Neural Network Implementations on the coastal water quality of Manora channel for the years 1996 to 2014.

Sidra Ghayas[1], Junaid Sagheer Siddiquie[1], Suboohi Safdar[1], Asif Mansoor[2] ,
[1] University of Karachi,
[2] National University of Science and Technology
74550
Pakistan

Abstract—- Neural Networks is an Important Part of Computational Intelligence, Systems Theory and Signal Processing and finds numerous important applications in Science and Engineering. Sea water quality contaminates due to the severe untreated domestic, sewage and industrial pollutants. Presence of ammonia in seawater causes the deterioration of coastal water in terms of diminution of oxygen levels which suffocates the marine lives, fishes and mangroves. Industrial, sewage and domestic effluents carried by Lyari River contaminate the Manora channel, Karachi. The aim of study is to make the clear and transparent step-wise use of Artificial Neural Networks for the data driven water quality parameters models of Manora channel (Lyari river outfall zone N 24-51-26, E 66-58-01), Karachi (Pakistan) as well as to compare the pollutant contaminant ratio with the national environmental quality standard limits and other sampling sites of Manora channel and southern east Karachi coast. In this study, Manora channel Physico-chemical water quality parameters are assessed by using Artificial Neural Network taking Biochemical Oxygen Demand (BOD), chemical oxygen Demand (COD), Bicarbonates, potential Hydrogen(pH) , Chloride(Cl) as input and Ammonia(NH_3)as output. Mean Square Error and R square are used for the model assessments statistical metrics. The computational work has been done by using R-studio. This is also found that Manora channel has the contaminated level of ammonia along the other sampling stations of both southern Karachi coast (N 24-47-03 E 67-08-39) as well as the other sampling site of Manora channel Karachi coast (N 24-50-15, E 66-58-01). In spite of all contamination Ammonia is found to be within National Environmental Quality Standards limits of Pakistan.

I. INTRODUCTION

Neural Networks is an Important Part of Computational Intelligence, Systems Theory and Signal Processing and finds numerous important applications in Science and Engineering [53], [54], [55], [56]. Sea water is the target of many human deteriorations like domestic, sewage and industrial effluents disposed finally in seawater. The water quality management is the most intense task to regulate the touring and business activities of traversing, shipping and fishing. Water quality modeling is an idea for the policy makers to follow the solutions for certain water quality hazardous situations. Water quality models are also evaluated by the statistical methods but their limitations are observed due to the various spatial and temporal variations of the water ecosystems. The non linearity among the relationships of water quality parameters is the potential reason for the non normal properties occurring by datasets.

Artificial Neural Network is that proposed methodology which absorbs and encounter all of the limitations of statistical and mathematical techniques by its capacity to generalize, learning through historical examples seeking no dependency on the seen models and to deal with non linear data sets.[1],[2] Artificial neural networks (ANN) are very famous for the parameters modeling having undefined relationships among the parameters of the considered problem as it is a data driven technique. ANN are composed of a number of neurons, hidden layers, algorithms and activation functions. Feed forward neural network is a very popular neural architect which is

developed in the early stages of the development of the neural network models. Feed forward artificial neural network is a system which starts from input towards output, moving forward. In feed forward neural network,

  1 Inputs entered the hidden layer.
  2 The hidden layer passes the weighted neuron and inputs to the activation function, generally known as transfer function or input activation function. Thus, derived a yield of the hidden layer.
  3 The new weight of hidden layer neurons is passed through the output layer which possesses the activation function. This output transfer function transforms the neuron into desired outputs in the output layer.

Back propagation, scale conjugate gradient, conjugate gradient, resilient back propagation algorithms, Marquardt algorithms are well known for the artificial neural networks algorithms. [3], [4], [5], [6], [7] some improvements and modifications to these algorithms are also made for efficient implementation of artificial neural network specially for back propagation algorithm [8], [9], [10], [11]. Numbers of neurons are also assessed by rules of thumb but there is no definite criterion for the selection of number of nodes and hidden layers. Weights and biases are also assigned to the network to get connected with inputs in training processes, usually a small amount of weight and bias is given to the network but there is no proper method for the allocation of weights. A simple rule of thumb maybe fulfills the purpose, that is random numbers must lie in the boundary between -2/I to 2/I, where I are inputs provided to the artificial neural network in any given node [12].

