rapidly due to the superior water quality of the East Bay [27]. On the other hand, Central Bay exhibits the highest proportion of indigenous fish in terms of catch composition and zooplankton ratio. The fishing ground allocation is approximately 1 fisher per 110 hectares. [27].

Previous studies have consistently reported a correlation between FC concentrations and levels of phosphorus and nitrogen [47][45][46] and [50]. A study also revealed the correlation between elevated phosphate levels and reduced FC concentrations. This implies that FC bacteria may have a greater affinity for feeding on phosphates than nitrogen. Consequently, high levels of phosphates could indicate the absence or low concentration of fecal coliform bacteria [50]. The same characteristics apply to the fecal coliform concentrations in Figure 4. The province of Laguna is a top rice-producing province and thus requires a huge volume of fertilizers. Such an economic feature is observable in JJA. This may have been apparent in the generated results. On the other hand, the glaring results of significantly high concentrations of FC in the monitoring stations near the NCR are also evident in all the seasonal quarters except for SON, where the clustering was random.

Fecal coliform contamination may originate from sewage treatment facilities or certain industrial sources. Frequently, it originates from numerous minor sources, each making a small contribution to the overall problem. Every individual residing within a watershed plays a role in creating these minor origins by managing on-site septic systems, livestock









waste, pet waste, and sanitary practices [51]. Unsurprisingly, the highly populated and industrialized West Bay has at least one monitoring station with High-High FC concentration. Regardless of the origin, an increase in the concentration of fecal coliform bacteria in water directly correlates with an elevated risk to public health [51].

