International Journal of
Agriculture, Environment and Food Sciences

JAEFS

www.jaefs.com

e-ISSN : 2618-5946 DOI: 10.31015/jaefs.2021.4.29

Research Article Int J Agric Environ Food Sci 5 (4): 675-688 (2021)

Exceedance probability assessment of bathing water quality standards in lake Van based on a geostatistical analysis

Mert Sanli[1] **Yasemin Dilsad Yilmazel**[1,*]

[1]Middle East Technical University, Engineering Faculty, Environmental Engineering Department, Ankara, Turkey

[*]Corresponding Author: dilsad@metu.edu.tr

Abstract

Monitoring bathing water quality (BWQ) is highly important in Turkey both for public health issues and tourism income. Lake Van is one of the largest lakes of Turkey and serves as one of the most important tourist attractions in the eastern part of Turkey. This study aims to assess critical bathing sites in Lake Van by using historical BWQ data that was collected twice a month during the swimming season from June 15 to August 31, between the years 2010 and 2020. To avoid public health hazards, it is very important to determine the spatial dimension of inland water pollution and provide visual tools for its presentation. Geostatistical data analysis and the determination of critical locations have been done by a spatial interpolation method, named Probability Kriging (PK) using Geographic Information System (GIS) based software ArcMap 10. Probability maps for exceeding the threshold values identified for the two microbiological water quality parameters of *Escherichia coli* and intestinal enterococci (IE) were generated, and used for the identification of four critical regions: İskele (4), west side of Gevaş (1), Edremit (3) and Muradiye/Erciş (7). Insufficient wastewater treatment plant capacity serving to high population in these regions may indicate the most pressing issues disturbing the BWQ.

Keywords: Geostatistics, Kriging, Bathing water quality, Lake Van

Introduction

Turkey is a semi-island Mediterranean country serving as the bridge between Europe and the Middle East, therefore, nationwide monitoring of the bathing water quality (BWQ) is highly important to protect public health and the aquatic environment and also to increase national tourism income. Along with the long coastline on the Aegean Sea, Mediterranean Sea, and the Black Sea, Turkey has several inland waters such as Lake Van, Lake Salda, Lake Sapanca, and Lake Hazar that are also used for recreational purposes. In Turkey, the Ministry of Environment and Urban Planning (MEUP, former Ministry of Environment and Forestry) is responsible to set the standards to be met by the bathing waters. Before 2006, BWQ standards were provided under the Water Pollution Control Regulation. In January 2006, Turkey released its first “Bathing Water Quality Regulation”, in line with the former European Council Directive 76/160/EEC. Recently, in May 2019 Turkey revised its BWQ standards and released “the Regulation on the Management of Bathing Water Quality” that is in compliance with the current European Council Directive 2000/60/EC. The regulation aims to protect human health, environment and prevent pollution of bathing waters from all kinds of contamination especially microbiological (MoH, 2019). A striking example of such contamination is the occurrence of mucilage, which may host a variety of microbial species that potentially could spread marine diseases and economically harm local businesses (Danovaro et al., 2009, Keleş et al., 2021).

Pollution enters water bodies through different sources and is found in various forms. One of the most prevalent pollutions observed in water environments is fecal contamination coming from sewage and animals. Fecal contamination occurs because of poorly treated sewage, overflow of sewage due to the stormwater accumulation, faulty or leaky septic systems, runoff from the urban areas and the other sorts of waste that comes from the animals (EEA, 2018a). Using contaminated waters for swimming or recreational purposes may cause spread of illnesses due to the presence of pathogens. Indicator microorganisms are non-pathogenic

Cite this article as:
Sanli, M. Yilmazel, Y.D. (2021). Exceedance probability assessment of bathing water quality standards in lake Van based on a geostatistical analysis. International Journal of Agriculture, Environment and Food Sciences, 5(4), 675-6688
Doi: https://doi.org/10.31015/jaefs.2021.4.29
Orcid: Mert Sanli: https://orcid.org/0000-0001-5719-5492 , Yasemin Dilsad Yilmazel: https://orcid.org/0000-0001-9223-9681
Received: 16 September 2021 **Accepted:** 20 December 2021 **Published Online:** 30 December 2021
Year: 2021 **Volume:** 5 **Issue:** 4 (December) **Pages:** 675-688
Available online at : http://www.jaefs.com - http://dergipark.gov.tr/jaefs

microorganisms that are present strictly in pathogen-contaminated water. They do not multiply in water, but can be reliably detected even at low concentrations and are more numerous than pathogens and have similar survival times to pathogens (Thomas et al., 2007). Quantification of the indicator microorganism concentrations is preferred, because it is easier, faster, and more reliable than examining pathogens in water samples for assessing their potentially harmful impacts (Ortega et al., 2009). With the knowledge about the indicator organisms, it is aimed to forecast the probability of a pathogenic organism's existence. Although the use of bacterial indicators to measure water quality is common, there is no universal agreement on which indicator organisms are most useful, and there are no federal regulations that require a single standard for bacterial indicators. Total coliform, fecal coliform, and fecal streptococci have been used commonly as microbiological BWQ parameters. Yet, with the newest European regulation, *Escherichia coli* (*E. coli*) and Intestinal enterococci (IE*)* were chosen as fecal indicator bacteria to be monitored (EC, 2006).

Lakes among other aquatic environments are comparatively more stagnant and thus the impact of human activities on their water quality is greater in comparison to rivers. While investigating the water quality in the lakes, it is important to evaluate the biological, physical, and chemical parameters of the environment and their changes over time (Kaymak et al., 2021). Numerous research studies have been conducted so far to understand the causes of coastal pollution around the recreational waters, to find measures for its prevention, and to accurately monitor BWQ. For instance, an integrated approach of load estimation was used to determine pollution loads from nearby sources of domestic wastewater, runoff, and industrial wastewaters to estimate the total pollution load of Cartagena Bay in Columbia, and the study emphasizes the importance of calculating confidence intervals for each load value by combining different load estimation methods for land-based pollution loads in coastal areas (Tosic et al., 2018). In another study, remote sensing, and geographic information system (GIS) technologies were used together to assess pollution load on Burullus Lake via Landsat images. (El-Zeiny and El-Kafrawy, 2017). Coastal water pollution is a very complex phenomenon that is impacted by many different parameters, such as the configuration of the coastal area, hydrodynamic features of the coastal sediment, and local weather conditions (El Mrini et al., 2012; Mali et al., 2018). The accuracy of field monitoring is extremely important for studying such a complex natural phenomenon, and a lot of effort has been put to develop sensors, and remote sensing tools (Hafeez et al., 2019; Zielinski et al., 2009). An equally very important task is to develop data analysis tools that will enable decision-makers to perform accurate and reliable analyses leading to proper, and responsible decisions. Especially for countries such as Turkey, where tourism is quite important, the development of inexpensive, widely adaptable, and simple tools for BWQ data analysis is extremely important.

Geostatistics has been widely used for environmental data such as groundwater, soil, water quality related spatial data analysis (Jang, 2018). Geostatistics provides an accurate tool of BWQ estimation, especially where monitoring actions are limited and a high number of observations are impossible to obtain due to financial reasons ("Geostatistical Appl. Precis. Agric.," 2010). Kriging is one of the most commonly applied geostatistical interpolation methods (Bostan, 2017) and can be used for estimating BWQ in non-monitored areas (Malcangio et al., 2018; Jang, 2018). Probability kriging is a special form of kriging, which provides estimations of spatial data based on the available field data and a comparison to a threshold value. It provides the exceedance probability of the threshold limit value in a given location which then can be put into visual aids such as probability maps and provide visual and easy to follow tools for the general public or decision-makers. This study aims to use a geostatistical tool to estimate the general BWQ of the largest inland bathing site of Turkey, Lake Van, and determine the most critical bathing sites in the coastal zone of Lake Van. To this purpose, for the first time in the literature, historical microbiological water quality data collected during a ten-year period (2010 – 2020) from the coastal line of Lake Van has been processed via ArcGIS 10 software and the probability maps along with error estimations are presented.