Many of the recent water quality studies are done by implying Artificial Neural Networks such as [13], [14], [15], [16] ,[17] ,[18] ,[19] ,[20] ,[21], [22] etc. In this study the Physico Chemical Water Quality of Manora Channel is assessed by the assessment of water quality parameter Ammonia. Ammonia (**NH₃**) is a unionized nutrient of Hydrogen and nitrogen. Ammonia emitted through sewage pollutants and also industrial pollutants of fertilizers majorly. Increased Ammonia causes severe issues to the fishes and other marine lives as well as disturbs the photosynthetic processes. Increment of algae is the eventual cause of increased Ammonia discharged by plants and animals and bring in nitrogen to the seawater causing diminution of oxygen levels in seawater [23]. Physico Chemical Water Quality of Ammonia is assessed by implying Resilient Back Propagation and Backpropagation Water Quality Artificial Neural Network. Statistical analysis is done by using R-Studio. Libraries of R-Studio used are: datasets, nueralnet, rminer, boot, plays.

The paper is organized in the following pattern:
In Section: 2 methods and materials is provided which is composed of these subsections: A) study area, B) Back propagation C) Resilient backpropagation D) Activation functions E) Standardization F) Cross validation G) Statistical analysis is provided. In Section: 3 Results on the evaluated technique are given providing the related tables. In Section: 4 conclusions are given with future work, In Section: 5 References are given.

II. METHOD AND MATERIAL

The selected water quality parameters of Manora channel where the Lyari river discharges in the Arabian sea from the year 1996 to 2014 are Bicarbonates(BCO_3),Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD),Potential Hydrogen(pH),Ammonia(NH_3), Chloride (Cl) and Sulphate (SO_4), these parameters are collected by implementing APHA (American Public Health Association) guidelines (APHA,1989)[24]. The descriptive statistics of the selected water quality parameters are given by table (1):

Table 1 DESCRIPTIVE STATISTICS

Parameters	Max	Min	Mean	Standard Deviation
BCO_3	483.	143.0	300.6	83.0032
BOD	380	191.8	263.0	56.915
COD	981.	207.2	636.8	199.164
Ph	8.47	6.13	7.529	0.40849
NH_3	11.0	6.07	8.903	1.03539
Cl	567	189.3	323.0	100.239
SO_4	192	95.09	140.8	28.44617

A. Study Area

Manora channel has been divided into two broad areas which are Karachi harbor and Kamari fish harbor, it is a navigational waterway having an area about 7.17 km^2, the importance of the Manora channel is evident from the fact that it connects the Karachi port to the Arabian sea Lyari river, Malir river and Creek river are the hub of effluents which surrounds the Manora channel. Effluents are directed to jump into the sea directly passing through these rivers. The Lyari River falls in the seawater at Manora channel. Major pollutants which are deposited in the Lyari, Malir and Creek rivers are industrial and sewage effluents. "Kala pani" or "Black Water" is another name for the Manora channel. The continual waste discharge of Lyari River turns the seawater in blackish grey color. Deterioration of Manora Channel is mainly responsible by the Lyari river effluents. Lyari river posses the sewage effluents from the discharges of West and North Karachi, Industrial effluents from S.I.T.E area, Fedral B.Area and Korangi areas. Sindh Industrial Trading Estate covers an area about 4000 Acres contributes about 80% of the effluents in Lyari river by manufacturing Basic Metals, Machinery, Food And Beverages, Tobacco , Paper ,Textile, Plastic , Rubber Products, Chemical, Non Metallic Minerals, Coal and Petroleum Products. Accumulation of 130000 tons of dissolved solid out of which Suspended Solids (12,000 tons), Organic Matters (16,000 tons), Nitrogen Compounds (800

tons), phosphate compounds (90 tons) into the sea via Lyari river. Silt dumps at the sea water are generated by loads of effluents and siltation deteriorations causing seawater being dredged every year. [25, 26, 27,28]. Biochemical oxygen demand and Chemical oxygen demand is evidently at the exceeding point from the water quality standards as reported by Alamigir et.al 2019[29], it has also been mentioned that the Lyari river effluents from industries and municipal purposes introduced the massive loads of organic and in organic contaminations to the channel. Eutrophication is a noticeable issue due to the nitrogen and phosphate contaminants. Phenol is also found in the Lyari river effluents, the Lyari river outfall is reported as the most contaminated area of the Karachi western backwaters.

B. Back propagation

Back propagation is the gradient decent feed forward neural network which is mostly used [30]. The simple method adopted by the back propagation is to minimize the error between the output of the back propagation neural network and the actual output. A step size is introduced to the gradient decent neural network termed as the learning rate. The significance of learning rate is that it allocates the variation in the weights. Back propagation algorithm is a gradual convergent and low efficient. Learning rate choices are very important to back propagation, low value learning rate slows down the learning processes where as high learning rates cause repetition of network in the weight domains. [31]. Low convergence and repetitive network limitation of back propagation neural network can be dealt with introducing the momentum value. Momentum allows the variation in weights in such a way that the direction of the last updated weight is almost similar to the new weight, which solves the issue of repetition of network with higher learning rates. [32]. the selection of learning rate and momentum is nominated through the experiments by many researchers. The values may take between 0 and1. Sharda and Patil., 1992[33]; Tang et.al.,1991[34]; Tang and Fishwick ,1993[35] worked on the network results with the variation in learning rate and momentum choices. Learning rate is also to be set for the convergence of network avoiding the local minima