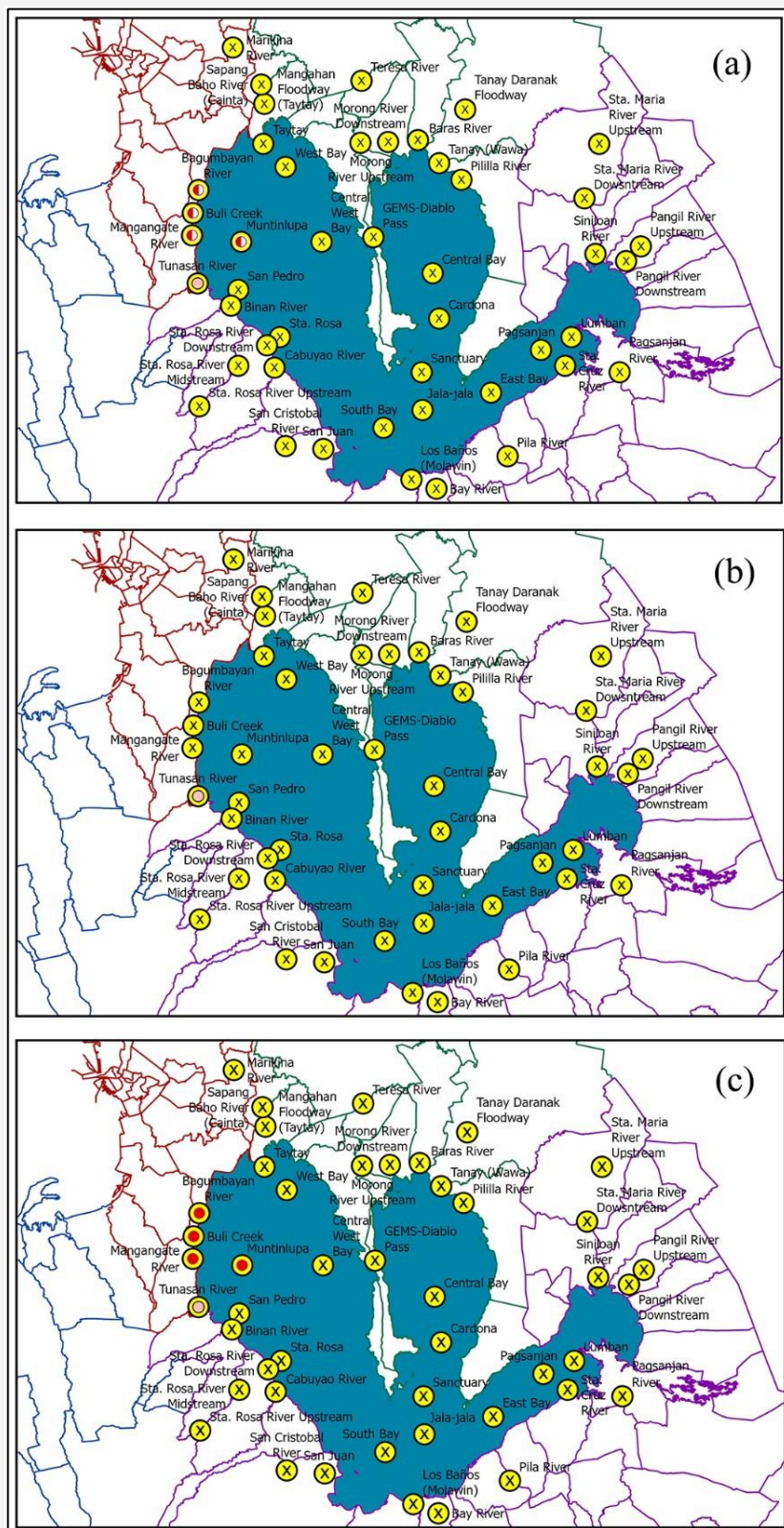
### 3.3 Comparison of Spots

With different levels of confidence at 90%, 95%, and 99%, the results of the Getis-Ord  $G_i^*$ , which were generated into individual GIS by seasonal quarter, were all significant analyses for all water quality parameters. Hence, the identified monitoring stations determined to be hot spots or cold spots were statistically supported. Figure 5, Figure 6, and Figure 7 show the changes in seasonal quarters of each spot by water quality parameter. The West Bay, particularly the monitoring stations within and near the NCR, concerning Figure 1 and depicted in red lines, are emphasized more. The consistency of EBK, Moran's  $I$ , and Getis-Ord  $G_i^*$  entails immediate actions to conserve Laguna Lake. Table 6 provides a guide on the colors used for the changes, as seen in Figure 5, Figure 6, and Figure 7. Figure 5 shows that by seasonal quarters, some monitoring stations vary from hot spots in DJF to not significant in MAM. Interestingly, some monitoring stations that are not significant in JJA became hot spots in SON. Consistent with the known inverse relationship between BOD and DO and the known effects of temperature on DO, the monitoring stations from the NCR changed from not significant in JJA to cold spots in SON.

