VOL. 16, NO. 15, AUGUST 2021 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

MACHINE LEARNING IMPLEMENTATIONS ON WATER QUALITY OF MANORA CHANNEL (PAKISTAN) FROM JANUARY 1996 TO DECEMBER, 2014

Sidra Ghayas[1], Junaid Sagheer Siddiquie[1], Suboohi Safdar[1] and Asif Mansoor[2]
[1]Department of Statistics, Faculty of Science, University of Karachi, Pakistan
[2]Department of Mathematics, National University of Science and Technology, Pakistan
E-Mail: sidrahghayas@gmail.com

ABSTRACT

Water quality deterioration leads to impairment of coastal lives, habitat and human health. Sewage, industrial and domestic anthropogenic pollutants deteriorates water quality when jumped untreated into the seawater. In this study, assessment of Water quality parameters at Manora channel Lyari river outfall zone (N 24-51-26, E 66-58-01) is carried out by implying Factor Analysis (FA) And Artificial Neural Network (ANN) and comparing them with National Environmental Water Quality Standards (NEQS) and other studies. Seven parameters Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Bicarbonates (BCO_3), potential Hydrogen (pH), Sulphate (SO4), Chloride (Cl) and Ammonia (NH3) is recorded for the study from January, 1996 to December, 2014. Water parameters responsible most for the water quality variation and their point sources are identified by implying FA. High factor loadings at FA identified the BOD and COD as the main contributor for the water quality deterioration as well as violating NEQS limits. BOD is predicted by implying ANN using Mean Square Error (MSE) and R square as Statistical Metrics showing promising results.VF1 (nutrient, agricultural industrial effluent and sewage effluent) and VF2 (industrial effluents) are pollutant sources resulted by FA.

Keywords: linear combinations, mean square error, over fitting, r-square, water quality.

1. INTRODUCTION

Coastal areas are suffering from multiple issues in all over the world including abasement of habitation, excessive utilization of natural endowments and tons of pollutants. (Lotze *et al*, 2006; EEA, 2006). Pakistan coastal line is 990 km long including Baluchistan and Sindh coastal areas which are mainly contaminated through a straight and treatment free ejection of industrial, drain and agricultural waste water effluents. Fortified organic substances cause Eutrophication, metallic and grease caustic substances, increased level of sea due to climatic fluctuations and sediments accumulation are the land based operations ruined 75% of the marine ecosystems in Pakistan. Approximate 8.8 billion revenue generated through fishery activities and exports, feed 80% of the population near coast in Pakistan (Saher *et al*, 2019). The Study area for this study Manora channel (Lyrai out fall zone) is from Karachi city, Karachi is a metropolitan city of Pakistan, Wastages and pollutants Mismanagement, poor dumping methods and arrangements leads Karachi city as the most contaminated city. About 158 million gallon per day of pollutants mainly municipal and industrial effluents meets the seawater at Karachi harbor via Lyari River and at Gizri via Malir River. Six thousand plus industries accumulated waste effluents and about 300 million gallon per day sewage waste is dumped into the seawater (Mashiatullah *et al*, 2016). Water quality modeling for Karachi coastal line is important to understand for present and future of Pakistan coastal Environment. Many of the studies are conducted on Pakistan's coastal line and Manora Chanel (Mashiat *et al*, 2020; Rashida *et al*, 2015; Jilani *et al*, 2018) but there is a lack of application of Machine Learning for Water Quality Assessment especially on the long span of dataset. Also statistically little number of observations are not a presentable sample size to formulate any problem solution. In this study Factor Analysis (FA) and BP ANN (Backpropogation Artificial Neural Network) is implied for the water quality assessments of Manora channel. FA using Varimax rotation are widely accepted for the contribution of uncorrelated linear combinations of water quality parameters in water quality variations as well as for the pollutants source identifications respectively (Oketola *et al*, 2013). Statistical methods are weak to deal with like non linear relationships among the water quality parameters as well as the datasets with non normal characteristics. These statistical models limitations are absorbed by ANN as it is capable to generalize, learning through past instances without any dependency on the observed models and also an expert for the non linearity in data sets (Schleiter, 1999, Taung, Yaseen 2020). Many of ANN water quality studies are recently done as Ahmed *et al* 2019; Ouma *et al*, 2020; Rajaee *et al*, 2020 to develop water quality models.