issues[36,37].small Learning rates brings local minima situations and learning is not done where as large learning rates bring weak answers one after another. Deviation in the back propagation algorithm is the only solution for the local minima issues. As there will no progress with the changes in weights with the previous learning parameters, once the network has been developed, the only way out is to increase the learning parameter which gives poor results .Dai and MacBeth,1997[38] worked on the variations in back propagation network with suitable weight adjustments during training.

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(1):

$$E=\frac{1}{2}\sum_{i=1}^{n} \lvert Ai - Di\rvert^2 \qquad (1)$$

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 (2) and Equation(3).

$$\Delta Ai = -\frac{\partial E}{\partial Ai} \qquad (2)$$

$$\Delta Di = -\frac{\partial E}{\partial Di} \in \qquad (3)$$

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

C. Resilient back propagation

Rprop is the abbreviation for "Resilient back propagation". Rprop is a batch learning for supervised neural networks based on the key rule of discarding the effect of magnitude of partial derivative on the updation of weights. Sign of partial derivatives is potential in this case as it provides the direction of the updation of weights.

Magnitude by which the changes in weight are observed is known as update value as given by Equation (4):

$$\Delta wij^{(t)}=\begin{cases} \Delta_{ij,}^{(t)} & \text{if } \frac{i\,\partial E^{(t)}}{\partial wij} > 0 \\ +\Delta_{ij,}^{(t)} & \text{if } \frac{i\,\partial E^{(t)}}{\partial wij} < 0 \\ 0, & \text{else} \end{cases} \qquad (4)$$

Where $\frac{i\,\partial E^{(t)}}{\partial wij}$ is batch learning total gradient information [36].

Rprop as compare to back propropagation searches the weight update value which can modify the weight up to the extent at which the error function's local minimum is achieved. To search root, derivative of error function with respect to weights (dE/dW) is taken

D. Significance of Rprop

Rprop is the fastest convergent algorithm but it bears some little of computational expenditures[36] .Weights are observed to be in the alternative direction for the improvements in

weights for the purpose of achievement of local minima, it means that weights are found to be increased if the partial gradient is negative and vice versa. A distinct property of Rprop is that a unique value for the learning rate during a training procedure is observed as well as it modifies the learning rate for each weight during training. In this way, Rprop is a solution towards the problem of determining an appropriate learning rate to define for network completely. use of signs for the weight updation also assured that there is an identical effect of rate of learning over the whole network [37].The magnitude of passage between the weights and the output layer of artificial neural network also influences the probability of weight to learn and adjust compatibly but there is no such consequences of this distance between weights and output layer as rprop only learn the order of signs throughout the training procedure of entire network. [36].

Weight adjustments are given by Equation (5) :

$$w_k^{(t+1)} = w_k^{(t)} - \eta_k^{(t)} \cdot \text{sign}\left(\frac{\partial E^{(t)}}{\partial w_k^{(t)}}\right) \qquad (5)$$

In order to attain the high convergent pace in shallow regions, learning rates are whether increased or decreased. The decision of increment or decrement of learning rates are clearly depends upon the consistent appearance of the sign of the partial derivates. If the sign of partial derivatives constantly remains same, it means that learning rate is required to be increased. However if the sign of partial derivatives continuously differs, it means that learning rate must have to be decreased as this continuous change of sign is actually occur due to the large learning rates which is the hurdle in the attainment of local minima. To avoid the local minima attainment frequently, a weight back tracking concept is used in which the last computed iteration is actually replaced by adding a tiny magnitude to the weight [39].

E. Activation functions

Step size of weight space is dependent of transfer function, as the derivative of transfer function is influence the updating of weights directly [40]. transfer functions are also termed as activation function, logistic transfer functions and hyperbolic transfer functions are the most popular and commonly used transfer functions. Many studies have been done to investigate the comparison among the various activation functions. Moody and Yarvin, (1992)[41] worked on the non linear and noisy datasets with various levels by comparing the logistic, polynomial, rational function and Fourier series activation functions. Results obtained shows clearly that sigmoid function does not perform well when the noise and non linearity occurs in the data. There are studies which are found to be in favor of hyperbolic tangent like Kalman and Kwasny, 1992[42] ; Maier and Dandy,1998a[40]; Kaastra and Boyd,1995[43] and Karunanithi et al.,1994[44] shows that in hidden layers sigmoid activation function and in output layer linear transfer function is used to anticipate outside the domain

of training datasets. Karlik and Olgac (2011) [45] discussed the significance of activation function for the improvement in performance of the neural network and found that the hyperbolic tangent function performed well.

