Integrated Intelligent Research (IIR) International Journal of Data Mining Techniques and Applications
Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

A Survey on the Analysis of Dissolved Oxygen Level in Water using Data Mining Techniques

R. Arunkumar[1]and T.Velmurugan[2]
[1]Research Scholar, [2]Associate Professor,
PG and Research Department of Computer Science,
D.G.Vaishnav College, Arumbakkam, Chennai-600106, India.

Abstract: Data Mining (DM) is a powerful and a new field having various techniques to analyses the recent real world problems. In DM, environmental mining is one of the essential and interesting research areas. DM enables to collect fundamental insights and knowledge from massive volume of environmental data. The water quality is determining the condition of water in the environment. It represents the concentration and state (dissolved or particulate) of some or all the organic and inorganic material present in the water, together with certain physical characteristics of the water. The Dissolved Oxygen (DO) is one of the important aspects of water quality. The DO is the quantity of gaseous oxygen (O_2) incorporated into the water. The DO is essential for keeping the water organisms alive. The amount of DO level in the water can be detected by various methods. The data mining techniques are properly used to find DO Level in the different types of water. A number of DM methods used to analyze the DO level such as Multi-Layer Perceptron, Multivariate Linear Regression, Factor Analysis, and Feed Forward Neural Network. This survey work discusses about such type of methods, particularly used for the analysis of DO level elaborately. Finally, this research suggests the best DM method to find DO level in water by means of a comparative analysis.

Keywords: Dissolved Oxygen, Multi-Layer Perceptron, Multivariate Linear Regression, Factor Analysis, Feed Forward Neural Network.

I. INTRODUCTION

Organism and non-biotic environment in ecosystem are one of the significant factors in water environment. The situations express, affection of humans using water resource on ecological environment. It also elaborates the relation between human activity and water environment. The water has various parameters such as Total Dissolved Solids (TDS), Fluoride, Chloride, Dissolved Oxygen, and Sulfate. These primary parameters decide the quality of water. In this analysis, DO was taken from one of these parameters. DO is the amount of free or non-compound oxygen gas (O_2) available in the water. The oxygen, which is not bonded with any other element, is called non-compound oxygen or free oxygen. DO is the presence of these free O_2 molecules within the water. The DO count does not considers bonded oxygen molecule in water (H_2O). Oxygen is a primary parameter for evaluating the water quality, because of its power on the organisms living within body. Oxygen enters into the water by direct inclusion from the environment or by plant photosynthesis. Figure 1 shows how DO enters into water from atmosphere. DO is necessary for the survival of fish, invertebrates, bacteria, and underwater plants. It is also needed for decomposing organic matter. The amount of DO needed varies from creature to creature. Bottom feeders, crabs, oysters and worms needed

Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

small amount of oxygen needed (1-6 mg/L), while shallow water fish requires higher levels (4-15 mg/L). Microorganisms (Microbes) such as bacteria and fungi also need DO. The DO level is one of the major scenarios to determine the water quality. At the bottom of water microorganisms decay the organic material with the use of DO.

Microbial decomposition is a significant contributor to nutrient recycling. However in a body of water with irregular or no turnover (also known as stratification), the oxygen at lower water levels will get used up quicker, if there is an excess of decaying organic material (from dying algae and other organisms).The DO level in water is relative to atmospheric pressure, water temperature and salinity and other factors. Concentrations of DO are constantly affected by diffusion and aeration, photosynthesis, respiration and decomposition. The DO enter into the water through the air or as a plant by product. From the surrounding atmosphere, oxygen can spread slowly across the water's surface or be combined in rapidly through aeration, whether natural or man-made. Natural aeration of water can be caused by wind (creating waves), fast-flowing water, and ground water discharge. Aeration of man-made causes differs from an aquarium air pump to a hand-turned waterwheel to a large dam.