The paper is organized as following sections:

In Section: 2 Methods and Materials are described with the sub-sections:2.1) Study area, 2.2) Data descriptions, 2.3) Factor Analysis, 2.4) Feed Forward Artificial Neural Network 2.6) Data Assessments are discussed. In Section 3, Results and Discussions are described. In Section 4, conclusion is provided, In Section 5 Future work is given and in section 6. References are provided.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

2. METHODS AND MATERIALS

2.1 Study Area

Pakistan's south eastern part is Sindh coastal region, situated near Indian border sharing eastern side with sir creek along the hub river and western side with Baluchistan coast. Length of Karachi coast is about 100 km situated in the middle of Indus delta on south east and Hub River on west. (Qureshi, 2011). Manora channel composed of Karachi Harbor and Kemari Fish Harbor. 7.17 km^{2}of area of Manora channel is the connection between the Karachi port and the Arabian Sea. Manora channel neighboring the Lyari River, Malir River and Creek River destined with Massive loads of manufacturing and municipal effluents. Unlimited pollutants loads and discharges from the Lyari river effluents turns the seawater of Manora channel into Blackish grey color as it falls directly falls in seawater at "Kala pani" or "Black Water" termed as Manora channel Lyari River Outfall Zone (N 24-51-26, E 66-58-01). Map of Manora Channel is displayed in Figure-1.

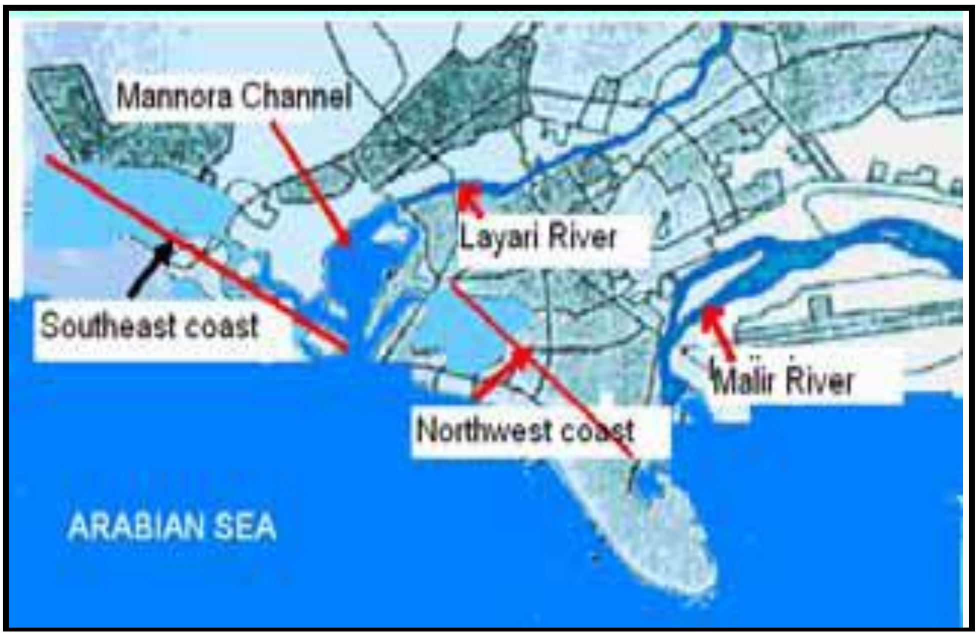

Figure-1. Map of Manora Channel.

2.2 Data Description

Physico-chemical water quality parameters are recorded from the January, 1996 to December 2014 for the assessment of water quality of Manora channel at the point where Lyari river discharges into the Arabian Sea termed as Lyari River Out Fall Zone. The parameters included in study are Bicarbonates (BCO_3), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Potential Hydrogen(pH), Ammonia(NH_3), Chloride (Cl) and Sulphate (SO_4). American public health association standard methods and procedures were used for the water sample collections in the polythene bottles and for laboratory analysis. (APHA, 1989).The descriptive statistics of the selected water quality parameters are given by Table-1 along with the literature evidence for the same sampling site.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Table-1. Descriptive Statistics.

Parameters	Mean	NEQS Limits Into Sea	Limits Criterion	Literature Evidence
BCO_3	**300.6108**	**-**	**N/A**	
BOD	263.0594	80[7]	Exceeds NEQS	Khan and khan (2001), shahzad, khan, shuakat and ahmed (2009), Aamir, Khan, Shaukat, and Kazmi, (2017), Alamgir, Fatima, Khan, Rehman and Shaukat, (2019).
COD	636.8322	400	Exceeds NEQS	Khan and khan (2001), shahzad, khan, shuakat and ahmed (2009), Alamgir, Fatima, Khan, Rehman and Shaukat, (2019).
Cl	323.0554	SC[8]	Within NEQS	
pH	7.259737	6 to 9	Within NEQS	Sirajuddin, Alamgir, Khan and Shaukat, (2016), Aamir, Khan, Shaukat and Kazmi, (2017), Alamgir, Fatima, Khan, Rehman and Shaukat, (2019).
NH_3	8.903816	40	Within NEQS	
SO_4	140.8227	SC[8]	N/A	Alamgir, Fatima, Khan, Rehman and Shaukat, (2019).