Materials and Methods

Study Area

Lake Van is the largest body of water in Turkey with a total area of 3,750 km^2 and the depth of the lake reaches up to 450 meters (Doğan et al., 2016). The eastern part of the lake is in the territory of Van province, while the western part of the lake is in the territory of Bitlis province. Lake Van houses the most important inland bathing water sites of Turkey, and there are more than 30 beaches ("General Directorate of Public Health," 2021). Lake Van is located between the geographical coordinates of 38.5° N 43° E (Degens et al., 1984), which come up to the Eastern part of Anatolia in Turkey (Figure 1). Lake Van is the largest soda lake on earth (Tomonaga et al., 2017) and among the closed lakes around the world, it takes fourth place in terms of volume (607 km^3) (Kadioğlu et al., 1997). The water level is 1,648-meter-high according to sea level. The long axis of Van Lake which lies between the southwest of the Tatvan Bay and the Erciş Gulf in the northeast is 130 km and the axis between Ahlat Bay and Gevaş Bay is 80 km. The lake is surrounded by mountains (Özalp et al., 2016). The lowest place on the edge of the lake is east of Reşadiye and it is 1800 meter-high. The shallowest parts of the lake are the Erciş Bay and Van Bay, where the depth of water is around 50 meters. A depth of 451 meters was measured between Ahlat and Adilcevaz. The water of Lake Van is bitter, salty, and soda rich. The main reason for this is the accumulation of salty water in the lake and the continuous condensation due to evaporation.

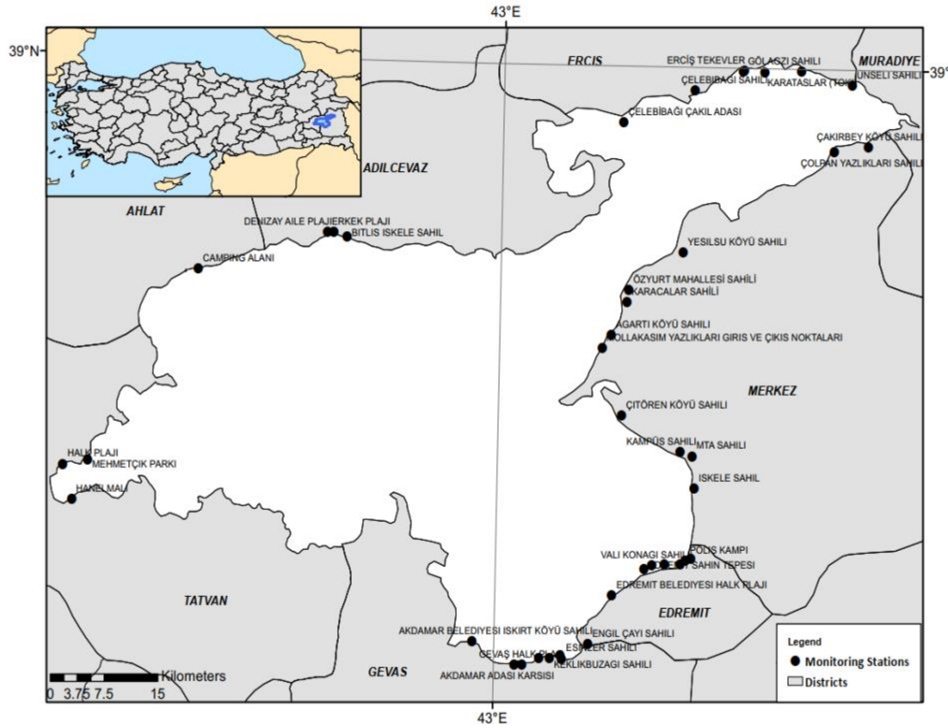

Figure 1. Map of monitoring stations and the study area, Lake Van

Sample Collection and Data Set

There are 40 monitoring stations in the Van Lake coastal line (Figure 1), and the data used in this study has been obtained from the Ministry of Health of the Republic of Turkey ("General Directorate of Public Health," 2021). Water samples were collected twice a month from each monitoring station during swimming season periods (June 15 - August 31) between 2010 and 2020. As stated in the Bathing Water Quality Regulation, samples were taken from each monitoring station, 15 days before the start of the swimming season (MoH, 2019). In order to represent the swimming conditions as much as possible, the sample must be taken at least 1 meter deep and 30 cm below the water surface. Within the scope of this study, 40 monitoring stations and a total number of 2,325 records were used.

Geostatistical Approach

The application of probabilistic methods to region-based variables has been known as geostatistics, which implies that any of these region-based variables have random and spatial properties (Varouchakis, 2019). The technique aims to develop a model for the spatial pattern, with the help of the variogram, since the variogram defines the spatial variability of the random variables between two points (Narany et al., 2014). The empirical variogram that is used for the analysis of the spatial variability between two points is defined as follows (Goovaerts et al., 2005);

In Equation 1, *N (h)* represents the number of data pairs, h is the lag vector and z (x) is the value of the spatial variables at the data collection points.

$$\boldsymbol{\gamma}\,(\boldsymbol{h}) = \frac{\mathbf{1}}{\mathbf{2}\,\boldsymbol{N}(\boldsymbol{h})}\sum_{\boldsymbol{\alpha}=\mathbf{1}}^{\boldsymbol{N}(\boldsymbol{h})}[\boldsymbol{z}(\boldsymbol{x}_{\boldsymbol{\alpha}} + \boldsymbol{h}) - \boldsymbol{z}(\boldsymbol{x}_{\boldsymbol{\alpha}})]^{\mathbf{2}} \qquad (1)$$

The experimental variogram needs to fit the theoretical variogram, which may be defined with eleven different functions.

Probability Kriging

The method used under this study for estimations of exceedance probability is called probability kriging (PK). PK is a modified version of co-kriging where only one variable is estimated using two spatial variables, indicator and uniform. The uniform value is defined as the z (x_α) in Equation 2 and the other variable is defined as I (x_α). The probability kriging equation is defined as (Adhikary et al., 2010);

$$\mathbf{I}(\mathbf{x_0}) = \sum_{\boldsymbol{\alpha}=\mathbf{1}}^{\mathbf{n}} \boldsymbol{\lambda}_{\boldsymbol{\alpha}}\,\mathbf{I}(\mathbf{x}_{\boldsymbol{\alpha}}) + \sum_{\boldsymbol{\alpha}=\mathbf{1}}^{\mathbf{n}} \boldsymbol{\xi}_{\boldsymbol{\alpha}}\,\mathbf{z}(\mathbf{x}_{\boldsymbol{\alpha}}) \qquad (2)$$

where λ_α and ξ_α are the weights associated with the two spatial variables.

For decreasing the variance of the error and ensure unbiased conditions the weights are defined as $\sum \lambda_\alpha = 1$ and $\sum \xi_\alpha = 0$. PK can be used to produce a probability map of the occurrence that exceeds the specified threshold value at a specified location. This method is appropriate to determine the critical bathing sites, which may exceed the threshold of BWQ set by the authorities, and also provides a standard error map. This information is useful to decision-makers especially when they need to make decisions on where and how many monitoring stations should be opened. Based on the provided maps, decision-makers can also decide whether they should close some bathing sites for remediation immediately, or to monitor slightly problematic areas with an increased frequency of sampling.