a. Hyperbolic tangent functions

Hyperbolic tangent function termed as bipolar sigmoid function as well, it ranges from -1 to 1 and occurs as S in shape. Significant characteristic of the hyperbolic tangent function is that it converges at very high rate. If the data in the hidden layers behaved non-linearly, it is very much appropriate to use hyperbolic tangent function. [47]

$$F_t(I_t) = \frac{2}{1+e^{-2it}} - 1 \qquad (6)$$

b. Logistic sigmoid function

Logistic activation functions are well known to introduce non linearity in the datasets, range of logistic sigmoid function is between 0 to 1, logistic function forms S shape. [46]

Logistic sigmoid function is expressed by:

$$F(I_t) = \frac{1}{1+e^{-it}} \qquad (7)$$

c. Linear function

Identity function is another name for linear transfer function, Variation in input variable is easily arranged by the linear transfer function as it works as a step function to generate bipolar or binary outputs. [46]

It is given by equation (8):

$$F_t(I_t) = I_t \qquad (8)$$

F. Division of datasets

Mostly ANN models are partitioned into two datasets: training partitions and testing partitions. To increase the precision predictability of ANN models, various methods are adapted such as Masters, (1995) [48] used holdout method to increase the usage of given data. Lacthermacher and Fuller, (1994) [48] considered a representative test partition having the same characteristics as of training partitions. Maier and Dandy(1998a) [40]; Khotanzad et al. (1997) [51] uses a test partition from the datasets to achieve number of epochs from the trail training process which also determines when training process should be stopped. Afterwards, this test partition is included to the training partition to train the network fully. The division of data sets for neural network into training set, validation set and testing set is well known. Training datasets are created to train various ANN topologies, validating datasets are actually for the assessment of training datasets in order to avoid over fitting as well as to asses which of the ANN topology is the smartest for the solution of given problem. Testing datasets are the independent datasets which are established to judge the selected ANN topology performance. [47].

a. Limitations for the division of data sets

These three datasets are encountered with few restrictions as discussed by Lachtermacher and Fuller [50]. First, the most evident issue for three independent partitions of the datasets is

the little accessibility of the data. The second issue comes up with the division of datasets is the method according to which these partitions are made , the method of data division can seriously influence the outcome of ANN topology performances. The third issue is the use of test datasets in order to stop the training is not a wise choice as the noise can introduce to the system at any point from the time series and the system remains unknown about the fact that at which point the system starts to learn this noise.[48]

G. Cross validation

Cross validation with the smart method of training across the datasets is the solution for the stated problems by Lachtermacher and Fuller (1994) in section F of this paper about the limitations for the divisions of datasets [50]. This technique is the most assessable criteria for stopping the training of the ANN models in order to select the best ANN model with respect to their performances. Dawson and wilby (2001)[48] ;Keskin and Terzi,(2006) [52] chooses the cross validation techniques for the stopping criteria of ANN, they evaluate many models and finally chooses the ANN model which has the most minimum mean Cross validation error (MSE).

Cross validation technique is the portioning of datasets into F identical magnitude fragments. These fragments are then used for the training and validating purposes such that the F-1 fragments data sets are trained for all the ANN models and validated for the left over fragment of the independent data. This process is recur F times so that there are F models which are required to be trained as well as validate for all the established ANN models. The purpose of this method is to have assurance to use the every data fragment only one time for the validation purpose. This is a simple technique which recollects the all of the datasets into the validation set which actually is the result of recombination of all the validation fragments. 5 and 10 are frequent values which are used for F [53]. The process of cross validation occurs actually in search of the minimum sum of squares, once the most minimized sum of squares is attained, the repetitive training procedure is allowed to stop and the outcome is compared with the observed data. As mentioned above, this training and validation process is observed to be repeated M times that is actually occurs single instance for every training partition. Cross validation error is given by the determination of the mean of remaining F fragments MSE. [52]

H. Statistical Analysis

Physico-Chemical Water Quality Model is developed by implementing Artificial Neural Network using Hyperbolic Tangent activation function for hidden layers and Linear Activation Function for output layer is used where BOD, COD, BCO_3, pH, Cl and SO_4 are taken as input variables and NH_3 is taken as output variable. Standardization of the parameters is done by using Range Standardization Method.