The Table-1 shows the exceeding BOD and COD levels as compare NEQS limits (NEQS 1999) in all studies as cited in table1 as well as in this study. Small sample sizes may be unable to give the true picture. This study emphasizes on the collection of datasets with long span of time to observe the water quality changes properly. All reported studies for the same sampling site in Table-1 including Khan and khan (2001), shahzad, khan, shuakat and ahmed (2009), Sirajuddin, Alamgir, Khan and Shaukat, (2016), Aamir, Khan, Shaukat, and Kazmi, (2017), Alamgir, Fatima, Khan, Rehman and Shaukat, (2019) did not cover a proper monthly observations from January to December of selected parameters as well as not for the large number of years. However this study has the limitation of spatial sites of the data but the study area sampling site is significant enough to study spatially and geographically as well the chosen physicochemical parameters are also have great significance for assessing the pollutants caused variability in the water quality. Lyari River out fall zone is the point at Manora channel where Lyari River empties into Arabian Sea with the sewage effluents from West and North Karachi and Industrial effluents from S.I.T.E area, Federal B. Area and Korangi areas. As well as 130000 tons of accumulation of dissolved solids distributed as Suspended Solids of 12,000 tons, Organic Matters of 16, 000 tons, Nitrogen Compounds of 800 tons, phosphate compounds of 90 tons jumped into the seawater through Lyari River. Seawater being dredged every year due to Silt dumps generated by loads of effluents and siltation deteriorations causes Seawater being dredged every year((Mashiatullah et.al, 2009; Mashiatullah *et al*, 2010; Beg, 1995; Abdullah *et al*, 2014). Alamigir *et al* 2019 reported the Biochemical Oxygen Demand (BOD) and Chemical Oxygen Demand (COD) at the above point from the water quality standards. Lyari river industries and municipal purposes effluents brought the immense tons of organic and in organic contaminations to the Manora channel. Eutrophication caused by the contaminants of nitrogen and phosphate. Lyari River effluents also include phenol, the Lyari river outfall is extremely contaminated part of the Karachi western back waters.

2.3 Factor Analysis (FA)

In FA, the variables with less importance are assigned to the reduction of the data and only considering those variables which are of significant importance coming after implication of PCA. These reductions of data structures are attained by using any determined methods of rotation developing new factors termed as varifactors such as varimax rotation and oblique rotation (Sahoo, 2014). FA is implied with the adequate results of Kaiser Mayer Olkin Test (KMO Test>0.5) and Barttlets Test of Sphercity (p-value<0.5)(Finkler *et al*, 2016; Oketola *et al*, 2013). Consider a set of ‘m’ observed random variables as $y_1,y_2,y_3\ldots\ldots\ldots\ldots.y_m$ and the means of these m set of observed variables are $u_1,u_2,u_3\ldots..u_m$. Also consider f_{ij} as a set of unknown constants and P_J underlying variables as i: 1,2,3…..s and j:1,2,3,….p and (p<s).

Then it can be written as mentioned in equation (1) and equation (2):

$$y_i - u_i = f_{i1}P_1 + f_{i2}P_2 + \cdots \ldots \ldots f_{SP}P_j + \epsilon_i \quad (1)$$

$$y - \mu = fP + \epsilon \quad (2)$$

Where ϵ_i are independently and identically distributed errors with zero mean and finite variance maybe vary from observation to observation.P and ϵ are independent, E(P)=0, Cov(P)=I ,which means all factors are uncorrelated. P are termed as factors and f is a loading matrix (Sahoo, 2014). Varimax rotation technique an orthogonal rotational method introduced by Kaiser (1958).All these efforts of Varimax criterion is actually for the smart interpretations of the new originated factor to

ARPN Journal of Engineering and Applied Sciences

link with less than or equal to three variables usually. These two or three variables can be grouped easily by using Varimax rotational technique. (Abdi, 2003).