In spatial data analysis, a semivariogram is used to illustrate the spatial correlation of collected data points and yet in kriging, the experimental variogram may not be used directly since the kriging algorithm requires a model fit to describe the continuity of the data. To this purpose, several positive definite models (*i.e.*, mathematical functions) are used in the modeling step. Geographic Information System (GIS) based software ArcMap 10 was used to get semivariogram parameters for a total of eleven theoretical models including but not limited to exponential, stable, spherical, and k-Bessel. Among these eleven model options available in the software, best fitting models were chosen considering error functions for both *E. coli* and IE datasets. Specifically, the best-fitting model for each data set was founded by the comparison of the mean standardized error (MSE) and root mean square standardized error (RMSSE) values. For choosing the best model, MSE should be closed to zero (0) and the RMSSE should be close to 1 (McCoy and Johnston, 2002). The MSE and RMSSE were defined as (Audu, 2015; Zhang and Wang, 2010);

$$\mathbf{MSE} = \frac{\mathbf{1}}{\mathbf{n}} \sum_{\alpha=1}^{\mathbf{n}} \left\{ \frac{\mathbf{z}^*(\mathbf{x}_\alpha) - \mathbf{z}(\mathbf{x}_\alpha)}{\mathbf{p}_\alpha} \right\} \quad (3)$$

$$\mathbf{RMSSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{n}} \sum_{\alpha=1}^{\mathbf{n}} \left\{ \frac{\mathbf{z}^*(\mathbf{x}_\alpha) - \mathbf{z}(\mathbf{x}_\alpha)}{\mathbf{p}_\alpha} \right\}^{\mathbf{2}}} \quad (4)$$

where z (x_α) are actual values, z* (x_α) are estimated values, n states the number of observation points and p_α is the standard error prediction at the corresponding location

Determination of the Critical Bathing Sites

Coastal and inland bathing waters in Turkey are monitored and regulated through the provisions of Regulation on the Management of Bathing Water Quality (MoH, 2019). In the previous regulation, the parameters of BWQ were chosen as Total coliform, fecal coliform, and fecal streptococci. However, based on the current regulation, Turkey switched to the *Intestinal Enterococci* (IE) and *Escherichia coli* as BWQ parameters. According to the most recent regulation, FC and FS are considered equivalent to *E. coli* and IE, respectively. As a result of this, in all data up to 2019, the BWQ was taken in terms of FC and FS, as of 2020, the data sets were switched to IE and *E. coli*. According to the current regulation, waters used for recreational and swimming purposes must meet BWQ criteria given in Table 1. The 95-percentile evaluation process used in this study is defined as per the regulation (MoH, 2019) (i) taking the log10 of the dataset and if the obtained logarithmic value is equal to zero, take the minimum log10 value which is equal to 1, (ii) calculation of arithmetic mean of the log10 values (μ), (iii) calculation of the standard deviation of log10 values (σ). After these three steps, 95-percentile is calculated with antilog (μ + 1.65 σ).

Table 1. BWQ standards to be satisfied for “excellent” quality inland bathing waters in Turkey (MoH, 2019)

Parameter (CFU/100 mL)	**Guide Value**
Escherichia coli	500*
Intestinal enterococci	200*

* 95-percentile should be taken into consideration.

With a conservative approach, the threshold values were selected as the “excellent” quality parameters defined by the most recent Bathing Water Quality regulation, which is 500 CFU/100 mL for *E. coli* and 200 CFU/100 mL for IE (Table 1). The probability maps for the exceedance of identified thresholds are plotted and lead to the identification of the critical bathing sites of the study area.

Results and Discussion

Data Set Analysis

Data were taken from the Ministry of Health between the years 2010 and 2020. Prior to data input to GIS, the collected data were listed according to their location, then the geographical location *i.e.* XY coordinates of each monitoring site was assigned. Each BWQ parameter was expressed in the terms of Colony Forming Units (CFU) per 100 mL *i.e.*

CFU/100 mL. Table 2 represents important properties such as maximum, minimum, and mean values of the *E. coli* and IE data sets used in this study. The distribution of the data shows that the null values and exceedance numbers of threshold limits for the study area were widespread (Figure 2 and Figure 3).

Table 2. Characteristics of raw datasets

Parameter	*E. coli*	IE
N	2,325	2,325
Maximum	10,000	10,000
Minimum	0	0
Mean	49	181
Stdev	385	751

N: total number of measurements
Stdev: standard deviation
All concentration units are in CFU/100 mL.

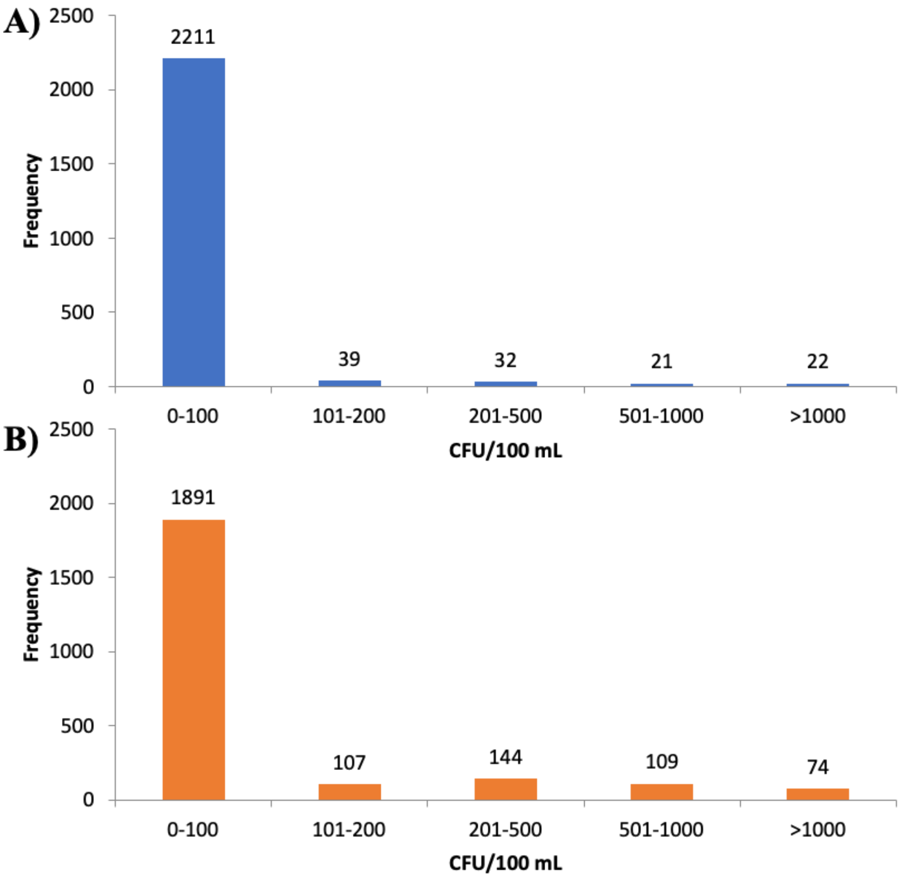

Figure 2. The histogram of the monitored parameters A) *E. coli* and B) IE

The complete dataset was used in the histogram (Figure 2) and the exceedance plot (Figure 3). As shown in Figure 2A, 98% of the data collected fall below the guide value of 500 CFU/100 mL for *E. coli*, indicating Lake Van water quality mostly was not a concern regarding this parameter. In the IE dataset, 86% of the data collected fall below the guide value of 200 CFU/ 100 mL specified for this parameter. The temporal variation of the exceedances provided in Figure 3B shows that most of the exceedances occurred in 2015 and 2016 for the IE parameter. For the *E. coli* data set (Figure 3A), a relatively higher number of exceedances were recorded between 2015 to 2018 over the 10-year period.