Datasets are split into two portions: 2/3rd of the data is used for training datasets and 1/3rd for the testing datasets. Neural network is applied by using *R-Studio version 1.1.456* along with the libraries included: *datasets, neuralnet, rminer, boot, plyr*. Resilient back prorogation algorithm and Back propagation algorithm is used as the learning algorithm. Grid search of Thresholds, weights and hidden nodes in three hidden layers is done for modeling the NH_3 Water Quality Neural Network Model. Over fitting and under fitting Neural Network modeling issues are countered by using k fold cross validation. Ten fold cross validation method is used to find the most accurate NH_3 Water Quality Neural Network Model. The minimum of mean square error of tenfold cross validation is considered to be the chosen Model, unless or until the repetitive iterations of training and testing NH_3 Water Quality Neural Network Model attained the training Mean Square Error very close to the Cross Validation Mean Square Error and optimum and close R squares of training and testing NH_3 Water Quality Neural Network Model.

III. RESULTS

Mean cross validation Mean Square Error is observed with three hidden layers for different hidden nodes combinations for the two ANN algorithms (Resillient Backpropagation Algorithm, Backpropagation Algorithm) as listed in the Table 2.

Table 2 MEAN CROSS VALIDATION MSE

HIDDEN NODES	MEAN CROSS VALIDATED MEAN SQAURE ERROR	
	RESILLIENT BACKPROPAGATION ALGORITHM	BACKPROPAGATION ALGORITHM
1,1,1	0.05001927	0.05176419
1,5,10	0.04206558	0.0575488
1,10,10	0.04234572	-
5,5,5	0.04679888	0.05526362
10,10,10	0.1534309	-

The most minimum Mean Cross Validated Mean Square Error is observed with the ANN architectures (1,10,10) and (1,1,1) hidden nodes with three hidden layers for NH_3 water quality model for Resilient Back propagation algorithm and for the Back propagation algorithm respectively shown in Table 2. To attain the acquired accuracy and precession with selected model, the selected model is to be iterated ten times to achieve the best accuracy and precision with the selected model.

The NH_3 Resilient Back propagation Water quality Neural Network Model is done with ten times repetitions reported in table (3)

Table 3(NH_3 RESILIENT BACKPROPAGATION WATER QUALITY NEURAL NETWORK MODEL ITERATIONS)

	MSE Cross	R²	MSE	R² test	MSE

	validation	train	train		test
1	0.072853	0.987	0.044	3.1842	0.19
2	0.032823	0.859	0.038	0.965	0.05
3	0.029989	0.868	0.039	0.9831	0.06
4	**0.03664**	0.859	**0.038**	0.8752	0.05
5	0.0348	0.842	0.037	0.834	0.05
6	0.0453	0.982	0.044	1.038	0.06
7	0.04103	0.890	0.040	0.8119	0.05
8	0.03406	0.955	0.043	0.8642	0.05
9	0.05053	0.911	0.041	1.1847	0.06
10	0.04255	1.001	0.045	1.0455	0.06

The NH_3 Resilient Back propagation Water quality Neural Network Model is attained at 4th iteration as MSE cross validation (0.03644) is found to be very close to MSE Train (0.038), also MSE Test (0.05) is not very far away from MSE cross validation along with the Rsqaure train (85.9%) and Rsqaure test (87.52%) as shown in the Fig.3 also. In Fig.3 MSE Train, MSE Test and MSE Cross validated lines are almost following the same pattern.
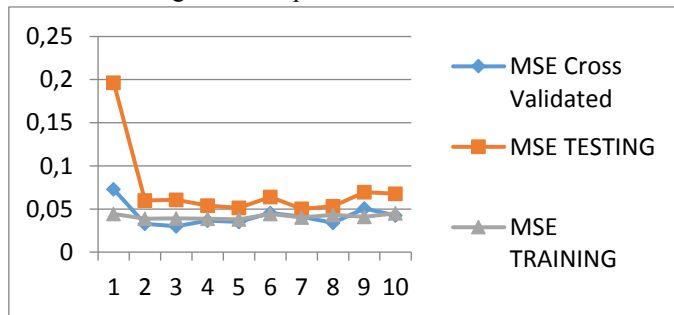

Fig. 3 Mean Square Errors for NH_3 Resilient Backpropagation Water Quality Neural Network Model Iterations

The NH_3 Back propagation Water quality Neural Network Model is done with ten times repetitions reported in table (4)

Table 4(NH_3BACKPROPAGATION WATER QUALITY NEURAL NETWORK MODEL ITERATIONS)

	MSE Cross validation	R^2 train	MSE train	R^2 test	MSE test
1	0.07093	0.9806	0.0441	1.09246	0.057
2	0.060532	1.0337	0.0465	0.99412	0.061
3	**0.047754**	0.9864	**0.0444**	1.0076	0.062
4	0.041015	1.0415	0.0469	0.9510	0.058
5	0.043639	1.0007	0.0450	1.0083	0.062
6	0.043029	1.0003	0.0455	1.0096	0.062
7	0.046001	1.0414	0.0469	1.03808	0.064
8	0.070127	1.01	0.0454	1.0005	0.061
9	0.057264	1.0004	0.0450	1.00796	0.062
10	0.037352	1.0003	0.0450	1.0086	0.062

