Archive of SID

Journal of Environmental Studies

Vol. 46, No. 2, Summer 2020

Journal Homepage: www.Jes.ut.ac.ir
Print ISSN: 1025-8620 Online ISSN 2345-6922

Water Quality Parameters of Tajan River Presenting a Proposed Method Based on Blind Kriging and Linear Regression

Document Type
Research Paper

Received
December 6, 2019

Accepted
March 26, 2020

Rana Kasem, Babak Omidvar*, Abdolreza Karbassi, Amin Sarang

Department of Environmental Engineering, School of Environment, College of Engineering, University of Tehran, Tehran, Iran

DOI: 10.22059/JES.2020.294959.1007960

Abstract

The ordinary kriging method has been widely used for interpolation and prediction of water quality parameters. One of its major weaknesses is that it assumes the mean of modelled variables to be constant. In this study, a modified kriging method is proposed, which has an unknown mean function based on blind kriging. It uses linear regression instead of Bayesian technique. The proposed method and ordinary kriging were used to model the spatial variability of heavy metals and water quality index and their results were compared. 21 water quality parameters were measured at ten stations on Tajan River, North Iran. It can be noticed that the water quality index had low values near Sari city and its value was increased by moving away from the city showing an improvement in the quality of water. The results showed that the proposed method was the most accurate model for estimating most of the parameters. Its improvement was between 8% and 169%. For Iron concentration that was highly correlated with the spatial structure, the ordinary kriging performed better.

Keywords: Geostatistics, Ordinary kriging, Tajan river, Universal kriging, Water quality index

* Corresponding author Email: bomidvar@ut.ac.ir

www.SID.ir

Archive of SID

Expanded Abstract

Introduction

The application of kriging in the field of environment is focused on four main sections; mapping of precipitation, quantitative and qualitative status of groundwater, quantitative and qualitative evaluation of surface water and spatial forecasting of air quality.

Different types of kriging have been developed, but the most popular is ordinary kriging (OK). The most important disadvantages of OK is that it assumes the mean of modelled variables to be constant and the prediction is only based on the spatial structure of the studied points. In addition, the effect of important parameters does not take into account the estimation result, and in some cases the predicted values by OK may be out of the studied range. To overcome these problems, the universal kriging and blind kriging have been developed. Blind kriging (BK) is a more complete version of universal kriging and is based on Bayesian variable selection technique which is complicated and taken a lot of time to identify the unknown mean function.

In this research, due to the complexity of the Bayesian computation, we will combine the regression technique with blind kriging as the unknown mean function is defined by variable selection techniques being used in linear regression analysis such as forward selection, backward elimination, and step-wise regression. The residuals at the known points are calculated from the difference of the observed values and the values of the selected function. Then, the residual mean at the unknown point is solved by OK method.

Tajan River is one of the most important rivers in Mazandaran province. This permanent river is about 140 km long and originates from the mountainous area on the northern slopes of Alborz Range. Tajan River runs to the plains carrying the water from various tributaries in the mountains and then drains into the Caspian Sea. Dodangeh, Lajim, Chahardangeh and Zarem rivers are the most important branches of it. The area of Tajan River watershed in the Aldehil region (before entering the Caspian Sea) is 4700 km^2. There are different land uses including agriculture, aquaculture, dam construction and industrial activities around the river. The necessity to perform a systematic study of the river water quality is, therefore, a need and of prime importance.

The objective of this study is to evaluate water quality parameters in different sites on Tajan River. This study also attempts to propose a modified kriging method, in which the unknown mean function is defined by using linear regression in order to simplify the computations of blind kriging. The proposed method and ordinary kriging were used to model the spatial variability of heavy metals and water quality index and their results were compared. This case study may be counted as an initial effort to study the spatial variability of water quality parameters, which may have many practical implications.

Material and Methods

The water sampling took place in spring of 2018. Sampling sites were selected based on natural conditions and accessibility to Tajan River by taking into account natural and human impacts, including river sub-branches, changes in the polluting sources such as agricultural lands, residential centers, existing industries, etc. Sampling sites were named 1 to 10 from downstream (near the Caspian Sea) to upstream (the Shahidrajae Dam). Water samples were collected by Nansen bottles and transferred into 1-liter bottles previously cleaned by nitric acid (0.1 N). The temperature, pH, electrical conductivity (EC), total dissolved solids (TDS), dissolved oxygen (DO), redox potential (Eh), and turbidity (Turb) of each water sample were directly measured at the sampling points. The biochemical oxygen demand (BOD_5) was determined by the Winkler Azide method and chemical oxygen demand

www.SID.ir

Archive of SID

(COD) by the dichromate reflux method. Phosphate (PO_4-P) and Nitrate nitrogen (NO_3-N) were analyzed by spectrophotometric method, and fecal coliform (Fcoli) were measured by multiple tube method.
For measuring metals, each water sample was filtered through Whatman filter (0.45 μm) and about 5 mL of HNO_3 (0.1 N) was added to the samples (until pH<2). Then, the samples were stored in the refrigerator at a temperature below 2°C until being transferred to the laboratory. The measured elements included Arsenic (As), Barium (Ba), Calcium (Ca), Cobalt (Co), Chromium (Cr), Copper (Cu), Iron (Fe), Nickel (Ni), Lead (Pb) and Thorium (Th). Metal measurements were done by inductively coupled plasma optical emission spectrometry (ICP-OES). Analysis of the samples was done based on the instructions recommended by (APHA, 2005).
All mathematical and statistical computations were made using Excel 2016, SPSS 22 and ArcGIS 4.10.1.
IRWQI is formed by nine selected water quality parameters namely as DO, BOD, COD, EC, PO_4-P, NO_3-N, Turb, Fcoli and pH. Its values range from 0 to100 and are calculated as following:

$$IRWQI = \left(\prod_{i=1}^{n} Ii^{wi}\right)^{\frac{1}{\sum wi}} \quad (1)$$

where n is the number of parameters, Ii the index value for the ith parameter (obtained from special curves) and wi is the weight of the ith parameter. IRWQI was calculated in all the sampling sites.
The ordinary kriging method (OK) is a linear unbiased geostatistical estimator and based on spatial correlation between sites. To indicate spatial correlation between the sampling sites, the empirical semivariogram is used as a function of distance between sampling sites. There are various models of the theoretical semivariogram in GIS system and one of these semivariograms must be selected based on the calculated errors. The prediction equation is a linear weighted combination of the form:

$$Z_{x_0} = \sum_{i=1}^{n} Z_{x_i} * W_i \quad (2)$$

where Z_{x_0} is the estimated value at the point x_0, Z_{x_i} and W_i represent the observed values and calculated weight at point x_i, and n is the number of observed values. Numerous studies have reported that conventional kriging is not suitable for phenomena that are highly nonlinear in nature. So, universal kriging and blind kriging have been suggested in other research. In universal kriging, a known mean function with unknown parameters is assumed. But blind kriging has an unknown mean function which is identified from experimental data using a Bayesian variable selection technique. One of its disadvantages is the timeliness and complexity of the Bayesian technique in comparison with simple kriging techniques.
In this research, due to the complexity of the Bayesian computation, we will combine the regression technique with blind kriging as the unknown mean function is defined by variable selection techniques being used in linear regression analysis. The residuals at the known points are calculated from the difference of the observed values and the values of the selected function. Then the residual mean at the unknown point is solved by ordinary. The regression equation combining with kriging blind is written as follows:

$$\hat{y}(x) = \overbrace{\mu_0 + \mu_1\mu_0 + \mu_2 v_2 + \cdots + \mu_m v_m}^{the\ mean\ function\ (regression)} + \overbrace{\omega(x\acute{)}\ \varphi^{-1}\ (y - V_m\ \hat{\mu}_m)}^{the\ residual\ functio(ordinary\ kriging)} \quad (3)$$

Where μ_0، μ_1 ، …، μ_m are regression coefficients, v_1 ، v_2 ، …، v_m are predictor parameters in regression, $\boldsymbol{V_m}$ is (n×n) matrix, n is the number of samples, m is the number of predictor parameters, $\hat{\mu}_m$ is (m+1) regression coefficients vector. $V_m\ \hat{\mu}_m$ is target value in n samples.

www.SID.ir

Archive of SID

The proposed method and ordinary kriging method were applied to interpole and predict IRWQI, Fe, Ni, Co, Th, Ba, As, Pb and Cr and their results were compared.

Results and Discussions

IRWQI had low values in midstream (sites 5, 6 and 7) near Sari City and its value increases by moving away from the city to show an improvement in the water quality status. This is due to the effect of anthropogenic pollutants in Sari city which are negatively associated with the water quality index. The average range of pH was from 7.6 to 8.11. pH values of water samples indicated that it was slightly alkaline. The relatively higher pH in water are probably due to the presence of pollution and eutrophication status. The mean values for Eh ranged from 142 mV to 201 mV. The water of Tajan River (excepting in site 1 near the sea) is non-saline.

The mean function for each variable (IRWQI, Fe, Ni, Co, Th, Ba, As, Pb and Cr) was created by multiple linear regression. The coefficient of determination (R^2) and adjusted R^2 were used to check regression model adequacy. t-Test and residual analysis were used in testing the regression coefficients verifying of the applicability of the regression model. The coefficient of determination (R^2), Coefficient of Variance (C.V.) and Root Mean Square Error (RMSE) were used to evaluate the results of the ordinary kriging and proposed method. The proposed method showed 8.8% improvement for Cr, 34 for Co, 56 for Pb, 62 for As, 44 for Ba and 169 for IRWQI. In addition, both methods for prediction of thorium and nickel were almost identical. While ordinary kriging performance was good in predicting iron and better than the proposed kriging, because for parameters whose correlation is strong with distance and spatial distribution, ordinary kriging method can work well in modeling them. the mean concentration of elements in the water followed the following pattern: Ca> Fe> Ba> Ni> As> Cr> Cu> Th> Co> pb. The mean concentration of elements (Co, Cr, Ni, Cu, pb and Th) demonstrate a similar pattern with a decreasing trend from the upstream to the downstream. This will strongly show a similar process and origin, while the mean concentration of Ba and As was increased at various sites from the upstream to the downstream.

Conclusions

This study was carried out to determine the water quality of Tajan River and appropriate method of estimation and interpolation of its qualitative parameters based on the combination of linear regression and blind kriging. Twenty one water quality parameters were measured at ten stations on Tajan River, an inflow to the Caspian Sea in Northern Iran.

Tajan water quality is in the middle and relatively poor class based on IRWQI. The parameters of dissolved oxygen, barium, biochemical oxygen demand and fecal coliform exceeded the guidelines of Iranian water quality standards for protection of aquatic ecosystems. Combining blind kriging and linear regression methods can be useful to improve the ordinary kriging model by reducing the error of prediction. The ordinary kriging is a powerful method for estimating parameters that are highly correlated with distance and spatial structure. One of the limitations of the proposed method is that the analyst must perform different steps in different software environments, both statistical software and GIS, and it requires an additional cost of measuring the parameters used to create the deterministic mean function (regression). Therefore, this limitation should be considered in future studies. The results of this study can be useful in formulating water quality monitoring program of Tajan river water.

.

www.SID.ir