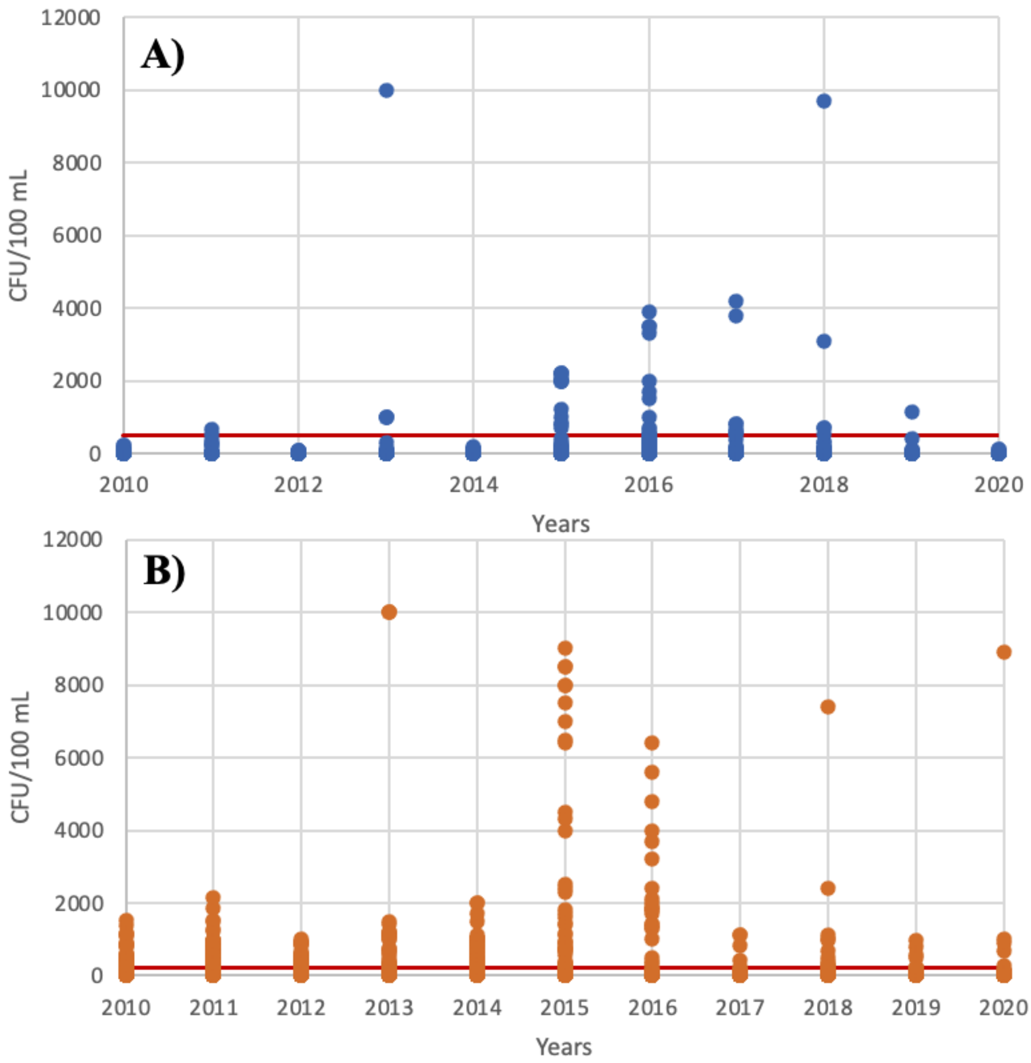

Figure 3. Temporal changes of exceedance number of threshold limits for A) *E. coli* and B) IE. The threshold limit value for each parameter is represented by the red line (Table 1).

Geographical Information System (GIS) Analysis

PK provides a value of probability that the regulatory limit of BWQ standards (the threshold limit for each BWQ parameter) is overcome at interpolation points. Probability maps that represent the prediction to exceed the threshold values were generated according to the semivariograms shown in Figure 4. As described in the Materials and Methods part, best-fitting model is chosen according to MSE and RMSSE values (Table 3).

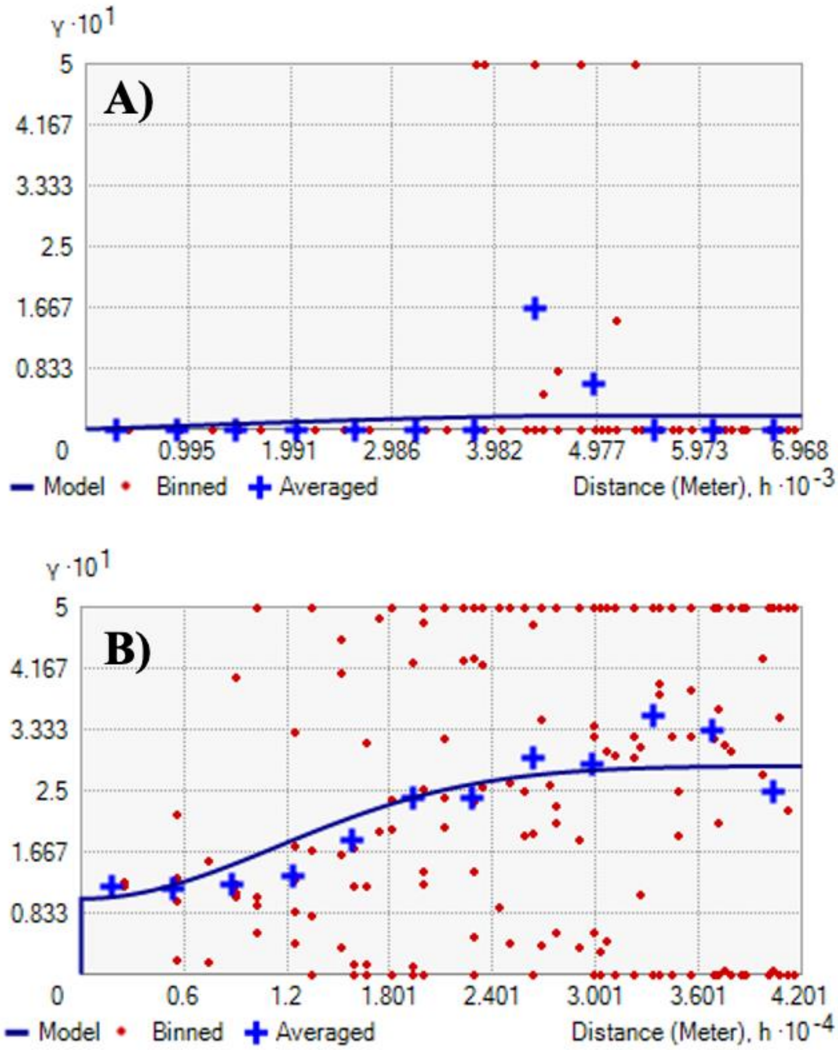

Figure 4. Variogram of datasets A) *E. coli* and B) IE

Table 2. Chosen semivariogram parameters and error values

BWQ parameter	Fitted Model	MSE[a]	RMSSE[b]	Range (m)	Nugget	Partial sill
E. coli	Circular	-0.100	1.008	4,645	0	0.0184
IE	Gaussian	0.044	1.010	5,000	0.1029	0.0180

a Mean standardized error
b Root mean square standardized error

Using PK, the probability maps for exceedance of thresholds are presented in Figure 5. For a better interpretation of results, the study area was divided into nine designated areas listed in Table 4, and also raw data properties that were presented for complete dataset in Table 2 was examined for each designated area in order to provide information on the raw datasets in addition to the 95-percentile evaluation. According to Table 4, mean values of *E. coli* were below the threshold limit for all of the designated areas, however, for the mean concentration of IE in four regions namely, west side of Gevaş (1), Edremit (3), İskele (4), and Muradiye/Erciş (7) were above the threshold limit. Based on this, there is an apparent high expectation to observe a higher probability of exceedance of IE threshold limit yet to confirm spatial analysis have been conducted. For the *E. coli* dataset, Figure 5A shows the most critical location as Region 4, İskele where the mean concentration is the highest with 132 CFU/100 ml (Table 4).

 DOI: 10.31015/jaefs.2021.4.29

Table 3. Data set analysis of designated areas in study area

	Designated Area	West side of Gevaş	East side of Gevaş	Edremit	İskele	Agartı/ Çıtören/ Mollakasım	Northern side of Merkez	Muradiye/ Erciş	Adilcevaz/ Ahlat	Tatvan
	Region Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
E. coli	N	186	196	546	268	282	246	370	138	93
	Max	10,000	2,100	4,200	3,900	300	300	3,500	700	400
	Min	0	0	0	0	0	0	0	0	0
	Mean	98	18	46	132	8	6	78	8	9
	Stdev	1,018	88	290	504	34	27	357	60	43
IE	N	186	196	546	268	282	246	370	138	93
	Max	10,000	810	7,000	10,000	2,500	1,000	9,000	500	550
	Min	0	0	0	0	0	0	0	0	0
	Mean	175	78	194	603	53	39	199	26	38
	Stdev	918	227	546	1592	198	112	898	52	70

N: total number of measurements

Stdev: standard deviation

All concentration units are in CFU/100 mL.