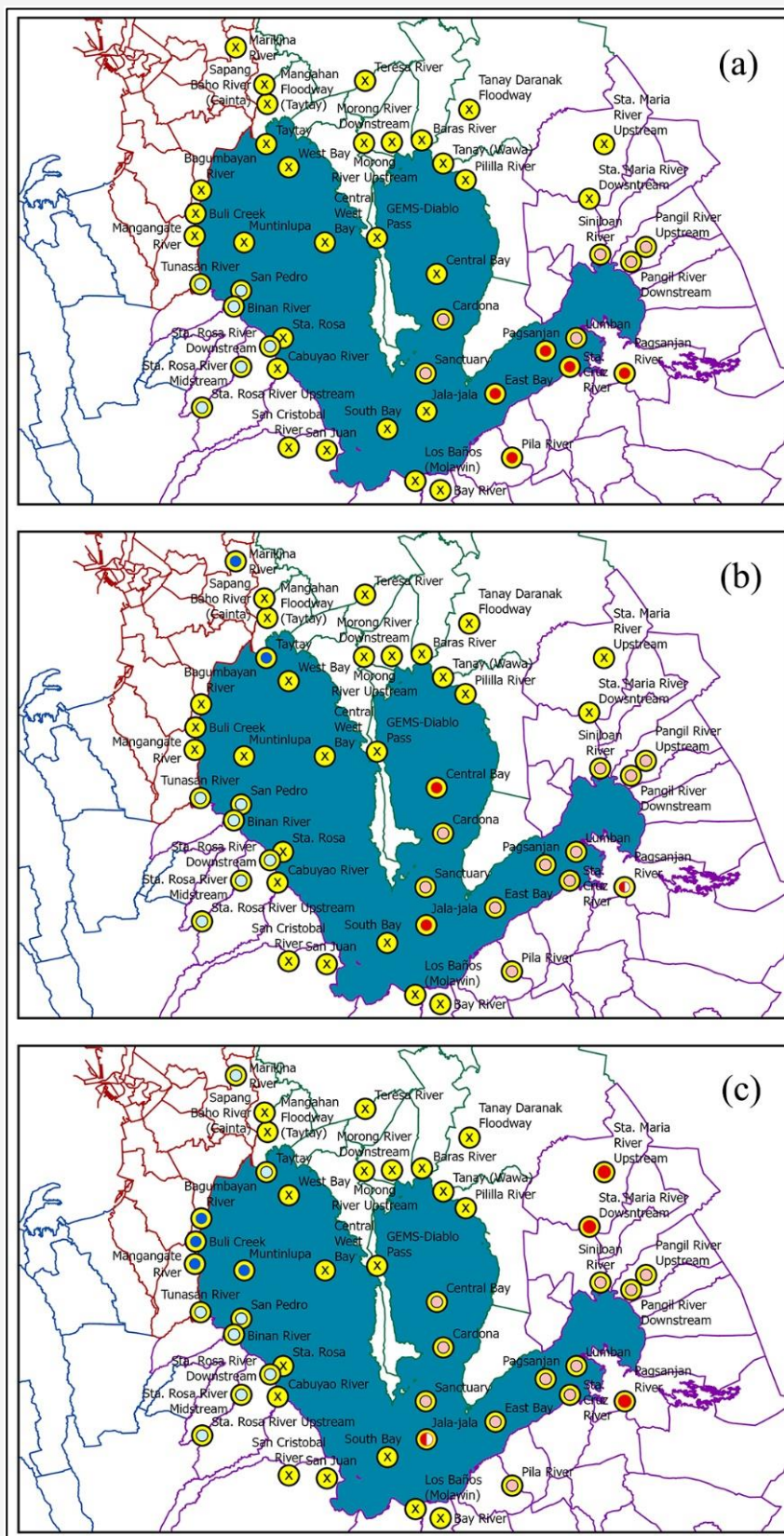
It appears that from MAM to JJA, the monitoring stations in the NCR were consistently identified as hot spots of FC. Note that the spots in JJA to SON occurred randomly.