2.4 Feed Forward Artificial Neural Network (FFANN)

FFANN starts with the division of datasets into training set, validation set and testing set. Training, validating and testing datasets are created to train ,assessment and judgment of the selected ANN topology performances respectively (Maier & Dandy, 2000) but these division of datasets comes up with the issues including over fitting and under fitting (Lachtermacher & Fuller (1994). These issues can be solved by using k fold Cross validation technique which also used as the stopping criterion of ANN by Keskin & Terzi (2006) choosing the most minimum Mean Cross Validation Error (MSE) for the final chosen model among other evaluated ANN model during training. In this study Gradient decent feed forward neural network is used also termed as Back Propogation neural network (BPNN). The main goal of the backpropogation is that the BPNN output and the actual output error differences must be at minimum. BPNN learning rate is a step size value introduced to the network. The variation in the weights is controlled by the learning rates. Right allocation of learning rate is significant for the backpropogation (Zhang *et al*., 1998). Error of Back propagation neural network may be described by the desired output value Di given by the back propagation neural network and the actual values A_i. The set of input Ii selected from the training combinations $\{(I_1,A_1),(I_2,A_2)\ldots\ldots,(I_n,A_n)\}$. There are identical values for each input and actual values from i=1,2,3, 4…. n.
Error is defined by Equation (3):

$$E=\frac{1}{2}\sum_{i=1}^{n}\left\|Ai-Di\right\|^2 \tag{3}$$

Reduced value of E is attained by an iterative method, steepest decent changes the Aij and Di by ΔAij and ΔDi respectively mentioned in Equation (4) and Equation (5).

$$\Delta Ai=-\frac{\partial E}{\partial Ai} \tag{4}$$

$$\Delta Di=-\frac{\partial E}{\partial Di}\in \tag{5}$$

Above Equation implies that $\Delta E<0$, hence E will decrease to local minimum.

2.5 Data Assessments

Assessment of water quality parameters at Manora channel is carried out by implying FA and ANN. Factor Analysis is implied to identify the water quality parameters which are mainly responsible for the water quality variation and their Point sources. Standardization of data is done to imply FA. The prediction of the resultant significant parameter is done by using artificial neural network. Data is divided into 80% for training and 20 % for testing. Range Standardization Method is used for Standardization of the parameters (Keskin and Terzi, 2006). Input parameters are chosen by using Pearson correlation, maximum redundancy and stepwise selection (May, Dandy and Maier, 2011). Ten fold cross validation is used to find the most accurate Water Quality Neural Network Model to encounter Over fitting and under fitting Neural Network modeling issues as well as to use the data for training and testing divisions efficiently. The minimum of Mean Square Error (MSE) of tenfold cross validation is considered to be the chosen Model with the optimum solution of training MSE and testing MSE along with the precision metrics R square for training and testing data. Statistical work is done by using *R-Studio version 1.1.456* along with the libraries included: *datasets, neuralnet, rminer, boot, plyr, psych*.

3. RESULTS AND DISCUSSIONS

Relationships of physicochemical water quality parameters and Histogram for each parameter are observed by the correlation diagram as shown in Figure-2. Pair of variables with high correlation(r>0.5), moderate correlation (0.3 < r < 0.45) and low correlation(r<0.3).