A maximum probability for exceedance of threshold value of 80% to 100% percent was found for the İskele region (4). In the other regions, the exceedance probability of the BWQ of Lake Van does not exceed 10%. North and south sections of the İskele region (4) are likely to exceed the threshold value by 30 to 60 percent due to their proximity to this identified critical area. For IE, critical site numbers and the exceedance probabilities were higher when compared to the *E. coli* dataset as expected based on the histogram (Figure 2B) frequency of the exceedances plot (Figure 3B) for this dataset. As shown on the probability map (Figure 5B) a significant portion of the coastal area of Lake Van including the west side of Gevaş (1), Edremit (3), İskele (4), Muradiye/Erciş (7) depict a probability of exceeding the threshold limit close to 100%. Among these critical sites, Iskele region (4) was also identified as a critical site according to the *E. coli* dataset. Therefore, definitely, more attention must be paid to this particular area. The authorities such as European Commission and World Health Organization (WHO) discuss whether monitoring of both *E. coli* and IE is necessary for monitoring of water quality in bathing sites is necessary or if one of these parameters is sufficient (WHO, 2018; Tiwari et al., 2021). Even though early work suggested monitoring of only *E. coli*, the results presented herein clearly indicated the importance of monitoring both *E. coli* and IE simultaneously for the case of Lake Van. There are also areas that show qood water quality in the Lake Van coastal zone, for instance, the probability of exceeding the threshold value in the east side of Gevaş (2) was relatively low when compared to the neighbors of this site. Although Mollakasım region (5) and northern side of Merkez (6) are located in between the critical areas, the probability of exceeding the threshold value does not exceed 10% in these regions. The other regions Agartı (5), northern side of Merkez (6), Adilcevaz (8), and Tatvan (9) regions depict a low probability of exceeding the threshold limits.

In summary, based on the information presented on the exceedance probability maps (Figure 5), the territory of İskele region (4) is identified as the most critical area, since it shows a high threshold overcoming probability for both parameters. This could be due to the presence of a chronic problem in this area. Regions of the west side of Gevaş (1), Edremit (3), and Muradiye/Erciş (7) also need attention since they show high overcoming probabilities based on the IE dataset. In another recently published work focusing on Lake Van water quality, similar results were attained for the year 2015 (Aydin et al., 2021). In their study, researchers collected samples between June and September 2015, and analyzed over 200 water samples for their total coliform, fecal coliform, and enterococcus content, which indicated lowered water quality in the same 4 regions and additionally in Region 5 namely, Agartı/Mollakasım/Çıtören (Aydin et al., 2021). Data used in our study covers years between 2010 – 2020 and the analysis conducted herein clearly indicates the problems with İskele (4), Gevaş (1), Edremit (3) and Muradiye/Erciş (7) were chronic; yet, over the ten-year period Region 5 conditions were improved since 2015.

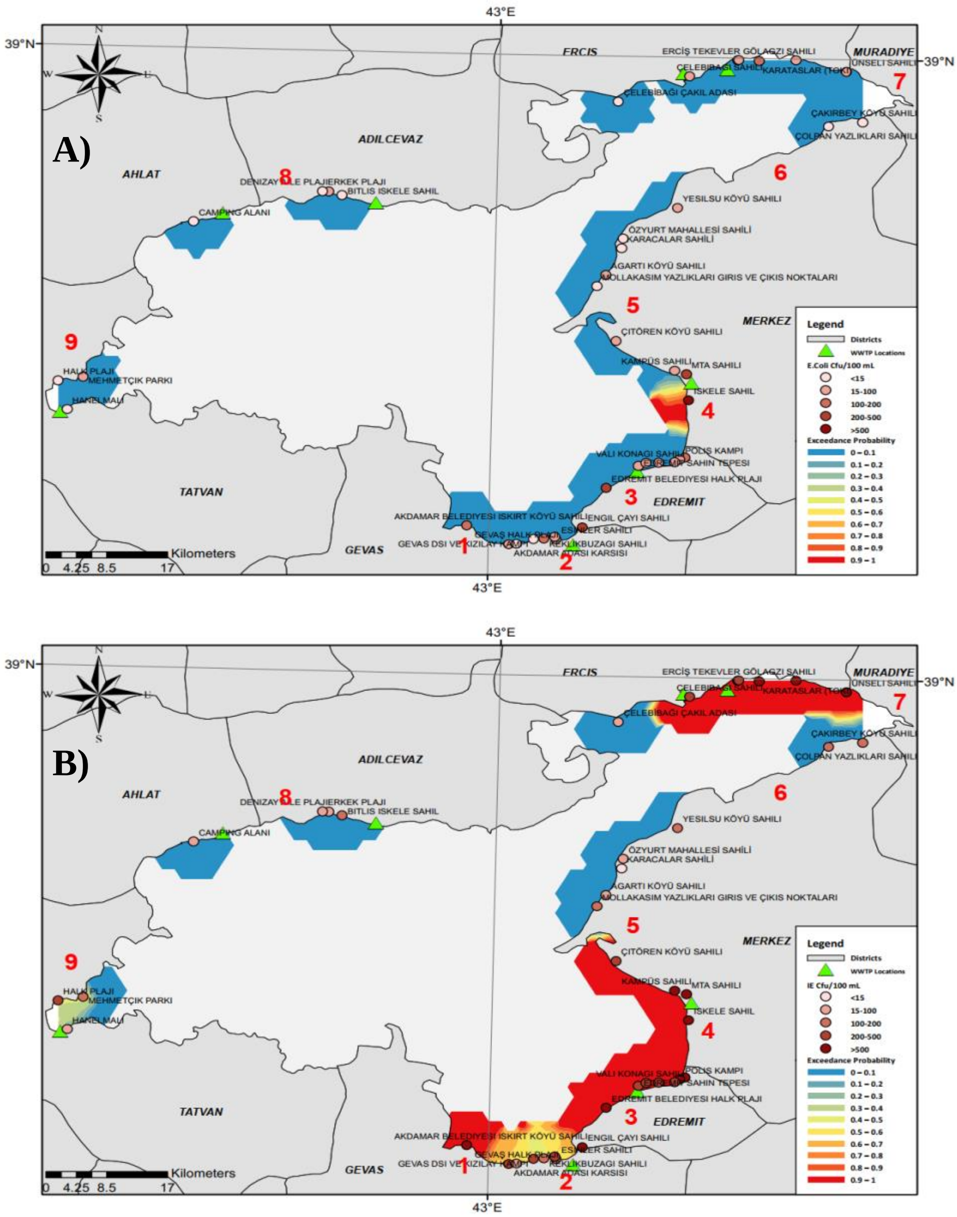

Figure 5. Probability map of threshold exceedance for datasets of A) *E. coli* and B) IE

GIS also provides a standard error map in the form of error variance. PK tool of GIS uses the methodology of an indicator variable, which is a binary, and computes maps of probability by classifying the dataset to 0 or 1 based on the threshold value. If the values are less than the threshold they are specified as 0, and if the values are higher than the threshold then they are set a value of 1. After the interpolation of the variable, the expected value of the variable is calculated by the prediction map. This expected value may be considered to exceed the threshold value of the expected variable. So, the error map represents the error of the probability that the threshold value is exceeded. Figure 6 shows the error maps for each BWQ parameter. As expected, the error percentage for two BWQ parameters is increasing while moving away from the coastal zones (Figure 6). Also, the percentage error is increasing with increasing distance between the two monitoring stations. Since our range was calculated as 4,645 meters for the *E. coli* dataset and 5,000 meters for the IE dataset in Table 3, if the number of monitoring stations within this range is high, the margin of error is minimized. This means that increasing the frequency of the monitoring stations will result in more reliable outputs. To give an example, sites like Camping Alanı in Region 8 and Çelebibağı Çakıl Adası (in Erciş, West of Region 7) has not enough monitoring stations within their range that is the reason why they have the highest error in the probability maps. Also, between the west side of Gevaş (1) and Tatvan (9) the frequency of monitoring station number is significantly less than the frequency of stations between the Edremit (3) and İskele (4), which has the lowest error percentage for all parameters. Since the frequency of the monitoring stations has higher in the İskele region, error for IE stands within a range of 32-36%. On the other hand, the error range is between 45 to 50% between Region 1 and Region 9 since there are not enough monitoring stations along this path. In the *E. coli* dataset, the standard error of the predictions is lower, because lower number of exceedances were recorded in comparison to the IE dataset. Nevertheless, İskele Region (4) is proven to be a problematic area since the error of prediction for this region is around 10 % and 30 % for *E. coli* and IE datasets, respectively.