**Table 6:** Colors used for the symbology of Hot Spots and Cold Spots comparison by seasonal quarter

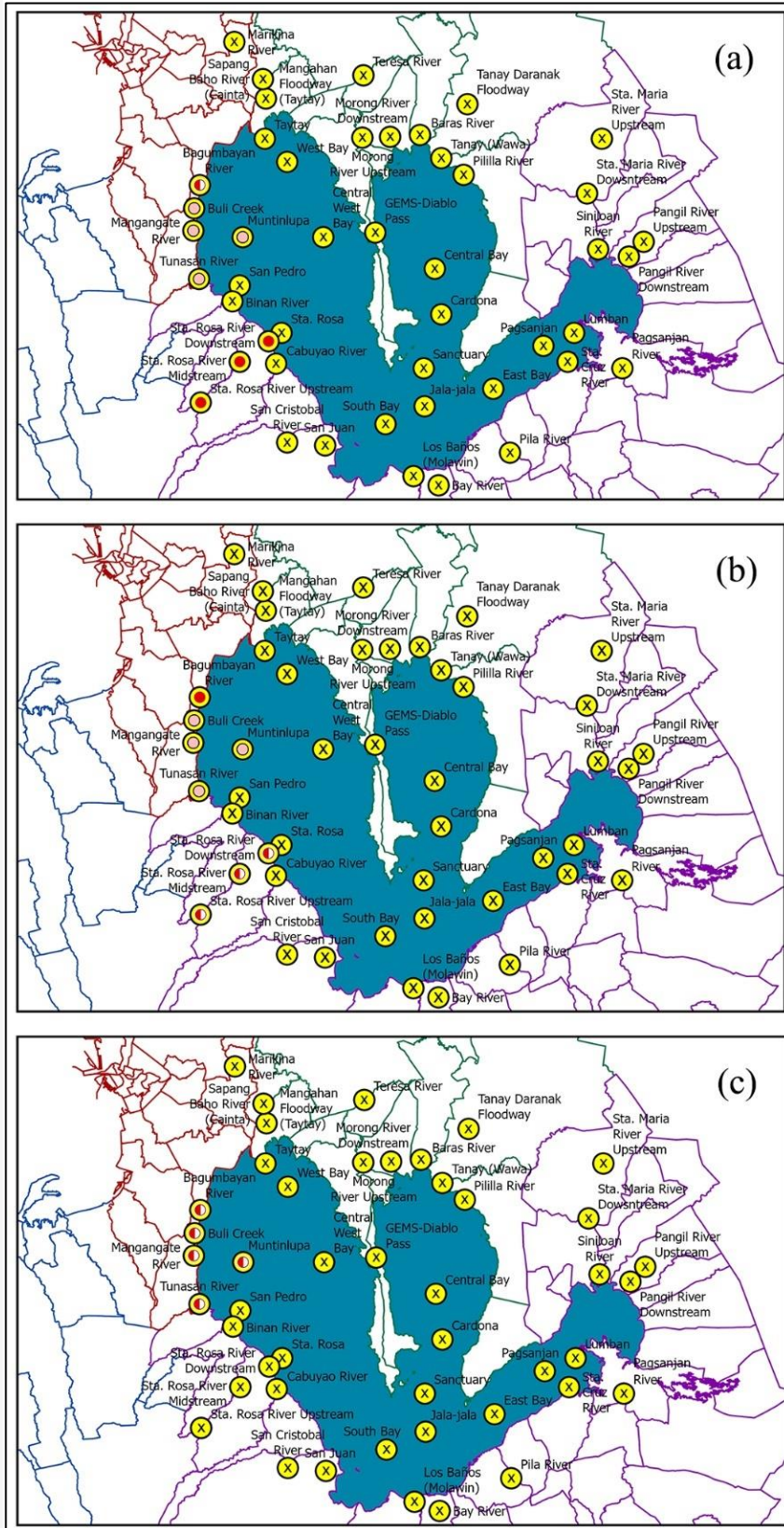
Color	Interpretation	Color	Interpretation	Color	Interpretation
	Cold to Hot		Hot to NS		Cold to Cold
	NS to Hot	X	NS to NS		NS to Cold
	Hot to Hot		Cold to NS		Hot to Cold



**Figure 5:** Changes in spots for BOD of (a) DJF to MAM, (b) MAM to JJA, and (c) JJA to SON



**Figure 6:** Changes in spots for DO of (a) DJF to MAM, (b) MAM to JJA, and (c) JJA to SON



**Figure 7:** Changes in spots for FC of (a) DJF to MAM, (b) MAM to JJA, and (c) JJA to SON

#### 4. Conclusion and Recommendations

This paper demonstrates the advantages of utilizing data visualization and geostatistical analyses to monitor the water quality parameters of Laguna Lake. The study revealed that employing quarterly seasonality based on the climate in the Philippines could yield more comprehension of the association between environmental changes and the fluctuations in the common water quality parameters obtained from the lake. Preliminary findings following the lifting of COVID-19 lockdowns show more details about potential sources of poor water quality parameters and the seasonal effects of climate change per quarter. The impacts of the increasing population and industrialization of the NCR present significant risks to the ecosystem, and the lake's economic benefits were also emphasized. More studies can be conducted to establish more conclusive scientific generalizations, especially including socioeconomic, climatic, chemical, and physical factors, when applicable, and the effects of lags on the present levels of water quality parameters. Models or algorithms may also be applied to predict the levels of water quality parameters, specifically after providing remedies for the missing data caused by the COVID-19 lockdowns.