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

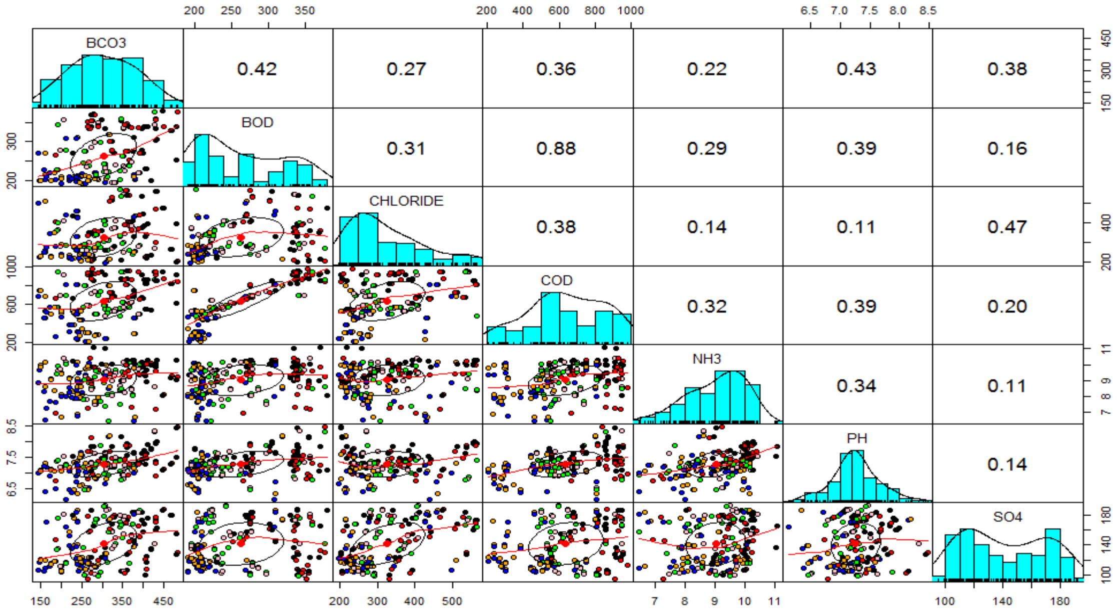

Figure-2. Correlation diagram of water quality parameters and correlation diagram of principal components.

Factor Analysis is implied to point out the sources responsible for the pollution content as well as to highlight the significant resultant factors; Varimax rotation is pursued to assess the rotated components. Factor analysis appropriateness tests are analyzed by carrying out Kaiser Mayer Olkins Test and Bartlett's Test of sphercity. KMO test value of 0.69(greater than 0.5), and significant p-value of Bartlett's Test of sphercity shows the suitability of Factor analysis. Rotated Varimax rotation results into two varifactors, varifactor1 contributed 37% of total variation in water quality and composed of BOD, COD, NH_3 pH and BCO_3 and varaifactor2 is responsible for the 23% of water quality variation and composed of SO_4,CHLORIDE. Varifactor1 accounts for the maximum variation in the water quality of Manora channel with the highest factor loadings of BOD (0.87) and COD (0.85) as shown in Table-2.

Table-2. Factor analysis output.

FACTOR ANALYSIS OUTPUT		
water quality parameters	**variFactor1**	**variFactor2**
BOD	0.87	0.15
COD	0.85	0.16
SO_4	0.10	0.86
BCO_3	0.53	0.41
Chloride	0.13	0.81
NH_3	0.53	0.14
pH	0.72	0.01
VARAIANCE PROPORTION	0.37	0.230

VARIFACTOR 1 is termed as" nutrient, sewage and agricultural industrial effluents" and vraifactor2 is termed as Industrial effluents. Highest factor loadings of BOD and COD at VF1 are the indication of domination of nutrient waste factor. Nutrient wastages from the domestic food wastages effluents and industrial food manufacturing processes & products effluents are the root cause for the high levels Biochemical Oxygen Demand and Chemical Oxygen Demand. Industrial effluents from *Paper and Pulp, Tannery, Textile (Cotton), Distillery and Brewery, Petrochemicals, Pharmaceuticals, Dairy and Sugar Mills* industries produces high amount of BOD and COD into the seawater, a very less amount of treated Sewage effluents enters to the Karachi harbor backwaters via treatment plants of Maripur and Site Areas, the rest of the Sewage Effluent enters the sea untreated causing increment in levels of BOD and Eutrophication. pH levels disturbances due to Industrial Effluents of Cement

 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Production, Petroleum Refineries, Chemical, Processing Plants of Natural Gas, Sugar Mills, Pharmaceuticals, and paper and pulp. NH_3 shows high loadings because of excessive introduction of Nitrogen to the water body by Biological Excretion and Fertilizer Making Industries. BOD is chosen for ANN modeling from the Significant parameters BOD and COD resulted by FA. For prediction of BOD minimum Mean cross validation MSE of BOD model with different combinations of hidden nodes in three hidden layers and Inputs are listed in the Table-3 along with the optimum solution of testing and training MSE and Precision metrics R^2in Table-4.

Table-3. Mean cross validated mean sqaure error for Bod artifical neural network.

Network Architecture	Number of neurons combination in three hidden layers	Mean cross validation MSE
Two Inputs (COD, pH)	555	0.101742
	111	0.05912
	1,5,10	0.0142
Two Inputs (BCO_3, COD)	555	0.01527
	111	0.05955
	1,5,10	0.01265
Three Inputs (pH, COD, BCO_3)	555	0.0305
	111	0.0128
	1,5,10	0.0314
Four Inputs (BCO_3, COD, pH,NH_3)	555	0.0218
	111	0.0689
	1,5,10	0.0344
Five Inputs (pH,BCO_3,COD,NH_3,CL)	555	0.03100
	111	0.0819
	1,5,10	0.03679
Six Inputs (pH,BCO_3,COD,NH_3,CL,SO_4)	555	0.03907
	111	0.0687
	1,5,10	0.05759

Minimum cross validated MSE is observed from Table-3 to select the précised prediction models among the different combinations of neurons in three hidden layers with the various network architecture. The statistical metrics for training and testing data for the chosen minimum cross validated MSE network architecture is reported in Table-4.