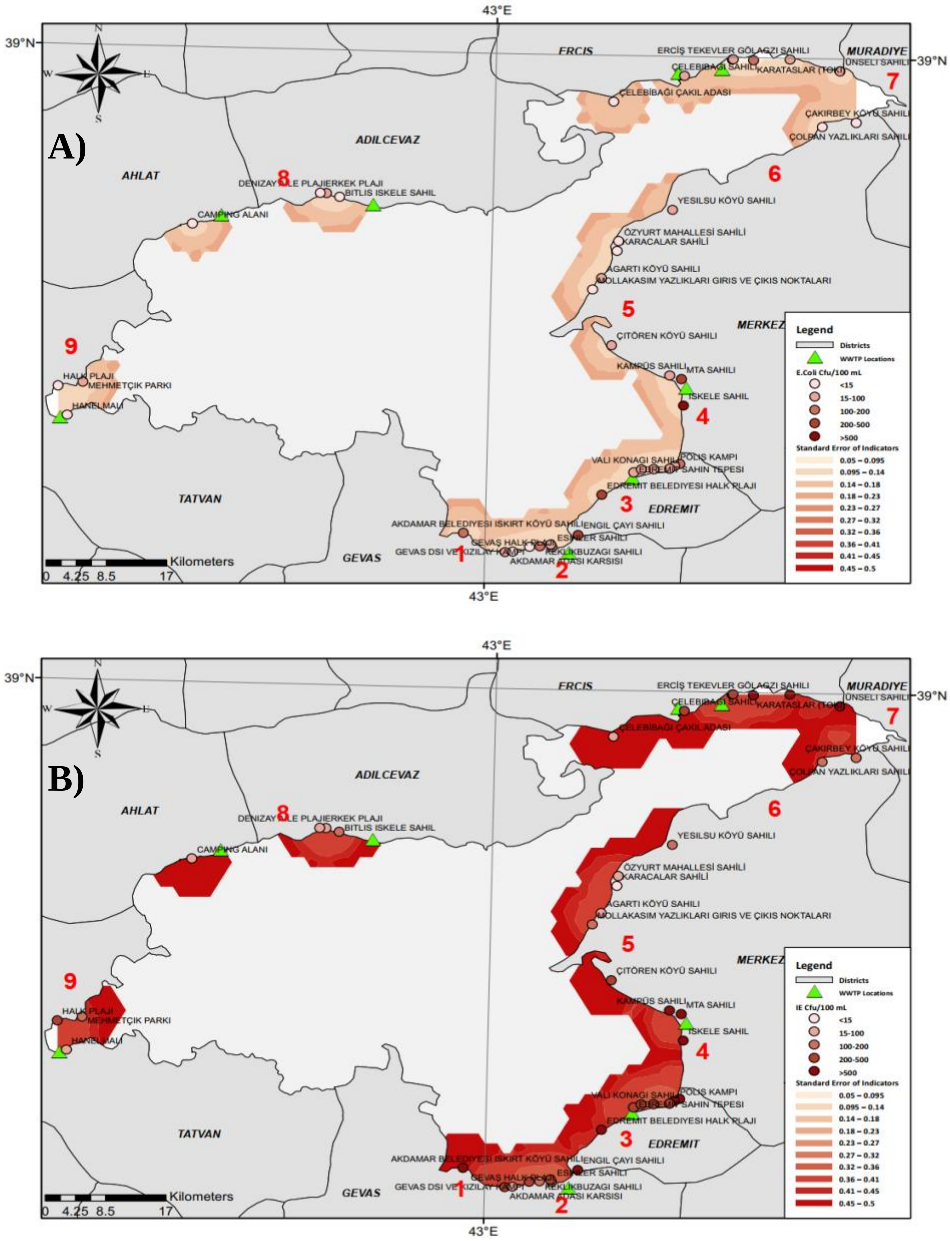

Figure 6. The standard error map for datasets of A) *E. coli* and B) IE

Even though the use of microbiological indicators for prediction of water quality in Lake Van provides information about the presence and the degree of fecal contamination; unfortunately, source tracking by just monitoring these parameters is impossible. For source tracking, culture-independent measurement methods such as real-time quantitative polymerase chain reaction (q-PCR) based assays are necessary, however, they are not preferred due to the economic burden on monitoring agencies (Tiwari et al., 2021). Lake Van is a closed water body, therefore, discharges and water draining from farms to the lake could make a significant contribution to its pollution. To evaluate the impact of the wastewater treatment plant (WWTP) discharges on the lake the discharge locations of WWTPs were also shown in probability maps generated (Figure 5). The treatment capacity of a WWTP is designed based on the population served, thus the capacity of each WWTP is different. Therefore, the load of the discharge to the lake from each WWTP may differ. The WWTP that is located in the İskele (4) district has the highest loads of discharge, and is serving to the highest population in the Lake Van area. Mostly, there is a strong correlation between the discharge, and the quality of nearby surface waters, thus, there is a need to assess the treatment performance of the WWTPs in the identified critical areas (Sanders er al., 2013). In fact, recently a monitoring study has been conducted in Edremit coastline to determine any potential negative impacts of WWTPs on Lake Van through impact analysis, and the results indicated that the discharge from Edremit WWTP significantly impacts lake water quality especially in terms of the provided organic pollution load (as measured by chemical oxygen demand and biological oxygen demand) (Ozguven and Yetis, 2020). Clearly, in the case of the Lake Van area, especially for the Edremit (3), İskele (4) and Erciş (7) regions, there might be a need for an additional WWTP or a capacity increase of the current WWTP. In another recent study, where heavy metal pollution over the Edremit coast (Region 3) of Lake Van is investigated, it was stated that the chromium and copper concentrations measured in the effluent water of Edremit WWTP are at a level that exerts pressure on Lake Van (Yetiş and Özgüven, 2020). Additionally, in a recent field study, the presence of animal-related pollution was reported for Karasu river (freshwater) at the specific location, where it is flowing into Lake Van, *i.e.* between Çıtören Köyü beach and Kampüs beach on Figure 5 (PEMAT, 2018). During the fieldwork, the measurements of water quality parameters confirmed the presence of pollution, and the lake water was classified under Class IV (the most polluted) according to the Turkish Water Pollution Control Regulation (MoEF, 2004).