This paper introduces two new areas for future research: (a) identifying the specific locations with high concentrations of water quality parameters and (b) examining the impact of different seasonal quarters in the Philippine climate. This could potentially contribute to achieving the targets outlined in SDG 6, as well as anticipatory actions in response to the impacts of climate change.

The stakeholders, economic managers, and health officials responsible for maintaining or regulating Laguna Lake can utilize this study. With additional conclusive evidence, proactive measures can be taken to prepare for the potential decline in the fish population during warmer seasons. This significantly impacts the economic and food security of the population that relies on the lake. Areas near the NCR frequently exhibit poor water quality parameters in general. The rapid and ongoing urbanization in the region may ultimately pose more serious risks to human health if no decisive intervention is undertaken, particularly due to the potential long-term accumulation of contaminants in the lake. The GIS also enhances the ability to interpret the clustering of high or low values of the water quality parameters.

#### References

- [1] United Nations. (2021). *Summary Progress Update 2021: SDG 6 - Water and Sanitation for All*. UN. <https://www.unwater.org/publications/summary-progress-update-2021-sdg-6-water-and-sanitation-all>.
- [2] WHO. (2016). *Water Quality and Health Strategy 2013-2020*. World Health Organization. <https://www.who.int/publications/m/item/water-quality-and-health-strategy-2013-2020>.
- [3] Oppedal, F., Dempster, T. and Stien, L. H. (2011). Environmental Drivers of Atlantic Salmon Behaviour in Sea-Cages: A Review. *Aquaculture*, Vol. 311(1–4), 1–18. <https://doi.org/10.1016/j.aquaculture.2010.11.020>.
- [4] Seager, J., Milne, I., Mallett, M. and Sims, I., (2000). Effects of Short-Term Oxygen Depletion on Fish. *Environmental Toxicology and Chemistry*, Vol. 19(12), 2937–2942. <https://doi.org/10.1002/etc.5620191214>.
- [5] Cerezo Valverde, J., Martínez López, F. J. and García García, B., (2006). Oxygen Consumption and Ventilatory Frequency Responses to Gradual Hypoxia in Common Dentex (Dentex Dentex): Basis for Suitable Oxygen Level Estimations. *Aquaculture*, Vol. 256(1–4), 542–551. <https://doi.org/10.1016/j.aquaculture.2006.02.030>.
- [6] Sun, L., Wang, B., Yang, P., Wang, X., Li, D. and Wang, J., (2022b). Water Quality Parameter Analysis Model Based on Fish Behavior. *Computers and Electronics in Agriculture*, Vol. 203, <https://doi.org/10.1016/j.compag.2022.107500>.
- [7] Bashir, I., Lone, F. A., Bhat, R. A., Mir, S. A., Dar, Z. A. and Dar, S. A., (1970). *Concerns and Threats of Contamination on Aquatic Ecosystems*. SpringerLink. [https://link.springer.com/chapter/10.1007/978-3-030-35691-0\\_1](https://link.springer.com/chapter/10.1007/978-3-030-35691-0_1).
- [8] Moxey, A., (2012). *Agriculture and Water Quality: Monetary Costs and Benefits Across OECD Countries*. OECD. <https://www.oecd.org/greengrowth/sustainable-agriculture/49841343.pdf>.
- [9] European Environment Agency. (2018). *The European Pollutant Release and Transfer Register (E-PRTR)*. [https://environment.ec.europa.eu/topics/industrial-emissions-and-safety/european-pollutant-release-and-transfer-register-e-prtr\\_en](https://environment.ec.europa.eu/topics/industrial-emissions-and-safety/european-pollutant-release-and-transfer-register-e-prtr_en).