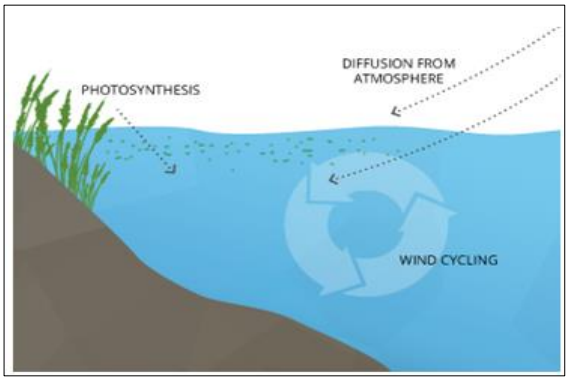

Figure 1: Dissolved Oxygen Enters into the Water

Photosynthesis also produces DO in water with the use of algae, seaweed, phytoplankton and other aquatic plants. Mostly it can occur in surface of the water but a large portion of the process happens in under water. During daylight time DO is produced as much as possible, because an aquatic photosynthesis is light-dependent and only smaller amount of DO is produced at night time. With the required light, carbon di-oxide (CO_2) is easily absorbed by water and the oxygen formed as a byproduct remains dissolved in water. The basic reaction of aquatic photosynthesis explains in Equation 1.

$$CO_2 + H_2O \rightarrow (CH_2O) + O_2 \qquad (1)$$

In stable state of water, dissolved oxygen will remain at 100% air saturation. It means water is keeping as many dissolved gas particles as it can in equilibrium. Until the water reaches complete equilibrium, it slowly absorbs oxygen and other gases from the atmosphere. This process is done by aeration. Due to the respiration of aquatic organism and microbial decomposition, DO levels come below 100% in deeper waters. Due to organic or in-organic substances water bodies are polluted. When the DO consumption rate is greater than the rate of oxygen in the water, the content of DO can be nearer to 0 and at the instant, the anaerobic bacteria breed actively increase and it get worse the water quality. Therefore, it is a great significance for environmental monitoring and the enlargement of the aquaculture industry to observe the content of DO in water.

Organization of the paper is described as follows. Section 2 discusses the different methods used in the environmental mining. Section 3 explores detection of dissolved oxygen in water. The suitable techniques in environmental mining for detecting DO in water were concluded in Section 4.

Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

II. METHODS OF ENVIRONMENTAL MINING

Data Mining is a process for determining hidden facts and understandable patterns from huge amounts of data. The goal of DM is to focus on choosing input database, develop new algorithms and distribute the statistical output for discovering potential knowledge. The novelty and comprehensibility of mining results are exhibitory in environmental domain. Environmental mining maintains the situation of the environment and keeps the ecosystem normal. The researchers used different methods to point out the state of environment and its parameters like DO, BOD (biochemical oxygen demand), wind temperature and etc., M. Kanevski et al. proposed a hybrid approach, which is a combination of MLA (Machine Learning Algorithm) and Geo-statistical based on decision-making process and probabilistic mapping [1]. This approach introduced new non-stationary SVRRSGS (support vector regression residual sequential gaussian simulation model) and NNRSGS (neural network residual sequential gaussian simulations model) for analysis and mapping spatial distributed data. The method used to removelarge-scale spatial structures and the computational cost. Chen Zhang et al. implemented a multi-hierarchy based fuzzy theory for assessing water environment security status of Luan – River [2]. They divided the water environment security in 3 aspects and assess the security level. The 3 aspects are 'water quality situation', 'ecology situation' and 'pollution situation and the result of these aspects are 'very safe', 'safe' and 'normal' respectively. The total evaluation result is "safe" in the water security. Pierre Accorsi et al. proposed a tool namely HydroQual, which is the combination of spatiotemporal data mining and visualization techniques [3]. The visual mapping focuses 3 kind of view that is geographical view, clustering view and temporal patterns view. The proposed method make easy to visual analysis of river water quality. The results are very promising.

Shoba G and Shobha G. discussed a major data mining approaches for predicting the water quality [4]. The researchers explores various data mining techniques like ANN, back propagation, MLP, decision tree etc., used to analyze the environment water quality. The research paper, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters" is proposed by HamidZare, Abyaneh. He implemented the multivariate linear regression (MLR) and ANN models for predicting the water quality [5]. The hybrid model predicted the BOD and chemical oxygen demand (COD) parameters in wastewater treatment plant. They find out the water quality using these parameters. D. M. Hamby proposed parameter sensitivity techniques, which is used to analyze the environment models [6]. The researcher discussed about various sensitivity analysis such as one-at-a-time method and direct method. He used the correlation analysis to find out the relationship between dependent and independent variables and regression analysis to complete sensitivity measures. Nitin Muttil and Kwok-Wing Chau discussed ANN and genetic programming (GP) algorithm based hybrid technique for predicting the algal bloom in Tolo Harbor, Hong Kong [7]. The result shows long-term trends of algal biomass reasonably well.