Table-4. Statistical metrics for different network architectures of training and testing data.

Network Architecture	Number of neurons combination in three hidden layers	Training Data		Testing Data	
		MSE	R^2	MSE	R^2
Two Inputs (COD, pH)	1,5,10	0.1511	0.832	0.0128	0.80
Two Inputs (BCO_3, COD)	1,5,10	0.027	0.7	0.0182	0.71
Three Inputs (pH, COD, BCO_3)	111	0.015	0.829	0.0115	0.82
Four Inputs (BCO_3, COD, pH,NH_3)	555	0.02	0.778	0.015	0.754
Five Inputs (pH,BCO_3,COD,NH_3,CL)	555	0.015	0.824	0.0126	0.8037
Six Inputs (pH,BCO_3,COD,NH_3,CL,SO_4)	555	0.013	0.855	0.0106	0.8352

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

The 3RD network architecture with the input parameters: pH, COD, BCO_3 is chosen as the best predictive model for the Biochemical Oxygen Demand with MSE of 0.015 and 0.0115 for training and testing data respectively. The R^2 of the BOD prediction model is attained 82.9% for training data and 82% for testing data. Estimated fitted line for actual train and ANN trained data shows in Figure-3 and for actual test and ANN test data shows in Figure-4 respectively.

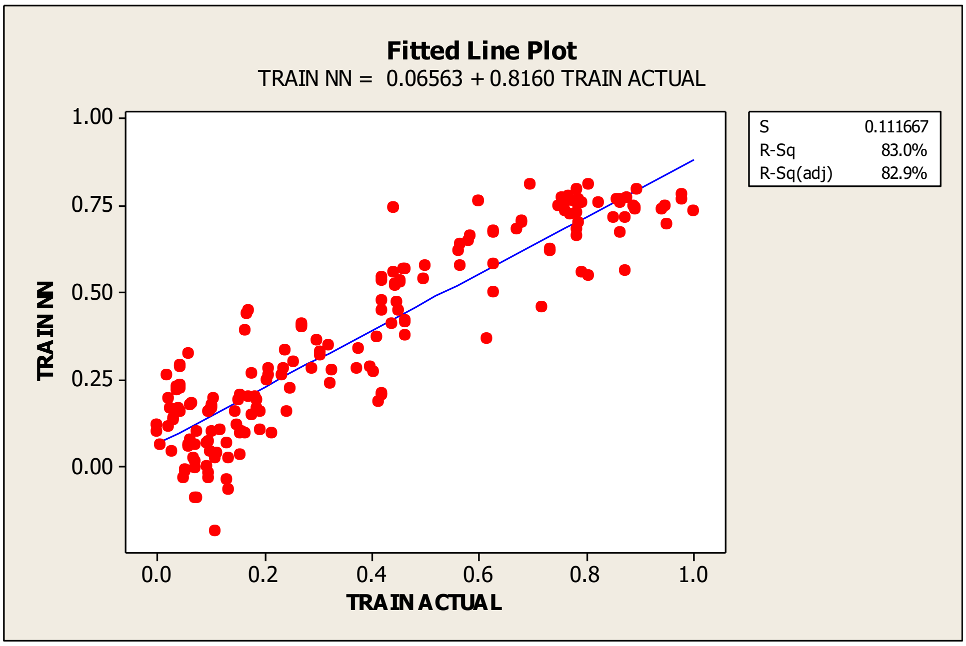

Figure-3. Regression line for BOD neural network training and BOD actual train data.