The NH_3 Back propagation Water quality Neural Network Model iterations shows that at third iteration the MSE Cross

validated(0.047) is closed to the MSE Train(0.044) but Rsqaure metric is not attained for the NH_3 Back propagation Water quality Neural Network Model as R square Train(98.6%) and R square Test (100%) showing over fitting of the Model. The Fig.4 shows that the MSE cross validation pattern is not following the MSE train pattern. MSE cross validated line is meeting MSE Train line at some point but their requirement is not satisfying of following each other smoothly along with a close MSE test line at some point at least.
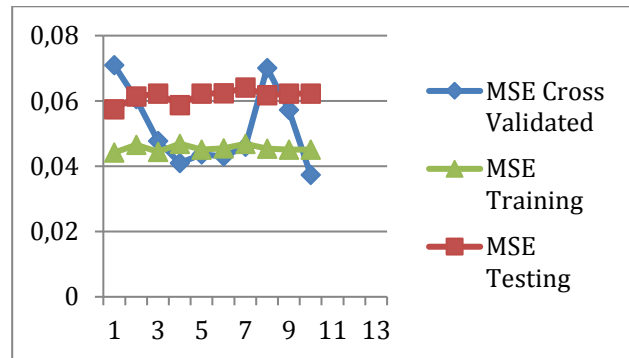

Fig. 4 Mean Square Errors for NH_3 back propagation Water quality Neural Network Model Iteration

The table 5 clearly indicates that the Lyari River out fall zone at Manora channel Ammonia is found to be within NEQS limits (the revised NEQS limits for the municipal and industrial liquid effluents which may enter to the sea according to the PEPA guidelines which were first given in 1997 and lately revised in1999). On average ammonia for this study is found (8.903816 mg/L) but this data sets are from January 1996 to December, 2014 taken from Lyari river out fall zone, with comparison of a study conducted in 2018(Ali et.al, 2018) worked on the other sampling stations (Ibrahim Haideri and Karachi Port Trust) near the Manora channel shows the ammonia, nitrogen levels on the average of 22.48 ppm. It shows the toxicity of ammonia levels in Lyari river out fall zone in 2014 must be increased in 2018 as other manaora channel site Ibrahim Haideri (N 24-47-03 E 67-08-39)and Karachi Port Trust(N 24-50-15, E 66-58-01) sampling stations ammonia contaminants as shown in table(5).

Table 5 MEAN OF AMMONIA AND NEQS LIMITS

	Lyari River Outfall Zone (mg/L)	Karachi Port Trust (ppm)	Ibarahim Hyderi (ppm)	NEQS Limits Into Sea	Limits Criterion
NH_3	8.9038	25.2	19.7	40	Within NEQS

IV. CONCLUSION

Neural Networks is an Important Part of Computational Intelligence, Systems Theory and Signal Processing and finds numerous important applications in Science and Engineering [53], [54], [55], [56]. Mean cross validated MSE is found to be minimum with the architecture of 6 inputs, 3 hidden layers with (1, 5 and 10) hidden nodes in the three hidden layers respectively and one output of NH_3 Resilient Back Propagation Water Quality Neural Network Model with R square metrics which shows fitness of Physico-Chemical Water Quality Neural Network Model is about 85.944 % for the training data as well as 87.521 % for the testing data set. Hence NH_3 Resilient Back Propagation Water Quality Neural Network Model is the best assessment Water Quality Neural Network Model for Ammonia. Manora channel is found to be contaminated sampling station of Karachi coast along other sampling stations of Karachi coast. Policy makers may use these water quality prediction models to estimate and forecast the Ammonia levels of years ahead of 2014. Predictive Coastal Modelling of Ammonia may help in dealing with the control increment in Ammonia levels causing algal blooms and stop turning these blooms into red tides. Algal blooms are the cause of depletion of oxygen for marine lives mainly. In future, the assessment of Ammonia and DO can be done in different seasons of Karachi coasts by implying Principal component analysis-Artificial neural network to assess the water quality of the marine habitats for marine lives.

References

[1] I. M.Schleiter, D. Borchardt, D., R. Wagner, T. Dapper, K. D.Schmidt, H. H.Schmidt, H. Werner. Modelling water quality, bioindication and population dynamics in lotic ecosystems using neural networks. *Ecological Modelling*, 120(2-3), 271-286 (1999).

[2] T. M.Tung and Z. M.Yaseen,. A survey on river water quality modelling using artificial intelligence models: 2000–2020. *Journal of Hydrology*, *585*, 124670. (2020).