- [10] Vigiak, O., Grizzetti, B., Udias-Moinelo, A., Zanni, M., Dorati, C., Bouraoui, F. and Pistocchi, A., (2019). Predicting Biochemical Oxygen Demand in European Freshwater Bodies. *Science of the Total Environment*, Vol. 666, 1089–1105. <https://doi.org/10.1016/j.scitotenv.2019.02.252>.
- [11] Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R. and Davies, P. M., (2010). Global Threats to Human Water Security and River Biodiversity. *Nature*, Vol. 467(7315), 555–561. <https://doi.org/10.1038/nature09440>.
- [12] Wen, Y., Schoups, G. and van de Giesen, N., (2018). Global Impacts of the Meat Trade on In-Stream Organic River Pollution: The Importance of Spatially Distributed Hydrological Conditions. *Environmental Research Letters*, Vol. 13(1). <https://doi.org/10.1088/1748-9326/aa94f6>.
- [13] Ferreira, A. R. L., Sanches Fernandes, L. F., Cortes, R. M. V. and Pacheco, F. A. L., (2017). Assessing Anthropogenic Impacts on Riverine Ecosystems Using Nested Partial Least Squares Regression. *Science of the Total Environment*, Vol. 583, 466–477. <https://doi.org/10.1016/j.scitotenv.2017.01.106>.
- [14] Timmons, M. B., (2002). *Recirculating Aquaculture Systems*. Cayuga Aqua Ventures.
- [15] Bagherzadeh Lakani, F., Sattari, M. And Falahatkar, B., (2012). *Effect of Different Oxygen Levels on Growth Performance, Stress Response and Oxygen Consumption in Two Weight Groups of Great Sturgeon Huso Huso*. *Iranian Journal of Fisheries Sciences*, Vol. 12(3), 533-549.0 [https://jifro.areeo.ac.ir/article\\_114297.html](https://jifro.areeo.ac.ir/article_114297.html).
- [16] Ritola, O., Tossavainen, K., Kiuru, T., Lindstrom-Seppa, P. and Molsa, H., (2002). Effects of Continuous and Episodic Hyperoxia on Stress and Hepatic Glutathione Levels in One-Summer-Old Rainbow Trout (*Oncorhynchus mykiss*). *Journal of Applied Ichthyology*, Vol. 18(3), 159–164. <https://doi.org/10.1046/j.1439-0426.2002.00324.x>.
- [17] Pichavant, K., Person-Le-Ruyet, J., Bayon, N. L., Severe, A., Roux, A. L. and Boeuf, G., (2001). Comparative effects of Long-Term Hypoxia on Growth, Feeding and Oxygen Consumption in Juvenile Turbot and European Sea Bass. *Journal of Fish Biology*, Vol. 59(4), 875–883. <https://doi.org/10.1111/j.1095-8649.2001.tb00158.x>.
- [18] Jobling, M., (1994). *Fish Bioenergetics*. SpringerLink. <https://link.springer.com/book/9780412580901>.
- [19] Samaras, A., Tsoukali, P., Katsika, L., Pavlidis, M. and Papadakis, I. E., (2023). Chronic Impact of Exposure to Low Dissolved Oxygen on the Physiology of *Dicentrarchus Labrax* and *Sparus Aurata* and its Effects on the Acute Stress Response. *Aquaculture*, Vol. 562. <https://doi.org/10.1016/j.aquaculture.2022.738830>.
- [20] Madigan, M. T., Martinko, J. M., Stahl, D. A. and Clark, D. P., (2012). *Brock Biology of Microorganisms*. B. Cummings.
- [21] Raña, J., Domingo, J., Opinion, A. G. and Cambia, F., (2017). Contamination of Coliform Bacteria in Water and Fishery Resources in Manila Bay Aquaculture Farms. *The Philippine Journal of Fisheries*, Vol. 24(2), 98–126. <https://doi.org/10.31398/tpjf/24.2.2016a0015>.
- [22] Forsythe, S. J., (2010). *The Microbiology of Safe Food, 2nd Edition*. John Wiley & Sons.
- [23] Henze, M., (2008). *Biological Wastewater Treatment: Principles, modelling and Design*. IWA Pub.
- [24] Yuan, F., Lu, L., Zhang, Y., Wang, S. and Cai, Y. D., (2018). Data Mining of the Cancer-Related Lncnas GO Terms and KEGG Pathways by using MRMR Method. *Mathematical Biosciences*, Vol. 304, 1–8. <https://doi.org/10.1016/j.mbs.2018.08.001>.
- [25] Sun, L., Wang, B., Yang, P., Wang, X., Li, D. and Wang, J., (2022a). Water Quality Parameter Analysis Model Based on Fish Behavior. *Computers and Electronics in Agriculture*, Vol. 203. <https://doi.org/10.1016/j.compag.2022.107500>.
- [26] Latitude.to. (n.d.). *GPS Coordinates of Laguna de Bay, Philippines*. Latitude: 14.3833 Longitude: 121.2500. Latitude.to, Maps, Geolocated Articles, Latitude Longitude Coordinate Conversion. <https://latitude.to/articles-by-country/ph/philippines/14491/laguna-de-bay>.
- [27] LLDA. (2016). *Laguna de Bay 2013 Ecosystem Health Report Card*. PEMSEA. <https://pemsea.org/publications/reports/laguna-de-bay-2013-ecosystem-health-report-card>.
- [28] LLDA. (n.d.-a). *Existing lake uses | LLDA*. Laguna Lake Development Authority. <https://llda.gov.ph/existing-lake-uses/>
- [29] LLDA. (n.d.-b). *Laguna de Bay and its Tributaries | LLDA*. Laguna Lake Development Authority. <https://llda.gov.ph/ldb-and-its-tributaries/>