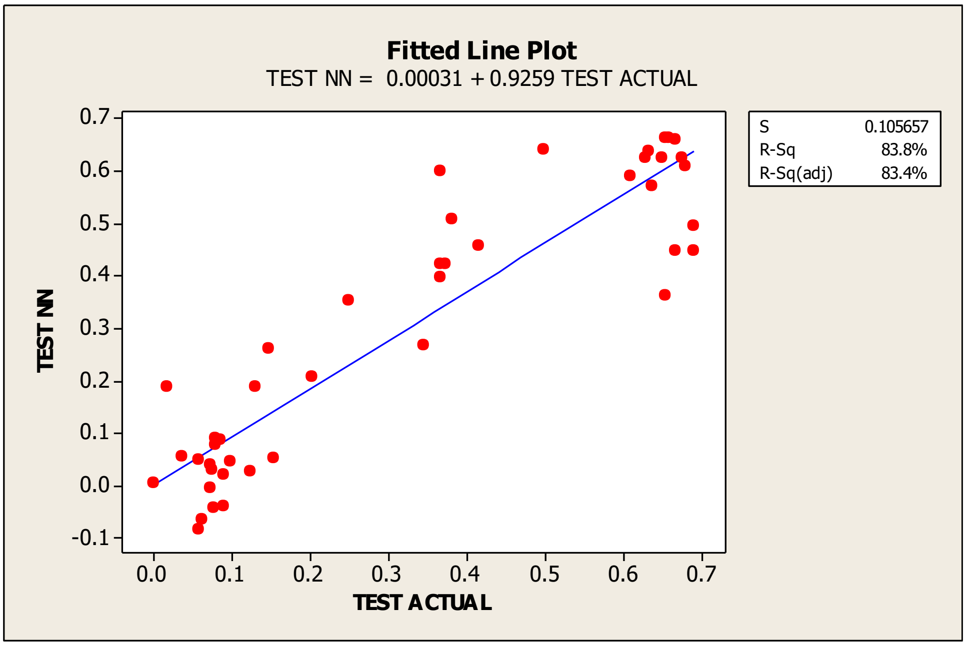

Figure-4. Regression line for BOD neural network test and BOD actual test data.

4. CONCLUSIONS

Control of all identified sources of water quality of Manora effluents: nutrient, sewage, industrial and agricultural is essential, it is observed that BOD and COD effluents prohibiting the NEQS limits mainly. The industries and the other sources which are exhibiting these uncontrollable effluents must be countered and managed. Policy makers can easily judge and evaluate the future of water quality parameters in Manora channel by using the estimated Water quality Models. Increased in BOD levels causes the depletion of oxygen which deteriorates the marine ecosystem. The significance of water quality future insight is essential for human health, mangroves, fishes and other aquatic lives and as well as for the touring and travelling economic challenges.

Future Work

In future, the other water quality parameters of the Manora channel will be predicted temporally by using ANN models.

REFERENCES

Aamir A., Khan M. A., Shaukat S. S. & Kazmi S. J. H. 2017. A quantitative appraisal of Lyari river effluent, Karachi, Pakistan. Desalination and Water Treatment. 62, 175-184.

Abdi H. 2003. Factor rotations in factor analyses. Encyclopedia for Research Methods for the Social Sciences. Sage: Thousand Oaks, CA. 792-795.

Ahmed A. N., Othman F. B., Afan H. A., Ibrahim R. K., Fai C. M., Hossain M. S. ... & Elshafie A. 2019. Machine learning methods for better water quality prediction. Journal of Hydrology. 578, 124084.

Alamgir A., Fatima N., Khan M. A., Rehman M. & Shaukat S. S. 2019. A preliminary pollution appraisal of western backwater at Karachi Coastal area. Applied Water Science. 9(7): 167.

APHA. 1989. Standard methods for the examination of water and waste water. American Public health Association, Washington, D.C

Azhar Mashiatullah N. A. & Mahmood R. 2016. Stable Isotope Techniques to Address Coastal Marine Pollution. Applied Studies of Coastal and Marine Environments. 227.

Beg, M. A. A. Ecological imbalances in the coastal areas of Pakistan and Karachi Harbour. (1995) Pakistan Journal of Marine Sciences 4(2):159-174.

Dawson C. W. & Wilby R. 1998. An artificial neural network approach to rainfall-runoff modelling. Hydrological Sciences Journal. 43(1): 47-66.

European Environment Agency. 2006. The changing faces of Europe's coastal areas (No. 6). Office for Official Publications of the European Communities.

Finkler Nícolas Reinaldo, Taison Anderson Bortolin, Jardel Cocconi, Ludmilson Abritta Mendes, and Vania Elisabete Schneider. 2016. Spatial and temporal assessment of water quality data using multivariate statistical techniques. Ciência e Natura. 38(2): 577.

Jilani, S. (2018). Present pollution profile of Karachi coastal waters. Journal of Coastal Conservation, 22(2), 325-332.

Kaiser H. F. 1958. The varimax criterion for analytic rotation in factor analysis. Psychometrika. 23, 187-200.

Karlik B. & Olgac A. V. 2011. Performance analysis of various activation functions in generalized MLP architectures of neural networks. International Journal of Artificial Intelligence and Expert Systems. 1(4): 111-122.

Keskin M. E. & Terzi Ö. 2006. Artificial neural network models of daily pan evaporation. Journal of Hydrologic Engineering. 11(1): 65-70.