3. DETECTION OF DISSOLVED OXYGEN IN WATER

Detection of dissolved oxygen is one of the important water quality parameter, which is used to determine the water body survival. The

Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

water parameters of BOD and COD are found out after the DO detection. These parameters are used to determine organic components in water. The researchers found DO in different types of water in the environment. Campo et al. proposed fiber-optic oxygen sensor based technique, which is used to measure the DO in water [8]. Using Stern-Volmer ratio, they determine the oxygen concentration in water. The detection limit is 0.05 mg/l and the sensor showed high stability, but the one big disadvantage of this method was longer response time and a significant limitation in some applications. The following researchers discussed the DO level in rivers. The research paper, "simulation of the concentration of dissolved oxygen in river waters using artificial neural networks" is proposed by Fabiana Costa de et al [9]. They proposed ANN technique, which is used to calculate the DO in river Algeria. The simulation method observes and predicts the concentration of DO in water. The DO concentration is depending on the number of variables present in the model. The results shows average of error is 11 and predicts the available oxygen in the waters of a river was 42%. V. Simeonov et al. proposed a hybrid technique, which is combination of component analysis (CA), PCA and multi regression analysis (MRA) techniques [10]. They evaluate the dissolved oxygen, BOD, and etc., level in surface water quality in the major rivers at Northern Greece. The better information can be consumed form this kind of multivariate statistical assessment. The hybrid method assesses the large and complex databases.

Ehsan Olyaie et al. discussed a combination of three different Artificial Intelligence (AI) methods and provided a new hybrid technique, which is used to estimate the concentration of DO in Delaware River station [11]. The methods are, 1. Two types of artificial neural network namely multi linear perceptron and radial based function (RBF). 2. Linear genetic programming (LGP) and 3. Support vector machine (SVM). ANNs, LGP and SVM displayed good forecasting accuracy for low values of DO but could not maintain their accuracy for high values of DO. The result shows, the analysis of SVM method is superior to the ANNs and LGP method in DO forecasting. Samira Nemati et al. proposed data-driven techniques such as MLR, ANFIS and ANN methods to evaluate the DO in Tai Po River [12]. They measured and estimated DO values to MLR model, ANFIS model and ANN model are 0.550 and 0.681, 0.791 and 0.645 and 0.796 and 0.798 respectively. Durdu Omer Faruk proposed a hybrid technique, which is a combination of ARIMA method, feed forward and BP network with optimized conjugated training algorithm [13]. The result shows the predicted value of boron, DO and WT are 0.902, 0.893 and 0.909 respectively. Using these water parameters he predicted the water quality at Buyuk Menderes River. The effective tool used by Kunwar P. Singh et al [14]. They predict the water quality in Gomti River in India. The researchers combined an ANN model that is FFNN, BP learning algorithm, which are used to measure the DO and biochemical oxygen demand concentration levels. The hybrid method improves the understanding of river pollution trends. Stewart A. Rounds et al. proposed MLR model and FFNN model, which are used to calculate the dissolved oxygen concentration in Tualatin River at Oswego Dam [15]. The MLR model failed to capture long term patterns in the DO data. This hybrid technique provides a better performance. The calculated mean absolute error (MAE) is less than 0.9 Mg/L.

The below researchers measured the dissolved oxygen level in stream water. Ozgur KISI and Murat AY proposed the radial basis neural network (RBNN) and adaptive neuro-fuzzy inference system (ANFIS) method, which is

Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

used to measure, DO concentration in Fountain Greek stream, Gauging Station [16]. The researchers evaluated DO water quality in daily basis, using RBNN and ANFIS methods. RBNN is slightly better than ANFIS method. ANFIS method can be good for DO estimation when only temperature data is available. The multilayer perceptron and radial basis neural network techniques are combined by Murat Ay and Ozgur Kisi [17]. This hybrid method used to measure the DO concentration in upstream and downstream station on Foundation Creek, Colorado. The researchers compared the MLP and RBNN models with multi linear regression model. Researchers also found the DO concentration in underwater, coastal water and reservoir. Yu Zhao et al. developed an optical sensor, which is based on the principle of fluorescence quenching [18]. The photoelectric detection technology, wired and wireless data communication technology and numbers of instruments placed in various locations are combined and the information is sent to primary monitoring upper machine. Underwater detected information is further processed by the primary machine. It can measure the concentration of DO level in water on-line at real time. This detection process was simple and fast. The researchers Kwok-wing Chau and Nitin Muttil proposed the box plots and multivariate statistical analysis, which is used to measure the DO level in coastal water area (Tolo Harbor) in Hong Kong [19]. H. Vicente et al. proposed the ANN based model that is feed-forward neural network model (FFNN) and back propagation (BP) learning algorithm, which is based on ANN [20]. They used both models to predict the DO, wind temperature, etc in the Monte Novo Reservoir. The training set value (0.995-0.998) and the test result value (0.994-0.996) shows the good match between them.