- [30] Tobler, W. R., (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, Vol. 46, <https://doi.org/10.2307/143141>.
- [31] Waller, L. A. and Gotway, C. A., (2004). *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons.
- [32] Anselin, L., (2020). *Geoda*. Contiguity-Based Spatial Weights. [https://geodacenter.github.io/workbook/4a\\_contig\\_weights/lab4a.html](https://geodacenter.github.io/workbook/4a_contig_weights/lab4a.html).
- [33] Pantolla, H. G. And Nacion, N. A., (2023). Spatial Analysis of Poverty Incidence and Road Networks in Eastern Visayas Region, Philippines. *Philippine Journal of Science*, Vol. 152(4). <https://philjournalsci.dost.gov.ph/publication/regular-issues/past-issues/122-vol-152-no-4-august-2023/1872-spatial-analysis-of-pov-erty-incidence-and-road-networks-in-eastern-visayas-region-philippines>
- [34] ESRI, (2024). *How High/Low Clustering (Getis-Ord General G) Works*. <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-high-low-clustering-getis-ord-general-g-spat.htm>.
- [35] Getis, A. and Ord, J. K., (1995). Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis*, Vol. 27(4), 286–306. <https://doi.org/10.1111/j.1538-4632.1995.tb00912.x>.
- [36] Ghodousi, M., Sadeghi-Niaraki, A., Rabiee, F. and Choi, S. M., (2020). Spatial-temporal Analysis of Point Distribution Pattern of Schools Using Spatial Autocorrelation Indices in Bojnourd City. *Sustainability*, Vol. 12(18), <https://doi.org/10.3390/su12187755>.
- [37] Gorr, W. L., Kurland, K. S. and Dodson, Z. M., (2018). *GIS Tutorial for Crime Analysis*. Esri Press.
- [38] Ali, G., Sajjad, M., Kanwal, S., Xiao, T., Khalid, S., Shoaib, F. and Gul, H. N., (2021). Spatial-temporal Characterization of Rainfall in Pakistan During the Past Half-Century (1961–2020). *Scientific Reports*, Vol. 11(1). <https://doi.org/10.1038/s41598-021-86412-x>.
- [39] Sahu, B., Ghosh, A. K. and Seema. (2021). Deterministic and Geostatistical Models for Predicting Soil Organic Carbon in a 60 ha Farm on Inceptisol in Varanasi, India. *Geoderma Regional*, Vol. 26. <https://doi.org/10.1016/j.geodrs.2021.e00413>.
- [40] Yang, R. and Xing, B., (2021). A Comparison of the Performance of Different Interpolation Methods in Replicating Rainfall Magnitudes Under Different Climatic Conditions in Chongqing Province (China). *Atmosphere*, Vol. 12(10), <https://doi.org/10.3390/atmos12101318>
- [41] Krivoruchko, K. and Gribov, A., (2013). *Pragmatic Bayesian Kriging for Non-Stationary and Moderately Non-Gaussian Data*. SpringerLink. [https://link.springer.com/chapter/10.1007/978-3-642-32408-6\\_15](https://link.springer.com/chapter/10.1007/978-3-642-32408-6_15).