[3] Y. LeCun, B. Boser, J. S.Denker, D. Henderson, R. E.Howard, W.Hubbard, and L. D.Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551 (1989).

[4] T. Steihaug. The conjugate gradient method and trust regions in large scale optimization. SIAM *Journal on Numerical Analysis*, 20(3), 626-637 (1983).

[5] S. Kamalasadan, D. Thukaram, and A. K.Srivastava, (2009). A new intelligent algorithm for online voltage stability assessment and monitoring. *International Journal of Electrical Power & Energy Systems*, 31(2-3), 100-110. (2009).

[6] M. T.Hagan and M. B.Menhaj. Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6), 989-993 (1994).

[7] M. F.Møller. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, 6(4), 525-533. (1993).

[8] M. C.Mozer. A focused backpropagation algorithm for temporal. Backpropagation: Theory, architectures, and applications, 137.(1995).

[9] J. Leonard and M. A.Kramer. Improvement of the backpropagation algorithm for training neural networks. *Computers & Chemical Engineering*, 14(3), 337-341 (1990).

[10] F. M.Silva and L. B.Almeida. Acceleration techniques for the backpropagation algorithm. In European Association for Signal Processing Workshop (pp. 110-119). Springer, Berlin, Heidelberg. (1990, February).

[11] X. H.Yu, G. A.Chen, and S. X.Cheng. Dynamic learning rate optimization of the backpropagation algorithm. IEEE Transactions on Neural Networks, 6(3), 669-677. (1995).

[12] C.W Dawson and R. Wilby. An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal,* 43(1), 47-66 (1998).

[13] Y. O.Ouma, C. O.Okuku and E. N.Njau. Use of Artificial Neural Networks and Multiple Linear Regression Model for the Prediction of Dissolved Oxygen in Rivers: Case Study of Hydrographic Basin of River Nyando, Kenya. *Complexity.*(2020).

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

[15] L. J.Stamenković, S. K..Mrazovac and U. V.Presburger. Prediction of nitrate concentration in Danube river water by using artificial neural networks. *Water Supply*. (2020).

[16] D Antanasijević,., V Pocajt,, A Perić-Grujić, and M Ristić. Multilevel split of high-dimensional water quality data using artificial neural networks for the prediction of dissolved oxygen in the Danube River. *Neural Computing and Applications*, 1-10.(2019).

[17] Y. Chen., X. Fang, L. Yang, Y. Liu, C. Gong and Y. Di. Artificial Neural Networks in the Prediction and Assessment for Water Quality: A Review. *In Journal of Physics: Conference Series* (Vol. 1237, No. 4, p. 042051). IOP Publishing. (2019, June).

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

[19] A. Csábrági, S. Molnár, P. Tanos, J. Kovács, M. Molnár, I. Szabó and I. G.Hatvani, I. G. (2019). Estimation of dissolved oxygen in riverine ecosystems: Comparison of differently optimized neural networks. *Ecological Engineering*, 138, 298-309 (2019).

[20] A. Š.Tomić, D. Antanasijević, M. Ristić, A.. Perić-Grujić and V. Pocajt. Application of experimental design for the optimization of artificial neural network-based water quality model: a case study of dissolved oxygen prediction. *Environmental Science and Pollution Research*, 25(10), 9360-9370 (2018).

[21] A. H.Haghiabi, A. H.Nasrolahi, and A. Parsaie. Water quality prediction using machine learning methods. *Water Quality Research Journal*, 53(1), 3-13 (2018).

[22] V. Vijayashanthar, J. Qiao, Z. Zhu, P. Entwistle and G. Yu. Modeling fecal indicator bacteria in urban waterways using artificial neural networks. *Journal of Environmental Engineering*, 144(6), 05018003 (2018).

[23] F. Ali, A. Hussain, Y. Nergis, M. Sharif, S. F.Shah and A. A.Khan. Impact of Eutrophication on Shallow Marine Water near Karachi Coast, Pakistan. *Proceedings of the Pakistan Academy of Sciences: B. Life and Environmental Sciences*, 55(2), 27-36.(2018).

[24] APHA (1989), "Standard methods for the examination of water and waste water," American Public health Association, Washington, D.C

[25] A. Mashiatullah, T. Javed, M. Z.Chaudhary, M. Fazil, and R. M. Qureshi. Stable carbon isotope ratio (δ13 c) of shallow marine bottom sediment as evidence of pollution in manora channel, karachi, pakistan. *Nucleus*, 47(3):233-238 (2010).

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

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

[28] A .M.Harun, R Dambul, M Mohamed, and S Harun. Spatial and seasonal variations in surface water quality of the Lower Kinabatangan River Catchment, Sabah, Malaysia. *Journal of Tropical Biology and Conservation* 11: 117-131. (2014).