Khan M. A. & Khan M. A. 2001. Impact of industrial discharge on Karachi coastal environment. Proc. Eight Stat. Sem, Karachi University. 205.

Lachtermacher G. & Fuller J. D. 1994. Backpropagation in hydrological time series forecasting. In Stochastic and statistical methods in hydrology and environmental engineering (pp. 229-242). Springer, Dordrecht.

Lotze H. K., Lenihan H. S., Bourque B. J., Bradbury R. H., Cooke R. G., Kay M. C. ... & Jackson J. B. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science. 312(5781): 1806-1809.

Maier H & D. R., andy G. C. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software. 15(1): 101-124.

Mashiatullah A., Chaudhary M. Z., Ahmad N., Javed T. & Ghaffar A. 2013. Metal pollution and ecological risk assessment in marine sediments of Karachi Coast, Pakistan. Environmental Monitoring and Assessment. 185(2): 1555-1565.

Mashiatullah A., Qureshi R. M., Ahmad N., Khalid F. & Javed T. 2009. Physico-chemical and biological water quality of Karachi coastal water. The Nucleus. 46(1-2): 53-59.

Mashiatullah A., Qureshi R. M., Javed T., Ahmad N. & Khalid F. 2020. Physico-chemical and biological water quality of Karachi coastal water. The Nucleus. 46(1-2): 53-59.

May R., Dandy G., & Maier H. 2011. Review of input variable selection methods for artificial neural networks. Artificial neural networks-methodological advances and biomedical applications. 10, 16004.

NEQS (National Environmental Quality Standards) Revised. 1999. The Gazette of Pakistan, Extraordinary, Published by Authority, Islamabad, Part II, Statutory Notification (SRO), Government of Pakistan; Environmental and Urban Affairs Division (Pakistan Environmental Protection Agency).

Oketola A. A., Adekolurejo S. M. & Osibanjo O. 2013. Water quality assessment of River Ogun using multivariate statistical techniques. Journal of Environmental protection. 4(05): 466.

Pakistan M. F. F. 2016. A Handbook on Pakistan’s Coastal and Marine Resources. MFF Pakistan: Karachi, Pakistan

Qureshi M. T. 2011. Integrated Coastal Zone Management Plan for Pakistan.

Rajaee T., Khani S. & Ravansalar, M. 2020. Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: A review. Chemometrics and Intelligent Laboratory Systems. 103978.

Rashida Q., Olufemi A., Rana M. & Rahim A. A. 2015. Seasonal variation in occurrence of heavy metals in Perna Viridis from Manora Channel of Karachi, Arabian Sea. International Journal of Marine Science. 5.

Saher N. U., Siddiqui A. S., Kanwal N., Narejo A. H., Gul A., Gondal M. A. & Abbass F. I. 2019. An Overview of Pollution Dynamics along the Pakistan Coast with Special Reference of Nutrient Pollution. Marine Ecology. 1, 136-172.

Sahoo, M. M. 2014. Analysis and modelling of surface water quality river basins. (Doctoral dissertation).

Shahzad, A., Shaukat, S. S., Kazmi, S. J. H., & Shahzad, K. Marine pollution monitoring at Lyari river outfall using satellite remote sensing. International Journal of Biology and Biotechnology (Pakistan).

Sidra Ghayas , Juniad Saghir Siddiquie and Asif Mansoor (2019).”An Application of Principal Component Analysis for the Temporal Variations in Water Quality Data of Manora Channel: Karachi” International Journal of Computer Science and Network Security. 19(5): 1-8.

Sirajuddin S., Alamgir A., Khan M. A. & Shaukat S. S. 2016. Estimation of commercially available pesticide residues from Lyari River outfall Karachi, Pakistan. Bulletin Environment Pharmacological Life Sciences. 5(4): 62-69.

T. A. Burke, J. S. Litt, M. A. Fox. 2000. Linking public health and the health of the Chesapeake Bay, Environ. Res. 82: 143-149.

T. M. Tung and Z. M. Yaseen. 2020. A survey on river water quality modelling using artificial intelligence models: 2000-2020. Journal of Hydrology. 585, 124670.

Vasseur P. & Cossu-Leguille C. 2006. Linking molecular interactions to consequent effects of persistent organic pollutants (POPs) upon populations. Chemosphere. 62(7): 1033-1042.

Vega M., Pardo R., Barrado E. and Debn L. 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res. 32, 3581-3592.

Zhang G., Patuwo B. E. & Hu M. Y. 1998. Forecasting with artificial neural networks:: The state of the art. International journal of forecasting. 14(1): 35-62.