- [42] Rajesh, R., Elango, L. and Brindha, K., (2019). Methods for Assessing the Groundwater Quality. *GIS and Geostatistical Techniques for Groundwater Science*, 57–78. <https://doi.org/10.1016/b978-0-12-815413-7.00006-7>.
- [43] IBM. (2024). What is the K-Nearest Neighbors (KNN) Algorithm? [https://www.ibm.com/topics/knn#:~:text=The%20k%2Dnearest%20neighbors%20\(KNN\)%20algorithm%20is%20a%20non,of%20an%20individual%20data%20point](https://www.ibm.com/topics/knn#:~:text=The%20k%2Dnearest%20neighbors%20(KNN)%20algorithm%20is%20a%20non,of%20an%20individual%20data%20point)
- [44] Jalilov, S. M., (2017). Value of Clean Water Resources: Estimating the Water Quality Improvement in Metro Manila, Philippines. *Resources*, Vol. 7(1). <https://doi.org/10.3390/resources7010001>.
- [45] United States Geological Survey. (2018). *Dissolved Oxygen and Water*. Dissolved Oxygen and Water | U.S. Geological Survey. <https://www.usgs.gov/special-topics/water-science-school/science/dissolved-oxygen-and-water#:~:text=Cold%20water%20can%20hold%20more,oxygen%20concentration%20is%20often%20lower>
- [46] Environmental Management Bureau. (n.d.). *Regional profile | EMB - National Capital Region*. <https://ncr.emb.gov.ph/regionalprofile/>
- [47] Bolster, C. H., Bromley, J. M. and Jones, S. H., (2005). Recovery of Chlorine-Exposed *Escherichia Coli* in Estuarine Microcosms. *Environmental Science & Technology*, Vol. 39(9), 3083–3089. <https://doi.org/10.1021/es048643s>.
- [48] Cabral, J. P. and Marques, C., (2006). Faecal Coliform Bacteria in Febros River (Northwest Portugal): Temporal Variation, Correlation with Water Parameters, and Species Identification. *Environmental Monitoring and Assessment*, Vol. 118(1–3), 21–36. <https://doi.org/10.1007/s10661-006-0771-8>.

- [49] Soo, C. L., Ling, T. Y., Lee, N. and Apun, K., (2014). Assessment of the Characteristic of Nutrients, Total Metals, and Fecal Coliform in Sibu Laut River, Sarawak, Malaysia. *Applied Water Science*, Vol. 6(1), 77–96. <https://doi.org/10.1007/s13201-014-0205-7>.
- [50] Aram, S. A., Saalidong, B. M. and Osei Lartey, P., (2021). Comparative Assessment of the Relationship between Coliform Bacteria and Water Geochemistry in Surface and Ground Water Systems. *PLOS ONE*, Vol. 16(9). <https://doi.org/10.1371/journal.pone.0257715>.
- [51] Washington State Department of Ecology. (2005). *Issue What's so Important About Fecal Coliform Bacteria? - Washington*. <https://apps.ecology.wa.gov/publications/documents/0210010.pdf>.