Following researchers determine the water quality by using DO level in waters. A. H. Pejman et al. used various multivariate statistical techniques such as cluster analysis, principal component analysis (PCA) and factor analysis (FA) to evaluate the spatial and seasonal variations of surface water body quality [21]. Using the cluster analysis eight sampling stations are grouped into three clusters based on the criteria. PCA and FA methods are responsible for water quality variations. The researchers S. Shrestha and F. Kazama discussed multivariate statistical techniques, which is used for measuring the temporal or spatial complex dataset [22]. They used different kinds of analysis such as cluster analysis, PCA, FA and discriminant analysis (DA), which are used to find various water quality parameters like DO, BOD, water temperature (WT) and etc. in Fuji River. The Hierarchical cluster analysis grouped the 13 samplings sites into 3 clusters depend on the similarity of water quality characteristics that is less polluted (LP), medium polluted (MP) and highly polluted (HP) sites in the river. The total variance in water quality datasets of LP, MP, and HP areas are 73.18, 77.61 and 65.39% respectively. DA gives the best result both spatially and temporally. F. Karimipour et al. proposed the geospatial information system (GIS) and geospatial data mining [23]. The researchers used these methods to pull out the information DO and find out the importance of water quality. They research DO, pH and TDS levels in North-West Iran. The result shows, if the pollution is increased, the DO level will be decreased. Minimum amount of DO is not good for aquatic life. Increase of TDS and decrease of DO imply the pollution of water. Sundarambal Palani et al. proposed the ANN technique, which is used to predict water quality [24]. The researchers discussed about water quantity parameters such as DO, salinity, temperature, etc. are evaluated using general regression neural network (GRNN) and multilayer perceptron methods. GRNN is better for DO forecasting.

The research paper, "Use of genetic algorithms to select input variables in decision tree models for the prediction of benthic macro invertebrates" proposed by Fang Huang et al [25]. They discussed fuzzy based technique to evaluate different water parameter (DO, BOD and fluoride and etc.) to determine the water quality in Qiantang River, china. Tom D'heygere et al. implemented genetic algorithm, which is used to predict the decision trees [26]. Different sets of input variables are (DO, pH, temperature and etc.) used to compare the predictive power of decision trees for macro invertebrates in deep water. They reduced the input variables from 15 to 2-8 variables without affecting the predictive power with the use of decision trees. Xiaohu Wen et al. proposed the three layer back propagation technique to find out the water quality parameters such as DO concentration, pH, calcium, chloride etc. [27]. The ANN model is better choice to evaluate the water quality parameters. The table 1 shows that the summary of various methods and techniques proposed by different researchers for finding the DO concentration in water and the quality of the water. Most of the proposed techniques integrate two or more different techniques into one single technique called hybrid.

Table 1: Results Comparison

Paper Ref. No.	Author Name	Proposed Method	Results	
			Accuracy	Output
[8]	J.C. Campo et al.	Fiber optic oxygen sensor	High stability with longer response time	
[9]	Fabiana Costa de Araujo Schtz et al.	ANN	--	42%
[10]	V. Simeonov et al.	CA, PCA, MRA	Better information can be consumed	
[11]	Ehsan Olyaie et al.	MLP, RBF, LGP, SVM	SVM method provide best performance	
[12]	Samira Nemati et al.	MLR, ANFIS, ANN	ANN model showed the good performance	
[13]	Durdu Omer Faruk	ARIMA, FF, BP with optimized conjugated training algorithm	The hybrid technology give the better accuracy compared with ARIMA and ANN models for water quality predictions.	
[14]	Kunwar P. Singh et al.	FFNN, BP	Find out water quality using DO and BOD concentration.	
[15]	Stewart A. Rounds et al.	MLR, FFNN	The ANN model provides better performance compared to MLR model.	
[16]	Ozgur KISI and Murat AY	RBNN, ANFIS	RBNN gave good result compared to ANFIS	
[17]	Murat Ay and Ozgur Kisi	MLP, RBNN	MLP and RBNN hybrid model gives good result compared to MLR model	
[18]	Yu Zhao et al.	fluorescence quenchin	Measure the DO level in under water at real	