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

[30] S. Haykin. Neural Networks: A Comprehensive Foundation, Mac. *Milan, New York*. (1994).

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

[32] D. E.Rumelhart, G. E.Hinton, and R. J.Williams. Learning representations by back-propagating errors. *nature*, 323(6088), 533-536. (1986).

[33] R. Sharda and R. B.Patil. Connectionist approach to time series prediction: an empirical test. *Journal of Intelligent Manufacturing*, 3(5), 317-323 (1992).

[34] Z. Tang, C. D.Almeida and P. A.Fishwick. Time series forecasting using neural networks vs. Box-Jenkins methodology. *Simulation*, 57(5), 303-310 (1991).

[35] Z. Tang and P. A.Fishwick. Feedforward neural nets as models for time series forecasting. *ORSA journal on computing*, 5(4), 374-385 (1993).

[36] H. Dai, and C. MacBeth. Effects of learning parameters on learning procedure and performance of a BPNN. *Neural networks*, 10(8), 1505-1521.(1997).

[37] G. D.Magoulas, M. N.Vrahatis and G. S.Androulakis. Effective backpropagation training with variable stepsize. *Neural networks*, 10(1), 69-82 (1997).

[38] M. Riedmiller and I. Rprop. Rprop-description and implementation details. (1994).

[39] F.Günther and S. Fritsch. neuralnet: Training of neural networks. *The R journal*, 2(1), 30-38 (2010).

[40] H. R.Maier and G. C.Dandy. Understanding the behaviour and optimising the performance of back-propagation neural networks: an empirical study. Environmental Modelling & Software, 13(2), 179-191. . (1998).

[41] J. Moody and N. Yarvin. Networks with learned unit response functions. *In Advances in neural information processing systems* pp. 1048-1055. (1992).

[42] B. L.Kalman and S. C.Kwasny. Why tanh: choosing a sigmoidal function. In [Proceedings 1992] IJCNN International Joint Conference on Neural Networks (Vol. 4, pp. 578-581). IEEE. (1992, June).

[43] Kaastra and M. S.Boyd. Forecasting futures trading volume using neural networks. Journal of Futures Markets, 15(8), 953-970. (1995).

[44] N. Karunanithi, W. J.Grenney, D. Whitley and K. Bovee. Neural networks for river flow prediction. *Journal of computing in civil engineering,* 8(2), 201-220. (1994).

[45] B. Karlik and A. V.Olgac. 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. (2011).

[46] K. Tsagkaris, A. Katidiotis, and P. Demestichas. Neural network-based learning schemes for cognitive radio systems. *Computer Communications*, 31(14), 3394-3404 (2008).

[47] H. R.Maier and G. C.Dandy. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental modeling & software*, 15(1), 101-124. (2000).

[48] W. C.Dawson and R. L.Wilby. Hydrological modeling using artificial neural networks. *Progress in physical Geography*, 25(1), 80-108 (2001).

[49] T. Masters. Advanced algorithms for neural networks: a C++ sourcebook. John Wiley & Sons, Inc. (1995).

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

[51] A. Khotanzad, R. Afkhami-Rohani, T. L.Lu, A. Abaye, M. Davis and D. J.Maratukulam. ANNSTLF-a neural-network-based electric load forecasting system. *IEEE Transactions on Neural networks*, 8(4), 835-846. (1997).

[52] M. E.Keskin and Ö. Terzi. Artificial neural network models of daily pan evaporation. *Journal of Hydrologic Engineering*, 11(1), 65-70 (2006).

[53] R. J.Schalkoff. Artificial neural networks. McGraw-Hill Higher Education. (1997).

[54] Luqman Hakim, Muhammad Ihsan Zul, Implementation of Discrete Wavelet Transform on Movement Images and Recognition by Artificial Neural Network Algorithm, WSEAS Transactions on Signal Processing, pp. 149-154, Volume 15, 2019

[55] Sherif G. Ahmad, Mohamed A. El-Gohary, Mohamed S. Elksas, Fayez G. Areed, Three Link Rigid Manipulator Control using Improved Neural Network based PID Controller

International Journal of Neural Networks and Advanced Applications, pp.60-68, Volume 6, 2019, E-ISSN: 2313-0563
[56] Abdulgani Albagul, Hafed Efheij, Bileid Abdulsalam, Comparison of Artificial Neural Network Controller and PID Controller in on Line of Real Time Industrial Temperature Process Control System, International Journal of Neural Networks and Advanced Applications, pp.69-74, Volume 6, 2019, E-ISSN: 2313-0563

Author Contributions:
SIDRA GHAYAS worked on the data analysis and programming in r studio.
Junaid sagheer siddiquie helped in screening the article
Asif mansoor helps in data collection.
Suboohi safdar helped in writing this article.
.