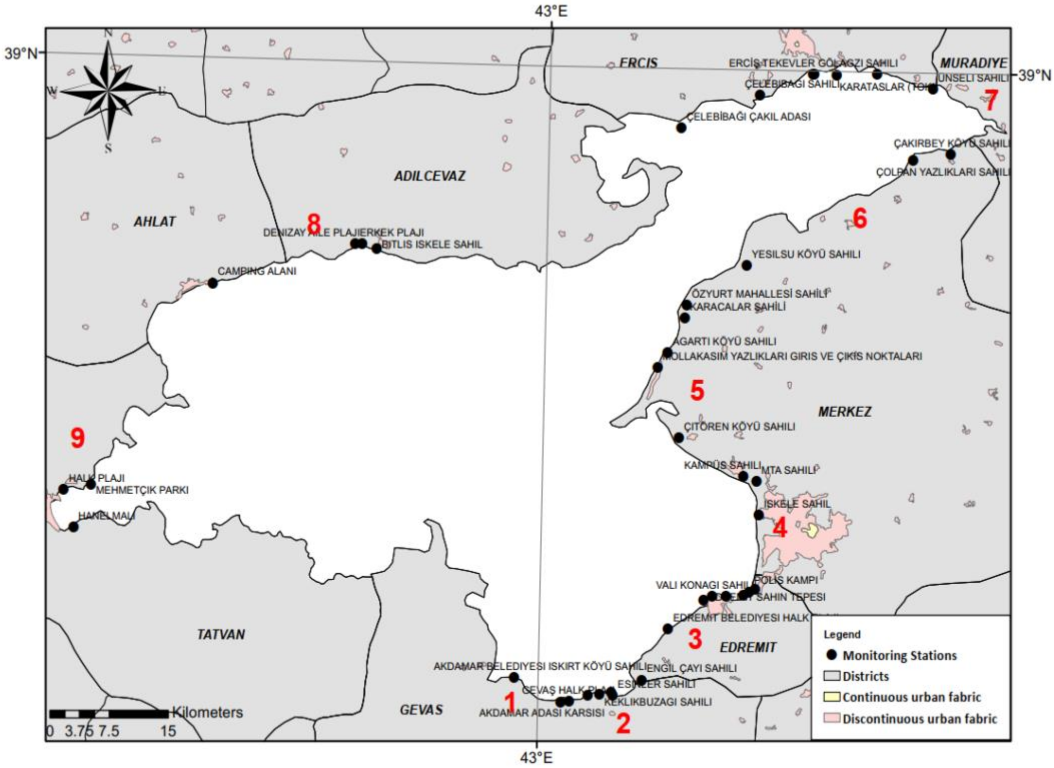

Figure 7. Urbanization Map around the Lake Van

Another supporting information is gathered by plotting the data about the urban and rural population densities that were obtained from the Coordination of Information on the Environment (CORINE) available as a free source. The data which was taken from CORINE (EEA, 2018b) illustrates the Continuous Urban Fabric and Discontinuous Urban Fabric (Figure 7). Continuous Urban Fabric is the class assigned to residential buildings, public service and commercial buildings and intercities of non-vegetated or bare surfaces. In addition to that, Discontinuous Urban Fabric illustrates individual houses, small and large blocks of flats, parking areas and such (EEA, 2020). Based on this information, it can be said that the urban fabric is a good tool for the assessment of the population around a given area. As it can be seen from Figure 7, there were 3 critical locations around Lake Van based on population. These 3 locations are İskele (4), Erciş (7) and Edremit (3) regions of the lake coastal zone. Apparently, there is a link between the disturbed BWQ parameters and population (Figure 7), which may indicate that there needs to be additional measures taken in order to prevent contamination of nearby water bodies in highly populated urban areas.

Conclusion

In this study, the critical bathing sites in the Lake Van coastal zone, that may pose risk to swimmers were identified via geostatistical evaluation of microbiological water quality data collected for ten years from 2010 – to 2020 by the Ministry of Health. Data collected from 40 monitoring stations were critically analyzed in terms of BWQ parameters, namely *E. coli* and IE. The results obtained in this study prove that PK is a good tool for evaluation of the critical sites of BWQ and can be effectively used by the decision-makers to determine the critical locations, where a higher number of monitoring stations are needed. In these critical sites, an increased frequencies of the sample collection may be necessary to ensure the safety of the swimmers and the locals. The most critical site of Lake Van is identified as İskele (4) region, with the highest percentage of exceedance probability of the regulatory standards observed in all BWQ parameters. The other three critical sites are regions of the west side of Gevaş (1), Edremit (3), and Muradiye/Erciş (7), which need more attention. The overlap between the exceedance probability maps, and increased urbanization underlines that highly populated areas may represent the high-risk areas especially related to the capacity of the WWTPs. Therefore, especially in highly populated areas the capacity of the WWTPs and their performance is quite important to lower the amount of contamination.

Compliance with Ethical Standards

Conflict of interest

The authors declared that for this research article, they have no actual, potential, or perceived conflict of interest.

Author contribution

Mert Sanli carried out data analysis and writing the original draft. Yasemin Dilsad Yilmazel designed the study, supervised, wrote, reviewed and edited the original draft. Both authors approve the final version of this manuscript. Both authors verify that the tables, figures and the main text are original and that they have not been published before.

Ethical approval

Ethics committee approval is not required.

Funding

No financial support was received for this study.

Data availability

Not applicable.

Consent for publication

Not applicable.

Acknowledgements

Authors are thankful to Ms. Almila Kindan at the Foundation for Environmental Education in Turkey (TURCEV) for discussion on nationwide BWQ data collection system.

References

Adhikary, P. P., Dash, C. J., Bej, R., and Chandrasekharan, H. (2010). Indicator and probability kriging methods for delineating Cu, Fe, and Mn contamination in groundwater of Najafgarh Block, Delhi, India. Environmental Monitoring and Assessment, 176(1–4), 663–676. https://doi.org/10.1007/s10661-010-1611-4

Audu, I. (2015). An Application of Geostatistics to Analysis of Water Quality Parameters in Rivers and Streams in Niger State, Nigeria. American Journal of Theoretical and Applied Statistics, 4(5), 373. https://doi.org/10.11648/j.ajtas.20150405.18

Aydın, E., Parlak, M., Güdücüoğlu, H. and Bayram, Y. (2021). Determination of Microbiological Pollution Level of Lake Van and Lake Erçek Situated within the Borders of Van Province. Türk Mikrobiyoloji Cemiyeti Dergisi. Published. https://doi.org/10.5222/tmcd.2021.02418

Bostan, P. (2017). Basic kriging methods in geostatistics. Yuzuncu Yil University Journal of Agricultural Sciences, 27, 10-20.

Danovaro, R., Fonda Umani, S. and Pusceddu, A. (2009). Climate Change and the Potential Spreading of Marine Mucilage and Microbial Pathogens in the Mediterranean Sea. PLoS ONE, 4(9), e7006. https://doi.org/10.1371/journal.pone.0007006

Degens, E. T., Wong, H. K., Kempe, S. and Kurtman, F. (1984). A geological study of lake van, Eastern Turkey. Geologische Rundschau. https://doi.org/10.1007/BF01824978

Peyzaj Mimarliği Akademik Topluluğu (PEMAT) Toplantısı Çalışma Sonuçları Edit. Şevket Alp, Emel Baylan, Pınar Bostan, Onur Şatır, Üyesi Feran Aşur, Gülçinay Başdoğan, Hilmi Ekin Oktay, XXI. PEMAT 2018 Yılı Toplantısı Van YYÜ/Türkiye, Vameda Ofset Maatbacılık Reklam Sn. Tic. Ltd. Şti, Van, ISBN 978-975-7616-68-9, Page 62.

Doğan, E., Kocamaz, U. E., Utkucu, M. and Yıldırım, E. (2016). Modelling daily water level fluctuations of Lake Van (Eastern Turkey) using Artificial Neural Networks. Fundamental and Applied Limnology, 187(3), 177–189. https://doi.org/10.1127/fal/2015/0736

EC. (2006). *Bathing Water Quality Directive 2006/7/EC*. (1882), 51.