		g	time	
[19]	Kwok-wing Chau and Nitin Muttil	Box plots and Multivariate Statistical Analysis	Steps to control the pollution	
[20]	H. Vicente et al.	FFNN, BP Learning Algorithm	--	During Training 0.995 - 0.998 Test Result 0.994 - 0.996
[21]	A. H. Pejman et al.	PCA, FA	DO and other parameters evaluated significantly	
[22]	S. Shrestha and F. Kazama	CA, PCA, FA, DA	6 parameters 85% 7 parameters 81%	--
[23]	F. Karimipour et al.	GIS and geospatial data mining	Analyze the DO, pH and TDS	
[24]	Sundarambal Palani et al.	GRNN, MLP	R^2 0.8 to 0.9	--
[25]	Fang Huang et al.	Fuzzy based FA and UNMIX	Pollution level categorized into LP, MP, and HP by various water	
		model	quality parameters.	
[26]	Tom D'heygere et al.	Genetic algorithm, Decision tree	The data trends are identify easily	
[27]	Xiaohu Wen et al.	BP	Find out various water quality parameters such as DO, pH	

From the table1 results show the performance of the applied techniques by various researchers. The results clearly explain, hybrid techniques are a good one to evaluate the DO level in water.

4. CONCLUSION

The content of aquatic organisms in water is characterized by the Dissolved Oxygen. The concentration of DO has been traditionally used to change water quality of water systems. The modeling of water quality parameters is a very important feature in the analysis of any aquatic systems. The aquatic life can be distress if DO level is too high or too low. In this research work, various DM methods and techniques proposed by different researchers are analyzed. Many existing DM methods and some of the hybrid methods are taken into account for the analysis. From the different researcher perspectives, it is identified that most of the hybrid techniques yields best results to find the concentration of DO level in water. Hence, this work concludes that the hybrid methods perform well in analyzing DO level in water compared with the existing methods.

References

[1] Kanevski, M., Roman Parkin, Aleksey Pozdnukhov, VadimTimonin, Michel Maignan, V. Demyanov, and

StéphaneCanu,"Environmental data mining and modeling based on machine learning algorithms and geostatistics", Environmental Modelling& Software, Vol. 19(9), 2004, pp.845-855.

[2] Zhang, Chen, Wanhong Liu, Zhaoshen Su, and Liyi Wang, "Index system and method for assessing water environment security of Luan River-Tianjin water diversion project". IEEE 4th Int. Conf. on Bioinformatics and Biomedical Engineering, 2010, pp. 1-5.

[3] Accorsi, Pierre, Nathalie Lalande, MickäelFabrègue, AgnèsBraud, Pascal Poncelet, Arnaud Sallaberry, Sandra Bringay, MaguelonneTeisseire, FlavieCernesson, and Florence Le Ber,"HydroQual: visual analysis of river water quality", IEEE Conf. onVisual Analytics Science and Technology, 2014, pp. 123-132.

[4] Shoba, G., and G. Shobha, "Water Quality Prediction Using Data Mining techniques: A Survey", Int. Journal of Engineering and Computer Science, Vol.3(6), 2014, pp. 6299-6306.

[5] Abyaneh, Hamid Zare, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters", Journal of Environmental Health Science and Engineering, Vol.12(1), 2014, pp. 40.

[6] Hamby, D. M., "A review of techniques for parameter sensitivity analysis of environmental models", Environmental monitoring and assessment, Vol. 32(2), 1994, pp. 135-154.

[7] Muttil, Nitin, and Kwok-Wing Chau, "Neural network and genetic programming for modeling coastal algal blooms", Int. Journal of Environment and Pollution, Vol. 28(3-4), 2006, pp. 223-238.

[8] Campo, J. C., M. A. Perez, M. Gonzalez, and F. J. Ferrero, "An optrode type sensor to measure dissolved oxygen in water", IEEE Conf. on Instrumentation and Measurement Technology, Vol. 3, 1999, pp. 1537-1540.

[9] AraujoSchtz, Fabiana Costa, Vera Lucia Antunes de Lima, Eduardo Eyng, Adriano de Andrade Bresolin, and Fernando Schtz, "Simulation of the concentration of dissolved oxygen in river waters using artificial neural networks", IEEE 11th Int. Conf. on Natural Computation, 2015, pp. 1252-1257.

[10] Simeonov, V., J. A. Stratis, C. Samara, G. Zachariadis, D. Voutsa, A. Anthemidis, M. Sofoniou, and ThKouimtzis, "Assessment of the surface water quality in Northern Greece", Water research, 2003, pp. 4119-4124.

[11] Olyaie, Ehsan, Hamid ZareAbyaneh, and Ali DanandehMehr,"A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River", Geoscience Frontiers, Vol. 8(3), 2017, pp. 517-527.

[12] Nemati, Samira, Mohammad HasanFazelifard, ÖzlemTerzi, and Mohammad Ali Ghorbani, "Estimation of dissolved oxygen using data-driven techniques in the Tai Po River, Hong Kong", Environmental Earth Sciences, Vol. 74(5), 2015, pp. 4065-4073.

[13] Faruk, DurduÖmer, "A hybrid neural network and ARIMA model for water quality time series prediction", Engineering Applications of Artificial Intelligence, Vol. 23(4), 2010, pp. 586-594.

[14] Singh, Kunwar P., AnkitaBasant, Amrita Malik, and Gunja Jain, "Artificial neural network modeling of the river water quality-a case study", Ecological Modelling, Vol. 220(6), 2009, pp. 888-895.

Volume: 06, Issue: 02, December 2017, Page No.43-51
ISSN: 2278-2419

[15] Rounds, Stewart A, "Development of a neural network model for dissolved oxygen in the Tualatin River, Oregon", Proc. of the 2nd Federal Interagency Hydrologic Modeling Conf., 2002.

[16] Kisi, Ozgur, and Murat Ay, "Comparison of ANN and ANFIS techniques in modeling dissolved oxygen", 16th Int. Water Technology Conference, Vol. 2, 2012, pp. 141.

[17] Murat Ay and Ozgur Kisi, "Modeling of dissolved oxygen concentration using different neural network techniques in Foundation Creek, El Paso County, Colorado", Journal of Environmental Engineering, Vol. 138(6), 2011, pp. 654-662.

[18] Zhao, Yu, Li Sun, and Meng-fei Li, "The research about detection of dissolved oxygen in water based on C8051F040", IEEE Int. Conf. on Information Engineering and Computer Science, 2009, pp. 1-4.

[19] Chau, Kwok-wing, and NitinMuttil, "Data mining and multivariate statistical analysis for ecological system in coastal waters", Journal of Hydroinformatics, Vol. 9(4), 2007, pp. 305-317.

[20] Vicente, Henrique, CatarinaCouto, José Machado, AntónioAbelha, and José Neves, "Prediction of water quality parameters in a reservoir using artificial neural networks", Int. Journal of Design & Nature and Ecodynamics, Vol. 7(3), 2012, pp. 310-319.

[21] Pejman, A. H., GR NabiBidhendi, A. R. Karbassi, N. Mehrdadi, and M. EsmaeiliBidhendi, "Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques", Int. Journal of Environmental Science & Technology, 2009, pp. 467-476.

[22] Shrestha, S., and F1 Kazama, "Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan", Environmental Modelling& Software, Vol. 22, 2007, pp. 464-475.

[23] Karimipour, Farid, Mahmoud Reza Delavar, and M. Kinaie, "Water quality management using GIS data mining", Journal of Environmental informatics, Vol. 5(2), 2005, pp. 61-71.

[24] Palani, Sundarambal, Shie-YuiLiong, and PavelTkalich, "An ANN application for water quality forecasting", Marine Pollution Bulletin, Vol. 56 (9), 2008, pp. 1586-1597.

[25] Huang, Fang, Xiaoquan Wang, Liping Lou, Zhiqing Zhou, and Jiaping Wu, "Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques", Water research, Vol. 44(5), 2010, pp. 1562-1572.

[26] D'heygere, Tom, Peter LM Goethals, and Niels DePauw, "Use of genetic algorithms to select input variables in decision tree models for the prediction of benthic macroinvertebrates", Ecological Modelling, Vol. 160(3), 2003, pp. 291-300.

[27] Wen, Xiaohu, Jing Fang, MeinaDiao, and Chuanqi Zhang, "Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China", Environmental monitoring asssnd assessment, Vol. 185(5), 2013, pp. 4361-4371.