EEA. (2018a). European bathing water quality in 2017. *European Environment Agengy Reports*, 1–28. https://doi.org/10.2800/116560

EEA. (Accessed 25 September 2020). CORINE Land Cover Nomenclature Guidelines. Retrieved from https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines

EEA. (2018b). Corine Land Cover. Retrieved from https://land.copernicus.eu/pan-european/corine-land-cover/clc2018

El-Zeiny, A. and El-Kafrawy, S. (2017). Assessment of water pollution induced by human activities in Burullus Lake using Landsat 8 operational land imager and GIS. *Egyptian Journal of Remote Sensing and Space Science*, *20*, S49–S56. https://doi.org/10.1016/j.ejrs.2016.10.002

El Mrini, A., Maanan, M., Anthony, E. J. and Taaouati, M. (2012). An integrated approach to characterize the interaction between coastal morphodynamics, geomorphological setting and human interventions on the Mediterranean beaches of northwestern Morocco. *Applied Geography*, *35*(1–2), 334–344. https://doi.org/10.1016/j.apgeog.2012.08.009

General Directorate of Public Health. (Accessed 01 July 2021). Retrieved from Yüzme Suyu Takip Sistemi website: https://yuzme.saglik.gov.tr

Geostatistical Applications for Precision Agriculture. (2010). In M. A. Oliver (Ed.), *Geostatistical Applications for Precision Agriculture*. https://doi.org/10.1007/978-90-481-9133-8

Goovaerts, P., AvRuskin, G., Meliker, J., Slotnick, M., Jacquez, G. and Nriagu, J. (2005). Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. *Water Resources Research*, *41*(7), 1–19. https://doi.org/10.1029/2004WR003705

Hafeez, S., Sing Wong, M., Abbas, S., Y. T. Kwok, C., Nichol, J., Ho Lee, K., Tang, D. and Pun, L. (2019). Detection and Monitoring of Marine Pollution Using Remote Sensing Technologies. In *Monitoring of Marine Pollution*. https://doi.org/10.5772/intechopen.81657

Jang, C. S. (2018). Probabilistic assessment of the water quality of bathing beaches on the Taiwanese coast using indicator-based geostatistics with one- dimensional cyclic coordinates. *Environmental Earth Sciences*, *77*(22), 1–10. https://doi.org/10.1007/s12665-018-7958-5

Kadioğlu, M., Şen, Z. and Batur, E. (1997). The greatest soda-water lake in the world and how it is influenced by climatic change. Annales Geophysicae, 15(11), 1489–1497. https://doi.org/10.1007/s00585-997-1489-9

Kaymak, G., Kayhan, F. E. and Yön Ertuğ, N. D. (2021). A biomonitoring study: Using the biomarkers in Cyprinus carpio for the evaluation of water pollution in Sapanca lake (Sakarya, Turkey). International Journal of Agriculture, Environment and Food Sciences, 5(1), 107–121. https://doi.org/10.31015/jaefs.2021.1.14

Keleş G., Yilmaz S. and Zengin M.2020. Possible Economic Effects of Musilage on Sea of Marmara Fisheries. Int. J. Agric. For. Life Sci.,4(2):173-177.

Mali, M., Malcangio, D., Dell' Anna, M. M., Damiani, L. and Mastrorilli, P. (2018). Influence of hydrodynamic features in the transport and fate of hazard contaminants within touristic ports. Case study: Torre a Mare (Italy). *Heliyon*. https://doi.org/10.1016/j.heliyon.2017.e00494

Malcangio, D., Donvito, C. and Ungaro, N. (2018). Statistical Analysis of Bathing Water Quality in Puglia Region (Italy). International Journal of Environmental Research and Public Health, 15(5), 1010. https://doi.org/10.3390/ijerph15051010

McCoy, J. and Johnston, K. (2002). Using Arcgis Spatial Analyst. Esri Press.

MoEF (Ministry of Environment and Forestry), 2004. Water Pollution Control Regulation. Official Gazette dated 31 December 2004 and numbered 25687.

MoH (Ministry of Health), 2019. Regulation on Concerning the Management of Bathing Water Quality. Official Gazette dated 25 September and numbered 30899.

Narany, T. S., Ramli, M. F., Aris, A. Z., Sulaiman, W. N. A. and Fakharian, K. (2014). Spatial assessment of groundwater quality monitoring wells using indicator kriging and risk mapping, Amol-Babol Plain, Iran. *Water (Switzerland)*, *6*(1), 68–85. https://doi.org/10.3390/w6010068

Ortega, C., Solo-Gabriele, H. M., Abdelzaher, A., Wright, M., Deng, Y. and Stark, L. M. (2009). Correlations between microbial indicators, pathogens, and environmental factors in a subtropical Estuary. *Marine Pollution Bulletin*. https://doi.org/10.1016/j.marpolbul.2009.04.015

Ozguven, A. and Demir Yetis, A. (2020). Assessment of Spatiotemporal Water Quality Variations, Impact Analysis and Trophic Status of Big Soda Lake Van, Turkey. Water, Air, & Soil Pollution, 231(6). https://doi.org/10.1007/s11270-020-04622-x

Özalp, S., Aydemir, B. S., Olgun, E., Şimşek, B., Elmacı, H., Evren, M., Emre, M., Aydın, M. B., Kurtuluş, O., Öcal, F., Can, A. Z., Yanmaz, M. N., Apa, R. and Duman, T. Y. (2016). Tectonic deformations in the quaternany deposits of the Lake Van (Edremit Bay),Eastern Anatolia, Turkey. Bulletin Of The Mineral Research and Exploration, 0(153). https://doi.org/10.19111/bmre.58034

Sanders, E. C., Yuan, Y. and Pitchford, A. (2013). Fecal coliform and E. coli concentrations in effluent-dominated streams of the upper santa cruz watershed. *Water (Switzerland)*, *5*(1), 243–261. https://doi.org/10.3390/w5010243

Thomas, J. C., Lutz, M. A., Bruce, J. L., Graczyk, D. J., Richards, K. D., Krabbenhoft, D. P., Westenbroek, S. M., Scudder, B. C., Sullivan, D. J. and Bell, A. H. (2007). Water-Quality Characteristics for Selected Sites Within the Milwaukee Metropolitan Sewerage District Planning Area, Wisconsin, February 2004-September 2005. Scientific Investigations Report. Published. https://doi.org/10.3133/sir20075084

Tiwari, A., Oliver, D. M., Bivins, A., Sherchan, S. P. and Pitkänen, T. (2021). Bathing Water Quality Monitoring

Practices in Europe and the United States. International Journal of Environmental Research and Public Health, 18(11), 5513. https://doi.org/10.3390/ijerph18115513

Tomonaga, Y., Brennwald, M. S., Livingstone, D. M., Kwiecien, O., Randlett, M. È., Stockhecke, M., Unwin, K., Anselmetti, F. S., Beer, J., Haug, G.H., Schubert, C. J., Sturm M. and Kipfer, R. (2017). Porewater salinity reveals past lake-level changes in Lake Van, the Earth's largest soda lake. *Scientific Reports*. https://doi.org/10.1038/s41598-017-00371-w

Tosic, M., Restrepo, J. D., Izquierdo, A., Lonin, S., Martins, F. and Escobar, R. (2018). An integrated approach for the assessment of land-based pollution loads in the coastal zone. *Estuarine, Coastal and Shelf Science*, *211*, 217–226. https://doi.org/10.1016/j.ecss.2017.08.035

Varouchakis, E. A. (2019). Geostatistics. Spatiotemporal Analysis of Extreme Hydrological Events, 1–38. https://doi.org/10.1016/b978-0-12-811689-0.00001-x

WHO. WHO Recommendations on Scientific, Analytical and Epidemiological Developments Relevant to the Parameters for Bathing Water Quality in the Bathing Water Directive (2006/7/EC); World Health Organization: Geneva, Switzerland, 2018.

Yetiş, A. D. and Özgüven, A. (2020). Investigation of Heavy Metal Pollution in Surface Waters of the Van Lake Edremit Coast. Uludağ University Journal of The Faculty of Engineering, 831–848. https://doi.org/10.17482/uumfd.752460

Zhang, H. and Wang, Y. (2009). Kriging and cross-validation for massive spatial data. Environmetrics, 21(3–4), 290–304. https://doi.org/10.1002/env.1023

Zielinski, O., Busch, J. A., Cembella, A. D., Daly, K. L., Engelbrektsson, J., Hannides, A. K. and Schmidt, H. (2009). Detecting marine hazardous substances and organisms: sensors for pollutants, toxins, and pathogens. *Ocean Science*, *5*(3), 329–349. https://doi.org/10.5194/os-5-329-2009